École des Ponts
ParisTech

THÈSE DE DOCTORAT
de l'École des Ponts ParisTech

Modèles d'épidémies en dimension infinie et optimisation des stratégies de vaccination

École doctorale N°532, Mathématiques et STIC (MSTIC)

Spécialité : Mathématiques

Thèse préparée au sein du :
CENTRE D'ENSEIGNEMENT ET DE RECHERCHE EN MATHEMATIQUES ET CALCUL SCIENTIFIQUE

Thèse soutenue le 26 novembre 2021, par :
Dylan DRONNIER

Composition du jury:

Mme. Marianne AKIAN INRIA, Centre de Recherche Saclay	*Examinatrice*
Mme. Isabelle CHALENDAR Université Gustave Eiffel, LAMA	*Examinatrice*
M. Jean-François DELMAS École des Ponts ParisTech, CERMICS	*Directeur de Thèse*
M. Jean-Stéphane DHERSIN Université Sorbonne Paris Nord, LAGA	*Président du jury*
M. Jean DOLBEAULT CNRS, CEREMADE	*Rapporteur*
M. Remco VAN DER HOFSTAD Eindhoven University of Technology, Eurandom	*Rapporteur*
Mme. Elisabeta VERGU INRAE, Unité MaIAGE	*Examinatrice*
M. Pierre-André ZITT Université Gustave Eiffel, LAMA	*Co-directeur de thèse*

Remerciements

La thèse n'est pas un exercice solitaire. C'est avant tout un travail d'équipe, des rencontres et tout ceci a été possible uniquement grâce au soutien de mes proches. À travers ces quelques lignes, j'aimerais rendre hommage à toutes les personnes qui m'ont tant apporté durant ces trois années.

Je voudrais commencer bien entendu par mes directeurs de thèse. Je remercie Jean-François Delmas pour ses qualités scientifiques et sa bienveillance. Toujours passionné, il a été un formidable interlocuteur pour discuter de la pertinence des modèles et de la minimalité des hypothèses choisies. Il m'a aussi grandement aidé à améliorer ma rédaction qui reste encore malgré tout beaucoup trop « sibylline ». Je tiens également à exprimer ma gratitude envers Pierre-André Zitt. Grâce à lui, beaucoup de démonstrations du manuscrit ont pu être simplifiées ou clarifiées. Enfin, je veux saluer l'approche qu'ont eu Jean-François et Pierre-André dans l'encadrement de ma thèse. Ils m'ont laissé la liberté d'explorer des sujets très divers même s'ils étaient éloignés de la théorie des probabilités et du sujet initial qu'ils m'avaient proposé en 2018. Mais cela m'a permis de m'épanouir et d'assouvir ma curiosité. Je leur dis une nouvelle fois merci !

Jean Dolbeault et Remco van der Hofstad m'ont fait l'honneur d'être mes rapporteurs de thèse. Je les remercie pour leur lecture minutieuse du manuscrit et pour leurs commentaires pertinents sur celui-ci. Je suis également très reconnaissant envers les membres du jury qui ont accepté de participer à ma soutenance : Marianne Akian, Isabelle Chalendar et Elisabeta Vergu. Je remercie tout particulièrement Jean-Stéphane Dhersin qui a été d'une grande aide au début de la thèse en apportant son point de vue aiguisé sur les différents modèles d'épidémies que j'ai étudiés.

J'aimerais aussi exprimer ma gratitude envers Michel Benaïm avec qui j'ai pu avoir des échanges très intéressants lors de ma visite en Suisse. Je le remercie pour son accueil à Neuchâtel et j'ai hâte de pouvoir commencer à travailler avec lui.

Je tiens à remercier l'ensemble du personnel du CERMICS. En particulier, je voudrais remercier Virginie Ehrlacher et Julien Reygner qui ont encadré mon stage de césure en 2017 et qui m'ont guidé dans mon orientation au cours de mes années aux Ponts en tant qu'élève ingénieur. Bien sûr, j'ai pu compter sur l'aide précieuse de Stéphanie Bonnel et Isabelle Simunic sur le plan administratif. Je leur suis très reconnaissant d'avoir été toujours réactives quand quelque chose n'allait pas et de m'avoir parfois remis dans le droit chemin quand il le fallait.

Je tiens également à remercier tous les doctorants et postdocs du CERMICS. Si ces trois années se sont si bien déroulées, c'est en grande partie grâce à eux. Merci donc aux anciens : Adrien, Benoît, Clément, Ezechiel, Inass, Mouad, Olga, Oumaima, Raed, Robert, Sébastien, Sofiane, Thomas et William. Un grand merci également aux nouvelles générations qui ont ramené la joie dans le labo après les confinements : Alfred, Clément, Coco, Cyrille, Éloïse, Emanuele, Gaspard, Guillaume, Hervé, Jean, Julien, Laurent, Louis, Maël, Nerea, Rémi, Rutger et Urbain. Mention spéciale aux « schaneurs » Edoardo, Michel et Roberta pour tous les bons moments passés au labo, dans les restaurants et bars parisiens et bientôt à la montagne. J'ai aussi une pensée pour mes collègues de l'autre côté de l'avenue Blaise Pascal, Elias et Quentin. Enfin je voudrais remercier Hélène pour ses relectures d'intro, son coaching MT180 mais surtout pour tout ce que nous avons partagé cette année.

Je salue aussi les « petits cons » de prépa Amnay et Joachim. C'est en grande partie grâce à eux si je me suis passionné pour les mathématiques. J'aimerais aussi remercier mes amis des Ponts : Amaury, Charlie, Eïda, Guillaume, Julien, Thomas et Youcef. Bien sûr, je n'oublie pas mon amie stéphanoise Mathilde que j'ai entraîné dans les situations les plus périlleuses mais qui continue d'être ma compagne d'aventures.

Enfin, je remercie chaleureusement tous les membres de ma famille. Mes parents bien sûr. Ma soeur Dolly et mon frère Kyeran. Ils ont fait ce que je suis aujourd'hui. Je tiens également à remercier ma grand-mère Bianca, Aline et Gérard pour leur soutien infaillible.

Résumé

Cette thèse est motivée par la modélisation mathématique de l'hétérogénéité des contacts dans les populations humaines et son impact sur la dynamique et le contrôle d'une maladie transmissible.

La première partie de la thèse porte sur l'étude d'un modèle SIS (*Susceptible/Infected/Susceptible*) déterministe en dimension infinie qui prend en compte l'hétérogénéité des contacts dans une population de grande taille. Grâce aux propriétés monotones que vérifient les solutions de ces équations différentielles, nous prouvons un résultat qui, comme pour les modèles en dimension finie, donne le comportement en temps long de la proportion d'infectés. En effet, le nombre de reproduction de base $\mathfrak{R}_0$, défini comme le rayon spectral d'un opérateur à noyau, détermine s'il existe un équilibre endémique stable ($\mathfrak{R}_0 > 1$) ou si toutes les solutions convergent vers l'état d'équilibre sans individus infectés ($\mathfrak{R}_0 \leq 1$).

Nous formalisons et étudions ensuite le problème de distribution optimale d'un vaccin qui immunise complètement les individus qui le reçoivent. Quand on suppose que les contacts sont homogènes et que $\mathfrak{R}_0 > 1$, il suffit de vacciner une proportion $1 - 1/\mathfrak{R}_0$ de la population atteindre l'immunité de groupe et éradiquer la maladie selon le théorème du seuil. Dans les modèles hétérogènes, ce théorème reste vrai mais avec une meilleure répartition des doses, on peut espérer atteindre l'immunité de groupe à moindre coût. Ainsi, nous étudions le problème où l'on cherche à minimiser à la fois le coût de la vaccination et une fonction perte qui peut être soit le nombre de reproduction effectif, soit la proportion totale d'infectés dans l'état endémique. En prouvant la continuité de ces deux fonctions pertes par rapport à une certaine topologie bien choisie, nous obtenons l'existence de stratégies Pareto optimales. Nous montrons également que si le nombre de reproduction de base est strictement supérieur à 1, alors la stratégie qui consiste à vacciner selon le profil des susceptibles dans l'état endémique est critique au sens où elle conduit à un nombre de reproduction effectif égal à 1.

Enfin, nous étudions les propriété du nombre de reproduction effectif et le problème de minimisation bi-objectif associé. Nous démontrons une généralisation de la conjecture de Hill-Longini sur la concavité et la convexité du nombre de reproduction effectif ainsi que d'autres résultats théoriques sur cette fonction perte. Ces derniers seront ensuite illustrés par de nombreux exemples. En particulier, les trois questions suivantes nous guideront notre analyse.

- Est-il possible de toujours vacciner optimalement quand les doses de vaccins ne sont disponibles qu'au fur et à mesure ?
- Quel est l'effet de l'assortativité (propension des individus à créer des liens avec des individus aux caractéristiques communes) sur le profils des vaccination optimale ?
- Que se passe-t-il quand tous les individus de la population ont le même nombre de contacts ?

Mots-Clés : Équation différentielle en dimension infinie, Équilibre endémique, Modèle SIS, Nombre de reproduction, Opérateur à noyau, Problème d'optimisation bi-objectif, Problème d'optimisation sous contraintes, Rayon spectral, Stabilité d'un équilibre, Stratégie de vaccination.

Abstract

This thesis is motivated by the mathematical modelling of heterogeneity in human contacts and the consequences on the dynamic and the control of contagious diseases.

In the first part of the thesis, we introduce and study an infinite-dimensional deterministic SIS (Susceptible/Infected/Susceptible) model which takes into account the heterogeneity of contacts within a large population. Thanks to the monotonic properties of the flow of these equations, we prove a result on the long-time behavior of the proportion of infected people. The basic reproduction number $\mathfrak{R}_0$, defined as the spectral radius of a kernel operator, determines whether there exists a stable endemic equilibrium ($\mathfrak{R}_0 > 1$) or if all the solutions tends to the disease-free equilibrium ($\mathfrak{R}_0 \leq 1$).

As an application, we formalize and study the problem of optimal allocation strategies for a vaccine that completely immunize from the disease those who received it. When we suppose that the contacts in the population are homogeneous, the threshold theorem states that the incidence of the infection will decrease if the proportion of vaccinated persons in the population is at least equal to $1 - 1/\mathfrak{R}_0$. In inhomogeneous models, this theorem remains true but with a better allocation of vaccine doses, we can hope for reaching herd immunity at lower cost. Hence, we study the problem where one tries to minimize simultaneously the cost of the vaccination, and a loss that may be either the effective reproduction number, or the overall proportion of infected individuals in the endemic state. By proving the continuity of these two loss functions, we obtain the existence of Pareto optimal strategies. We also show that vaccinating according to the profile of the endemic state is a critical allocation, in the sense that, if the initial reproduction number is larger than 1, then this vaccination strategy yields an effective reproduction number equal to 1.

The last part of the thesis is a detailed study of the effective reproduction number and the bi-objective minimization problem associated. We prove a generalization of the Hill-Longini conjecture on the concavity and convexity of the effective reproduction number along with other theoretical results on this loss function. We then illustrate with multiple examples those properties. In particular, we investigate the three following questions.

- Is it possible to always vaccinate optimally when the vaccine doses are given one at a time?
- What is the effect of assortativity (the tendency to have more contacts with similar individuals) on the shape of optimal vaccination strategies?
- What happens when every individuals have the same number of neighbors?

Keywords: Bi-objective optimization problem, Constrained optimization problem, Endemic equilibrium, Equilibrium stability, Infinite-dimensional differential equation, Kernel operator, SIS model, Reprocuction number, Spectral radius, Vaccination strategy.

Contents

List of Figures

Chapitre I

Introduction

Contenu du chapitre :

I.1 Bref historique de la modélisation mathématique des épidémies

Dans cette première section, nous présentons une série de modèles fondateurs en épidémiologie qui illustrent comment la compréhension que nous avons de la dynamique des maladies transmissibles a évolué au cours de l'histoire. Nous nous intéresserons en particulier à la modélisation des effets de la vaccination.

I.1.1 Les origines : Bernoulli et l'étude de la variolisation

L'histoire de la modélisation des maladies transmissibles[1] remonte aux travaux de Daniel Bernoulli qui montre mathématiquement en 1760 les bénéfices de l'inoculation malgré les risques liés à cette pratique. Revenons rapidement ici sur le modèle de Bernoulli et le contexte dans lequel a il a été publié. Nous commençons par une présentation rapide de la variole. Pour plus de détails, nous renvoyons le lecteur à la monographie *Smallpox and its Eradication* [59] qui retrace l'histoire de cette maladie et en donne la description clinique.

La variole, aussi appelée petite vérole, est une maladie virale fortement contagieuse vraisemblablement apparue au IV^e millénaire avant Jésus Christ en Inde ou en Égypte. Elle se transmet principalement par voie respiratoire rapprochée (postillons, aérosols). La variole se caractérise par des éruptions cutanées, l'apparition de pustules sur le corps et une forte fièvre. Au cours de l'histoire, de nombreuses vagues épidémiques ont touché durement la plupart des civilisations humaines jusqu'à l'éradication de la maladie en 1980 grâce à la vaccination. Au XVIII^e siècle, on estime à plus de 400 000 le nombre de victimes de la variole chaque année en Europe.

En mars 1718, la femme de l'ambassadeur d'Angleterre en Turquie, Lady Mary Mortley Montagu, inquiète des ravages de la variole qui a déjà emporté son frère, décide de faire inoculer son fils à Constantinople. Cette technique importée de Chine et relativement répandue dans l'empire Ottoman consiste à introduire dans les plaies un morceau de coton imbibé de pus variolique prélevé sur des malades. Suite au succès de l'opération, elle fait varioliser sa fille 3 ans plus tard alors qu'elle est de retour à la cour royale d'Angleterre. La technique va peu à peu se populariser en Angleterre puis dans le reste de l'Europe malgré les vives controverses qu'elle suscite.

En France, la variolisation est introduite par le docteur Théodore Tronchin qui inocule son fils et les enfants du Duc d'Orléans en 1756. Charles Marie de la Condamine, célèbre savant français, fera campagne pour son usage généralisé dans plusieurs *memoranda*. Cependant de nombreux médecins et scientifiques de l'époque s'opposent à cette pratique en raison de ses risques. Ses détracteurs l'accusent aussi de provoquer des épidémies, les personnes inoculées étant contagieuses.

Poussé par son ami Pierre Louis Moreau de Maupertuis, Daniel Bernoulli, neveu de Jacques Bernoulli et professeur de physique à l'université de Bâle, s'intéresse à ces controverses et tente d'y apporter une réponse mathématique. En 1760, il envoie à l'Académie des Sciences de Paris un manuscrit dont l'objectif est de calculer le gain d'espérance de vie en cas de variolisation de toute la population. Le 16 avril, son manuscrit est présenté en lecture publique et ne sera publié qu'en 1766 [16]. Nous reprenons ci-dessous le modèle de Bernoulli revisité par Dietz et Heesterbeek et présenté sous un formalisme moderne [44].

La population est structurée par âge et est divisée en deux catégories : les individus susceptibles et les individus infectés ou inoculés qui sont désormais immunisés. Le taux de mortalité des individus d'âge a hors variole est noté $\nu(a)$. La force d'infection $\lambda(a)$ (en anglais *infection force*) est définie comme le taux de contamination des susceptibles et dépend de leur âge. Enfin, on note $c(a)$ la probabilité de mourir en attrapant la variole à l'âge a. Ainsi, un individu d'âge a guérira de la variole et en sera immunisé avec une probabilité $1 - c(a)$. La probabilité $u(a)$ pour un nouveau-né

[1] On parle plus communément de modélisation mathématique des maladies infectieuses mais cette terminologie est imprécise. En effet, certaines maladies comme le tétanos sont bien infectieuses mais ne peuvent pas être transmises.

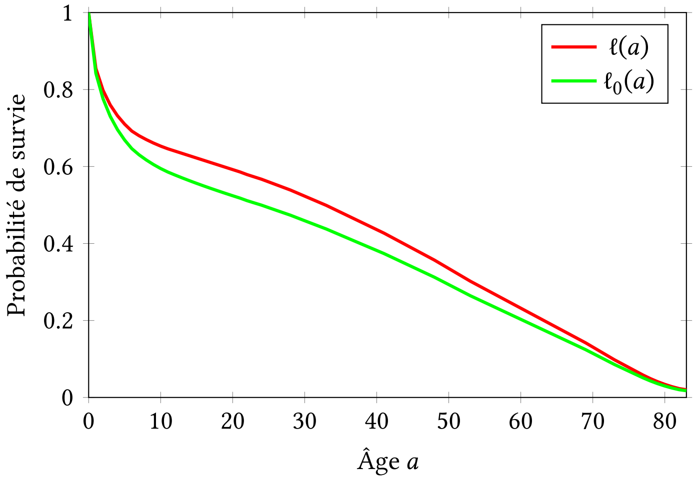


FIGURE I.1 : Evolution de la probabilité de survie des individus avec et sans la variole.

de vivre jusqu'à l'âge a sans avoir attrapé la variole vérifie l'équation différentielle :

$$\frac{\mathrm{d}u}{\mathrm{d}a} = -(\lambda(a) + \nu(a))u, \tag{I.1}$$

avec pour condition initiale : $u(0) = 1$. La probabilité $w(a)$ pour un nouveau né de vivre jusqu'à l'âge a en ayant attrapé la variole est donnée par la dynamique :

$$\frac{\mathrm{d}w}{\mathrm{d}a} = (1 - c(a))\lambda(a)u(a) - \nu(a)w. \tag{I.2}$$

À la naissance, le nouveau né n'est pas contaminé et donc $w(0) = 0$. Notons Λ et M les primitives de λ et ν respectivement s'annulant en 0. En intégrant ces équation linéaires, nous obtenons que la probabilité de survie jusqu'à l'âge a, $\ell(a) = u(a) + w(a)$, s'exprime ainsi :

$$\ell(a) = \ell_0(a)\left(\mathrm{e}^{-\Lambda(a)} + \int_0^a (1 - c(t))\lambda(t)\mathrm{e}^{-\Lambda(t)}\mathrm{d}t\right), \tag{I.3}$$

où et ℓ_0 est la probabilité de survie d'un individu jusqu'à l'âge a en l'absence de variole :

$$\ell_0(a) = \mathrm{e}^{-M(a)}. \tag{I.4}$$

À partir de ce modèle, Bernoulli veut calculer le gain d'espérance de vie obtenu en cas d'éradication de la variole (par l'inoculation de toute la population). Pour cela, il s'appuie sur les travaux d'Edmond Halley. En analysant les données démographiques de la ville de Breslau de 1687 à 1691, cet astronome anglais a estimé la quantité $\ell(a)$, pour des âges a compris entre 0 et 83 ans [73]. N'ayant aucune donnée sur l'incidence de la variole et son taux de létalité, Bernoulli suppose simplement que λ et c sont constants et égaux à $1/8$. En utilisant l'Équation (I.3), il compare dans un tableau les valeurs de ℓ_0 et de ℓ pour des valeurs de a comprises entre 0 et 25 ans. Nous avons tracé les courbes correspondantes dans la Figure I.1 que nous avons étendues à toute la plage d'âge donnée dans la table de Halley. Finalement, Bernoulli estime, par intégration, que l'espérance de vie progresserait de 3 ans et 2 mois en cas d'éradication de la variole.

Au-delà de leur intérêt historique, les travaux de Bernoulli nous montrent que, dès le départ, la modélisation mathématique des maladies transmissibles a été utilisée pour essayer d'apporter une justification rigoureuse à l'application d'une politique sanitaire. Avec la pandémie de Covid-19 qui sévit aujourd'hui, il est désormais acquis que la modélisation mathématique tient une place très importante dans la protection des populations contre les maladies contagieuses.

I.1.2 Les travaux fondateurs de Kermack et McKendrick

Le modèle de Bernoulli est révolutionnaire pour l'époque mais il suppose que la force d'infection λ ne dépend que de l'âge de l'individu ce qui ne permet pas de modéliser un certain nombre de phénomènes qui apparaissent dans la dynamique des maladies transmissibles (vagues épidémiques, croissance exponentielle du nombre de cas à l'apparition d'une nouvelle infection). Au début du XX$^{\text{e}}$ siècle, le médecin et épidémiologiste écossais McKendrick s'inspire de la loi d'action de masse qui stipule que deux réactifs chimiques réagissent à une vitesse proportionnelle à leur concentration. En supposant que les individus d'une population sont comme les particules d'un gaz qui se déplacent aléatoirement dans l'espace et que les collisions entre les individus sains et les individus infectés aboutissent à des infections, il propose de considérer une force d'infection de la forme $\lambda = KI$, où K est une constante appelée le taux de transmission de la maladie et I est la proportion d'individus infectés. Nous renvoyons à [78] pour une chronologie détaillée de l'introduction du principe d'action de masse en épidémiologie.

Dans leur article de 1927 [95], McKendrick et le biochimiste écossais Kermack proposent de diviser la population en classes épidémiologiques et de décrire leurs évolutions grâce à des équations différentielles – la force d'infection étant donnée par le principe d'action de masse. Cette approche fut ensuite étendue dans deux autres articles scientifiques quelques années plus tard [96, 97]. Dans leur premier article, Kermack et McKendrick présentent un système d'équations différentielles qui est maintenant bien connu sous le nom de modèle SIR (de l'anglais *Susceptibles/Infected/Recovered*). Ils supposent que la population est divisée en trois compartiments : les personnes susceptibles (ou sensibles), les personnes infectées et les personnes guéries. La proportion d'individus de la population dans chacune de ces trois classes est notée S, I et R respectivement. L'évolution de ces quantités au cours du temps est alors donnée par les équations :

$$\begin{cases} \frac{\mathrm{d}S}{\mathrm{d}t} = -KSI, \\ \frac{\mathrm{d}I}{\mathrm{d}t} = KSI - \gamma I, \\ \frac{\mathrm{d}R}{\mathrm{d}t} = \gamma I, \end{cases} \tag{I.5}$$

En effet, rappelons que la force d'infection – définie comme le taux de contamination des susceptibles – est donnée à chaque instant par $\lambda(t) = KI(t)$ selon la loi d'action de masse. Ainsi, la dérivée par rapport au temps du nombre de nouvelles infections divisée par le nombre total d'individus dans la population au temps t est égale à $\lambda(t)S(t) = KS(t)I(t)$. On fait également l'hypothèse que les individus infectés guérissent à vitesse γ et qu'une fois guéris, ils sont définitivement immunisés contre la maladie.

Il apparaît que la quantité $S + I + R$ est conservée au cours du temps et égale à $S(0) + I(0) + R(0) = 1$. Cela est cohérent avec le modèle qui suppose que chaque individu peut être classé dans exactement un compartiment épidémiologique à tout instant t. On ne peut pas résoudre analytiquement le système (I.5). On peut néanmoins montrer que le système converge en temps long vers un équilibre $(S_\infty, 0, R_\infty)$ qui est l'unique solution de l'équation implicite :

$$S_\infty = 1 - R_\infty = S(0)\mathrm{e}^{-\mathfrak{R}_0(R_\infty - R(0))}, \tag{I.6}$$

où $\mathfrak{R}_0$, le nombre de reproduction de base, est défini par la formule :

$$\mathfrak{R}_0 = \frac{K}{\gamma}\cdot \tag{I.7}$$

Il correspond au nombre moyen de cas directement générés par un individu infecté au cours de sa maladie dans une population où toutes les autres personnes sont sensibles à l'infection. Si $\mathfrak{R}_0 \leq 1$, le nombre d'infectés dans la population décroît pour toute condition initiale $(S(0), I(0), R(0)) \in [0,1]^3$ telle que $S(0) + I(0) + R(0) = 1$. Si $\mathfrak{R}_0 > 1$ et la quantité de susceptible est suffisamment grande (*i.e.*, $S(0) \geq 1 - 1/\mathfrak{R}_0$), alors le nombre d'infectés croît au début de l'épidémie. Pour beaucoup d'autres modèles comme le modèle SIS (*Susceptible/Infected/Susceptible*) défini ci-dessous, la valeur que prend le nombre de reproduction de base a une importance encore plus grande car elle détermine si l'infection est endémique ou si elle régresse jusqu'à son éradication totale.

Certaines maladies ne confèrent pas une protection aux personnes ayant déjà attrapé la maladie. Si l'on suppose que les individus infectés ne sont pas immunisés une fois qu'ils sont guéris mais qu'ils redeviennent susceptibles, alors il ne reste plus que deux compartiments dans la population et on obtient le modèle SIS :

$$\begin{cases} \frac{\mathrm{d}S}{\mathrm{d}t} = -KSI + \gamma I, \\ \frac{\mathrm{d}I}{\mathrm{d}t} = KSI - \gamma I. \end{cases}$$

Une nouvelle fois, la quantité $I + S$ est conservée au cours du temps et égale à 1 et donc on peut réécrire la dynamique de l'épidémie en donnant seulement l'évolution de la proportion de personnes infectées :

$$\frac{\mathrm{d}I}{\mathrm{d}t} = (1 - I)KI - \gamma I. \tag{I.8}$$

Pour ce modèle, il est possible de résoudre ces équations analytiquement. Si $\mathfrak{R}_0$ est différent de 1, alors on a la formule :

$$I(t) = \frac{\mathfrak{R}_0 - 1}{\mathfrak{R}_0 + ((1 - \mathfrak{R}_0)/I(0) - \mathfrak{R}_0)\mathrm{e}^{(1-\mathfrak{R}_0)\,t/\gamma}}.$$

Si $\mathfrak{R}_0$ est égale à 1, alors la proportion d'individus infectés dans la population au temps t est donnée par l'équation :

$$I(t) = \frac{1}{(1/I(0)) + t/\gamma}.$$

Ainsi, le comportement en temps long de la dynamique de l'épidémie est déterminé par la valeur que prend le nombre de reproduction.

Régime sous-critique Si $\mathfrak{R}_0 < 1$, $I(t)$ converge à une vitesse exponentielle vers le seul équilibre $I = 0$, qu'on appelle état d'équilibre sans infection.

Régime critique Si $\mathfrak{R}_0 = 1$, alors $I(t)$ converge encore vers $I = 0$ mais pas exponentiellement vite. L'état d'équilibre sans infection est encore l'unique équilibre.

Régime sur-critique Si $\mathfrak{R}_0 > 1$, alors le point fixe $I = 0$ devient instable et un autre équilibre $G = 1 - 1/\mathfrak{R}_0$ apparaît. Il est dit endémique car il correspond à la situation où la maladie devient persistante dans la population. L'état d'équilibre endémique est globalement stable au sens où pour toute condition initiale $I(0) > 0$, $I(t)$ converge vers G.

Supposons que $\mathfrak{R}_0 > 1$ et que l'on dispose d'un vaccin qui immunise les individus contre l'infection à $t = 0$. Il y a désormais une nouvelle classe épidémiologique dans notre population : les personnes vaccinées dont la proportion – qui reste constante au cours du temps – est notée V. Le taux de contamination des susceptibles est toujours égal à KI mais, maintenant, $S(t) = 1 - I(t) - V$. Ainsi, la nouvelle dynamique de la proportion d'individus infectés est alors donnée par :

$$\frac{\mathrm{d}I}{\mathrm{d}t} = (1 - V - I)KI - \gamma I. \tag{I.9}$$

En divisant l'équation précédente par $1 - V$, on retombe sur l'Équation (I.8) où K est remplacé par $K(1 - V)$:

$$\frac{\mathrm{d}J}{\mathrm{d}t} = (1 - J)K(1 - V)J - \gamma J, \tag{I.10}$$

le nombre $J = I/(1 - V)$ correspondant à la proportion d'infectés dans la population non-vaccinée. On en déduit la formule du nouveau nombre de reproduction $\mathfrak{R}_e$ que l'on qualifie d'effectif :

$$\mathfrak{R}_e = \frac{K(1 - V)}{\gamma}. \tag{I.11}$$

Maladie	Transmission	$\mathfrak{R}_0$	Seuil d'immunité de groupe
Rougeole	Aérosols	12-18	92-95%
Coqueluche	Aérosols	12-17	92-94%
Varicelle	Aérosols	10-12	90-92%
Diphtérie	Gouttelettes respiratoires	6-7	83-86%
Rubéole	Gouttelettes respiratoires	6-7	83-86%
Poliomyélite	Voie fécale-orale	5-7	80-86%
Variole	Gouttelettes respiratoires	5-7	80-86%
Oreillons	Gouttelettes respiratoires	4-7	75-86%

TABLE I.1 : Nombres de reproduction estimés et seuils de l'immunité grégaire pour des maladies transmissibles pouvant être prévenues par vaccination [61].

Supposons que le nombre de reproduction de base $\mathfrak{R}_0$ est strictement supérieur à 1. Pour que le nombre de reproduction effectif soit inférieur à 1, c'est-à-dire pour pouvoir mettre fin à l'épidémie, il faut vacciner au moins une proportion $V \geq 1 - 1/\mathfrak{R}_0$ d'individus. Ce résultat constitue ce qu'on appelle parfois le théorème de seuil qui a été formulé pour la première fois par Smith en 1970 [141] et Dietz en 1975 [43]. Il s'agit d'une formulation mathématique du phénomène d'immunité de groupe (aussi appelé immunité grégaire) en épidémiologie. En vaccinant un individu, non seulement on le protège directement mais on protège également indirectement les personnes en contact avec lui car il ne pourra plus leur transmettre la maladie. Ainsi, on peut espérer éradiquer une maladie sans pour autant avoir à immuniser toute la population.

Le nombre $1 - 1/\mathfrak{R}_0$ est appelé seuil d'immunité grégaire. L'objectif général des politiques de santé publique est d'établir l'immunité de groupe au sein des populations en atteignant ce seuil. C'est ce qui a été accompli au cours du XX$^{\text{e}}$ siècle, avec la vaccination de masse contre la variole [59]. Le seuil d'immunité grégaire (situé aux environs de 80-85 %) a été ainsi atteint dans la plupart des pays du monde, permettant d'éliminer cette maladie. Dans la Table I.1, nous donnons les nombres de reproduction de base et les seuils d'immunité de groupe associés pour des maladies pouvant être prévenues par vaccination [61]. Comme de nombreux facteurs peuvent faire fluctuer $\mathfrak{R}_0$ (par exemple, la saison, la densité de population, etc), ces valeurs ne doivent pas être considérées comme des constantes précises propres à chaque maladie mais plutôt comme des grandeurs approximatives qui varient en fonction de l'environnement. Ajoutons à cela qu'en pratique, aucun vaccin n'offre une protection complète contre un agent infectieux. Ainsi, pour calculer la proportion de personnes à vacciner, il faut aussi estimer l'efficacité du vaccin en réalisant des études sérologiques et cliniques. Notons qu'à part la poliomyélite qui est quasiment éradiquée mondialement et la variole qui a totalement disparu, toutes les autres maladies de la Table I.1 sont encore endémiques dans beaucoup de pays, y compris des pays développés.

I.1.3 L'hétérogénéité des contacts

Dans la loi d'action de masse, il est supposé que chacune des espèces réactives est répartie de manière homogène dans le soluté. En épidémiologie, cette hypothèse ne semble pas refléter la complexité des chaînes de contamination. Selon le mode de transmission de la maladie, les collisions que l'on évoquait plus haut entre les individus sains et les individus infectés sont de nature différentes : relations sexuelles pour les infections sexuellement transmissibles, contacts rapprochés pour les maladies se transmettant par voie aérienne, etc. Dans tous les cas, ces contacts ne sont pas homogènes. Les chercheurs ont d'abord essayé de modéliser l'hétérogénéité spatiale [11, 94, 118] ou par âge [86] des populations. Ici, nous nous intéressons plutôt à une approche qui consiste à diviser la population en groupes d'individus ayant des propriétés similaires. Ainsi, on peut considérer toute sorte de groupes (sociaux, âge, spatiaux) en fonction de ce qui est adapté pour la maladie considérée. Pour les épidémies de type SIS, ce modèle dit de métapopulation a été proposé pour la première fois par Lajmanovich et Yorke en 1976 [102] pour étudier la propagation de la gonorrhée. On suppose que la population est divisée en $N \geq 2$ groupes et l'on note $\gamma_1, \gamma_2, \ldots, \gamma_N$

les taux de guérison des individus les composant. Le taux de transmission du groupe j au groupe i est noté $K_{i,j}$ pour $1 \leq i, j \leq N$. La dynamique de la proportion d'individus infectés dans le groupe $i \in \{1, \dots, N\}$ est alors donnée par l'équation :

$$\frac{\mathrm{d}I_i}{\mathrm{d}t} = (1 - I_i) \sum_{j=0}^{N-1} K_{i,j} I_j - \gamma_i I_i. \tag{I.12}$$

Ainsi, le modèle de Lajmanovich et Yorke suppose que le groupe i possède son propre taux de guérison γ_i et sa propre force d'infection λ_i où λ_i est égal à une combinaison linéaire des proportions d'infectés dans chaque groupe. Il est relativement courant de décomposer la matrice des taux de transmission de la manière suivante $K = \mathrm{Diag}(\beta) \times M \times \mathrm{Diag}(\theta)$, où β et θ sont des vecteurs donnant respectivement la susceptibilité et l'infectiosité des individus des différentes sous-populations et, pour $i, j \in \{1, 2, \dots, N\}$, le coefficient $M_{i,j}$ est égal à la fréquence de contacts d'un individu du groupe j avec les individus du groupe i.

Le système (I.12) ne peut pas être résolu analytiquement. Lajmanovich et Yorke arrivent cependant à montrer que le comportement en temps long des solutions dépend essentiellement d'un unique paramètre : le nombre de reproduction de base $\mathfrak{R}_0$ défini comme la plus grande valeur propre de la matrice $K/\gamma = (K_{i,j}/\gamma_j)_{0 \leq i,j \leq N-1}$ dans ce modèle hétérogène :

$$\mathfrak{R}_0 = \rho(K/\gamma), \tag{I.13}$$

où l'on note $\rho(A)$ le rayon spectral d'une matrice A, c'est-à-dire le plus grand module de ces valeurs propres. Rappelons que, d'après le théorème de Perron-Froebenius, $\mathfrak{R}_0$ est en fait une valeur propre de K/γ et qu'il existe un vecteur propre à droite G^0 associé à $\mathfrak{R}_0$ dont tous les coefficients sont positifs. Comme dans le modèle homogène, le vecteur sans individu infecté est un équilibre du système et il y a deux régimes possibles pour le comportement en temps long des solutions du système différentiel.

1. Si $\mathfrak{R}_0 \leq 1$, alors $\lim_{t \to \infty} I(t) = (0, 0, \dots, 0)$.
2. Si $\mathfrak{R}_0 > 1$, alors il existe au moins un équilibre G endémique, c'est-à-dire, tel que $G \neq (0, 0, \dots, 0)$. Si on suppose aussi que K est irréductible, alors cet équilibre endémique est unique et tous ses coefficients sont strictement positifs : $G \in]0, 1]^N$ et pour toute condition initiale $I(0) \neq (0, 0, \dots, 0)$, on a $\lim_{t \to \infty} I(t) = G$.

Dans une population hétérogène, l'interprétation biologique du nombre de reproduction de base est plus subtile que dans le modèle homogène : il correspond au nombre moyen de personnes qu'infecte un individu « typique » au début de l'épidémie [42]. Cela signifie que l'individu malade à l'instant initial est tiré au hasard selon la loi donnée par G^0, le vecteur propre de Perron à droite de la matrice K/γ dont la somme des coefficients vaut 1. En effet, en approximant le début de l'épidémie par un processus de branchement multitype, on peut démontrer, sous certaines conditions sur la matrice K, qu'une proportion G_i^0 des premiers individus infectés appartient au groupe étiqueté $i \in \{1, 2, \dots, N\}$ [10].

Supposons, comme pour le modèle SIS homogène, que l'on dispose d'un vaccin qui immunise parfaitement ceux qui le reçoivent. Comme pour les autres classes épidémiologiques, il faut spécifier la proportion totale d'individus vaccinés dans chaque groupe. Notons donc V_0, V_1, ..., V_{N-1} la proportion de personnes vaccinées dans chaque groupe. Le vecteur des forces d'infection reste le même mais la proportion de susceptibles devient $S_i = (1 - V_i - I_i)$. Ainsi, la nouvelle dynamique s'écrit :

$$\frac{\mathrm{d}I_i}{\mathrm{d}t} = (1 - V_i - I_i) \sum_{j=1}^{N} K_{i,j} I_j - \gamma_i I_i. \tag{I.14}$$

Pour les calculs, il est en fait plus pratique de donner la proportion de personnes non-vaccinées dans chaque groupe, notée $\eta_i = 1 - V_i$. En divisant l'Équation (I.14) par η_i, on obtient que la proportion $J_i = I_i/\eta_i$ d'individus infectés dans la population non-vaccinée évolue selon :

$$\frac{\mathrm{d}J_i}{\mathrm{d}t} = (1 - J_i) \sum_{j=1}^{N} K_{i,j} \eta_j J_j - \gamma_i J_i. \tag{I.15}$$

Ainsi, on retrouve la dynamique SIS sans vaccination dans la population effective avec un nombre de reproduction :

$$\mathfrak{R}_e(\eta) = \rho(K\eta/\gamma), \tag{I.16}$$

où l'on note $K\eta/\gamma$ la matrice $K \times \mathrm{Diag}(\eta) \times \mathrm{Diag}(1/\gamma)$.

Notons $\mu_1, \mu_2, \ldots, \mu_N$ la taille relative des différentes sous-populations. Le nombre total de doses de vaccin administrées est proportionnel à la proportion totale d'individus immunisés :

$$C(\eta) = \sum_{i=1}^{N} (1 - \eta_i)\mu_i = 1 - \sum_{i=1}^{N} \eta_i \mu_i. \tag{I.17}$$

Il est alors naturel de chercher, étant donnée une quantité de vaccin, une distribution des doses qui « freine » le plus l'épidémie ou l'éradique au plus vite. Pour évaluer l'efficacité d'une stratégie, la littérature en modélisation mathématiques des maladies transmissibles utilise la plupart du temps le nombre de reproduction effectif comme critère et ce, pour plusieurs raisons dont nous donnons une liste non exhaustive :

- $\mathfrak{R}_e$ est défini pour une large classe de modèles,
- $\mathfrak{R}_e$ détermine le comportement en temps long de l'épidémie pour beaucoup de modèles,
- bien qu'on ne puisse pas en donner une expression explicite en fonction de η dans la plupart des cas, il est possible d'étudier la fonction $\mathfrak{R}_e$ en utilisant les propriétés bien connues du rayon spectral.

Il est alors naturel de chercher une stratégie qui donne lieu à un nombre de reproduction inférieur ou égal à $r \in [0, \mathfrak{R}_0]$ et qui minimise le nombre de doses à administrer :

$$\begin{cases} \textbf{Minimiser :} & C(\eta) \\ \textbf{Avec :} & \mathfrak{R}_e(\eta) \leq r \end{cases} \tag{I.18}$$

En particulier, si $\mathfrak{R}_0 > 1$, le problème (I.18) avec $r = 1$ correspond à chercher la stratégie de coût minimal qui permet d'atteindre l'immunité grégaire. Remarquons qu'en prenant $\eta = (1/\mathfrak{R}_0, 1/\mathfrak{R}_0, \ldots, 1/\mathfrak{R}_0)$ on obtient, par homogénéité du rayon spectral, $\mathfrak{R}_e(\eta) = 1$ pour une proportion $C(\eta) = 1 - 1/\mathfrak{R}_0$ de la population vaccinée. Ainsi, le théorème de seuil reste également vrai dans le modèle hétérogène. La question sous-jacente est donc la suivante : peut-on abaisser le coût pour atteindre l'immunité de groupe en profitant de l'inhomogénéité de la population ?

Il est également intéressant d'étudier le problème inverse de minimisation du nombre de reproduction effectif quand la société ne dispose que d'une quantité limitée de doses de vaccins :

$$\begin{cases} \textbf{Minimiser :} & \mathfrak{R}_e(\eta) \\ \textbf{Avec :} & C(\eta) \leq c \end{cases} \tag{I.19}$$

Il existe une littérature abondante traitant du Problème (I.18) et/ou du Problème (I.19), voir par exemple [29, 46, 52, 81, 116, 130, 161]. Nous en discuterons en détail dans les chapitres 3, 4 et 5 de la thèse.

I.2 Les réseaux sociaux comme limites de grands graphes

L'objectif de cette section est de motiver de manière informelle le modèle déterministe SIS que l'on étudie dans cette thèse. Pour cela, dans la section I.2.1, nous définissons le processus SIS stochastique qui permet de décrire à l'échelle locale l'évolution de l'épidémie. Nous exposons ensuite les résultats de convergence connus pour ces processus quand la taille de la population tend vers l'infini. Dans la section I.2.2, nous présentons rapidement les objets limites que considère la littérature qui étudie la convergence des grands graphes. Ces objets limites sont à la base de notre modèle SIS en dimension infinie.

I.2.1 La modélisation stochastique des épidémies

Dans les modèles décrits dans la section précédente, on a supposé qu'il y avait suffisamment d'individus dans la population (ou dans les sous-populations) pour appliquer le principe d'action de masse. Pour une description locale des épidémies, on utilise généralement des processus stochastiques. Dans cette section, nous présentons le modèle SIS markovien et en donnons quelques propriétés.

Soit $n \geq 2$ la taille de la population. Soient $A^n = (a^n_{i,j})_{1\leq i,j\leq n}$ une matrice positive de taille $n \times n$ et b^n un vecteur de taille n dont les coefficients sont strictement positifs. Pour $1 \leq i, j \leq n$, $a^n_{i,j}$ représente l'intensité des contacts susceptibles d'aboutir à une transmission de l'infection de l'individu j à l'individu i. La quantité $1/b^n_i$ est égale au temps moyen de guérison de l'individu i. Notons $X^n_i(t)$ la variable aléatoire égale à 1 si i est infecté au temps t et 0 sinon. L'évolution du processus se fait selon les règles suivantes :

- Guérison : si l'individu i est infecté alors il guérit à taux b^n_i.
- Contamination : si l'individu i est sain au temps t alors il devient infecté à taux :

$$\lambda^n_i(t) = \sum_{j=1}^{n} a^n_{i,j} X^n_j(t). \tag{I.20}$$

Le processus X^n est markovien et prend ses valeurs dans un espace d'états de taille 2^n. L'état $(0, 0, \dots, 0)$ où tous les individus sont sains, est absorbant.

En partant de ce modèle SIS stochastique local, on peut retrouver les dynamiques déterministes de la section précédente.

On considère une population de plus en plus grande en faisant tendre n vers l'infini. Pour tout $n \geq 2$, supposons que tous les coefficients $a^n_{i,j}$ sont égaux à une constante K/n et que les b^n_i sont égaux à une constante $\gamma > 0$. Notons $Y^n(t) = \sum_i X^n_i(t)/n$ la proportion d'individus infectés. Si $Y^n(0)$ converge en probabilité vers $y \in [0, 1]$, alors, pour tout $t \in \mathbb{R}_+$ et tout $\delta > 0$, on a :

$$\mathbb{P}\left(\sup_{s\leq t} | Y^n(s) - I(s) | \geq \delta\right) = 0, \tag{I.21}$$

où $I : \mathbb{R}_+ \to [0, 1]$ est l'unique solution de l'équation différentielle (I.8) vérifiant la condition initiale $I(0) = y$.

Supposons maintenant que les coefficients de la matrice de transmission et du vecteur des taux de guérison sont périodiques de période $N \geq 2$:

$$a_{i,j} = K_{i^*,j^*}/n, \quad b^n_i = \gamma_{i^*}, \qquad \text{où} \quad i = i^* \mod N, \; j = j^* \mod N. \tag{I.22}$$

Définissons pour tout $i \in \{1, 2, \dots, N\}$, la proportion d'individu dans la sous-population i :

$$Y^n_i(t) = \frac{N}{n} \sum_m X^n_{mN+i}(t). \tag{I.23}$$

En utilisant les résultats de Kurtz [101], on peut montrer la convergence du processus stochastique vers la solution des équations (I.12) de Lajmanovich et Yorke à N groupes. Plus précisément, si pour tout $i \in \{1, 2, \dots, N\}$, $Y^n_i(0)$ converge vers en probabilité vers $y_i \in [0, 1]$, alors, pour tout $t \in \mathbb{R}_+$ et tout $\delta > 0$, on a :

$$\mathbb{P}\left(\sup_{s\leq t} \| Y^n(s) - I(s) \| \geq \delta\right) = 0, \tag{I.24}$$

où $I : \mathbb{R}_+ \to [0, 1]^N$ est l'unique solution de l'équation différentielle (I.12) vérifiant la condition initiale $I(0) = y$. Le cas $N = 1$ est rigoureusement traité dans la monographie d'Andersson et Britton [4] et la généralisation de ce résultat pour un nombre fini de groupes est évoquée dans [13].

Ces résultats de convergence permettent de justifier les Équations (I.8) et (I.12). Néanmoins, les coefficients de la matrice A^n et du vecteur b^n avait une forme bien particulière. On a en fait classé les individus par groupe. Le nombre N de groupe était fixé et le nombre d'individus dans chaque groupe tendait vers l'infini. Cela nous limite dans la modélisation. Par exemple, si l'on veut décrire une population formée de ménages, il faudrait faire tendre le nombre de groupes N vers l'infini et les tailles des groupes resteraient constante.

Dans la deuxième partie de la section, nous utilisons la théorie des graphons pour tenter de répondre de manière non-rigoureuse aux questions suivantes : que se passe-t-il dans le cas général ? Peut-on écrire une équation limite vers lequel converge le processus stochastique ?

I.2.2 La convergence des graphes denses

Comprendre les grands réseaux est un problème fondamental en théorie moderne des graphes. Au cours des 15 dernières années, la théorie des graphons développée entre autres par Borgs, Chayes, Lovász, Janson et Szegedy [18, 20, 90, 111, 112] a connu un important succès. Elle décrit les limites de graphes denses, c'est-à-dire les graphes dont le nombre d'arêtes est proportionnel au carré du nombre de nœuds. Nous donnons dans cette section quelques éléments de cette théorie qui permettront de comprendre notre modèle SIS en dimension infinie.

Un graphon est une fonction mesurable $k : [0,1]\times[0,1] \to \mathbb{R}_+$ bornée et symétrique : $k(x,y) = k(y,x)$, pour tout $x, y \in [0,1]$. Étant donné deux graphons k_1 et k_2, on peut définir la norme de coupe de $k_1 - k_2$:

$$\| k_1 - k_2 \|_{\square} = \sup_{A,\, B\subset[0,1]} \left| \int_{A\times B} (k_1(x,y) - k_2(x,y))\, \mathrm{d}x\mathrm{d}y \right|. \tag{I.25}$$

Il est facile de voir que cette norme est plus faible que les normes L^p. Soit une matrice symétrique $A = (a_{i,j})_{1\le i,j\le n}$ aux coefficients positifs. On peut lui associer le graphon suivant :

$$k[A] = \sum_{1\le i,j\le n} a_{i,j} \mathbb{1}_{[(i-1)/n, i/n[\times[(j-1)/n, j/n[}. \tag{I.26}$$

Tout comme les matrices $n \times n$ agissent sur les vecteurs, on peut faire agir les graphons sur les fonction $f : [0,1] \to \mathbb{R}$ par l'opérateur intégral :

$$T_k(f)(x) = \int_0^1 k(x,y) f(y)\, \mathrm{d}y. \tag{I.27}$$

Reprenons notre suite de matrices $A^n = (a^n_{i,j})_{1\le i,j\le n}$ de taille $n\times n$ et aux coefficients positifs de la section précédente. Supposons que les coefficients des matrices sont bornés par une constante M (indépendante de n) et que les matrices sont symétriques : $a_{i,j} = a_{j,i}$ pour tout $1 \le i < j \le n$. D'après la théorie des graphons [111], quitte à re-numéroter les individus, on peut trouver une sous-suite de matrices A^n qui converge pour la norme de coupe vers un graphon. Plus précisément, il existe :

- une fonction $\psi : \mathbb{N} \to \mathbb{N}$ strictement croissante,
- pour toute $n \in \mathbb{N}$, une permutation σ_n de $\{1, 2, \dots, \psi(n)\}$,
- un graphon $k : [0,1]\times[0,1] \to [0, M]$,

tels que :

$$\lim_{n\to\infty} \left\| k - k\left[A^{\psi(n)}_{\sigma_n}\right] \right\|_{\square} = 0, \tag{I.28}$$

où $A^{\psi(n)}_{\sigma_n}$ est la matrice $A^{\psi(n)}$ dont on a permuté les coefficients selon σ_n. Réciproquement, il est facile de montrer que tout graphon s'exprime comme la limite d'une suite de matrices. La convergence d'une suite de matrice en norme de coupe est illustrée dans la Figure I.2.

Si l'on suppose que les coefficients des matrices sont donnés par (I.22), alors il est facile de montrer que la suite $k[A^n]$ converge en norme de coupe (à permutation près) vers le noyau $k[K]$.

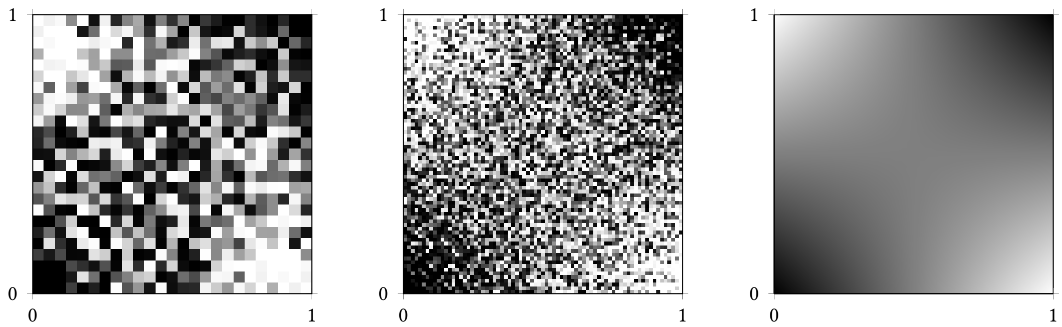


FIGURE I.2 : Illustration de la convergence d'une suite de matrices vers le graphon $k(x, y) = 1 + (2x - 1)(2y - 1)$ en norme de coupe. Les niveaux de gris correspondent aux valeurs prises par les noyaux $k[A_i]$ et k.

Ainsi, la théorie des graphons donne un sens à la convergence d'une suite de matrices de tailles différentes et la généralise.

Comme dans le cas où le nombre de groupe est fini, le système de particules converge, on est amené à penser que cela reste encore vrai quand les matrices de transmissions convergent en norme de coupe vers un graphon k. Par analogie avec l'équation (I.12), il est naturel de considérer une équation de « champ moyen » du type :

$$\partial_t u(t, x) = (1 - u(t, x)) \int_0^1 k(x, y) u(t, y)\, \mathrm{d}y - \gamma(x) u(t, x), \tag{I.29}$$

où :

- les individus de la population sont étiquetés par $[0, 1]$ au lieu de $1, 2, \dots, N$,
- $u(t, x)$ représente la probabilité qu'un individu étiqueté par x soit infecté au temps t,
- k est le graphon qui représente les contacts entre les individus et qui joue le même rôle que la matrice K,
- γ donne le taux de guérison en fonction de l'étiquette de l'individu.

Récemment, plusieurs articles généralisent la théorie des graphons et définissent les limites de graphes non denses, possiblement orientés [12, 19, 21, 100]. Ainsi, dans cette thèse, nous considérons des objets un peu plus généraux que des graphons pour décrire l'hétérogénéité des contacts dans les populations.

I.3 Contribution de la thèse

Cette section rassemble les résultats principaux obtenus pendant la thèse.

I.3.1 Un modèle SIS en dimension infinie

Inspirés par la théorie des graphons, nous définissons dans le chapitre II un nouveau modèle SIS en dimension infinie.

Remarquons en premier lieu que les graphons ne sont pas obligatoirement définis sur le carré $[0, 1] \times [0, 1]$ munie de la mesure de Lebesgue. Ils peuvent être définis sur un espace de probabilité général $(\Omega, \mathscr{F}, \mu)$. Cet espace représente les étiquettes/caractères des individus de la population. Pour un ensemble d'étiquettes $A \in \mathscr{F}$, le nombre $\mu(A)$ correspond à la proportion d'individus de la population dont l'étiquette appartient à A. Soient $\gamma : \Omega \to \mathbb{R}_+$ et $k : \Omega \times \Omega \to \mathbb{R}_+$. Pour $x, y \in \Omega$, $\gamma(x)$ correspond au taux de guérison d'un individu étiqueté par $x \in \Omega$ et $k(x, y)$ donne le taux de transmission de la maladie d'un infecté étiqueté y vers les individus étiquetés x. En généralisant l'équation (I.29), nous supposons que l'évolution de la probabilité $u(t,x)$ qu'un individu de type x soit infecté au temps t suit la dynamique :

$$\partial_t u(t,x) = (1-u(t,x)) \int_\Omega k(x,y)u(t,y)\,\mu(\mathrm{d}y) - \gamma(x)u(t,x). \tag{I.30}$$

En particulier, le théorème 1 assure que le nombre $u(t,x)$ appartient à $[0,1]$ pour tout $t \in \mathbb{R}_+$ et $x \in \Omega$ et donc peut bien être interprété comme une probabilité d'être infecté.

Notons qu'en prenant $\Omega = \{1,2,\dots,N\}$ et $k(x,y) = K_{x,y}/\mu(\{y\})$ pour tout $x, y \in \Omega$, on retrouve le système (I.12) (la fonction γ sur un espace discret pouvant être interprété comme un vecteur). Ce modèle est donc une généralisation des équations différentielles proposées par Lajmanovich et Yorke.

Notons $\mathbf{k}(x,y) = k(x,y)/\gamma(y)$, pour tout $x, y \in \Omega$.

Hypothèse 1. *Il existe $q \in]1,+\infty[$ tel que :*

$$\sup_{x\in\Omega} \int_\Omega \mathbf{k}(x,y)^q\,\mu(\mathrm{d}y) < \infty. \tag{I.31}$$

Cette hypothèse servira à montrer l'existence de solutions de l'équation différentielle (I.30). Même si dans les chapitres III, IV et V cette condition d'intégrabilité est allégée, nous supposerons, sauf mention contraire, que, dans cette introduction, tous les noyaux $\mathbf{k}$ vérifient cette condition pour simplifier la présentation des résultats.

Notons $p = q/(q-1)$. L'hypothèse 1 assure également la compacité de l'opérateur à noyau défini sur $L^p(\Omega,\mu)$:

$$T_\mathbf{k}(f)(x) = \int_\Omega \mathbf{k}(x,y)f(y)\,\mu(\mathrm{d}y), \qquad f \in L^p(\Omega,\mu). \tag{I.32}$$

Cet opérateur étant également positif, on en déduit, grâce au théorème de Krein-Rutman, que le rayon spectral est une valeur propre. Ce rayon spectral définit le nombre de reproduction de base du modèle de manière analogue à la dimension finie :

$$\mathfrak{R}_0 = \rho(T_\mathbf{k}). \tag{I.33}$$

Pour énoncer le résultat principal, nous avons besoin d'une condition qui assure qu'il n'existe pas un groupe d'individus isolés qui ne peuvent pas se faire contaminer quand le reste de la population est infecté. Il s'agit de l'analogue de l'hypothèse d'irréductibilité de la matrice K qui garantit l'unicité de l'équilibre endémique dans le théorème de Lajmanovich et Yorke (voir le point 2 page 7). Ainsi, on dit que le noyau $\mathbf{k}$ est irréductible si pour tout $A \in \mathscr{F}$:

$$\int_{A\times A^c} \mathbf{k}(x,y)\,\mu(\mathrm{d}y)\mu(\mathrm{d}x) = 0 \implies \mu(A) \in \{0,1\}. \tag{I.34}$$

Nous montrons alors dans le chapitre II le théorème suivant.

Théorème 1. *Supposons que l'hypothèse 1 est satisfaite. Nous avons les propriétés suivantes :*

(i) ***Existence de solution globale :*** *il existe une unique solution u à l'équation* (I.30). *Cette solution vérifie pour tout $(x,t) \in \Omega \times \mathbb{R}_+$, $u(t,x) \in [0,1]$.*

(ii) ***Équilibre sans infectés stable dans les régimes critique et sous-critique :*** *Si $\mathfrak{R}_0 \le 1$, alors la maladie régresse jusqu'à être éradiquée :*

$$\lim_{t\to\infty} u(t,x) = 0, \qquad x \in \Omega.$$

(iii) ***Équilibre endémique stable dans le régime sur-critique :*** *Si $\mathfrak{R}_0 > 1$, alors il existe un équilibre $\mathfrak{g} : \Omega \to [0,1]$, appelé équilibre endémique, dont l'intégrale sur Ω est strictement positive et qui domine tout autre équilibre. Si nous supposons en plus que le noyau k est irréductible, alors les seuls équilibres du système sont $\mathbb{0}$ et $\mathfrak{g}$. En outre, pour toute condition initiale u_0 telle que :*

$$\int_\Omega u_0(x)\,\mu(\mathrm{d}x) > 0,$$

la solution u de l'équation (I.30) *converge simplement vers l'équilibre endémique :*

$$\lim_{t\to\infty} u(t,x) = \mathfrak{g}(x), \quad x \in \Omega.$$

Si $u_0 = \mathbb{0}$ μ-p.s., alors u converge simplement vers $\mathbb{0}$.

Pour démontrer ce théorème, nous suivons l'approche proposée par Hirsch et Smith pour traiter le système de Lajmanovich et Yorke [82, 142]. Nous prouvons que le semi-flot défini par l'équation (I.30) est monotone. Nous obtenons alors la stabilité des équilibres en montrant les deux propriétés ci-dessous.

- En partant de l'état où tout le monde est infecté, le système converge vers 0 dans les régimes critique et sous-critique.
- Il existe une famille de conditions initiales arbitrairement proche de $\mathbb{0}$ (dans un sens à définir), pour lesquelles le semi-flot converge vers $\mathfrak{g}$.

Dans le régime sur-critique, quand k n'est pas supposé irréductible, il peut y avoir des équilibres différents de $\mathfrak{g}$ et $\mathbb{0}$. Nous montrons dans le chapitre III que tous les équilibres différents de $\mathfrak{g}$ sont instables (voir le théorème 3). Si $\mathfrak{R}_0 \leq 1$, on pose $\mathfrak{g} = \mathbb{0}$. En effet, avec cette convention, $\mathfrak{g}$ reste le plus grand équilibre (car c'est le seul) dans les régimes critiques et sous-critiques et il a encore la propriété d'être stable.

Dans la seconde partie du chapitre II, nous appliquons notre modèle SIS en dimension infinie à l'étude des effets du confinement et de la vaccination sur la dynamique de l'épidémie et sur le nombre de reproduction. En particulier, donnons ici la forme de la dynamique quand on dispose d'un vaccin qui immunise parfaitement les personnes qui le reçoivent. Pour $x \in \Omega$, notons $\eta(x)$ la probabilité pour un individu x de ne pas être vacciné. L'évolution de la probabilité $u(t,x)$ qu'un individu de type x soit infecté au temps x suit alors la dynamique :

$$\partial_t u(t,x) = (\eta(x) - u(t,x)) \int_\Omega k(x,y)u(t,y)\,\mu(\mathrm{d}y) - \gamma(x)u(t,x). \tag{I.35}$$

Comme pour le cas fini-dimensionnel, on peut diviser l'équation précédente par η pour retrouver l'équation (I.30) avec un nouveau noyau de transmission :

$$\partial_t \tilde{u}(t,x) = (1 - \tilde{u}(t,x)) \int_\Omega k(x,y)\eta(y)\tilde{u}(t,y)\,\mu(\mathrm{d}y) - \gamma(x)\tilde{u}(t,x), \tag{I.36}$$

où $\tilde{u}(t,x)$ est la probabilité pour un individu non-vacciné en x d'être infecté au temps t. En particulier, le nombre de reproduction effectif est défini par la formule :

$$\mathfrak{R}_e(\eta) = \rho(T_{\mathrm{k}\eta}), \tag{I.37}$$

où $\mathrm{k}\eta$ est le noyau qui à $(x,y) \in \Omega \times \Omega$ associe $\mathrm{k}(x,y)\eta(y)$. Remarquons que, comme dans le cas fini-dimensionnel, le théorème de seuil reste vrai par homogénéité du rayon spectral.

On définit également $\mathfrak{g}[\eta]$ comme l'équilibre maximal associé à la dynamique (I.36) du nombre d'infectés chez les non-vaccinés.

I.3.2 Stratégies de vaccination optimales

Le chapitre III traite du problème de vaccination optimale pour le modèle SIS en dimension infinie du chapitre précédent. Rappelons que l'on suppose que l'hypothèse 1 est satisfaite pour que la dynamique (I.30) fasse sens.

Formalisation du problème de vaccination optimale

On considère l'ensemble des stratégies de vaccination :

$$\Delta = \{\eta : \Omega \to [0,1] \quad \text{mesurable}\}. \tag{I.38}$$

Rappelons que pour η dans Δ, $\eta(x)$ correspond à la proportion d'individus de type x non-vaccinés. Ainsi, la proportion $C(\eta)$ de personnes vaccinées est donnée par :

$$C(\eta) = \int_\Omega (1 - \eta(x))\, \mu(\mathrm{d}x). \tag{I.39}$$

Cette fonction peut également être interprétée comme le coût de la vaccination. La plupart des résultats énoncés ici ne reposent en réalité que sur la continuité de C par rapport à la topologie faible (voir la fin de la section) et sur sa décroissance stricte (*i.e.*, si $\eta_2 - \eta_1$ est une fonction positive qui n'est pas égale à 0 presque sûrement, alors $C(\eta_1) > C(\eta_2)$). Ainsi on pourrait considérer des fonctions coût plus générales vérifiant ces deux propriétés.

La société cherche alors à minimiser ce coût et une autre fonction « perte » qui mesure la gravité de l'épidémie et que nous notons L (pour *loss function*). Nous formalisons cela grâce au problème bi-objectif :

$$\min_{\eta \in \Delta} (C(\eta), \mathrm{L}(\eta)). \tag{I.40}$$

Dans ce chapitre, nous considérons deux fonctions pertes L :

- le nombre de reproduction effectif $\mathfrak{R}_e$,
- la proportion totale $\mathfrak{I}$ d'infectés dans l'état endémique $\mathfrak{g}[\eta]$ maximal de (I.36) :

$$\mathfrak{I}(\eta) = \int_\Omega \mathfrak{g}[\eta](x)\eta(x)\, \mu(\mathrm{d}x). \tag{I.41}$$

En général le profil η ne résulte pas d'une décision prise par la société mais plutôt de l'ensemble des volontés des individus de la population. Il peut alors être intéressant d'avoir une estimation de la « pire » distribution des doses de vaccin que l'on peut faire. Ainsi nous considérons également le problème opposé :

$$\max_{\eta \in \Delta} (C(\eta), \mathrm{L}(\eta)). \tag{I.42}$$

Il n'est généralement pas possible de minimiser (ou maximiser) à la fois la fonction coût et la fonction perte. En optimisation bi-objectif, on cherche donc les stratégies dont on ne peut pas améliorer un objectif sans dégrader l'autre.

Definition I.3.1. *Une stratégie $\eta_\star \in \Delta$ est dite Pareto optimale pour le Problème bi-objectif* (I.40) *si pour tout $\eta \in \Delta$:*

$$C(\eta) < C(\eta_\star) \implies \mathrm{L}(\eta) > \mathrm{L}(\eta_\star) \quad \text{et} \quad \mathrm{L}(\eta) < \mathrm{L}(\eta_\star) \implies C(\eta) > C(\eta_\star).$$

Une stratégie $\eta^\star \in \Delta$ est dite anti-Pareto optimale pour le Problème bi-objectif (I.42) *si pour tout $\eta \in \Delta$:*

$$C(\eta) > C(\eta^\star) \implies \mathrm{L}(\eta) < \mathrm{L}(\eta^\star) \quad \text{et} \quad \mathrm{L}(\eta) > \mathrm{L}(\eta^\star) \implies C(\eta) < C(\eta^\star).$$

Les frontières de Pareto et d'anti-Pareto correspondent à l'image des ensembles des stratégies Pareto et anti-Pareto optimales par les fonctions objectives :

$$\mathcal{F}_\mathrm{L} = \{(C(\eta), \mathrm{L}(\eta) : \eta \text{ Pareto optimal}\}, \tag{I.43}$$

$$\mathcal{F}_\mathrm{L}^{\mathrm{Anti}} = \{(C(\eta), \mathrm{L}(\eta) : \eta \text{ anti-Pareto optimal}\}. \tag{I.44}$$

Remarquons qu'une stratégie Pareto optimale $\eta_\star$ est solution des problèmes sous contraintes suivants :

$$\begin{cases} \textbf{Minimiser :} & C(\eta) \\ \textbf{Avec :} & \mathrm{L}(\eta) \leq \ell \end{cases} \tag{I.45}$$

et :

$$\begin{cases} \textbf{Minimiser :} & \mathrm{L}(\eta) \\ \textbf{Avec :} & C(\eta) \leq c \end{cases} \tag{I.46}$$

où $\ell = \mathrm{L}(\eta_\star)$ et $c = C(\eta_\star)$. Nous avions considérés initialement ces deux problèmes de minimisation sous contraintes dans la section I.1.3 avec $\mathrm{L} = \mathfrak{R}_e$. Nous préférons le formalisme bi-objectif car il fait jouer les mêmes rôles à la fonction perte et à la fonction coût.

De la même manière, si $\eta^\star$ est une stratégie anti-Pareto optimale, alors elle est solution des problèmes :

$$\begin{cases} \textbf{Maximiser :} & C(\eta) \\ \textbf{Avec :} & \mathrm{L}(\eta) \geq \ell \end{cases} \tag{I.47}$$

et :

$$\begin{cases} \textbf{Maximiser :} & \mathrm{L}(\eta) \\ \textbf{Avec :} & C(\eta) \geq c \end{cases} \tag{I.48}$$

où $\ell = \mathrm{L}(\eta^\star)$ et $c = C(\eta^\star)$.

Maintenant que l'on a défini les objets principaux nécessaires pour étudier l'optimalité des stratégies de vaccination, nous allons énoncer les résultats principaux du chapitre III.

Régularité des fonctions pertes

On munit l'espace Δ de la topologie faible : on dit qu'une suite $(\eta_n,\, n \in \mathbb{N})$ converge faiblement vers $\eta \in \Delta$ si pour toute fonction bornée $h \in L^\infty(\Omega, \mu)$:

$$\lim_{n\to\infty} \int_\Omega h(x)\eta_n(x)\, \mu(\mathrm{d}x) = \int_\Omega h(x)\eta(x)\, \mu(\mathrm{d}x). \tag{I.49}$$

Il est facile de voir que la fonction coût C est continue pour cette topologie. On obtient dans le chapitre III la continuité des fonctions pertes.

Théorème 2. *Les fonctions $\mathfrak{R}_e$ et $\mathfrak{I}$ sont continues par rapport à la topologie faible.*

La continuité de la fonction $\mathfrak{R}_e$ ne découle pas directement des résultats connus en analyse fonctionnelle sur la convergence des spectres des opérateurs compacts. En effet, la convergence faible de η_n vers η n'implique pas la convergence de $T_{\mathrm{k}\eta_n}$ vers $T_{\mathrm{k}\eta}$ en norme d'opérateur mais seulement la convergence forte, c'est-à-dire :

$$\lim_{n\to\infty} \left\| T_{\mathrm{k}\eta_n}(g) - T_{\mathrm{k}\eta}(g) \right\|_p = 0, \qquad g \in L^p(\Omega, \mu), \tag{I.50}$$

où l'on rappelle que $p = q/(q-1)$ où $q \in]1, +\infty[$ est défini dans l'hypothèse 1. En général, la convergence forte n'implique pas la convergence du spectre ou du rayon spectral, même pour des opérateurs compacts. Néanmoins la famille $(T_{\mathrm{k}\eta_n},\, n \in \mathbb{N})$ vérifie une propriété plus forte. Elle est collectivement compacte, c'est-à-dire que l'ensemble suivant est relativement compact pour la topologie induite par la norme $\|\cdot\|_p$:

$$\bigcup_{n\in\mathbb{N}} T_{\mathrm{k}\eta}(B),$$

où B est la boule unité de $L^p(\Omega, \mu)$. Cette propriété permet d'obtenir la convergence des spectres grâce à un résultat d'Anselone [6].

Criticalité de la stratégie consistant à vacciner selon le profil endémique maximal

La continuité de la fonctionnelle $\mathfrak{I}$ découle du résultat ci-dessous.

Théorème 3. *Supposons que $\mathfrak{R}_0 > 1$. La stratégie qui consiste à vacciner selon le profil $\mathfrak{g} = \mathfrak{g}[1]$ est critique, c'est-à-dire qu'on a $\mathfrak{R}_e(1 - \mathfrak{g}) = 1$. Réciproquement, si un équilibre h associé à la dynamique* (I.30) *vérifie $\mathfrak{R}_e(1 - h) = 1$, alors il est égal à $\mathfrak{g}$.*

En plus de son utilité pour démontrer la continuité de $\mathfrak{I}$, ce théorème nous donne une stratégie critique différente de la stratégie qui consiste à vacciner uniformément à un niveau $1 - 1/\mathfrak{R}_0$. Il est alors naturel de se demander laquelle de ces deux vaccinations est la moins coûteuse. Nous n'avons pas obtenu de résultats généraux mais dans un prochain article [40], nous allons exhiber une classe de noyau pour lesquels $C(1-\mathfrak{g}) \leq C(1/\mathfrak{R}_0)$.

Le théorème 3 est à mettre en parallèle avec la section 4.5 du livre [80] de Hethcote et Yorke qui propose de vacciner les individus après qu'ils ont été infectés par la gonorrhée.

Propriétés des problèmes bi-objectifs

Nous avons vu que les fonctions pertes considérées sont continues par rapport à la topologie faible. En outre, d'après le théorème de Banach-Alaoglu, l'espace Δ muni de la topologie faible est compact. Ainsi, son image par L est compact. Par connexité de Δ, il vient que $\mathrm{L}(\Delta)$ est un intervalle $[0, \ell_{\max}]$ où $\ell_{\max} \in \mathbb{R}_+$. On déduit également du théorème 2 que les Problèmes (I.45) et (I.46) admettent des solutions pour tout $c \in [0,1]$ et $\ell \in [0, \ell_{\max}]$. Nous notons les valeurs de ces problèmes :

$$\mathrm{L}_\star(c) = \min\{\mathrm{L}(\eta) \ :\ \eta \in \Delta,\, C(\eta) \leq c\}, \tag{I.51}$$

$$C_{\star,\mathrm{L}}(\ell) = \min\{C(\eta) \ :\ \eta \in \Delta,\, \mathrm{L}(\eta) \leq \ell\}. \tag{I.52}$$

Notons que d'après le théorème 1, nous avons $C_{\star,\mathfrak{I}}(0) = C_{\star,\mathfrak{R}}(1)$. En outre, grâce au théorème 3, nous obtenons que :

$$C_{\star,\mathfrak{I}}(0) = C_{\star,\mathfrak{R}}(1) \leq \min\left(1 - \frac{1}{\mathfrak{R}_0}, \int_\Omega \mathfrak{g}\,\mathrm{d}\mu\right).$$

Nous avons vu que les stratégies Pareto optimales sont solutions des problèmes (I.46) et (I.45) avec des contraintes bien choisies. Réciproquement, nous pouvons nous demander pour quelles contraintes les solutions de ces problèmes sont Pareto optimales. Le résultat suivant répond à cette question.

Théorème 4. *Pour* $\mathrm{L} \in \{\mathfrak{R}_e, \mathfrak{I}\}$*, la fonction* $C_{\star,\mathrm{L}}$ *est continue et strictement décroissante sur l'intervalle* $[0, \ell_{\max}]$*. La fonction* $\mathrm{L}_\star$ *est continue sur l'intervalle* $[0,1]$*, strictement décroissante sur* $[0, C_{\star,\mathrm{L}}(0)]$ *et constante égale à 0 sur* $[C_{\star,\mathrm{L}}(0), 1]$*. Enfin, la frontière de Pareto est donnée par :*

$$\mathcal{F}_\mathrm{L} = \{(c, \mathrm{L}_\star(c)) \ :\ c \in [0, C_{\star,\mathrm{L}}(0)]\} = \{(C_{\star,\mathrm{L}}(\ell), \ell) \ :\ \ell \in [0, \ell_{\max}]\}.$$

En particulier, on déduit de ce théorème que, pour tout $\ell \in [0, \ell_{\max}]$, il existe une stratégie Pareto optimale $\eta_\star$ telle que $\mathrm{L}(\eta_\star) = \ell$. Dans le chapitre III, nous prouvons également une propriété de stabilité de la fonction $\mathrm{L}_\star$ par rapport à des petites variations de k ou de la fonction γ. Cette propriété est importante car, même si nous n'avons qu'une estimation du noyau de transmission et des temps de guérison, il est quand même possible de donner des stratégies proches de l'optimalité.

La description de la frontière anti-Pareto est légèrement plus compliquée. Ici, nous ne donnons que les résultats pour la fonction perte $\mathrm{L} = \mathfrak{R}_e$ et en supposant une condition supplémentaire sur le noyau. Notons les valeurs des problèmes de maximisation sous contraintes (I.47) :

$$\mathfrak{R}_e^\star(c) = \max\{\mathrm{L}(\eta) \ :\ \eta \in \Delta,\, C(\eta) \geq c\}, \tag{I.53}$$

$$C^\star(r) = \min\{C(\eta) \ :\ \eta \in \Delta,\, \mathfrak{R}_e(\eta) \geq r\}. \tag{I.54}$$

Nous avons obtenu le résultat suivant qui donne la forme de la frontière d'anti-Pareto pour la fonction perte $\mathrm{L} = \mathfrak{R}_e$.

Théorème 5. *Supposons que* $\mathbf{k}$ *est irréductible (voir l'équation* (I.34)*). La fonction* $C^\star$ *est continue et strictement décroissante sur l'intervalle* $[0, \mathfrak{R}_0]$*. La fonction* $\mathfrak{R}_e^\star$ *est continue et strictement décroissante sur l'intervalle* $[0,1]$*. Enfin, la frontière d'anti-Pareto est donnée par :*

$$\mathcal{F}^{\mathrm{Anti}} = \{(c, \mathfrak{R}_e^\star(c)) \ :\ c \in [0,1]\} = \{(C^\star(r), r) \ :\ r \in [0, \mathfrak{R}_0]\}.$$

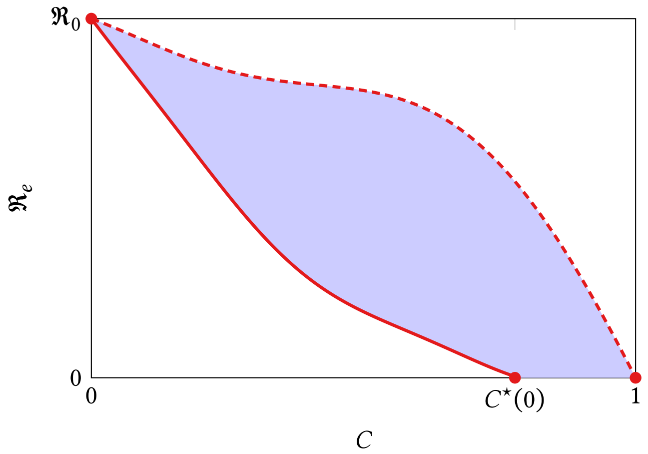


FIGURE I.3 : Allure typique de la frontière de Pareto et de la frontière d'anti-Pareto pour $\mathrm{L} = \mathfrak{R}_e$ et $\mathbf{k}$ est irréductible. Ligne rouge continue : la frontière de Pareto $\mathcal{F}$; Ligne en pointillé : l'anti-frontière de Pareto $\mathcal{F}^{\text{Anti}}$; zone bleutée : résultats réalisables.

L'expression de la frontière d'anti-Pareto pour la fonction de perte $\mathfrak{I}$ est plus compliquée car discontinue. Dans la figure I.3, nous avons représenté l'allure typique des frontières de Pareto et d'anti-Pareto quand $\mathrm{L} = \mathfrak{R}_e$ et le noyau $\mathbf{k}$ est irréductible.

Nous prouvons également dans le chapitre III que, pour $\mathrm{L} \in \{\mathfrak{R}_e, \mathfrak{I}\}$, l'ensemble des résultats réalisables (en bleu dans la figure I.3) :

$$\{(C(\eta), \mathrm{L}(\eta)) \ : \ \eta \in \Delta\},$$

est simplement connexe, c'est-à-dire, qu'il n'a pas de « trou ».

I.3.3 Propriétés du nombre de reproduction effectif

L'objectif du chapitre IV est d'étudier les propriétés de la fonctionnelle $\mathfrak{R}_e$. Cela nous permettra d'avoir une meilleure compréhension du problème de minimisation bi-objectif associé.

La conjecture de Hill-Longini

En optimisation, la convexité ou la concavité des fonctions objectives joue un rôle très important. Dans un article de 2003 [81], Hill et Longini formulent dans le cadre fini-dimensionnel une conjecture qui fournit une condition nécessaire pour que la fonction $\mathfrak{R}_e$ soit convexe et une autre condition qui implique la concavité de cette même fonction. Nous montrons que cette conjecture telle qu'elle est formulée dans l'article de 2003 est fausse. Néanmoins, en rajoutant une condition de symétrie, nous parvenons à prouver un résultat proche qui se généralise dans le cadre infini-dimensionnel.

Notons $\sigma(T_\mathbf{k})$ le spectre de l'opérateur $T_\mathbf{k}$. D'après le théorème de Krein-Rutman, $\mathfrak{R}_0 \in \sigma(T_\mathbf{k})$. Si on suppose que le noyau $\mathbf{k}$ est symétrique, c'est-à-dire $\mathbf{k}(x,y) = \mathbf{k}(y,x)$ $\mu \otimes \mu$-presque sûrement, alors l'opérateur $T_\mathbf{k}$ est auto-adjoint et donc son spectre appartient à la droite réelle : $\sigma(T_\mathbf{k}) \subset \mathbb{R}$. Dans le chapitre IV, nous obtenons le résultat suivant.

Théorème 6. *Supposons* $\mathbf{k}$ *symétrique et* $\int_{\Omega\times\Omega} \mathbf{k}(x,y)^2 \mu(\mathrm{d}x)\mu(\mathrm{d}y) < \infty$.

(i) Si $\sigma(T_\mathbf{k}) \subset \mathbb{R}_+$*, alors la fonction* $\mathfrak{R}_e$ *est convexe.*

(ii) Si $\sigma(T_\mathbf{k}) \backslash \{\mathfrak{R}_0\} \subset \mathbb{R}_-$ *et la valeur propre* $\mathfrak{R}_0$ *est simple, alors la fonction* $\mathfrak{R}_e$ *est concave.*

Le point (i) avait déjà été obtenu par Cairns dans le cadre fini-dimensionnel [29]. Pour le démontrer en dimension infinie, nous suivons la même stratégie qui consiste à exprimer $\mathfrak{R}_e(\eta)$ en fonction de la racine carrée de l'opérateur $T_\mathbf{k}$ grâce à la formule de Courant-Fisher. Sous cette forme, $\mathfrak{R}_e(\eta)$ est égal au supremum d'une famille de fonctions linéaires en η, d'où sa convexité.

Pour montrer le point (ii), on calcule la hessienne de $\mathfrak{R}_e$ grâce à la formule de Kloeckner [98].

La politique du cordon sanitaire n'est pas anti-Pareto optimal

Si on suppose que le noyau $\mathbf{k}$ est irréductible, on dit qu'une stratégie $\eta \in \Delta$ est un cordon sanitaire si η n'est pas presque-sûrement nulle et le noyau $\mathbf{k}$ restreint à l'ensemble $\{\eta > 0\}$ n'est pas irréductible. Un cordon sanitaire est donc une stratégie qui permet de diviser la population non-vaccinée en deux groupes A et B tels que les individus du groupe A n'infectent pas ceux du groupe B. Il n'existe pas toujours de cordon sanitaire (prendre par exemple $\mathbf{k} > 0$ $\mu \otimes \mu$-presque sûrement). En revanche, quand une telle stratégie existe, ce n'est jamais la pire.

Théorème 7. *Si la stratégie $\eta \in \Delta$ est un cordon sanitaire alors elle n'est pas anti-Pareto optimal.*

Une caractérisation du coût minimal pour arrêter toute contamination

Parmi les cordons sanitaires, il peut exister des stratégies qui stoppent complètement les transmissions. Elles correspondent aux stratégies pour lesquelles tous les individus non-vaccinés ne sont en contact qu'avec des personnes vaccinées et sont donc complètement protégés. Décrivons formellement ces stratégies.

En théorie des graphes, un ensemble indépendant – appelé aussi stable – est un ensemble de sommets deux à deux non adjacents. Le nombre d'indépendance d'un graphe G, noté $\alpha(G)$, est alors défini comme au rapport $\#A/\#G$ où A est un ensemble indépendant de taille maximale. Par analogie, nous appelons ensemble indépendant un ensemble mesurable A tel que $\mathbf{k}(x, y) = 0$ pour presque tout $x, y \in A$. Nous montrons dans le chapitre IV, que tout noyau admet un ensemble indépendant maximal, c'est-à-dire un ensemble A tel que pour tout autre ensemble indépendant $B \in \mathscr{F}$, $\mu(B) \leq \mu(A)$. On note alors $\alpha(\mathbf{k}) = \mu(A)$. Nous obtenons le résultat suivant.

Théorème 8. *Supposons que $\mathbf{k}$ est symétrique. Si A est un ensemble indépendant maximal alors $\mathbb{1}_A$ est Pareto optimal pour la fonction perte $\mathrm{L} = \mathfrak{R}_e$. En particulier, on a :*

$$C_\star(0) = 1 - \alpha(\mathbf{k}). \tag{I.55}$$

I.3.4 Quelques exemples de vaccinations optimales

Dans le chapitre V, nous présentons des exemples de noyaux où l'on peut résoudre analytiquement les problèmes (I.40) et (I.42) avec $\mathrm{L} = \mathfrak{R}_e$. En particulier, ces exemples permettent de construire des intuitions autour des trois questions suivantes.

- Est-il possible de toujours vacciner optimalement quand les doses de vaccins ne sont disponibles qu'au fur et à mesure, autrement dit, est-ce que l'algorithme glouton parcourt l'ensemble des solution Pareto optimales ?
- Quel est l'effet de l'assortativité (propension des individus à créer des liens avec des individus aux caractéristiques communes) sur les profils des vaccination optimale ?
- Que se passe-t-il quand tous les individus de la population ont le même nombre de contacts ?

Revenons rapidement sur la dernière question dans cette introduction. On dit qu'un noyau $\mathbf{k}$ est de degré constant si les fonctions ci-dessous sont presque-sûrement constante :

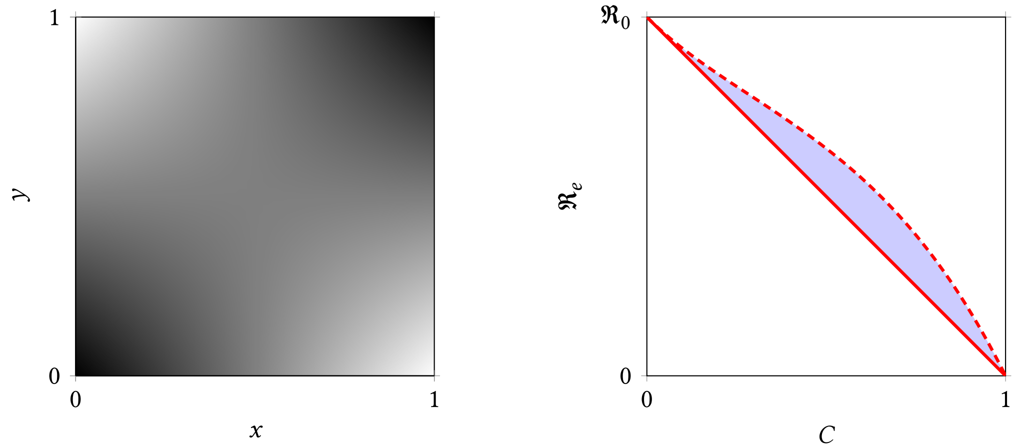
$$x \mapsto \int_\Omega \mathbf{k}(x, y)\, \mu(\mathrm{d}y) \qquad \text{et} \qquad x \mapsto \int_\Omega \mathbf{k}(y, x)\, \mu(\mathrm{d}y). \tag{I.56}$$

Dans le chapitre V, nous démontrons le résultat suivant.

Théorème 9. *Supposons que $\mathbf{k}$ est de degré constant.*

(i) Si $\mathfrak{R}_e$ est convexe, alors les stratégies uniformes sont Pareto optimales,

(ii) Si $\mathfrak{R}_e$ est concave, alors les stratégies uniformes sont anti-Pareto optimales.



(A) Représentation du noyau $\mathbf{k}(x, y) = 1 + (2x - 1)(2y - 1)$ défini sur $[0, 1]$ muni de la mesure de Lebesgue.

(B) Ligne rouge continue : la frontière de Pareto $\mathcal{F}$ correspondant aux stratégies de vaccination uniformes ; Ligne en pointillé : l'anti-frontière de Pareto $\mathcal{F}^{\text{Anti}}$; zone bleutée : résultats réalisables.

FIGURE I.4 : Exemple d'optimisation pour la perte $\mathrm{L} = \mathfrak{R}_e$.

Ainsi, malgré l'hétérogénéité de la population, pour certaines configurations, on ne peut pas faire mieux (ou pire) que la vaccination uniforme. La figure I.4 illustre cela avec un noyau $\mathbf{k}$ de degré constant pour lequel $\mathfrak{R}_e$ est convexe d'après le théorème 6.

Dans le chapitre V, on construit un noyau de degré constant pour lequel l'ensemble des stratégies Pareto optimales a une infinité de composantes connexes. En particulier, l'algorithme glouton ne parcourt pas cet ensemble.

Summary of the main assumptions on the kernels

Let $(\Omega, \mathscr{F}, \mu)$ be a probability space. A kernel is a measurable fonction $\mathbf{k} : \Omega \times \Omega \to \mathbb{R}_+$. In order to define the reproduction number associated to a kernel $\mathbf{k}$, we make the following assumption.

Assumption A. *There exists $p \in (1, +\infty)$ such that:*

$$\int_\Omega \left(\int_\Omega \mathbf{k}(x, y)^q \, \mu(\mathrm{d}y) \right)^{p/q} \mu(\mathrm{d}x) \quad \textit{with } q \textit{ given by} \quad \frac{1}{p} + \frac{1}{q} = 1.$$

A kernel satisfying Assumption A is said to have finite double-norm. Assumption A corresponds to Assumption III.1. In Chapter IV, we mainly consider kernels with finite double-norm. In Chapter V, all the kernels satisfy Assumption A with $p = q = 2$; see Equation (V.22). Assumption A ensures that the integral operator associated to the kernel $\mathbf{k}$ is bounded from L^p to L^p and compact. It is needed to apply the Krein-Rutman theorem which is an important tool used many times throughout the thesis.

In the dense case, the parameters of the SIS dynamic studied in Chapter II are the transmission kernel $k : \Omega\times\Omega \to \mathbb{R}_+$ and the recovery rate function $\gamma : \Omega \to \mathbb{R}_+$. In order to prove the existence and the uniqueness of the trajectory of the SIS differential equation along with their long-time behavior, we make the following assumption on the parameters.

Assumption B. *The function γ is bounded and positive and there exists $q \in (1, +\infty)$:*

$$\sup_{x\in\Omega} \int_\Omega \frac{k(x, y)^q}{\gamma(y)^q} \, \mu(\mathrm{d}y) < +\infty.$$

Assumption B corresponds to Assumption II.2 and Assumption III.2. Note that if Assumption B is satified, then Assumption A is satisfied for the kernel $\mathbf{k}(x, y) = k(x, y)/\gamma(y)$.

Finally, we consider the following assumption which is the analogue of the irreducibility of matrices but for kernels.

Assumption C. *The kernel $\mathbf{k}$ is such that*

$$\int_{A\times A^{\mathrm{c}}} \mathbf{k}(x, y) \, \mu(\mathrm{d}x)\mu(\mathrm{d}y) > 0.$$

for all measurable set $A \in \mathscr{F}$ such that $\mu(A) > 0$ and $\mu(A^{\mathrm{c}}) > 0$.

A kernel satisfying Assumption C is called (strongly) connected or irreducible; see Assumption II.3, Sections III.5.4 and IV.3.4. It ensures the uniqueness of the endemic equilibrium in the SIS dynamic. In Chapter IV, we prove that it also implies the regularity of the anti-Pareto frontier of vaccination.

Chapter II

An Infinite-Dimensional Metapopulation SIS Model

Chapter Abstract

In this chapter, we introduce an infinite-dimensional deterministic metapopulation SIS model which takes into account the heterogeneity of the infections and the social network among a large population. We study the long-time behavior of the dynamic. We identify the basic reproduction number $\mathfrak{R}_0$ which determines whether there exists a stable endemic steady state (super-critical case: $\mathfrak{R}_0 > 1$) or if the only equilibrium is disease-free (critical and sub-critical case: $\mathfrak{R}_0 \leq 1$). As an application of this general study, we prove that the so-called "leaky" and "all-or-nothing" vaccination mechanism have the same effect on $\mathfrak{R}_0$. This framework is also very natural and intuitive to model lockdown policies and study their impact.

The material for this chapter has been released in [35] and is currently under review

Chapter Content:

II.1 Introduction

II.1.1 Motivation

Our goal in this chapter is to provide a generalization of the classical SIS model to infinite-dimensional metapopulation and study its properties, under very weak assumptions. For pedagogical purposes, we start by discussing in detail how one can arrive at this generalization, starting from the one-dimensional model, and building our way through the finite dimensional model of Lajmanovich and Yorke, before introducing the general framework. The huge literature concerning these models, how our framework is situated in this picture, and comparison with related works, are discussed below in Section II.1.6.

The SIS model

Some infections do not confer any long-lasting immunity. With such infections, individuals become susceptible again once they have recovered from the disease. The simplest deterministic way to model this kind of epidemics in a constant size population is the following system of ordinary differential equations, introduced by Kermack and McKendrick in [95] and known as the SIS (susceptible/infected/susceptible) model:

$$\begin{cases} \frac{\mathrm{d}S}{\mathrm{d}t} = -\frac{K}{N} IS + \gamma I, \\ \frac{\mathrm{d}I}{\mathrm{d}t} = \frac{K}{N} IS - \gamma I, \end{cases}$$

where $S = S(t)$ and $I = I(t)$ are the number of susceptible and infected individuals, the total size $N = S(t) + I(t)$ of the population is constant in time, and K and γ are two positive numbers which represent the infectiousness and the recovery rate of the disease. The proportion $U(t) = I(t)/N$ of infected individuals in the population evolves autonomously, according to:

$$\frac{\mathrm{d}U}{\mathrm{d}t} = (1-U)KU - \gamma U. \tag{II.1}$$

Looking at a time change of U given by $V(t) = U(t/\gamma)$ and setting $\mathfrak{R}_0 = K/\gamma$, one gets that $\mathrm{d}V/\mathrm{d}t = (1-V)\mathfrak{R}_0 V - V$. The parameter $\mathfrak{R}_0$ can be interpreted as the number of infected individuals one infected individual generates on average over the course of its infectious period, in an otherwise uninfected population. This basic reproduction number was first introduced by Macdonald [113], and appears in a large class of models in epidemiology, see the monograph [24] from Brauer and Castillo-Chavez . The ordinary differential equation in V is well-posed and admits an explicit solution. If $V(0) = 0$, then $V(t) = 0$ for all t: as V represents the proportion of infected individuals, this constant solution is called the *disease-free equilibrium*. Now assume $V(0) = V_0 \in (0, 1]$. If $\mathfrak{R}_0 \neq 1$, the proportion of infected individuals in the population for $t \geq 0$ is given by:

$$U(\gamma t) = V(t) = \frac{\mathfrak{R}_0 - 1}{\mathfrak{R}_0 + ((1-\mathfrak{R}_0)/V_0 - \mathfrak{R}_0)\mathrm{e}^{(1-\mathfrak{R}_0)t}}.$$

If $\mathfrak{R}_0 = 1$, then the proportion of infected individuals in the population is given by:

$$U(\gamma t) = V(t) = \frac{1}{(1/V_0) + t}.$$

Hence, one can identify three possible longtime behaviors for the dynamical system:

Sub-critical regime If $\mathfrak{R}_0 < 1$, $U(t)$ converges exponentially fast to 0, and the only equilibrium is the disease-free solution $U(t) = 0$.

Critical regime If $\mathfrak{R}_0 = 1$, $U(t)$ still converges to 0 but not exponentially. The disease-free equilibrium is still the only one.

Super-critical regime If $\mathfrak{R}_0 > 1$, the constant solution 0 becomes unstable and another equilibrium appears, $G^* = 1 - \mathfrak{R}_0^{-1}$. This equilibrium is called *endemic*, and is globally stable in the sense that $U(t)$ converges towards G^* for all initial positive conditions.

The multidimensional Lajmanovich Yorke extension

In a pioneering paper [102], Lajmanovich and Yorke introduced an extension of the SIS model for the propagation of gonorrhea, which takes into account the fact that the propagation of the virus is highly non homogeneous among the population — we refer to the survey [125, Section V.A.2] from Pastor-Satorras, Castellano, van Mieghem and Vespignani, and more precisely Section 2 therein, for broader context and more details.

In this model the population is divided into n groups and the transmission rates of the disease between these groups are not equal, leading to a system of coupled ODEs:

$$\frac{\mathrm{d}U_i}{\mathrm{d}t} = (1 - U_i)\sum_{j=1}^{n} K_{i,j}U_j - \gamma_i U_i, \qquad \forall i \in \{1, 2, \dots, n\} \tag{II.2}$$

where U_i is the proportion of infected individuals in group i with $U_i(0) \in [0, 1]$ for all $1 \leq i \leq n$, $K = (K_{i,j})_{1\leq i,j\leq n}$ is a non-negative matrix that represents the transmission rates of the infection between the different groups, and the non-negative number $\gamma_i > 0$ is the recovery rate of group i. Since the matrix $K/\gamma = (K_{i,j}/\gamma_j)_{1\leq i,j\leq n}$ has non-negative entries, we recall it has a Perron eigenvalue, that is, an eigenvalue $\mathfrak{R}_0 \in \mathbb{R}_+$ such that all other complex eigenvalues λ of K/γ satisfy $|\lambda| \leq \mathfrak{R}_0$. The following result is given in [102].

1. There exists a unique solution $(U_i(t) \,:\, t \geq 0)_{1\leq i\leq n}$ of Equation (II.2) and $U_i(t) \in [0, 1]$ for all $t \in \mathbb{R}_+$.

2. If $\mathfrak{R}_0 \leq 1, U_i(t)$ converges to 0 for all $1 \leq i \leq n$, so that the disease-free equilibrium $(0, 0, \dots, 0)$ is globally stable.

3. If K is irreducible and $\mathfrak{R}_0 > 1$, then there exists an endemic equilibrium $G^* = (G_i^*)_{1\leq i\leq n}$ such that for $i = 1 \dots n$:
$$\lim_{t\to\infty} U_i(t) = G_i^* \in (0, 1),$$
provided that $U(0) \neq (0, 0, \dots, 0)$.

Thus, under the assumption that people are connected enough, the epidemic has two possible behaviors exactly like in the one-dimensional model:

> Biotheorem 1, [102] Either the epidemic will die out naturally for every possible initial stage of the epidemic, or when it is not true and the initial number of infectives of at least one group is nonzero, the disease will remain endemic for all the future time. Moreover, the number of infectives and susceptibles of each group will approach nonzero constant levels which are independent of the initial levels.

In 1957, before Lajmanovich and Yorke's seminal work, a similar type of behavior was formally derived by Kendall [94] for a heterogeneous SIR epidemic model but with strong assumptions on the transmission rates (see Equation (4) therein).

Towards a generalization

The epidemiological models discussed so far assume a large population, possibly made of a few groups with different behaviors, so that the epidemics is deterministic. At the opposite side of the modelling spectrum, some probabilistic models of interacting particles may be seen as modelling epidemics.

In 1974, Harris [76] introduced the so-called contact process on $\mathbb{Z}^d$. The contact process is a continuous-time Markov process often used as a model for the spread of an infection. Nodes of the graph represent the individuals of a population. They can either be infected or healthy. Infected individuals become healthy after an exponential time, independently of the configuration. Healthy individuals become infected at a rate which is proportional to the number of infected neighbors. The contact process share a numerous properties with the multigroup SIS equations: the existence of an upper invariant measure, a disease-free invariant measure and a monotone coupling [105, 106]. This proximity is not surprising since Equation (II.2) can be obtained from a mean-field approximation of the contact process [125, Section V.A]. Notice that Equation (II.2) can also be obtained as a limit of individual based models, see [13].

We refer to [125], and the numerous references therein, for a survey on epidemic processes in complex networks. Since social networks are very large graphs, it is natural to consider epidemic processes on limits of large graphs using theories developed during the last two decades. The first type of limiting objects are called *graphings*, and are used to deal with very sparse graphs, namely those with bounded degree, see [2, 50, 111]. At the other end of the spectrum, *graphons* are flexible objects that define a limit for dense graphs where the mean degree is of the same order as the number of vertices, see, for example, [111, 112]. We refer to [19, 21, 100] for further attempts at defining a limit theory for all kind of graphs.

The SIS equation that we propose in the present chapter has to be thought as the limit of the mean-field approximations of the contact processes defined on a convergent sequence of large graphs. Thus, the solutions take values in an abstract space Ω (the set of vertices), which can be interpreted as the set of features of the individuals, the transmission of the disease is given by a kernel κ and the recovery rate by a function γ (see Examples II.1.3 and II.1.2), see the infinite-dimensional evolution Equation (II.3) below.

The two main goals of this chapter are the following:

- introduce an infinite-dimensional SIS model, generalizing the model developed by Lajmanovitch and Yorke (see Equation (II.3) below), and prove a result similar to [102, Biotheorem 1] in that general setting;

- argue that this general setting is flexible enough to take into account not only the topology of the social network, or the disparities between different subgroups of the population, but also the effect of vaccination policies (see Section II.5), or the effect of lockdown (see Section II.6), in the spirit of the policies used to slow down the propagation of Covid-19 in 2020.

II.1.2 The model

It is natural to extend the Lajmanovich and Yorke model (II.2) to a population with an infinite number of groups. We choose to present this extension in an abstract setting, as this allows us to include general vaccination and lockdown policies. We denote by Ω the set of the features of the individuals in a given population. Since Ω might not be countable, we shall consider a σ-field $\mathscr{F}$ on Ω so that $(\Omega, \mathscr{F})$ is a measurable space. We represent the transmission rate from an infinitesimal part of the population $\mathrm{d}y$ to x by a non-negative kernel $\kappa(x, \mathrm{d}y)$: κ is a function from $\Omega \times \mathscr{F}$ to $\mathbb{R}_+$ such that, for all $A \in \mathscr{F}$, the mapping $x \mapsto \kappa(x, A)$ is measurable and, for all $x \in \Omega$, the mapping $A \mapsto \kappa(x, A)$ is a non-negative measure defined on $(\Omega, \mathscr{F})$. We model the recovery rate of individuals with feature x by $\gamma(x)$, where γ is a non-negative measurable function defined on $(\Omega, \mathscr{F})$. The number $1/\gamma(x)$ can be thought as the typical time of recovery for individuals with feature x. For $x \in \Omega$ and $t \geq 0$, we denote by $u(t, x)$ the probability for an individual (or the proportion of individuals) with feature x to be infected at time t. So the intensity of infection attempts on x coming from infected individuals in $\mathrm{d}y$ is given by $u(t, y)\kappa(x, \mathrm{d}y)$. Recall that in a SIS model, the probability for an infection attempt to succeed is proportional to the number of susceptible individuals, *i.e.*, those who are not already infected; this explains the term $(1 - u(t, x))$ in front of the integral in the next equation. The evolution equation of the function u for the SIS model of the probability for being infected is given by the following differential equation (in infinite dimension):

$$\begin{cases} \partial_t u(t,x) = (1-u(t,x)) \displaystyle\int_\Omega u(t,y)\,\kappa(x,\mathrm{d}y) - \gamma(x)u(t,x), \quad x \in \Omega, t \in [0,\tau), \\ u(0,x) = u_0(x), \quad x \in \Omega, \end{cases} \tag{II.3}$$

where the measurable function $u_0 : \Omega \to [0,1]$ is the so-called initial condition and the solution u is defined up to time $\tau \in (0,\infty]$. We shall prove that Equation (II.3) is well defined up to $\tau = +\infty$, and we will mainly focus our study on the long-time behavior of the solutions to this equation and on the study of existence of equilibria. Once again, we refer to Section II.1.6 for a discussion on related work, and in particular the work by Thieme [150] on a spatial SIR model and by Busenberg, Iannelli and Thieme [28] on long-time behavior of an age-structured SIS infection.

Example II.1.1 (Lajmanovich and Yorke model). Consider a finite set of features, $\Omega = \{1, 2, \dots, n\}$ (with the σ-field $\mathscr{F} = \mathscr{P}(\Omega)$ of all sub-sets of Ω), a finite kernel κ and a positive recovery rate γ. We set for all $i, j \in \Omega$ and $t \geq 0$:

$$K_{i,j} = \kappa(i, \{j\}), \quad \gamma_i = \gamma(i) \quad \text{and} \quad U_i(t) = u(t,i),$$

where u is the solution to Equation (II.3). The functions U_i, for $1 \leq i \leq n$, clearly solve the finite-dimensional model (II.2).

There are two natural extensions of Example II.1.1 to large bounded degree graphs and large dense graphs, which is a first approach to model large complex social networks.

Example II.1.2 (Graph model). Consider a representation of the social interaction of a population by a simple graph G, with set of vertices $V(G) = \Omega$ which is at most countable, and set of edges $E(G) \subset \Omega \times \Omega$. For $x \in \Omega$, let $\mathcal{N}(x) = \{ y \in G : (x,y) \in E(G) \}$ stands for the neighborhood of x in G and $\deg_G(x) = \operatorname{Card}(\mathcal{N}(x))$ for its degree. If the degree of the vertices of G are finite, we may consider a kernel with the following form:

$$\kappa(x, \mathrm{d}y) = \beta(x) \sum_{z \in \mathcal{N}(x)} \theta(y) \delta_z(\mathrm{d}y), \tag{II.4}$$

where β and θ are non-negative functions, which represent the susceptibility and the infectiousness of the individuals respectively, and δ_z is the Dirac mass at z. Then Equation (II.3) represents the evolution equation for a SIS model on a graph. The strength of the formalism of (II.3) is that one can consider limit of large bounded degree undirected graphs called graphings, see Section 18 in [111] for the definition of a graphings.

Example II.1.3 (Graphon form). One of the initial motivation of this work, was to consider a SIS model on *graphons*, which are limit of large dense graphs, see the monograph [111] from Lovàsz. In a recent paper [156], Vizuete, Frasca and Garin studied the stability of deterministic SIS epidemics over a large network generated by a Lipschitz graphon.

Recall the set of features of the individuals in the population is given by a set Ω. In this approach, the typical form of the transmission kernel κ we may consider is:

$$\kappa(x, \mathrm{d}y) = \beta(x) W(x,y) \theta(y)\, \mu(\mathrm{d}y), \tag{II.5}$$

where β represents the susceptibility and θ the infectiousness of the individuals; W models the graph of the contacts within the population and the quantity $W(x,y) \in [0,1]$ is interpreted as the probability that x and y are connected, or as the density of contacts between individuals with features x and y; μ is a probability measure on $(\Omega, \mathscr{F})$ and $\mu(\mathrm{d}y)$ represents the infinitesimal proportion of the population with feature y. Formally, β and θ are non-negative measurable functions, and the function $W : \Omega \times \Omega \to [0,1]$ is symmetric measurable. The quadruple $(\Omega, \mathscr{F}, \mu, W)$ is called a graphon. The degree $\deg_W(x)$ of $x \in \Omega$ (*i.e.* the average number of his contacts) and the mean degree d_W for a graphon W are defined by:

$$\deg_W(x) = \int_\Omega W(x,y)\, \mu(\mathrm{d}y) \text{ and } \mathrm{d}_W = \int_\Omega \deg_W(x)\, \mu(\mathrm{d}x) = \int_{\Omega^2} W(x,y)\, \mu(\mathrm{d}y)\, \mu(\mathrm{d}x). \tag{II.6}$$

Let us give a bit more detail in three particular cases.

(i) **Constant graphon.** One elementary example, is the constant graphon, $W = p \in [0, 1]$. In this case, the degree function is constant, equal to the mean degree and thus equal to the parameter p. We recall this constant graphon appears as the limit, as n goes to infinity, of Erdös-Rényi random graphs with n vertices and parameter p (that is: independently, for each pair of vertices, there is an edge between those two vertices with probability p). If furthermore the functions β, θ and γ from (II.3) are constant on Ω, then we recover the SIS model (II.1) with $K = p\beta\theta$ and $U(t) = \int_\Omega u(t, x)\mu(\mathrm{d}x)$.

(ii) **Stochastic block model.** The stochastic block model of communities, introduced by [85] (and referred to as step graphons in [111], see also [1] for further references), corresponds to the case where W is constant by block, *i.e.* there exists a finite partition $(\Omega_i : 1 \leq i \leq n)$ of Ω such that W is constant on the blocks $\Omega_i \times \Omega_j$ for all i, j, and equal say to $W_{i,j}$. If furthermore, the functions β, θ and γ from (II.3) are also constant on the partition, then we recover the Lajmanovich and Yorke model, see (II.2), with: $K_{i,j} = \beta_i W_{i,j} \theta_j \mu(\Omega_j)$; β_i, θ_i and γ_i are the constant values of β, θ and γ on Ω_i; and $U_i(t) = \int_{\Omega_i} u(t, x)\mu(\mathrm{d}x)/\mu(\Omega_i)$.

(iii) **Geometric graphon.** In this case, which is a natural generalization of the Random Geometric Graph (see [129] for a survey and [32] and references therein for related models), the probability of contact between x and y depends on their relative distance. For example, consider the population uniformly spread on the unit circle: $\Omega = [0, 2\pi]$ and $\mu(\mathrm{d}x) = \mathrm{d}x/2\pi$. Let f be a measurable non-negative function defined on $\mathbb{R}$ which is bounded by 1 and 2π-periodic. Define the corresponding geometric graphon W_f by $W_f(x, y) = f(x - y)$ for $x, y \in \Omega$. In this case, the degree of $x \in [0, 1]$ is constant with:

$$\deg_{W_f}(x) = \mathrm{d}_{W_f} = \frac{1}{2\pi} \int_{[0,2\pi]} f(y)\, \mathrm{d}y.$$

II.1.3 Main assumptions and definition of the reproduction rate

In order for Equation (II.3) to make sense, we will need the following assumption. It will always be in force throughout this chapter without supplementary specification.

Assumption II.1. *The function γ is positive, bounded and the non-negative kernel κ is uniformly bounded:*

$$\sup_{x\in\Omega} \kappa(x, \Omega) < \infty. \tag{II.7}$$

Assuming the recovery rate γ to be bounded is equivalent to require the time of recovery $1/\gamma$ to be bounded from below by a positive constant. The function $1/\gamma$ is also finite for all individuals because γ is supposed to be positive. It is possible with Assumption II.1 to have individuals with arbitrary large time of recovery, though. Finally, Equation (II.7) limits the maximal force of infection that can be put upon a susceptible individual.

In Examples II.1.1, II.1.2 and II.1.3, we observe that the kernel has a density with respect to a reference measure (the counting measure in the first two examples and the probability measure μ in the third one). From an epidemiological point of view, the reference measure μ can be seen as a way to quantify the size of the population and its sub-groups (defined by a given feature such as sex, spatial coordinates, social condition, health background, ...). If the measure μ is finite, then for every measurable set A, the number $\mu(A)/\mu(\Omega)$ is the proportion of individuals in the population whose features belong to A. We shall consider the case where the density k of κ with respect to the reference measure μ satisfies some mild integrability condition. We emphasize that the space Ω is not equipped with a topology and, as a consequence, we do not assume any smoothness condition on the density k. By a slight abuse of language, we will also call the density k a kernel.

Assumption II.2. *There exists a finite positive measure μ on $(\Omega, \mathscr{F})$, a non-negative measurable function $k : \Omega \times \Omega \to \mathbb{R}_+$ such that for all $x \in \Omega$, $\kappa(x, \mathrm{d}y) = k(x, y)\mu(\mathrm{d}y)$. Besides, there exists $q > 1$ such that:*

$$\sup_{x\in\Omega} \int_\Omega \frac{k(x, y)^q}{\gamma(y)^q}\, \mu(\mathrm{d}y) < \infty. \tag{II.8}$$

Note that the kernel $k(x,y)/\gamma(y)$ that appears in Assumption II.2 is the analogue of the ratio K/γ in the multi-dimensional Lajmanovich–Yorke model.

Since we assume that γ is bounded, Equation (II.8) implies the following integrability condition for the kernel k:

$$\sup_{x\in\Omega} \int_\Omega k(x,y)^q \, \mu(\mathrm{d}y) < \infty. \tag{II.9}$$

We shall study in Section II.4.5 an example which does not satisfy the integrability condition (II.8) nor (II.9).

Finally, some results in the supercritical regime will only hold under the following connectivity assumption.

Assumption II.3 (Connectivity). *The kernel k is* connected, *that is,*

$$\int_{A\times A^c} k(x,y)\, \mu(\mathrm{d}x)\mu(\mathrm{d}y) > 0 \tag{II.10}$$

for any measurable set A such that $\mu(A) > 0$ and $\mu(A^c) > 0$.

The sociological interpretation of the connectivity assumption is that we cannot separate the population into two groups of individuals with no interaction. Contrary to Assumption II.1 which is assumed throughout the text, we will specify each time whether Assumptions II.2 or II.3 are needed.

Remark II.1.4 (The finite dimensional case). Assumption II.2 is automatically satisfied in the finite-dimensional model of Example II.1.1, where we supposed Assumption II.1. We can indeed take μ to be the counting measure and Equation (II.8) is true because k is bounded from above and γ is bounded from below by a positive constant as it is positive. Notice Assumption II.3 is equivalent to the matrix of transmission rates $K = (K_{i,j})_{1\leq i,j\leq n}$ being irreducible.

The basic reproduction number of an infection, denoted by $\mathfrak{R}_0$, has originally been defined as the number of cases one typical individual generates on average over the course of its infectious period, in an otherwise uninfected population. This number plays a fundamental role in epidemiology as it provides a scale to measure how difficult to control an infectious disease is. More importantly, in many models, the particular value $\mathfrak{R}_0 = 1$ turns out to be a threshold: the disease will die out if $\mathfrak{R}_0 < 1$, and invade the population if $\mathfrak{R}_0 > 1$.

In mathematical epidemiology, Diekmann, Heesterbeek and Metz [42] define rigorously the basic reproduction number for a class of models with heterogeneity in the population. They propose to consider the next-generation operator which gives the distribution of secondary cases arising from an infected individual picked randomly according to a certain distribution — the population being assumed uninfected otherwise. In our model, under Assumption II.2, following [42, Equation (4.2)], we define the next generation operator, denoted by $T_{k/\gamma}$, as the integral operator:

$$T_{k/\gamma}(g)(x) = \int_\Omega \frac{k(x,y)}{\gamma(y)} g(y)\, \mu(\mathrm{d}y) \quad \text{for all } x \in \Omega, \tag{II.11}$$

which is, thanks to (II.8), a bounded positive operator on the space $\mathscr{L}^\infty(\Omega)$ of bounded real-valued measurable functions defined on Ω. Following [42, Definition of $\mathfrak{R}_0$ in Section 2], the basic reproduction number is then defined by, :

$$\mathfrak{R}_0 = \rho(T_{k/\gamma}), \tag{II.12}$$

where ρ is the spectral radius, whose exact definition in our general setting will be recalled below (Equation (II.34)). These definitions of the next-generation operator and the basic reproduction number are consistent with the finite dimensional SIS model given in [153].

II.1.4 Long time behavior of solutions to the evolution equation (II.3)

We now state our main result concerning solutions of the evolution equation (II.3). Recall the initial condition of (II.3), u_0, takes values in $[0, 1]$.

Theorem II.1.5. *We have the following properties.*

(i) ***(Equation** (II.3) **is well defined and $\tau = +\infty$.)** Under Assumption II.1, there exists a unique solution u to Equation* (II.3). *This solution is such that, for all $(x, t) \in \Omega \times \mathbb{R}_+$, $u(t, x) \in [0, 1]$.*

(ii) ***(Disease free equilibrium in the critical and sub-critical case.)** Assume that Assumptions II.1 and II.2 are in force. Let $\mathfrak{R}_0$ be defined by* (II.12). *If $\mathfrak{R}_0 \leq 1$, then the disease dies out: for all $x \in \Omega$,*

$$\lim_{t\to\infty} u(t, x) = 0.$$

(iii) ***(Stable endemic equilibrium in the super-critical case.)** Assume that Assumptions II.1, II.2 and II.3 are in force. If $\mathfrak{R}_0 > 1$, then there exists a unique equilibrium $\mathfrak{g} : \Omega \to [0, 1]$ different from* 0 *and it is positive μ-a.e. For all initial condition u_0 such that its integral is positive:*

$$\int_\Omega u_0(x)\, \mu(\mathrm{d}x) > 0,$$

the solution u to (II.3) *converges pointwise to $\mathfrak{g}$, i.e., for all $x \in \Omega$:*

$$\lim_{t\to\infty} u(t, x) = \mathfrak{g}(x).$$

If $u_0 = 0$ μ-a.e. then the solution u to (II.3) *converges pointwise to* 0.

For property (i), see Proposition II.2.7; property (ii) is a consequence of Theorems II.4.6 and II.4.7; and property (iii) follows from Corollary II.4.9 and Theorem II.4.13.

Remark II.1.6 (Uniform convergence). The convergence of $u(t, \cdot)$ towards 0 in (ii) or towards $\mathfrak{g}$ in (iii) in Theorem II.1.5 is uniform on any measurable subset $A \subset \Omega$ such that $\inf_A \gamma > 0$, see Theorem II.4.17. In particular these convergences hold in uniform norm if the recovery rate γ is bounded from below.

II.1.5 Modelling vaccination policies, vaccination mechanisms and lockdown

Vaccination

In Section II.5, we propose extensions of Equation (II.3) which take into account the effect of a vaccination policy. Vaccination confers a direct protection on the targeted individuals but also acts indirectly on the rest of the population through herd immunity. However, all vaccinated individuals will not be totally immune to the disease. In [146], Smith, Rodrigues and Fine propose two possible models to explain vaccine efficacy. In the first model, the vaccine offers complete protection to a portion of the vaccinated individuals but does not take in the remainder of vaccinated individuals. The second model supposes that the vaccination confers a partial protection to every vaccinated individual. In [140], Halloran, Lugini and Struchiner called the former mechanism the *all-or-nothing* vaccination and the latter one the *leaky* vaccination. We define below one infinite-dimensional SIS model for each of these two mechanisms.

In order to write down the vaccination model, we first adapt the one-group SIR models proposed by Shim and Galvani in [140] to the one-group SIS model. Let us denote by η_v the proportion of vaccinated individuals in the total population, and let $\eta_\mathrm{u} = 1 - \eta_\mathrm{v}$. Let U_v and U_u be the proportion of infected individuals in the vaccinated and unvaccinated population respectively, so that $\eta_\mathrm{v} U_\mathrm{v} + \eta_\mathrm{u} U_\mathrm{u}$ is the proportion of infected individuals in the total population. For both models, we assume that vaccinated individuals who are nevertheless infected by the disease become less contagious (see [131, 139] for instance). We will denote the vaccine efficacy for infectiousness, that is, the relative reduction of infectiousness for vaccinated individuals by a parameter $\delta \in [0, 1]$.

In what follows, K and γ represent the transmission rate and the recovery rate of the disease as in the model (II.1) or (II.2) and are assumed to be the same for the vaccinated and unvaccinated population. We now introduce two models for the so-called vaccine efficacy e, see [70, 74, 140, 146] for discussion on this parameter.

In the leaky vaccination, we define the efficacy $e \in [0, 1]$ as the relative reduction of susceptibility for vaccinated individual. Following [140, Equations (1)-(8)], with the parameters δ and e corresponding to σ and α in [140], the evolution equations for the leaky vaccination are then given by:

$$\begin{cases} \frac{\mathrm{d}U_\mathrm{v}}{\mathrm{d}t} = (1 - U_\mathrm{v})(1 - e)K((1 - \delta)\eta_\mathrm{v}U_\mathrm{v} + \eta_\mathrm{u}U_\mathrm{u}) - \gamma U_\mathrm{v}, \\ \frac{\mathrm{d}U_\mathrm{u}}{\mathrm{d}t} = (1 - U_\mathrm{u})K((1 - \delta)\eta_\mathrm{v}U_\mathrm{v} + \eta_\mathrm{u}U_\mathrm{u}) - \gamma U_\mathrm{u}. \end{cases} \tag{II.13}$$

In the all-or-nothing vaccination, we denote the proportion of vaccinated individuals immunized to the disease (people who can neither contract not transmit the disease) by the parameter $1 - e \in [0, 1]$. Following [140, Equations (13)-(20)], the evolution equations for the all-or-nothing vaccination in the SIS setting are given by:

$$\begin{cases} \frac{\mathrm{d}U_\mathrm{v}}{\mathrm{d}t} = (1 - e - U_\mathrm{v})K((1 - \delta)\eta_\mathrm{v}U_\mathrm{v} + \eta_\mathrm{u}U_\mathrm{u}) - \gamma U_\mathrm{v}, \\ \frac{\mathrm{d}U_\mathrm{u}}{\mathrm{d}t}(1 - U_\mathrm{u})K((1 - \delta)\eta_\mathrm{v}U_\mathrm{v} + \eta_\mathrm{u}U_\mathrm{u}) - \gamma U_\mathrm{u}. \end{cases} \tag{II.14}$$

Since vaccinated individuals that are immunized cannot get the disease, we have $U_\mathrm{v}(t) \leq 1 - e$ for all $t \in \mathbb{R}_+$.

Remark II.1.7. Notice that, in both models, the unvaccinated population can be viewed as a population inoculated with a vaccine of efficacy equal to 0.

In Section II.5, we derive in Equations (II.66) and (II.69) the analogue of (II.13) and (II.14) in the infinite-dimensional setting. Those two equations can be seen as a particular case of Equation (II.3). We also prove that, as far as the basic reproduction number is concerned the two different vaccination mechanisms, the all-or-nothing and leaky mechanisms, have the same effect in the infinite dimensional model, see Proposition II.5.2. This result was already observed in a one-group model by Shim and Galvani [140]. In the case of a perfect vaccine, where vaccinated people cannot be infected nor infect others, the evolution equation of the proportion of infected among the non vaccinated population is also given by Equation (II.3) with the kernel $\kappa(x, \mathrm{d}y)$ replaced by $\eta^0(y)\kappa(x, \mathrm{d}y)$ where $\eta^0(y)$ is the proportion of individuals with feature $x \in \Omega$ which are not vaccinated, see Equation (II.73). We shall study in a future work the optimal vaccination in this setting with the basic reproduction number as an objective function to minimize.

Effect of lockdown policies

Eventually, we model the effect of lockdown (see Section II.6) for graphon models presented in Example II.1.3, in the spirit of the policies used to slow down the propagation of Covid-19 in 2020, see for example the study [41]. In particular, we prove that a lockdown which bounds the number of contacts of the individuals (this roughly corresponds to reduce significantly the number of contacts for highly connected groups) is enough to reduce the basic reproduction number, see Proposition II.6.3. Recall the definition of the degree $\deg_W(x)$ of x and the mean degree d_W for a graphon W defined in (II.6). Following Remark II.6.4, we get that the heterogeneity in the degree for the graphon model implies larger value of the basic reproduction number. In this direction, see also [108, Section 1.1] on the SIS model from Pastor-Satorras and Vespignani, where the basic reproduction number increases with the variance of the degrees of the nodes in a finite graph.

Corollary II.1.8. *Consider the SIS model* (II.3) *with transmission kernel given in a graphon form* (II.5) *(so that $k(x, y) = \beta(x)W(x, y)\theta(y)$). Assume that the susceptibility β, the infectiousness θ and the recovery rate γ are constant and positive. The weakest value of the basic reproduction number $\mathfrak{R}_0$*

defined by (II.12) *among all graphons W with mean degree* $d_W \geq p$ *for some threshold* $p \in [0, 1]$ *is obtained for graphons with constant degree equal to p (i.e. graphon W such that* $\deg_W(x) = p$ *for all* $x \in \Omega$*).*

We recall from Example II.1.3 (i) and (iii), that the constant graphon and the geometric graphons have constant degree. Considering a geometric graphon with (mean) degree p, we get that $\mathfrak{R}_0 = \gamma^{-1}\beta\theta p$, and for $\mathfrak{R}_0 > 1$, we deduce (directly or from Proposition II.2.17), that the equilibrium $\mathfrak{g}$ is constant equal to $1 - \mathfrak{R}_0^{-1}$ (compare with model (II.1) with $K = \beta\theta p$). Furthermore, the example of the geometric graphon with a given mean degree, indicates that, if the parameters β, θ and γ are constant, then the contamination distance (or support of the function f, see end of Remark II.6.4) from an infected individual is not relevant for the value of the basic reproduction number nor for the equilibria.

II.1.6 Discussion and related results

The dichotomy of possible dynamics described in [102, Biotheorem 1] has been established for many other compartmental models and possibly their multigroup version by using Lyapunov function techniques (see for instance [14, 107]). For a survey, we refer to Fall, Iggidr, Sallet and Tewa [55]. In [82, Section 6] and [142], Hirsch and Smith proved the long-time behavior of Equation (II.2) thanks to their theory of order preserving systems, thereby giving a completely new perspective to the study of mathematical epidemic models. Their work greatly inspired Li and Muldowney [104] in their important proof of the global stability of the endemic equilibrium of the SEIR model (susceptible-exposed-infected-recovered) which was a long-standing conjecture at that time.

Continuous models involving transmission rates that depend on the localization of the individuals [11, 94, 118] or their age [86] have long been studied. Both of these can be thought of as multigroup models with a continuous set of groups and therefore lead to differential equations in infinite-dimensional space. In this setting, results about global stability of the endemic or the disease-free equilibrium have also been obtained. We outline some of them below and highlight how they differ from our framework.

In [28], Busenberg, Iannelli and Thieme established the long-time behavior of an age-structured SIS infection. They proved, thanks to semi-group theory and positive operators methods, that the system converges to a unique endemic equilibrium if it exists. Otherwise, it converges to the disease-free equilibrium. In this work the transmission kernel is assumed to be bounded from above and below by product kernels (see Equation (2.9) therein). This represents a restriction (see the discussion at the end of [28]) as it is not possible to forbid contacts between some but not all groups. By contrast, in the setting of Example II.1.3, it is easy and natural to model the absence of contact between individuals with feature x and y by imposing that $W(x, y) = 0$, without imposing conditions on the probability of contact between x and other features than y.

In [58] Feng, Huang and Castillo-Chavez considered a similar dynamic for a multigroup age-structured SIS model, but where the endemic equilibrium exists but is not globally stable. They assume that the system has a quasi-irreducibility property (see Definition 3.1 therein) which is a weaker assumption than Assumption II.3, but impose bounds on the transmission kernel.

In [152], Thieme also used an operator approach to study a SIR model with variable susceptibility (see Section 4 therein). In particular, he studied the close relation between the spectral bound of the operator $T_k - \gamma$ and of the basic reproduction number $\mathfrak{R}_0$ which is the spectral radius of the operator $T_{k/\gamma}$. In [150], Thieme analyzed a space-structured SIR model with birth. In this model, the incidence term, *i.e.* the equivalent of $(1 - u(t, x))u(t, y)\kappa(x, \mathrm{d}y)$ in Equation (II.3), is replaced by a non bilinear term $f(x, y, 1 - u(t, x), u(t, y))\,\mathrm{d}y$, where the function f is continuous, locally Lipschitz continuous and increasing in its third and fourth argument. Imposing also that the recovery rate γ is bounded away from 0, he proved an analogue of [102, Biotheorem 1] (see Theorems 7.1, 8.2, 9.1 and 12.1 therein) using Lyapunov functions. Part of those results would not hold in general if $\inf \gamma = 0$. In contrast to these works, we consider very few regularity assumptions on the parameters, and in particular allow that $\inf \gamma = 0$.

Other works introduce movement of populations, either between discrete patches, see for example [8] and the biological examples therein, or in a continuous space. In [134], Ruan and Xiao obtained the global stability of the steady states for a general spatial SIS model with delay, a diffusion term for the infected population and a non-local kernel governing the transmission of the disease. It is assumed that the transmission kernel is smooth and satisfies a constant-degree assumption, so that the endemic equilibrium is unique and constant. Here, we do not consider the constant degree assumption as this condition intrinsically does not hold when considering vaccination strategies, see Section II.5. In the recent [3], Almeida, Bliman, Nadin, Perthame and Vauchelet studied a spatial SEIR model with or without diffusion. The case without diffusion is formally very close to our model, and makes no smoothness assumptions on the infection kernel; however, the infection kernel is supposed to be bounded and positive everywhere, and the various rates bounded above below. Once again, the positiveness assumption of the infection kernel breaks down when taking into account vaccination strategies or lockdown policies. Let us mention also that removing the boundedness assumption on the infection kernel could lead to the existence of an infinite number of positive equilibria, see Section II.4.5. For a deeper discussion on spatial epidemic models and a detailed review, we refer the reader to [133].

The principal tools we use to prove Theorem II.1.5, see also the key Lemma II.3.7 and Proposition II.2.5, can be summarized as follows.

Cooperative systems The function $g \mapsto F(g) = (1-g)T_\kappa(g) - \gamma g$ is *cooperative* (see Definition II.2.1 and Remarks II.2.2 and II.2.3), which implies that the solution of (II.3) are well defined and the corresponding dynamical system is order preserving. For approaches based on cooperation (or quasi-monotonicity) and monotone dynamical systems on various models, see [13, 82, 83, 142, 143, 144].

Positive operators Under Assumption II.2 and Equation (II.8), the integral operator $T_{k/\gamma}$ can be seen as an Hille-Tamarkin operator on $L^p(\mu)$ with the corresponding compactness property see [159, Theorem 41.6]. Then the positivity of the operator $T_{k/\gamma}$ allows to use Krein-Rutman theorem to get that its spectral radius is an eigenvalue with a non-negative eigenfunction. This argument has been widely used, see for example [28] (where the operator is of rank one, and thus is compact) and also [150, 152].

Connectivity Under Assumption II.3 on the connectivity of kernel k (which in finite dimension corresponds to the irreducibility of non-negative matrices and is related to the Perron-Frobenius theorem), we can consider the unique corresponding eigenvector, thanks to the Perron-Jentzsch theorem (see [136, Theorem V.6.6] or [69, Theorem 5.2]). This eigenvector is an essential tool to study the long-time behavior of the solution to Equation (II.3) in the super-critical regime. In finite dimension, see [102], where the matrix K from (II.2) is assumed to be irreducible, or [13] for a more general finite-dimensional model. In infinite dimension, see [58] for a weaker quasi-irreducibility condition.

Finally let us remark that we do not use the standard tool of Lyapunov functions, in contrast with many previous works, see for example [14, 107, 150].

II.1.7 Structure of the chapter

In Section II.2 we construct the semi-flow associated to the infinite dimensional SIS model (II.3), and prove its main regularity and monotonicity properties. We introduce in Section II.3 some important tools of spectral analysis in Banach lattices. This allows us to define in Section II.4.1 the basic reproduction number $\mathfrak{R}_0$. The convergence of the system towards an equilibrium is established in Section II.4. In Section II.5, we take into account the effect of a vaccination policies on the propagation of the disease. Eventually, in Section II.6, we model the impact of lockdown policies on the propagation of the disease when κ takes the graphon form of Example II.1.3.

II.2 Model analysis

II.2.1 Preamble

In this paragraph, we recall some definitions of functional analysis. Most of them can be found in [33]. Let $(X, \|\cdot\|)$ be a Banach space. The topological dual $X^\star$ of X is the space of all bounded linear forms and we use the notation $\langle x^\star, x\rangle$ for the value of an element $x^\star \in X^\star$ at $x \in X$. We consider K a proper cone on X, *i.e.*, a closed convex subset of X such that $\lambda K \subset K$ for all $\lambda \geq 0$ and $K \cap (-K) = \{0\}$. The proper cone K defines a partial ordering $\leq$ on X: $x \leq y$ if $y - x \in K$. It is said to be reproducing if $K - K = X$ (any element $x \in X$ can be expressed as a difference of elements of K). The dual cone of K is the set $K^\star \subset X^\star$ consisting of all $x^\star$ such that $\langle x^\star, x\rangle \geq 0$ for all $x \in K$. If the proper cone K is reproducing then the set $K^\star$ is a proper cone (see beginning of [33, Section 19.2]).

We denote by $\mathcal{L}(X)$ the space of bounded linear operators from X to X. The operator norm of a bounded operator $A \in \mathcal{L}(X)$ is given by:

$$\| A \| = \sup\{ \| Ax\| \ : \ x \in X,\ \|x\| \leq 1 \}.$$

The topology associated to $\|\cdot\|$ in $\mathcal{L}(X)$ is called the uniform operator topology. A linear bounded operator $A \in \mathcal{L}(X)$ is said to be positive (with respect to the proper cone K) if $AK \subset K$.

Let F be a function defined on an open domain $D \subset X$ and taking values in X. The function F is said to be Fréchet differentiable at $x \in D$, if there exists a bounded linear operator $\mathcal{D}F[x]$ such that:

$$\lim_{y\to 0} \| F(x+y) - F(x) - \mathcal{D}F[x](y) \| / \| y \| = 0.$$

The operator $\mathcal{D}F[x]$ is called the Fréchet derivative of F at point x.

We define the cooperativeness property which is related to the definition of quasimonotony firstly introduced by Volkmann [157] for abstract operators.

Definition II.2.1 (Cooperative function)**.** *Let $D_1, D_2 \subset X$. A function $F : X \to X$ is said to be cooperative on $D_1 \times D_2$ (with respect to K) if, for all $(x, y) \in D_1 \times D_2$ such that $x \leq y$ and for all $z^\star \in K^\star$, we have the following property:*

$$\langle z^\star, x - y\rangle = 0 \implies \langle z^\star, F(x) - F(y)\rangle \leq 0. \tag{II.15}$$

We shall mainly consider the cases $D_1 = X$ or $D_2 = X$.

Remark II.2.2. For a better understanding of the cooperativeness property, let us examine the finite dimensional case. Let $d \geq 2$, $X = \mathbb{R}^d$ and $K = \mathbb{R}^d_+$. Then, for a smooth function $F = (F_1, F_2, \dots, F_d)$, it is easy to see that F is cooperative on $X \times X$ with respect to K if and only if:

$$\frac{\partial F_j}{\partial x_i}(x) \geq 0 \quad \text{for all } x \in \mathbb{R}^d \text{ and all } i \neq j. \tag{II.16}$$

We recover the definition of cooperativeness introduced by Hirsch [82]. Suppose the vector x represents the utilities of a group of agents $\{1, 2, \dots d\}$ and F is the dynamics of the system, that is, $\mathrm{d}x/\mathrm{d}t = F(x)$. Then, the higher the utilities of agents $j \neq i$ are, the more beneficial the situation is for agent i, as it increases the value of the time derivative of x_i. For this reason, the function F satisfying (II.16) is called cooperative.

We extend the differential version of cooperativeness of Remark II.2.2 to infinite dimension in the next remark.

Remark II.2.3. Let X be a Banach space, D an open domain and $F : X \to X$ be a Fréchet differentiable function. Assume that F is cooperative on $D \times X$. Let $(x, z) \in D \times K$ and let $z^\star \in K^\star$ such that $\langle z^\star, z\rangle = 0$. Since F is cooperative on $D \times X$, we have:

$$\langle z^\star, (F(x + \lambda z) - F(x))/\lambda\rangle \geq 0,$$

for all $\lambda > 0$. Letting λ go to 0, we obtain the following inequality:

$$\langle z^\star, \mathcal{D}F[x](z)\rangle \geq 0. \tag{II.17}$$

Using path integrals in Banach space, we can prove the reverse implication in the case $D = X$. Indeed, for all $x, y \in X$ and all $z^\star \in X^\star$, we have:

$$\langle z^\star, F(x) - F(y)\rangle = -\int_0^1 \langle z^\star, \mathcal{D}F[(1-\lambda)x + \lambda y](y - x)\rangle \, \mathrm{d}\lambda. \tag{II.18}$$

Assume (II.17) holds for $z^\star \in K^\star$ and $z \in K$. Then, if $x \leq y$, $z^\star \in K^\star$ and $\langle z^\star, y - x\rangle = 0$, we get that $\langle z^\star, F(x) - F(y)\rangle$ is non-positive thanks to Equation (II.17). Thus the function F is cooperative.

Ordinary differential equations (ODEs) driven by cooperative vector fields enjoy a number of nice properties that we now review. Let us first recall a few definitions and classical properties of ODEs. Let $F : X \to X$ be a locally Lipschitz function. The Picard–Lindelöf theorem ensures the existence of $0 < \tau \leq \infty$ and of a continuously differentiable function y from $J = [0, \tau)$ to X which is the unique solution of the Cauchy problem:

$$\begin{cases} y'(t) = F(y(t)) & t \in J, \\ y(0) = y_0, \end{cases} \tag{II.19}$$

where $y_0 \in X$ is the so-called initial condition (see [34, Section 1.1]). A solution y defined on an interval $[0, \tau)$ is said to be maximal if there is no solution of Equation (II.19) defined on $[0, \tau')$ with $\tau' > \tau$. A solution is said to be global if it is defined on $[0, \infty)$.

All comparison properties will be derived from the following key result.

Theorem II.2.4 (Comparison Theorem). *Let K be a proper cone of X with non-empty interior. Denote by $\leq$ the corresponding partial order. Let $F : X \to X$ be locally Lipschitz, $D_1, D_2 \subset X$, $\tau > 0$, and let $a : [0, \tau) \to D_1$ and $b : [0, \tau) \to D_2$ be $\mathcal{C}^1$ paths. Suppose that F is cooperative on $D_1 \times X$ or on $X \times D_2$, and that:*

$$a'(t) - F(a(t)) \leq b'(t) - F(b(t)) \quad \forall t \in [0, \tau). \tag{II.20}$$

If $a(0) \leq b(0)$, then $a(t) \leq b(t)$ for all $t \in [0, \tau)$.

This result, in the spirit of [34, Theorem 5.2], is a generalization to infinite dimensional systems of classical comparison theorems for ODEs. Note in particular that (II.20) holds if a and b solve the ODE $u' = F(u)$, yielding the monotony of the flow of cooperative vector fields as a corollary, see Proposition II.2.8 below. For the sake of completeness, a proof of Theorem II.2.4 is given in Section II.7.

II.2.2 Notations

In this section, we will work in the Banach space $\mathscr{L}^\infty(\Omega)$ of measurable bounded real-valued functions defined on Ω equipped with the supremum norm $\|\cdot\|$. We shall write $\mathscr{L}^\infty$ when there is no ambiguity on the underlying space. The set:

$$\mathscr{L}^\infty_+ = \{ f \in \mathscr{L}^\infty : \ f(x) \geq 0 \quad \forall x \in \Omega \}, \tag{II.21}$$

is a proper cone in $\mathscr{L}^\infty$ with non-empty interior. The order defined by this proper cone is the usual order: $g \leq h$ if $g(x) \leq h(x)$ for all $x \in \Omega$.

We denote by $\mathscr{L}^{\infty,\star}$, the topological dual of $\mathscr{L}^\infty$. It can be identified as the space of bounded and finitely additive signed measures on Ω equipped with the total variation norm (see [158, Section 2]). Since $\mathscr{L}^\infty_+$ is reproducing, the dual cone $\mathscr{L}^{\infty,\star}_+$ is a proper cone. It consists of the continuous linear positive forms on $\mathscr{L}^\infty$.

Let κ be a non-negative kernel on $\mathscr{L}^\infty$ (endowed with its Borel σ-field) satisfying Assumption II.1. We denote by T_κ the operator:

$$\begin{aligned} T_\kappa \quad : \quad & \mathscr{L}^\infty \to \mathscr{L}^\infty \\ & g \mapsto \left(x \mapsto \int_\Omega g(y)\, \kappa(x, \mathrm{d}y) \right). \end{aligned} \tag{II.22}$$

According to Assumption II.1, the operator T_κ is a bounded linear operator with:

$$\| T_\kappa \| = \sup_{x\in\Omega} \kappa(x, \Omega) < \infty. \tag{II.23}$$

Since, for all $x \in \Omega$, $\kappa(x, \mathrm{d}y)$ is a positive measure, the operator T_κ is moreover positive. Defining now a function F from $\mathscr{L}^\infty$ to $\mathscr{L}^\infty$ by:

$$F(g) = (1 - g)T_\kappa(g) - \gamma g, \tag{II.24}$$

we may rewrite Equation (II.3) as an ODE in the Banach space $(\mathscr{L}^\infty, \|\cdot\|)$:

$$\begin{cases} \partial_t u = F(u), \quad t \in [0, \tau) \\ u(0, \cdot) = u_0, \end{cases} \tag{II.25}$$

where $u_0 \in \mathscr{L}^\infty$ and $\tau \in (0, \infty]$. Let Δ be the set of non-negative functions bounded by 1:

$$\Delta = \{ f \in \mathscr{L}^\infty \ : \ 0 \leq f \leq 1 \}. \tag{II.26}$$

Since the solution $u(t, x)$ of Equation (II.3) defines the proportion of x-type individuals being infected at time t, it should remain below 1 and above 0. Hence, for (II.3) to make a biological sense, the initial condition should belong to Δ and the solution, if it exists, should remain in Δ. This will be checked in Proposition II.2.7.

II.2.3 Properties of the vector field

Recall that Assumption II.1 is in force. The main results of this section are gathered in the following proposition.

Proposition II.2.5 (Properties of F)**.** *The function F defined in* (II.24) *has the following properties.*

(i) F is of class $\mathscr{C}^\infty$ on $\mathscr{L}^\infty$ in the Fréchet sense.

(ii) F and its repeated derivatives are bounded on bounded sets.

(iii) F is continuous on Δ with respect to the topology of pointwise convergence.

(iv) F is cooperative both on $(1 - \mathscr{L}^\infty_+) \times \mathscr{L}^\infty$ and on $\mathscr{L}^\infty \times (1 - \mathscr{L}^\infty_+)$, where:

$$1 - \mathscr{L}^\infty_+ = \{ g \in \mathscr{L}^\infty \ : \ g \leq 1 \}. \tag{II.27}$$

Proof. The bilinear map $(g, h) \mapsto gh$ and the linear maps $g \mapsto \gamma g$ and $g \mapsto T_\kappa(g)$ are bounded on $\mathscr{L}^\infty$ (hence, smooth as they are linear). Since the function F is a sum of compositions of the previous maps, properties (i) and (ii) are proved.

In order to prove property (iii), consider $(g_n,\ n \in \mathbb{N})$, a sequence of functions in Δ converging pointwise to $g \in \Delta$. Let $x \in \Omega$. The functions g_n are dominated by the function equal to 1 everywhere. The latter is integrable with respect to the measure $\kappa(x, \mathrm{d}y)$ since $\kappa(x, \Omega) < \infty$ according to (II.7). Therefore, we can apply the dominated convergence theorem and obtain:

$$\lim_{n\to\infty} \int_\Omega g_n(y)\kappa(x, \mathrm{d}y) = \int_\Omega g(y)\kappa(x, \mathrm{d}y).$$

Thus, the operator T_κ is continuous on Δ with respect to the pointwise convergence topology. The maps $(h_1, h_2) \mapsto h_1 h_2$ and $(h_1, h_2) \mapsto h_1 + h_2$ are also continuous with respect to the pointwise convergence topology. Hence, property (iii) is proved since F is a composition of these functions.

Finally, let us prove property (iv). Let $g, h \in \mathscr{L}^\infty$ such that $g \leq 1$ and $g \leq h$, and let $v \in \mathscr{L}_+^{\infty,\star}$ such that $\langle v, g - h\rangle = 0$. We have:

$$\begin{aligned}\langle v, F(g) - F(h)\rangle &= \langle v, (1-g)T_\kappa(g-h) + (h-g)(T_\kappa(h) + \gamma)\rangle \\ &= \langle v, (1-g)T_\kappa(g-h)\rangle\,,\end{aligned}$$

where we used Lemma II.2.6 below (with g replaced by $h - g$ and h by $T_\kappa(h) + \gamma$) in order to get that $\langle v, (h-g)(T_\kappa(h) + \gamma)\rangle$ is equal to 0. Since T_κ is a positive operator and $g \leq 1$, the function $(1-g)T_\kappa(g-h)$ is non-positive. The number $\langle v, (1-g)T_\kappa(g-h)\rangle$ is also non-positive because $v \in \mathscr{L}_+^{\infty,\star}$. Hence, we get that $\langle v, F(g) - F(h)\rangle \leq 0$. This proves that F is cooperative on $(1-\mathscr{L}_+^\infty)\times\mathscr{L}^\infty$, thanks to Definition II.2.1. If $(g,h) \in \mathscr{L}^\infty \times (1 - \mathscr{L}_+^\infty)$ satisfy $g \leq h$, then g is itself bounded above by 1, and the exact same proof applies, showing that F is also cooperative on $\mathscr{L}^\infty \times (1 - \mathscr{L}_+^\infty)$. ☐

The proof of Property (iv) of Proposition II.2.5 uses the following lemma.

Lemma II.2.6. *Let $g \in \mathscr{L}_+^\infty$ and $v \in \mathscr{L}_+^{\infty,\star}$ such that $\langle v, g\rangle = 0$. Then, for all $h \in \mathscr{L}^\infty$, we have $\langle v, hg\rangle = 0$.*

Proof. Let $g \in \mathscr{L}_+^\infty$ and $v \in \mathscr{L}_+^{\infty,\star}$ such that $\langle v, g\rangle = 0$. Since g is everywhere non-negative, we have:

$$-\|h\|\, g \leq hg \leq \|h\|\, g.$$

Since $v \in \mathscr{L}_+^{\infty,\star}$, the previous inequalities give:

$$-\|h\|\,\langle v, g\rangle \leq \langle v, hg\rangle \leq \|h\|\,\langle v, g\rangle\,.$$

By assumption, $\langle v, g\rangle$ is equal to 0. Hence, the lemma is proved. ☐

II.2.4 Properties of the ODE semi-flow

The aim of this subsection is to define a semi-flow associated to Equation (II.3) and to study its main properties. Proposition II.2.5 (ii) enables to apply the Picard-Lindelöf theorem and show the existence of local solutions of in $\mathscr{L}^\infty$ of Equation (II.3). We can actually prove a stronger result. Recall that $\Delta = \{ f \in \mathscr{L}^\infty \ :\ 0 \leq f \leq 1 \}$.

Proposition II.2.7. *Let F defined by* (II.24).

(i) The domain Δ is forward invariant*: if $u_0 \in \Delta$ and u solves* (II.25) *on $[0, \tau)$, then $u(t) \in \Delta$ for all $0 \leq t < \tau$.*

(ii) Maximal solutions of Equation (II.25) *such that $u_0 \in \Delta$ are global, i.e., they are defined on $\mathbb{R}_+$.*

Proof. We first prove property (i). Let $u_0 \in \Delta$, and suppose that u solves (II.25) on $[0, \tau)$. Let $a(t) \in (1 - \mathscr{L}_+^\infty)$ be equal to the constant function 0, for all t; let $b(t) = u(t)$. Since $F(0) = 0$, and $b(t) = u(t)$ solves the ODE, we have for all $t < \tau$:

$$a'(t) - F(a(t)) = 0 = b'(t) - F(b(t)).$$

Since $0 = a(0) \leq u(0) = u_0$, we may apply the comparison principle from Theorem II.2.4, noting that F is locally Lipschitz and cooperative on $(1 - \mathscr{L}_+^\infty) \times \mathscr{L}^\infty$ by Lemma II.2.5. We deduce that $0 \leq u(t)$ for all $t < \tau$.

Similarly, letting now $a(t) = u(t) \in \mathscr{L}^\infty$ and $b(t) \in (1 - \mathscr{L}_+^\infty)$ be the constant function 1 for all t, and remarking that $F(b(t)) = -\gamma \leq 0$, we may apply Theorem II.2.4 again, using this time the cooperativeness on $\mathscr{L}^\infty \times (1 - \mathscr{L}_+^\infty)$, to get $u(t) \leq 1$ for all $t < \tau$.

Now we prove property (ii). Let $(y, [0,\tau))$ be a solution of Equation (II.3) with $y(0) \in \Delta$. Assume that τ is a positive finite number. Property (i) asserts that $y(t) \in \Delta$, for all $0 \leq t < \tau$. Since F is bounded on Δ (see Proposition II.2.5 (ii)), $s \mapsto F(y(s))$ is integrable and:

$$\lim_{t\to\tau^-} y(t) = y(0) + \lim_{t\to\tau^-} \int_0^t F(y(s))\,\mathrm{d}s = y(0) + \int_0^\tau F(y(s))\,\mathrm{d}s.$$

The solution y can be extended up to its right boundary, *i.e.*, on $[0,\tau]$. By the Picard–Lindelöf theorem, it may thus be extended to $[0,\tau')$ for a $\tau' > \tau$. This shows that y is not maximal. We deduce that the maximal solution is defined on $\mathbb{R}_+$. □

Thanks to Proposition II.2.7, it is possible to define the semi-flow associated to the autonomous differential equation (II.3) on Δ, *i.e.*, the unique function $\phi : \mathbb{R}_+ \times \Delta \to \Delta$ solution of:

$$\begin{cases} \partial_t\phi(t,g) = F(\phi(t,g)), \\ \phi(0,g) = g. \end{cases} \tag{II.28}$$

It satisfies the semi-group property, that is, for all $g \in \Delta$ and for all $t, s \in \mathbb{R}_+$, we have:

$$\phi(t+s, g) = \phi(t, \phi(s,g)).$$

The result below is a fundamental property about the semi-flow of the SIS model. It expresses the intuitive idea that if an epidemics is worse everywhere compared to a reference state, it will remain worse compared to the evolution of this reference state in the future.

Proposition II.2.8 (Order-preserving flow)**.** *If $0 \leq g \leq h \leq 1$, then we have $\phi(t,g) \leq \phi(t,h)$ for all $t \in \mathbb{R}_+$.*

Proof. Since $\partial_t\phi(t,g) - F(\phi(t,g)) = 0$ and $\partial_t\phi(t,h) - F(\phi(t,h)) = 0$, the inequality (II.20) is satisfied on $\mathbb{R}_+$ for the paths $a : t \mapsto \phi(t,g)$ and $b : t \mapsto \phi(t,h)$. By assumption, we have also that $g = \phi(0,g) \leq \phi(0,h) = h$. Furthermore F is locally Lipschitz (see Proposition II.2.5 (ii)) and cooperative on $(1 - \mathscr{L}^\infty_+) \times \mathscr{L}^\infty$ (see Proposition II.2.5 (iv)), and $a(t) = \phi(t,g) \in (1 - \mathscr{L}^\infty_+)$ by Proposition II.2.7. Hence, we can apply Theorem II.2.4 to obtain that $\phi(t,g) \leq \phi(t,h)$ for all $t \in \mathbb{R}_+$. □

As a consequence of the previous proposition, we have the following result.

Corollary II.2.9 (Local Monotony implies Global Monotony)**.** *Let $g \in \Delta$. Suppose that there exist $0 \leq a < b$ such that, for all $t \in [a,b)$, the inequality $\phi(a,g) \leq \phi(t,g)$ (resp. $\phi(a,g) \geq \phi(t,g)$) holds. Then, $t \mapsto \phi(t,g)$ is non-decreasing (resp. non-increasing) on $[a,\infty)$.*

Proof. It is sufficient to show that $t \mapsto \phi(t,g)$ is non-decreasing on all subintervals of $[a,\infty)$ whose lengths are bounded from above by $b-a$. Let $t > s \geq a$ such that $t - s < b - a$. By assumption, we have: $\phi(a,g) \leq \phi(a+t-s, g)$. Thus, Proposition II.2.8 gives:

$$\phi(s-a, \phi(a,g)) \leq \phi(s-a, \phi(a+t-s, g)).$$

By the semi-group property of the semi-flow, this implies that $\phi(s,g) \leq \phi(t,g)$. □

Proposition II.2.10. *Let $g \in \Delta$. The path $t \mapsto \phi(t,g)$ is non-decreasing (resp. non-increasing) if and only if $F(g) \geq 0$ (resp. $F(g) \leq 0$).*

Proof. Let $g \in \Delta$, and suppose $F(g) \geq 0$. Let $a(t) = g$ for all t, and let $b(t) = \phi(t,g)$. Since $(a(t), b(t)) \in (1 - \mathscr{L}^\infty_+) \times \mathscr{L}^\infty$ for all t, and

$$a'(t) - F(a(t)) = -F(g) \leq 0 = b'(t) - F(b(t)),$$

we may apply the comparison Theorem II.2.4: for all $t \geq 0$,

$$g = a(t) \leq b(t) = \phi(t, g).$$

We may now apply Corollary II.2.9, proving that $t \mapsto \phi(t, g)$ is non-decreasing.

Now, suppose that $t \mapsto \phi(t, g)$ is non-decreasing. For all $t > 0$, the function $(\phi(t, g) - g)/t$ belongs to $\mathscr{L}_+^\infty$. Since $\mathscr{L}_+^\infty$ is closed, it follows that $F(g) = \lim_{t\to 0+}(\phi(t, g) - g)/t$ also belongs to $\mathscr{L}_+^\infty$.

The equivalence between $F(g) \leq 0$ and the fact that $t \mapsto \phi(t, g)$ is non-increasing is proved the same way. □

Now we give some results about the regularity of the semi-flow.

Proposition II.2.11 (Flow regularity). *Let $\phi : \mathbb{R}_+ \times \Delta \to \Delta$ be the semi-flow defined by Equation* (II.28).

(i) For all $g \in \Delta$, $t \mapsto \phi(t, g)$ is $\mathcal{C}^\infty$ and its repeated derivatives are bounded.

(ii) For all $t \in \mathbb{R}_+$, $g \mapsto \phi(t, g)$ is Lipschitz with respect to $\|\cdot\|$.

(iii) For all $t \in \mathbb{R}_+$, $g \mapsto \phi(t, g)$ is continuous with respect to the pointwise convergence topology.

Remark II.2.12. Stronger regularity property than (ii) could be proved as in finite dimension. Since we use only the Lipschitz continuity property, we didn't go further in this direction.

Proof. We begin with property (i). The smoothness of the semi-flow with respect to the time variable can be shown by recurrence in a classical way. We have indeed:

$$\partial_t\phi(t, g) = F(\phi(t, g)), \qquad \partial_t^2\phi(t, g) = \mathcal{D}F[\phi(t, g)](\partial_t\phi(t, g)), \qquad \dots$$

Since F is of class $\mathcal{C}^\infty$ and its repeated derivatives are bounded on Δ (see (i) and (ii) in Proposition II.2.5), the function $t \mapsto \phi(t, g)$ is of class $\mathcal{C}^\infty$ and its repeated derivatives are bounded for all $g \in \Delta$.

We prove (ii). Recall that, since ϕ is the semi-flow associated to Equation (II.25), the following equality holds for all $g \in \Delta$ and $t \in \mathbb{R}_+$:

$$\phi(t, g) = g + \int_0^t F(\phi(s, g))\, \mathrm{d}s. \tag{II.29}$$

Let $g, h \in \Delta$. We have the following control:

$$\begin{aligned}
\| \phi(t, g) - \phi(t, h) \| &\leq \| g - h \| + \int_0^t \| F(\phi(s, g)) - F(\phi(s, h)) \| \,\mathrm{d}s \\
&\leq \| g - h \| + C \int_0^t \| \phi(s, g) - \phi(s, h) \| \,\mathrm{d}s,
\end{aligned}$$

where C is the Lipschitz coefficient of F on Δ (see Proposition II.2.5 (ii)). We conclude by applying Grönwall's inequality.

We prove property (iii). Let $(g_n,\ n \in \mathbb{N})$ be a sequence of functions in Δ converging pointwise toward $g \in \Delta$. We define for $n \in \mathbb{N}$:

$$\overline{g}_n = \sup_{j\geq n} g_j \quad \text{and} \quad \underline{g}_n = \inf_{j\geq n} g_j.$$

The sequence $(\overline{g}_n,\ n \in \mathbb{N})$ is non-increasing while $(\underline{g}_n,\ n \in \mathbb{N})$ is non-decreasing. We also have $\underline{g}_n \leq g_n \leq \overline{g}_n$ for all natural number n. Since the semi-flow is order-preserving by Proposition II.2.8, the following inequalities hold for all $(t, x) \in \mathbb{R}_+ \times \Omega$ and all $n \in \mathbb{N}^*$:

$$\phi(t, \underline{g}_{n-1})(x) \leq \phi(t, \underline{g}_n)(x) \leq \phi(t, g_n)(x) \leq \phi(t, \overline{g}_n)(x) \leq \phi(t, \overline{g}_{n-1})(x). \tag{II.30}$$

Thus, we can define two measurable functions $v, w : \mathbb{R}_+ \times \Omega \to [0,1]$ by:

$$v(t,x) = \lim_{n\to\infty} \phi(t, \underline{g}_n)(x), \qquad w(t,x) = \lim_{n\to\infty} \phi(t, \overline{g}_n)(x),$$

for all $(t,x) \in \mathbb{R}_+ \times \Omega$. Notice that $v(t,x) \leq w(t,x)$ by construction.

Fix $x \in \Omega$ and $t \geq 0$. We have:

$$\phi(t, \overline{g}_n)(x) = \overline{g}_n(x) + \int_0^t F(\phi(s, \overline{g}_n))(x)\,\mathrm{d}s.$$

The sequence of functions $(\overline{g}_n(x),\ n \in \mathbb{N})$ converges to $g(x)$ while the sequence of functions $(\phi(s, \overline{g}_n),\ n \in \mathbb{N})$ converges pointwise to $w(s,\cdot) \in \Delta$ for all $s \geq 0$. By continuity (see Proposition II.2.5 (iii)), $F(\phi(s, \overline{g}_n))(x)$ converges to $F(w(s,\cdot))(x)$. Furthermore, the functions $s \mapsto F(\phi(s, \overline{g}_n))(x)$ are uniformly bounded since F is bounded on Δ (see Proposition II.2.5 (ii)). Hence, we deduce from the dominated convergence theorem that:

$$w(t,x) = g(x) + \int_0^t F(w(s,\cdot))(x)\,\mathrm{d}s.$$

The previous equality is true for all $x \in \Omega$ and $t \geq 0$. Since $t \mapsto \phi(t,g)$ is the only solution of (II.3) having g as initial condition, we have necessarily $w(t,\cdot) = \phi(t,g)$. We prove that $v(t,\cdot) = \phi(t,g)$ the same way. Letting n go to infinity in (II.30) proves that $\phi(t, g_n)$ converges pointwise to $\phi(t,g)$, for all $t \geq 0$. □

II.2.5 Equilibria

A function $g \in \Delta$ is an equilibrium of the dynamical system (Δ, ϕ) (also called a stationary point) if for all $t \in \mathbb{R}_+$, $\phi(t,g) = g$. The latter assertion is equivalent to $F(g) = 0$. The function equal to 0 everywhere is a trivial stationary point. In mathematical epidemiology, the other equilibria, if they exist, are called *endemic states* because they model a situation where the infection is constantly maintained at a baseline level in the population.

The following result gives an easy way to identify those special states in the system. It is a well-known fact in dynamical system theory, we give a short proof for completeness.

Proposition II.2.13 (Limit points are equilibria)**.** *Let $g \in \Delta$. If $t \mapsto \phi(t,g)$ converges pointwise to a limit $h^* \in \Delta$ when t goes to ∞, then the function h^* is an equilibrium.*

Proof. For all $x \in \Omega$ and $s \geq 0$, we have

$$\phi(s, h^*)(x) = \lim_{t\to\infty} \phi(s, \phi(t,g))(x) = \lim_{t\to\infty} \phi(s+t, g)(x) = h^*(x),$$

where the first inequality follows from the continuity of ϕ with respect to the pointwise convergence topology given in Proposition II.2.11 (iii). Thus, h^* is an equilibrium. □

In the next remark, we check that any equilibrium is continuous with respect to an intrinsic distance on Ω based on κ and γ.

Remark II.2.14 (Continuity of the equlibria). We consider for all $x, y \in \Omega$:

$$r(x,y) = \|\kappa(x,\cdot) - \kappa(y,\cdot)\|_{\mathrm{TV}} + |\gamma(x) - \gamma(y)|,$$

where $\|\cdot\|_{\mathrm{TV}}$ is the total variation norm. The function r defines a pseudo-metric on the space Ω. This pseudo-metric can be thought as an extension of the neighborhood distance on graphons (see [111, Section 13.3]). Notice that the Borel σ-field associated to the topology defined by r is included in $\mathscr{F}$ since γ is measurable and κ is a kernel.

We have that if h^* is an equilibrium of the dynamical system (Δ, ϕ), then it is continuous with respect to r. Indeed, we have for all $x \in \Omega$:

$$h^*(x) = \frac{\lambda(x)}{\lambda(x) + \gamma(x)} \quad \text{with} \quad \lambda(x) = \int_\Omega h^*(z)\,\kappa(x, \mathrm{d}z).$$

Both λ and γ are continuous with respect to r and the function $(a,b) \mapsto a/(a+b)$ is continuous on $\mathbb{R}_+ \times \mathbb{R}_+^*$. This implies that h^* is continuous.

II.2.6 The maximal equilibrium

As a consequence of Proposition II.2.7 and Corollary II.2.9, the path $t \mapsto \phi(t, 1)$ is non-increasing and bounded below by 0. Thus, the path $t \mapsto \phi(t, 1)$ converges pointwise to a limit say $\mathfrak{g}$ when t goes to infinity:

$$\mathfrak{g}(x) = \lim_{t \to +\infty} \phi(t, 1)(x), \qquad \forall x \in \Omega. \tag{II.31}$$

Proposition II.2.15. *Let $\mathfrak{g}$ be defined by* (II.31). *We have the following properties.*

(i) The function $\mathfrak{g}$ is the maximal equilibrium of the dynamical system (Δ, ϕ), i.e., if h^* is an equilibrium, then $h^* \leq \mathfrak{g}$.

(ii) For all $\mathfrak{g} \leq g \leq 1$, $\phi(t, g)$ converges pointwise to $\mathfrak{g}$ as t goes to infinity.

Proof. We first prove property (i). The function $\mathfrak{g}$ is an equilibrium according to Proposition II.2.13. Let h^* be another equilibrium in Δ. By Proposition II.2.8, $h^* = \phi(t, h^*) \leq \phi(t, 1)$ for all t; sending t to infinity yields $h^* \leq \mathfrak{g}$. The function $\mathfrak{g}$ is thus the maximal equilibrium.

To prove property (ii), we consider $\mathfrak{g} \leq g \leq 1$. By Proposition II.2.8, we know that:

$$\mathfrak{g} = \phi(t, \mathfrak{g}) \leq \phi(t, g) \leq \phi(t, 1).$$

Since the rightmost term converges to $\mathfrak{g}$ by (II.31), this implies that $\phi(t, g)$ converges to $\mathfrak{g}$ as t tends to infinity for the pointwise convergence. □

Remark II.2.16. Since $\gamma(x) > 0$ for all $x \in \Omega$ according to Assumption II.1 and $F(\mathfrak{g}) = 0$, we have that $\mathfrak{g}(x) < 1$ for all $x \in \Omega$.

There is no closed-form formula for $\mathfrak{g}$ in the general case, even in a finite dimensional model. However, if the function $x \mapsto \kappa(x, \Omega)/\gamma(x)$ is constant, then the formula used for the one-group model can be extended.

Proposition II.2.17. *Suppose that there exists $C \in \mathbb{R}_+$ such that $\kappa(x, \Omega)/\gamma(x) = C$ for all $x \in \Omega$. Then, $\mathfrak{g}$ is a constant function equal to* $\max(0, 1 - 1/C)$.

Proof. It is straightforward to check that the function $x \mapsto \max(0, 1 - 1/C)$ is an equilibrium. Now, we prove that it is maximal. Let $h^* \in \Delta$ be an equilibrium. From $F(h^*)/\gamma = 0$, we obtain the inequality:

$$h^* \leq C(1 - h^*) \, \| h^* \| \, .$$

Taking a sequence $(x_n, n \in \mathbb{N})$ such that $h^*(x_n)$ converges to $\| h^* \|$, we obtain at the limit that $\| h^* \| \leq C(1 - \| h^* \|) \| h^* \|$. It follows that $\| h^* \| \leq \max(0, 1 - 1/C)$. □

Since we cannot determine $\mathfrak{g}$ in the general case, the important question that naturally arises is to find out whether the epidemic can survive in the population or if it will die out whatever the initial condition is, *i.e.*, we have to determine if $\mathfrak{g}(x) = 0$ for all $x \in \Omega$. In the following, we answer this question with Assumption II.2 which imposes further conditions on the transmission kernel κ and the recovery rate γ.

II.3 Tools from operator theory

In this section, we introduce some tools that we will use in Section II.4.

II.3.1 Compactness, weak compactness and the Dunford Pettis property

Let X and Y be Banach spaces and $T : X \to Y$ be a bounded linear operator. Recall that T is *compact* if the image of the unit ball in X is relatively compact for the strong topology on Y. Similarly it is *weakly compact* if the image of the unit ball in X by T is relatively compact for the weak topology on Y (that is, its weak closure is weakly compact). We recall the following results on weak compactness ([48, Corollary VI.4.3, Theorem VI.4.5]).

Theorem II.3.1 (Weak compactness)**.** *If X and Y are Banach spaces, one of which is reflexive, and T is a bounded operator from X to Y, then T is weakly compact.*

The composition of a bounded operator and a weakly compact operator, in any order, is weakly compact.

The following important property is given in [136] (the space $\mathscr{L}^\infty$ is a so-called "abstract-max space" by [136, II §7, example 3 p. 103], so that [136, Theorem II.9.9] applies).

Theorem II.3.2 (Dunford Pettis property)**.** *If Y is a Banach space, and if $T : \mathscr{L}^\infty \to Y$ is weakly compact, then T is absolutely continuous, that is, it maps weakly convergent sequences in $\mathscr{L}^\infty$ to strongly convergent sequences in Y.*

Corollary II.3.3. *If $T : \mathscr{L}^\infty \to \mathscr{L}^\infty$ is weakly compact then T^2 is compact.*

Proof. Let x_n be a bounded sequence in $\mathscr{L}^\infty$. Since T is weakly compact there exists a subsequence such that (Tx_n) converges weakly. By the Dunford Pettis property, T is absolutely continuous so $T(Tx_n)$ converges strongly along the subsequence. Therefore T^2 is compact. □

II.3.2 Banach lattices

Let us first recall standard definitions on Banach lattices; we refer the reader to the standard texts [159] and [136] for a more detailed introduction to the subject. Banach lattices provide a convenient framework to study positive operators and generalizations of the Perron–Frobenius theorem; the two examples we have in mind are the space $\mathscr{L}^\infty$ of bounded functions and the space $L^p(\mu)$.

Let $(X, \leq)$ be a set equipped with a partial order. The set X is a *lattice* if for any x and y in X, there exist two elements i and s in X such that for all z,

$$(z \leq x \text{ and } z \leq y) \implies z \leq i; \qquad (x \leq z \text{ and } y \leq z) \implies s \leq z.$$

The infimum i and supremum s are customarily denoted $x \wedge y$ and $x \vee y$ respectively. A *Riesz space* is a vector space X endowed with a lattice structure (denoted $\leq$), such that the two following compatibility conditions are satisfied:

Translation invariance For all x, y and z in X, if $x \leq y$ then $x + z \leq y + z$.

Positive homogeneity For all x, y in X, if $x \leq y$, then $\lambda x \leq \lambda y$, for all non negative scalar $\lambda \geq 0$.

The *absolute value* $|x|$ of an element x of a Riesz space is defined by $|x| = x \vee (-x)$. We proceed with some further definitions.

Definition II.3.4. *A Banach lattice $(X, \leq, \|\cdot\|)$ is a Riesz space $(X, \leq)$ equipped with a complete norm $\|\cdot\|$ and such that, for any $x, y \in X$, we have:*

$$|x| \leq |y| \implies \|x\| \leq \|y\|. \tag{II.32}$$

In the Banach lattice X, the positive cone:

$$X_+ = \{ x \in E : x \geq 0 \}.$$

is a proper cone, as it is a closed (see Theorem 15.1 (ii) in [159]) convex set such that $\lambda X_+ \subset X_+$ for all $\lambda \in \mathbb{R}_+$, and $X_+ \cap (-X_+) = \{0\}$. It is also a reproducing cone ($X = X_+ - X_+$) as every element x in X can be decomposed as $x = (x \vee 0) - ((-x) \vee 0)$ and $y \vee 0 \in X_+$ for all $y \in X$.

II.3.3 Spectral analysis in Banach lattices

The main result of this section concerns the spectrum of operators on Banach lattices. Let us first recall a few classical definitions of spectral theory in Banach spaces. Let $(X, \|\cdot\|)$ be a Banach space. The *spectrum* $\sigma(A)$ of a bounded operator A on X is the set of all complex numbers λ such that $A - \lambda\mathrm{Id}$ does not have a bounded inverse operator. It is well known that the spectrum of a bounded operator is a compact set in $\mathbb{C}$. The *essential spectrum* of an operator is the part that cannot be removed by a compact perturbation:

$$\sigma_{\mathrm{ess}}(A) = \bigcap_{P \text{ compact operator}} \sigma(A + P).$$

Note that there are several conflicting definitions of the essential spectrum, see [49, Section 1.4]; by [49, Theorem 9.1.4, p 422], our definition corresponds to $\sigma_{e4}(A)$ defined p 37 in [49]. Let us remark that other definitions would lead to the same essential spectral radius in the definition below, see [49, Corollary 1.4.11].

For a bounded operator A on X, the spectral bound, the spectral radius and the essential spectral radius are defined as:

$$s(A) = \sup\{\operatorname{Re}(\lambda) \,:\, \lambda \in \sigma(A)\}, \tag{II.33}$$

$$\rho(A) = \sup\{|\lambda| \,:\, \lambda \in \sigma(A)\} = \lim_{n\to+\infty} \|A^n\|^{1/n} = \inf_{n\in\mathbb{N}^*} \|A^n\|^{1/n}, \tag{II.34}$$

$$r_{\mathrm{ess}}(A) = \sup\{|\lambda| \,:\, \lambda \in \sigma_{\mathrm{ess}}(A)\}, \tag{II.35}$$

respectively, with the convention that $\sup\emptyset = 0$. Since $\sigma_{\mathrm{ess}}(A) \subset \sigma(A)$, we get:

$$r_{\mathrm{ess}}(A) \leq \rho(A) \leq \|A\|. \tag{II.36}$$

The spectral theory of positive bounded operator on Banach lattice extends the Perron-Frobenius theory in infinite dimension. Let A be a positive operatoron a Banach lattice $(X, \leq, \|\cdot\|)$, that is, $AX_+ \subset X_+$, such that its spectral radius $\rho(A)$ is positive. Recall $X_+^\star$ is the dual cone of X_+. A vector $x \in X_+\setminus\{0\}$ (resp. $x^\star \in X_+^\star\setminus\{0\}$) such that $Ax = \rho(A)x$ (resp. $A^\star x^\star = \rho(A)x^\star$) is called a right (resp. left) Perron eigenvector. We have the following important result.

Theorem II.3.5. *Let $(X, \leq, \|\cdot\|)$ be a Banach lattice. Let A, B be positive bounded operators on X. We have the following properties.*

(i) If $B - A$ is a positive operator, then $\rho(A) \leq \rho(B)$.

(ii) The spectral radius $\rho(A)$ belongs to $\sigma(A)$ and thus $\rho(A) = s(A)$.

(iii) If $r_{\mathrm{ess}}(A) < \rho(A)$, then, there exists $x \in X_+\setminus\{0\}$ such that: $Ax = \rho(A)x$.

Proof. Property (i) is proved in [114, Theorem 4.2]. Property (ii) is proved in [137] (notice that (II.32) implies that X_+ is normal in the setting of [137]), see also [159, Lemma 41.1.(ii)]. Property (iii) was shown by Nussbaum in [123, Corollary 2.2] (notice that a reproducing cone is total), where the essential spectrum in [123] is defined in [124] and corresponds to $\sigma_{e5}(A)$ in [49, p. 37]. However, the essential spectral radius of $\sigma_{e5}(A)$ is equal to $r_{\mathrm{ess}}(A)$ the essential spectral radius of $\sigma_{e4}(A)$, according to [49, Theorem I.4.10]. □

It A is assumed to be a compact operator, then Theorem II.3.5 (iii) is the so called Krein-Rutman theorem, see [159, Theorem 41.2]. We will also need the following result proved in [62, Propositions 2.1-2.2].

Proposition II.3.6 (Collatz-Wielandt inequality)**.** *Let $(X, \leq, \|\cdot\|)$ be a Banach lattice and A be a positive bounded operator on X. We have:*

$$\sup\{\lambda \in \mathbb{R} \,:\, \exists x \in X_+\setminus\{0\},\, Ax \geq \lambda x\} \leq \rho(A).$$

II.3.4 The Banach lattice of bounded measurable functions

The Banach space $(\mathscr{L}^\infty, \|\cdot\|)$ equipped with the partial order $\leq$ defined by the proper cone $\mathscr{L}^\infty_+$ from (II.21) is a Banach lattice.

Let ν be a finite signed measure on $(\Omega, \mathscr{F})$. For $g \in \mathscr{L}^\infty$, we write $\langle \nu, g\rangle = \int_\Omega g(x)\, \nu(\mathrm{d}x)$ and thus identify ν as an element of $\mathscr{L}^{\infty,\star}$, the dual space of $\mathscr{L}^\infty$ (recall that $\mathscr{L}^{\infty,\star}$ can be identified as the space of bounded and finitely additive signed measures on $(\Omega, \mathscr{F})$.)

Let μ be a given finite positive measure on $(\Omega, \mathscr{F})$. For $q \in (1, +\infty)$, denote by $(L^q(\mu), \|\cdot\|_q)$ the usual Banach space of real-valued measurable functions f defined on $(\Omega, \mathscr{F})$ such that $\|f\|_q = \left(\int_\Omega |f(x)|^q\, \mu(\mathrm{d}x)\right)^{1/q}$ is finite and where we have identified functions which agree μ-almost everywhere.

Let ι be the natural linear application ι from $\mathscr{L}^\infty$ to $L^p(\mu)$, with $p = q/(q-1)$ the conjugate of q, and $\iota^\star$ its dual. For $f \in L^q(\mu)$, we can see $\iota^\star(f)$ as the bounded σ-finite signed measure $f(x)\mu(\mathrm{d}x)$ elements of $\mathscr{L}^{\infty,\star}$. By convention, for $f \in L^q(\mu)$ and g in $\mathscr{L}^\infty$, we write:

$$\langle f, g\rangle = \langle \iota^\star(f), g\rangle = \int_\Omega f(x) g(x)\, \mu(\mathrm{d}x).$$

Let k be a non-negative measurable function defined on $(\Omega \times \Omega, \mathscr{F} \otimes \mathscr{F})$ such that :

$$\sup_{x\in\Omega} \int \mathrm{k}(x, y)\, \mu(\mathrm{d}y) < \infty.$$

We define the integral operator T_k as the operator T_κ defined by (II.22) with kernel $\kappa(x, \mathrm{d}y) = \mathrm{k}(x, y)\, \mu(\mathrm{d}y)$. Let $q \in (1, +\infty)$. We assume the following condition holds:

$$\sup_{x\in\Omega} \int \mathrm{k}(x, y)^q\, \mu(\mathrm{d}y) < \infty. \tag{II.37}$$

Then, we can also define the bounded operator:

$$\begin{aligned} \tilde{T}_\mathrm{k} \quad : \quad & L^p(\mu) \to \mathscr{L}^\infty \\ & g \mapsto \left(x \mapsto \int_\Omega g(y)\, \mathrm{k}(x, y)\, \mu(\mathrm{d}y)\right). \end{aligned}$$

With this notation, $T_\mathrm{k} = \tilde{T}_\mathrm{k}\iota$. We also define a bounded operator $\hat{T}_\mathrm{k}$ from $L^p(\mu)$ to $L^p(\mu)$:

$$\hat{T}_\mathrm{k} = \iota \tilde{T}_\mathrm{k}. \tag{II.38}$$

To sum up we have the following commutative diagram:

$$\begin{tikzcd} \mathscr{L}^\infty \arrow[r, "\iota"] \arrow[d, "T_\mathrm{k}"'] & L^p(\mu) \arrow[dl, "\tilde{T}_\mathrm{k}"'] \arrow[d, "\hat{T}_\mathrm{k}"] \\ \mathscr{L}^\infty \arrow[r, "\iota"] & L^p(\mu) \end{tikzcd}$$

The following lemma has a fundamental importance for the development of Section II.4. The last property on connected integral operator is part of the Perron-Jentzsch theorem, see [136, Theorem V.6.6 and Example V.6.5.b].

Lemma II.3.7. *Let* k *be a non-negative measurable function defined on* $(\Omega \times \Omega, \mathscr{F} \otimes \mathscr{F})$ *such* (II.37) *holds for some* $q \in (1, +\infty)$. *Then, the positive bounded operators* $T_\mathrm{k} : \mathscr{L}^\infty \to \mathscr{L}^\infty$ *and* $\hat{T}_\mathrm{k} : L^p(\mu) \to L^p(\mu)$, *with* $p = q/(q-1)$, *satisfies:*

(i) If $g = 0$ μ*-a.e. then we have* $T_\mathrm{k} g = 0$.

(ii) *The operator* T_{k} *is weakly compact.*

(iii) *The operators* T_{k}^2 *and* $\hat{T}_{\mathrm{k}}$ *are compact.*

(iv) *The operators* T_{k} *and* $\hat{T}_{\mathrm{k}}$ *have the same spectrum, and thus* $\rho(T_{\mathrm{k}}) = \rho(\hat{T}_{\mathrm{k}})$.

(v) *If* $\rho(T_{\mathrm{k}}) > 0$, *then the operator* T_{k} *has a right Perron eigenvector in* $\mathscr{L}^\infty_+ \setminus \{0\}$ *and a left Perron eigenvector in* $L^q_+(\mu)\setminus\{0\} \subset \mathscr{L}^{\infty,\star}_+ \setminus \{0\}$.

(vi) *If* k *is connected in the sense of Assumption II.3, then* $\rho(T_{\mathrm{k}}) > 0$ *and the right and left Perron eigenvector are unique (up to a multiplicative constant) and are* μ*-a.e. positive, with the left Perron eigenvector seen as an element of* $L^q_+(\mu)\setminus\{0\}$.

Proof. Property (i) is straightforward.

To prove property (ii), one may write T_{k} as the composition $T_{\mathrm{k}} = \tilde{T}_{\mathrm{k}} \circ \iota$, where $\tilde{T}_{\mathrm{k}}$ is bounded, and ι is weakly compact by the first part of Theorem II.3.1 since $L^p(\mu)$ is reflexive ([48, Corollary IV.8.2]). By the second part of Theorem II.3.1, T_{k} is weakly compact.

The first part of Property (iii), that is, the compactness of T_{k}^2, follows directly from Corollary II.3.3. Consider now the operator $\hat{T}_{\mathrm{k}} = \iota \circ \tilde{T}_{\mathrm{k}}$. By Theorem II.3.1, both $\tilde{T}_{\mathrm{k}}$ and ι are weakly compact. From any bounded sequence x_n in $L^p(\mu)$, we may extract a subsequence such that $\tilde{T}_{\mathrm{k}} x_n$ converges weakly in $\mathscr{L}^\infty$. By the Dunford Pettis property (Theorem II.3.2), the weakly compact operator ι maps this weakly convergent subsequence to a strongly convergent subsequence in $L^p(\mu)$. Therefore $\hat{T}_{\mathrm{k}}$ is compact.

Let us prove property (iv). If Ω is finite then the operators T_{k} and $\hat{T}_{\mathrm{k}}$ coincide and there is nothing to prove. So, we assume that Ω is infinite. In this case, $\sigma_{\mathrm{ess}}(T_{\mathrm{k}})$ and $\sigma_{\mathrm{ess}}(\hat{T}_{\mathrm{k}})$ are non empty according to [92, Footnote 2, p. 243]. As $\hat{T}_{\mathrm{k}}$ and T_{k}^2 are compact, we deduce from [48, Theorems VII.4.5 and VII.4.6] respectively, that the essential spectra of $\hat{T}_{\mathrm{k}}$ and T_{k} are reduced to $\{0\}$, and that the non-null elements of their spectrum are eigenvalues. Then, use that $\iota T_{\mathrm{k}} = \hat{T}_{\mathrm{k}} \iota$ and property (i), to deduce that if $f \in \mathscr{L}^\infty \setminus \{0\}$ is an eigenvector of T_{k}, then $\iota(f)$ belongs to $L^p(\mu)\setminus\{0\}$ thanks to property (i) and that $\iota(f)$ is thus an eigenvector of $\hat{T}_{\mathrm{k}}$ corresponding to the same eigenvalue. If $v \in L^p(\mu)\setminus\{0\}$ is an eigenvector of $\hat{T}_{\mathrm{k}}$ corresponding to the eigenvalue λ, then $f = \tilde{T}_{\mathrm{k}}(v)$ belongs to $\mathscr{L}^\infty$ and $f \neq 0$ (as $\iota(f) = \hat{T}_{\mathrm{k}}(v) = \lambda v$). We have $T_{\mathrm{k}}(f) = \tilde{T}_{\mathrm{k}} \iota \tilde{T}_{\mathrm{k}}(v) = \tilde{T}_{\mathrm{k}}(\hat{T}_{\mathrm{k}}(v)) = \lambda \tilde{T}_{\mathrm{k}}(v) = \lambda f$, thus λ is also an eigenvalue of T_{k}. We deduce that $\sigma(T_{\mathrm{k}}) = \sigma(\hat{T}_{\mathrm{k}})$.

Let us now prove property (v). We have seen that $\sigma_{\mathrm{ess}}(T_{\mathrm{k}}) \subset \{0\}$ and thus $r_{\mathrm{ess}}(T_{\mathrm{k}}) = 0$. According to Theorem II.3.5 (iii) (or the Krein-Rutman theorem) there exists a right Perron eigenvector for T_{k}. Since $\hat{T}_{\mathrm{k}}^\star$ is a compact operator, thanks to Schauder Theorem [48, Theorem VI.5.2], with the same spectrum as $\hat{T}_{\mathrm{k}}$, thanks to [48, Lemma VII.3.7], and which is clearly positive, we deduce from Theorem II.3.5 (iii) that there exists a right Perron eigenvector, $v^\star \in L^q_+(\mu)\setminus\{0\}$, for $\hat{T}_{\mathrm{k}}^\star$. Since $T_{\mathrm{k}}^\star \iota^\star = \iota^\star \hat{T}_{\mathrm{k}}^\star$, we deduce that $\iota^\star(v^\star)$, and thus $v^\star$ by convention, is also a left Perron eigenvector for T_{k}. This gives property (v).

Finally let us prove property (vi). Set $\lambda = \rho(T_{\mathrm{k}}) = \rho(\hat{T}_{\mathrm{k}})$, see property (iv). According to the Perron-Jentzsch theorem [136, Theorem V.6.6 and Example V.6.5.b], since k is connected in the sense of Assumption II.3, we have $\lambda > 0$ and there exists a unique (up to a multiplicative constant) eigenvector v of $\hat{T}_{\mathrm{k}}$ associated to the eigenvalue λ, and it can chosen such that μ-a.e. $v > 0$. According to the proof of property (iv), we get that $f = \tilde{T}_{\mathrm{k}}(v)$ is an eigenvector of T_{k} associated to λ. Notice that $f \geq 0$ as $\mathrm{k} \geq 0$. Since $\iota(f) = \hat{T}_{\mathrm{k}}(v) = \lambda v$, we deduce that μ-a.e. $f > 0$. Assume that $g \in \mathscr{L}^\infty \setminus \{0\}$ is a right Perron eigenvector of T_{k}, then $\iota(g)$ is a right Perron eigenvector of $\hat{T}_{\mathrm{k}}$ and thus (up to a multiplicative constant chosen to be equal to λ), we have μ-a.e. $\iota(g) = \lambda v = \iota(f)$. We deduce that μ-a.e. $g - f = 0$ and thanks to property (i), we deduce that $\lambda(f - g) = T_{\mathrm{k}}(f - g) = 0$. So the right Perron eigenvector of T_{k} is unique and μ-a.e. positive.

Let $f^\star$ be a left Perron eigenvector of T_{k}. Then $v^\star = \tilde{T}_{\mathrm{k}}^\star(f^\star)$ is an eigenvector of $\hat{T}_{\mathrm{k}}^\star$ associated to λ and $v^\star \in L^q_+(\mu)\setminus\{0\}$ as $\mathrm{k} \geq 0$. By the Perron-Jentzsch theorem, we get that $v^\star$ is unique (up to a multiplicative constant) and that μ-a.e $v^\star > 0$. Since $\iota^\star(v^\star) = \lambda f^\star$, we deduce that μ-a.e $f^\star > 0$ and that $f^\star$ is unique (up to a multiplicative constant). $\square$

Remark II.3.8. As a consequence of Lemma II.3.7 (i), under Assumption II.2, if h^* is an equilibria which is μ-a.e. equal to 0, then it is equal to 0 everywhere.

II.4 Infinite-dimensional SIS model when the kernel has a density

The objective of this section is to study the long time behavior of the solutions of (II.3) under Assumption II.1 and Assumption II.2 (but for Section II.4.2 where the latter is not assumed). Recall the definition of the spectral bound given in (II.33). We will consider the spectral bound $s(T_k - \gamma)$ of the bounded operator $T_k - \gamma$ on $\mathscr{L}^\infty$ to characterize three different regimes: sub-critical, critical and super-critical, corresponding to the cases $s(T_k - \gamma) <, =, > 0$ respectively. In the first part of the section, we establish a link between $s(T_k - \gamma)$ and the basic reproduction number $\mathfrak{R}_0 = \rho(T_{k/\gamma})$ associated to (II.3).

II.4.1 Basic reproduction number and spectral bound

Recall that Assumption II.1 is in force. If we assume $\inf \gamma > 0$, then the operator $T_{\kappa/\gamma}$, where κ/γ is the kernel defined by $(\kappa/\gamma)(x, \mathrm{d}y) = \kappa(x, \mathrm{d}y)/\gamma(y)$ is bounded. The following result is a direct consequence of a theorem of Thieme [152, Theorem 3.5].

Proposition II.4.1 (Equivalent conditions for criticality)**.** *If* $\inf \gamma > 0$*, then* $\rho(T_{\kappa/\gamma}) - 1$ *has the same sign as* $s(T_\kappa - \gamma)$ *(i.e. these two numbers are simultaneously negative, zero, or positive).*

Proof. Consider the operators $A = T_\kappa - \gamma$ and $B = -\gamma$, where $-\gamma$ is the operator corresponding to the multiplication by $-\gamma$. It is clear from [152, Definition 3.1] that the operator B is a resolvent-positive operator, as the operator $\lambda - B = \lambda + \gamma$ is invertible and its inverse is positive for $\lambda > 0$. We also get that $s(B) = s(-\gamma) = -\inf \gamma < 0$. Let $Q = A + \|\gamma\|$. The operator Q is positive. For $\lambda > \rho(Q)$, the resolvent operator $(\lambda - Q)^{-1}$ is also positive since, thanks to the Neumann series expansion, we have:

$$(\lambda - Q)^{-1} = \sum_{i=0}^{\infty} \frac{1}{\lambda^{i+1}} Q^i \geq 0.$$

We deduce that $(\lambda - A)$ is invertible and its inverse is positive for $\lambda > \rho(Q) - \|\gamma\|$, thus A is resolvent-positive. Applying [152, Theorem 3.5] (notice it is required that $\mathscr{L}^\infty_+$ is normal, which is the case, see [33, Proposition 19.1], as the norm $\|\cdot\|$ is monotonic: $0 \leq f \leq g$ implies $\|f\| \leq \|g\|$), we deduce that $s(A)$ has the same sign as $\rho(-(A - B)B^{-1}) - 1 = \rho(T_{\kappa/\gamma}) - 1$. □

Notice that under Assumption II.2, we have by definition $\mathfrak{R}_0 = \rho(T_{\kappa/\gamma}) = \rho(T_{k/\gamma})$, as k is the density of κ with respect to μ. In what follows, we also write T_k for T_κ. According to Assumption II.2 (see (II.8)), the operator $T_{k/\gamma}$ defined by (II.11) is a bounded operator on $\mathscr{L}^\infty$ which satisfies the integrability condition of Lemma II.3.7 with $\mathrm{k}(x, y) = k(x, y)/\gamma(y)$.

We wish to prove a result similar to Proposition II.4.1 without assuming that $\inf \gamma > 0$. In this case however, $\mathfrak{R}_0 - 1$ and $s(T_k - \gamma)$ may have different signs: For instance, if one takes $T_k = 0$ and $\inf \gamma = 0$, then we clearly have $s(T_k - \gamma) = s(-\gamma) = 0$ and $\mathfrak{R}_0 = 0 < 1$. In order to get a result similar to Proposition II.4.1, we must therefore settle for a weaker conclusion, which will however be sufficient for our purposes.

Proposition II.4.2 (Equivalent conditions for the supercritical regime)**.** *Suppose Assumption II.2 is in force. Then, the following assertions are equivalent:*

(i) $s(T_k - \gamma) > 0$.

(ii) $\mathfrak{R}_0 > 1$.

(iii) There exists $\lambda > 0$ *and* $w \in \mathscr{L}^\infty_+ \setminus \{0\}$ *such that:*

$$T_k(w) - \gamma w = \lambda w. \tag{II.39}$$

Proof. It is immediate that property (iii) implies property (i).

We assume property (i) and prove (ii). Let $a \in (0, s(T_k-\gamma))$, so that $s(T_k-(\gamma+a)) = s(T_k-\gamma)-a > 0$. Using Proposition II.4.1 (with γ replaced by $\gamma + a$), we get that $r\left(T_{k/(\gamma+a)}\right) > 1$. Since $\rho(T_{k/\gamma}) \geq r\left(T_{k/(\gamma+a)}\right)$ according to Theorem II.3.5 (i), property (ii) is shown.

We assume property (ii) and prove property (iii). For any non-negative real number $a \geq 0$, let $\psi(a) = \rho(T_{k/(\gamma+a)})$. Property (ii) exactly means that:

$$\psi(0) > 1. \tag{II.40}$$

Moreover, it follows from the inequality $\rho(T_{k/(\gamma+a)}) \leq \left\| T_{k/(\gamma+a)} \right\| \leq \| T_k \| / a$ (use (II.36) for the first inequality), that:

$$\lim_{a\to\infty} \psi(a) = 0. \tag{II.41}$$

Equation (II.8) of Assumption II.2 enables to apply Lemma II.3.7 (iii) and we obtain that all the operators $T_{k/(\gamma+a)}$, for $a \in \mathbb{R}_+$, are power compact (as $T^2_{k/(\gamma+a)}$ is compact). According to [99, Theorem p. 21], their spectra are totally disconnected. Moreover, the function $a \mapsto T_{k/(\gamma+a)}$ mapping $\mathbb{R}_+$ to $\mathcal{L}(\mathscr{L}^\infty)$ is continuous as (II.8) holds. Indeed, for all $0 \leq a_1 \leq a_2$, thanks to Hölder's inequality, we have:

$$\left\| T_{k/(\gamma+a_1)} - T_{k/(\gamma+a_2)} \right\| \leq \left\| \frac{a_2 - a_1}{\gamma + a_2} \right\|_p \sup_{x\in\Omega} \left(\int_\Omega k(x,y)^q / \gamma(y)^q \, \mu(\mathrm{d}y) \right)^{1/q},$$

and by dominated convergence $\left\| \frac{a_2-a_1}{\gamma+a_2} \right\|_p$ converges to 0 as $|a_2 - a_1|$ converges to 0. Thanks to [119, Theorem 11], we get that the application $a \mapsto \sigma(T_{k/(\gamma+a)})$ mapping $\mathbb{R}_+$ to the set $\mathcal{K}(\mathbb{C})$ of non-empty compact subsets endowed with the Hausdorff distance (see Section II.8 for the definition of the Hausdorff distance) is continuous. Hence, the function ψ is continuous according to Lemma II.8.1. From the continuity of ψ and Equations (II.40) and (II.41), we conclude that there exists $\lambda > 0$ such that $\psi(\lambda) = 1$. According to Lemma II.3.7 (v), there exists a function $v \in \mathscr{L}^\infty_+ \setminus \{0\}$ such that:

$$T_k\left(\frac{v}{\gamma + \lambda}\right) = v.$$

Then, Equation (II.39) holds with $w = v/(\gamma + \lambda)$ which proves property (iii). □

Remark II.4.3. Using Lemma II.3.7 (i), it is easy to show that w in Proposition II.4.2 (iii) should satisfy $\int_\Omega w(x)\, \mu(\mathrm{d}x) > 0$.

The next result is stronger than the implication (i) $\implies$ (ii) in Proposition II.4.2.

Lemma II.4.4. *Under Assumption II.2, the following inequality holds:*

$$s(T_k - \gamma) \leq \max(\| \gamma \| (\mathfrak{R}_0 - 1), 0). \tag{II.42}$$

Proof. If $s(T_k - \gamma) \leq 0$, the result is obviously true. Suppose $s(T_k - \gamma) > 0$. Since $T_k - \gamma + \|\gamma\|$ is a positive operator, Theorem II.3.5 (ii) implies that $\rho(T_k - \gamma + \|\gamma\|) = s(T_k - \gamma + \|\gamma\|)$. Since $s(T_k - \gamma + \|\gamma\|) = s(T_k - \gamma) + \|\gamma\| > \|\gamma\|$, we obtain that:

$$\rho(T_k - \gamma + \|\gamma\|) > \|\gamma\|.$$

Besides, we have $\|\gamma\| \geq \rho(\|\gamma\| - \gamma)$ according to Theorem II.3.5 (i) and $\rho(\|\gamma\| - \gamma) \geq r_{\mathrm{ess}}(\|\gamma\| - \gamma)$ according to Equation (II.36). We deduce that:

$$\rho(T_k - \gamma + \|\gamma\| > r_{\mathrm{ess}}(\|\gamma\| - \gamma).$$

The operator T_k is weakly compact thanks to Lemma II.3.7 (ii) since k satisfies (II.37), see Assumption II.2 and more precisely (II.9). Since $\mathscr{L}^\infty$ has the Dunford-Pettis property, see Theorem II.3.2, we deduce from [103, Theorem 3.1] (where $\sigma_{\text{ess}}(A)$ in our setting corresponds to $\sigma_{e5}(A)$ in [103]) that $r_{\text{ess}}(\|\gamma\| - \gamma) = r_{\text{ess}}(T_k - \gamma + \|\gamma\|)$. Therefore, we get that:

$$\rho(T_k - \gamma + \|\gamma\|) > r_{\text{ess}}(T_k - \gamma + \|\gamma\|).$$

Hence, we can apply Theorem II.3.5 (iii) with the positive operator $T_k - \gamma + \|\gamma\|$, to get the existence of a function $w \in \mathscr{L}^\infty_+ \setminus \{0\}$ such that:

$$T_k(w) - \gamma w = s(T_k - \gamma)w, \tag{II.43}$$

where we used $\rho(T_k - \gamma + \|\gamma\|) = s(T_k - \gamma + \|\gamma\|) = s(T_k - \gamma) + \|\gamma\|$ for the equality. We have shown that one can actually take $\lambda = s(T_k - \gamma)$ in Equation (II.39). Thus, we obtain:

$$T_{k/\gamma}(\gamma w) = T_k(w) = (\gamma + s(T_k - \gamma))w \geq \left(1 + \frac{s(T_k - \gamma)}{\|\gamma\|}\right)\gamma w.$$

According to Proposition II.3.6, we conclude that:

$$\mathfrak{R}_0 = \rho(T_{k/\gamma}) \geq 1 + \frac{s(T_k - \gamma)}{\|\gamma\|}. \tag{II.44}$$

We deduce that Equation (II.42) holds. □

The next result gives information about the behavior of $\mathfrak{R}_0$ when we modify the susceptibility of individuals, and will be needed below in the proof of Theorem II.4.13. For $g \in \Delta$, we define $\mathfrak{R}_0(g) = \rho(gT_{k/\gamma})$.

Proposition II.4.5 (A continuity property of $\mathfrak{R}_0$)**.** *Suppose Assumption II.2 holds. The function $g \mapsto \mathfrak{R}_0(g)$ defined on Δ is non-decreasing and continuous with respect to the $L^1(\mu)$ topology.*

Proof. The fact that $g \mapsto \mathfrak{R}_0(g)$ is non-decreasing is a direct consequence of Theorem II.3.5 (i).

For $g \in \Delta$, the bounded operator $A_g = \hat{T}_{\mathrm{k}}$ on $L^p(\mu)$ defined in Equation (II.38) with the kernel $\mathrm{k}(x, y) = g(x)k(x, y)/\gamma(y)$ is compact according to Lemma II.3.7 (iii). According to Lemma II.3.7 (iv), we have that for all $g \in \Delta$:

$$\mathfrak{R}_0(g) = \rho(A_g). \tag{II.45}$$

Besides, the function $g \mapsto A_g$ mapping Δ to $\mathcal{L}(L^p(\mu))$ is continuous with respect to the $L^p(\mu)$ norm. We deduce from [119, Theorem 11], that the function $g \mapsto \sigma(A_g)$ from $(\Delta, \|\cdot\|_p)$ to $(\mathcal{K}(\mathbb{C}), d_{\mathrm{H}})$ is continuous, where $\mathcal{K}(\mathbb{C})$ is the set of non-empty compact subsets and d_{H} is the Hausdorff distance (see Section II.8 for the definition of the Hausdorff distance). Using Lemma II.8.1 and then Equation (II.45), we get that $g \mapsto \mathfrak{R}_0(g)$ defined on $(\Delta, \|\cdot\|_p)$ is continuous. In order to conclude, we notice that the topologies induced by $L^p(\mu)$ and $L^1(\mu)$ are equal on Δ because Δ is a bounded subset of $L^\infty(\mu)$. This proves that $g \mapsto \mathfrak{R}_0(g)$ defined on $(\Delta, \|\cdot\|_1)$ is continuous. □

II.4.2 The subcritical regime: $s(T_\kappa - \gamma) < 0$

We show here that in the subcritical regime, the solutions of Equation (II.25) converge exponentially fast to 0 in norm.

Theorem II.4.6 (Uniform exponential extinction)**.** *Suppose that $s(T_\kappa - \gamma) < 0$. Then, for all $c \in (0, -s(T_\kappa - \gamma))$, there exists a finite constant $\theta = \theta(c)$ such that, for all $g \in \Delta$, we have:*

$$\|\phi(t, g)\| \leq \theta \|g\| \, \mathrm{e}^{-ct}. \tag{II.46}$$

In particular, the maximal equilibrium $\mathfrak{g}$ is equal to 0 everywhere.

Proof. Recall $T_\kappa - \gamma$ is a bounded operator. For all $t \in \mathbb{R}_+$, define:

$$v(t) = \mathrm{e}^{t(T_\kappa - \gamma)} 1 = \sum_{n \in \mathbb{N}} \frac{t^n}{n!} (T_\kappa - \gamma)^n \, 1. \tag{II.47}$$

Since multiplication by the constant $\| \gamma \|$ commutes with $T_\kappa - \gamma$, we also have:

$$\mathrm{e}^{\| \gamma \| t} v(t) = \mathrm{e}^{t(T_\kappa - \gamma + \| \gamma \|)} 1 = \sum_{n \in \mathbb{N}} \frac{t^n}{n!} (T_\kappa - \gamma + \| \gamma \|)^n \, 1.$$

As $T_\kappa - \gamma + \| \gamma \|$ is positive, we deduce that $v(t) \geq 0$. As T_κ is positive, we deduce that:

$$v'(t) - F(v(t)) = (T_\kappa - \gamma)(v(t)) - F(v(t)) = v(t) T_\kappa(v(t)) \geq 0.$$

Thus, the following inequality holds for all $g \in \Delta$ and all $t \geq 0$:

$$0 = \partial_t \phi(t, g) - F(\phi(t, g)) \leq v'(t) - F(v(t)).$$

As F is cooperative on $(1 - \mathscr{L}^\infty_+) \times \mathscr{L}^\infty$, see Proposition II.2.5 (iv), we can apply Theorem II.2.4 with the positive cone $K = \mathscr{L}^\infty_+$, $D_1 = 1 - \mathscr{L}^\infty_+$, $D_2 = \mathscr{L}^\infty$, $a(t) = \phi(t, g)$ and $b(t) = v(t)$, to obtain that:

$$\phi(t, g) \leq v(t) \quad \text{for all } t \in \mathbb{R}_+. \tag{II.48}$$

Besides, since $T_k - \gamma$ is a bounded operator, its growth bound (*i.e.*, the left member of the equality below) is equal to its spectral bound according to [31, Theorem I.4.1]:

$$\inf \left\{ \eta \in \mathbb{R} \ : \ \sup_{t \in \mathbb{R}_+} \mathrm{e}^{-\eta t} \| \exp(t(T_\kappa - \gamma)) \| < \infty \right\} = s(T_\kappa - \gamma). \tag{II.49}$$

We deduce from Equations (II.47), (II.48) and (II.49), that for all $c \in (0, -s(T_\kappa - \gamma))$, there exists a finite constant θ such that Equation (II.46) is true. In particular, $t \mapsto \phi(t, 1)$ converges uniformly to 0. It then follows from Equation (II.31) that $\mathfrak{g}$ is equal to 0 everywhere. $\square$

II.4.3 Critical regime: $s(T_k - \gamma) = 0$

We suppose here that Assumption II.2 holds, so that the kernel κ has a density k, and we write T_k for T_κ. We give the main result of this section.

Theorem II.4.7 (Extinction at criticality)**.** *Suppose Assumption II.2 is in force and $s(T_k - \gamma) = 0$. Then the maximal equilibrium $\mathfrak{g}$ is equal to 0 everywhere. In other words, for all $g \in \Delta$ and all $x \in \Omega$, we have that:*

$$\lim_{t \to \infty} \phi(t, g)(x) = 0.$$

Proof. Suppose, to derive a contradiction, that $\mathfrak{g}$ is not equal to 0 μ-almost everywhere. We know according to Remark II.2.16 that $1 - \mathfrak{g}$ is positive everywhere. Hence, we get:

$$T_{k/\gamma}(\gamma \mathfrak{g}) = T_k(\mathfrak{g}) = \left(1 + \frac{\mathfrak{g}}{1 - \mathfrak{g}} \right) \gamma \mathfrak{g} \geq \gamma \mathfrak{g}. \tag{II.50}$$

According to Proposition II.3.6, $\mathfrak{R}_0$ is then greater than or equal to 1.

We now prove that the inequality is strict: $\mathfrak{R}_0 > 1$. Consider the support of $\mathfrak{g}$: $A = \{ x \in \Omega \ : \ \mathfrak{g}(x) > 0 \}$. Equation (II.50) remains true by replacing k by $k' = \mathbb{1}_A k \mathbb{1}_A$ (*i.e.* $k'(x, y) = \mathbb{1}_A(x) k(x, y) \mathbb{1}_A(y)$):

$$T_{k'/\gamma}(\gamma \mathfrak{g}) = \left(1 + \frac{\mathfrak{g}}{1 - \mathfrak{g}} \right) \gamma \mathfrak{g} \geq \gamma \mathfrak{g}. \tag{II.51}$$

Using Proposition II.3.6, we get that $\rho(T_{k'/\gamma}) \geq 1$. Since Assumption II.2 is in force, $T_{k'/\gamma}$ has a left Perron eigenvector h in $L^q_+(\mu) \backslash \{ 0 \}$ (see Lemma II.3.7 (v)). By multiplying both members of Equation (II.51) by h and integrating with respect to μ, we obtain:

$$(\rho(T_{k'/\gamma}) - 1) \langle h, \gamma \mathfrak{g} \rangle = \langle h, (\mathfrak{g})^2 \gamma / (1 - \mathfrak{g}) \rangle \tag{II.52}$$

It is clear that $h\mathbb{1}_{A^c} = 0$. Since $h \in L^q_+(\mu)\backslash\{0\}$, we have necessarily:

$$\int_A h(x)\,\mu(\mathrm{d}x) > 0.$$

Hence, both brackets in Equation (II.52) are positive. Thus, we get that $\rho(T_{k'/\gamma}) > 1$. Using Theorem II.3.5 (i) and that the operator $T_{k/\gamma} - T_{k'/\gamma}$ is positive, we deduce that $\mathfrak{R}_0 \geq \rho(T_{k'/\gamma}) > 1$. This is in contradiction with Proposition II.4.2 which asserts that $\mathfrak{R}_0 \leq 1$ as $s(T_k - \gamma) = 0$. Thus, we obtain that μ-a.e. $\mathfrak{g} = 0$. We conclude using Remark II.3.8. □

II.4.4 Supercritical regime: $s(T_k - \gamma) > 0$

Assumption II.2 is in force in this section, where we consider the case $s(T_k - \gamma) > 0$. We begin by proving that $\mathfrak{g}$ is different from 0, then we show the convergence of the system to $\mathfrak{g}$.

Informally, the main idea in the following is to write the linear approximation of the dynamics (II.25) near the zero equilibrium: $\partial_t u = (T_k - \gamma)u$, and apply the results from Section II.3 (Theorem II.3.5 (iii) and Lemma II.3.7) in order to identify a dominant component of the evolution near 0 as a Perron eigenvector of $T_k - \gamma$. A particularly nice feature is that this procedure yields a monotonous trajectory.

In order to control the non-linearity, we give ourselves a little room by choosing ε small enough so that $(1-\varepsilon)\mathfrak{R}_0 > 1$. Since $\rho((1-\varepsilon)T_{k/\gamma}) = (1-\varepsilon)\rho(T_{k/\gamma}) = (1-\varepsilon)\mathfrak{R}_0 > 1$, Proposition II.4.2 ensures the existence of a vector $w_\varepsilon \in \mathscr{L}^\infty_+\backslash\{0\}$ and a positive real number $\lambda(\varepsilon) > 0$, such that:

$$(1-\varepsilon)T_k(w_\varepsilon) = (\gamma + \lambda(\varepsilon))w_\varepsilon. \tag{II.53}$$

We can take w_ε such that $\|w_\varepsilon\| < \varepsilon$. Moreover, according to Remark II.4.3:

$$\int_\Omega w_\varepsilon(x)\,\mathrm{d}x > 0. \tag{II.54}$$

Then, we get the following proposition.

Proposition II.4.8 (Increasing trajectory)**.** *Suppose Assumption II.2 is in force and that* $s(T_k-\gamma) > 0$*, and let* w_ε *be defined by* (II.53)*. The trajectory starting from* w_ε *is monotonous: the map* $t \mapsto \phi(t, w_\varepsilon)$ *is non-decreasing.*

Proof. Using Equation (II.53) and the fact that $\|w_\varepsilon\| < \varepsilon$, we have:

$$0 \leq \lambda(\varepsilon)w_\varepsilon = (1-\varepsilon)T_k(w_\varepsilon) - \gamma w_\varepsilon \leq (1-w_\varepsilon)T_k(w_\varepsilon) - \gamma w_\varepsilon = F(w_\varepsilon).$$

This implies the stated monotony by Proposition II.2.10. □

Proposition II.4.8 shows that the equilibrium 0 is not asymptotically stable, in the sense that we can find initial conditions arbitrarily close to 0 in norm such that $\phi(t, g)$ does not converge to 0 pointwise (note that $\|w_\varepsilon\|$ may be chosen arbitrarily small). Since $w_\varepsilon \leq 1$, we get by monotony and the comparison Theorem that $w_\varepsilon \leq \phi(t, w_\varepsilon) \leq \phi(t, 1)$, which implies that $w_\varepsilon \leq \mathfrak{g}$. In particular we get the following corollary.

Corollary II.4.9. *If Assumption II.2 is in force and* $s(T_k - \gamma) > 0$*, then we have:*

$$\int_\Omega \mathfrak{g}(x)\,\mu(\mathrm{d}x) > 0.$$

We deduce from Proposition II.4.8 that $t \mapsto \phi(t, w_\varepsilon)$ converges pointwise as t tends to infinity since $\phi(t, w_\varepsilon) \leq 1$ for all t. According to Proposition II.2.13, the limit is an equilibrium. It is not 0 but it might be different from $\mathfrak{g}$. We will use Assumption II.3 to ensure that 0 and $\mathfrak{g}$ are the only equilibria. In order to prove this result, we need the following lemma.

Lemma II.4.10 (Instantaneous propagation of the infection)**.** *Suppose Assumptions II.2 and II.3 are in force. If $g \in \Delta$ satisfies $\int_\Omega g(x)\, \mu(\mathrm{d}x) > 0$, then, for all $t > 0$, $\phi(t, g)$ is μ-a.e. positive.*

Proof. Since the flow is order-preserving (see Proposition II.2.8), it is sufficient to show the proposition for g such that $\| g \| < 1/2$. It follows from Equation (II.29) that:

$$\phi(t, g) \leq \| g \| + t \| T_k \| .$$

Thus, for all $t \in [0, c)$, with $c = (1 - 2\| g \|)/2\| T_k \|$ (and $c = +\infty$ if $\| T \| = 0$), we have that $\phi(t, g) < 1/2$. Now, we define the function:

$$u(t) = \mathrm{e}^{-\|\gamma\| t}\, \mathrm{e}^{tT_k/2} g.$$

Taking $c > 0$ smaller if necessary, we get $u(t) \leq \| g \| (1 + t\| T_k \|) < 1/2$ for $t \in [0, c)$. Then we get for $t \in [0, c)$:

$$\begin{aligned} u'(t) - F(u(t)) &= (T_k/2 - \|\gamma\|)u(t) - (1 - u(t))T_k u(t) + \gamma u(t) \\ &= (u(t) - 1/2)T_k u(t) - (\|\gamma\| - \gamma)u(t) \\ &\leq 0 = b'(t) - F(b(t)), \end{aligned}$$

where $b(t) = \phi(t, g)$. Using the comparison Theorem II.2.4, we deduce

$$\phi(t, g) \geq u(t). \tag{II.55}$$

Now, we fix $t \in [0, c)$. We denote by $A = \{ x \in \Omega \,:\, u(t)(x) > 0 \}$ the support of $u(t)$. We have:

$$0 = \langle \mathbb{1}_{A^{\mathrm{c}}}, u(t) \rangle = \mathrm{e}^{-\|\gamma\| t} \sum_{n \in \mathbb{N}} \frac{1}{n!} \langle \mathbb{1}_{A^{\mathrm{c}}}, (tT_k/2)^n(g) \rangle .$$

This implies that $\langle \mathbb{1}_{A^{\mathrm{c}}}, (tT_k/2)^n(g) \rangle = 0$ for all n, and thus that $\langle \mathbb{1}_{A^{\mathrm{c}}}, T_k u(t) \rangle = 0$. We deduce that:

$$\int_{A^{\mathrm{c}} \times A} k(x, y)\, \mu(\mathrm{d}x)\mu(\mathrm{d}y) = 0.$$

Since the set A contains the support of g, we get $\mu(A) > 0$. It follows from Assumption II.3 that $\mu(A^{\mathrm{c}}) = 0$. This means that $u(t)$ is μ-a.e. positive. Hence, from Equation (II.55), we get that, for $t \in [0, c)$, $\phi(t, g)$ is μ-a.e. positive. Using the semi-group property of the semi-flow this results propagates on the whole positive half-line and the result is proved. □

Remark II.4.11. One can check from its proof, that Lemma II.4.10 does not require the integrability condition (II.8) in Assumption II.2 to be true.

Now we can show the following important result.

Proposition II.4.12 (Uniqueness of the endemic state)**.** *Under Assumptions II.2 and II.3, the maximal equilibrium $\mathfrak{g}$:*

(i) is positive μ-a.e.,

(ii) is the unique equilibrium different from 0.

Proof. From Lemma II.4.10 together with Remark II.3.8, we deduce that every equilibrium different from 0 is positive μ-a.e. This proves Point (i) as $\int \mathfrak{g} \mathrm{d}\mu > 0$ in the supercritical regime according to Corollary II.4.9.

We now prove Point (ii). Let h^* be another equilibrium different from 0. Since $\mathfrak{g}$ is the maximal equilibrium, we have $h^* \leq \mathfrak{g}$. We shall prove that h^* is equal to $\mathfrak{g}$ almost everywhere. Let us define the non-negative kernel k by:

$$\mathrm{k}(x, y) = (1 - \mathfrak{g}(x))\, \frac{k(x, y)}{\gamma(y)} \quad \text{for } x, y \in \Omega.$$

Notice that k satisfies (II.37). Since $T_{\mathrm{k}}(\gamma\mathfrak{g}) = \gamma\mathfrak{g}$, we deduce from Proposition II.3.6 that $\rho(T_{\mathrm{k}}) \geq 1$. Let $v \in L^q(\mu)_+ \setminus \{0\}$ be a left Perron vector of the operator T_{k} (given by Lemma II.3.7 (v)). The kernel k satisfies Assumption II.3 as k does and $1 - \mathfrak{g}$ is positive everywhere (see Remark II.2.16). Hence, v can be chosen positive μ-a.e. according to Lemma II.3.7 (vi). The following computation:

$$\langle v, \gamma\mathfrak{g}\rangle = \langle v, T_{\mathrm{k}}(\gamma\mathfrak{g})\rangle = \rho(T_{\mathrm{k}})\,\langle v, \gamma\mathfrak{g}\rangle\,,$$

shows that $\rho(T_{\mathrm{k}})$ is actually equal to 1 since $\langle v, \gamma\mathfrak{g}\rangle > 0$. Now we compute:

$$\begin{aligned} 0 &= \langle v, F(h^*)\rangle \\ &= \langle v, T_{\mathrm{k}}(\gamma h^*) - \gamma h^*\rangle + \langle v, (\mathfrak{g} - h^*)T_{k/\gamma}(\gamma h^*)\rangle \\ &= \langle v, (\mathfrak{g} - h^*)T_k(h^*)\rangle\,, \end{aligned}$$

where we used that $\langle v, T_{\mathrm{k}} f - f\rangle = 0$ as $\rho(T_{\mathrm{k}}) = 1$ and v is a left Perron eigenvector. According to the first part of the proof, h^* is μ-a.e. positive. Since we have $T_k(h^*) = \gamma h^*/(1 - h^*)$, the function $T_k(h^*)$ is also μ-a.e. positive. Hence $\mathfrak{g}$ and h^* are equal μ-a.e. since v is μ-a.e. positive, see Lemma II.3.7 (vi). This implies in particular that $T_k(h^*) = T_k(\mathfrak{g})$ by Lemma II.3.7 (i). We deduce that, for all $x \in \Omega$:

$$h^*(x) = T_k(h^*)(x)/(\gamma(x) + T_k(h^*)(x)) = T_k(\mathfrak{g})(x)/(\gamma(x) + T_k(\mathfrak{g})(x)) = \mathfrak{g}(x).$$

Therefore $\mathfrak{g}$ is then unique equilibrium different from 0. □

Now we can prove the main result of this section on the pointwise convergence of $\phi(t, g)$. If g is μ-a.e. equal to 0, then clearly, as γ is positive, we get that $\lim_{t\to\infty} \phi(t, g) = 0$ pointwise, so we only need to consider the case where g is not μ-a.e. equal to 0.

Theorem II.4.13 (Convergence towards the endemic equilibrium)**.** *Suppose that Assumptions II.2 and II.3 are in force. Let $g \in \Delta$ such that $\int_\Omega g(x)\,\mu(\mathrm{d}x) > 0$. Then,we have that for all $x \in \Omega$:*

$$\lim_{t\to\infty} \phi(t, g)(x) = \mathfrak{g}(x).$$

Proof. By Lemma II.4.10, it is enough to show the result for g μ-a.e. positive. The idea is similar to the proof of Proposition II.4.8, that is, to try and find a monotonous trajectory; the difference here is that we look for a trajectory that is below $\phi(t, g)$, and we have to adapt the proof accordingly. For such a g, the functions $(1 - \varepsilon)g\mathbb{1}_{g\geq\varepsilon}$ converge in $L^1(\mu)$ to g when ε goes to zero. Besides, $\mathfrak{R}_0$ is greater than 1 by Proposition II.4.2. Hence, according to Proposition II.4.5, for ε small enough, we get

$$\mathfrak{R}_0\left((1 - \varepsilon)\,\mathbb{1}_{g\geq\varepsilon}T_{k/\gamma}\right) > 1.$$

By Proposition II.4.2 (iii), applied to the kernel $(1 - \varepsilon)\mathbb{1}_{g(x)\geq\varepsilon}k(x, y)$, there exists $w_\varepsilon \in \mathscr{L}^\infty_+ \setminus \{0\}$ and $\lambda(\varepsilon) > 0$ such that:

$$(1 - \varepsilon)\,\mathbb{1}_{g\geq\varepsilon}T_k(w_\varepsilon) = (\gamma + \lambda(\varepsilon))w_\varepsilon. \tag{II.56}$$

We may and will assume additionally that $\|w_\varepsilon\| \leq \varepsilon$. Since (II.56) implies that $w_\varepsilon(x) = 0$ when $g(x) < \varepsilon$, we know that $w_\varepsilon \leq g$. The monotony of the semi-flow (see Proposition II.2.8) then implies that, for all $t \in \mathbb{R}_+$:

$$\phi(t, w_\varepsilon) \leq \phi(t, g) \leq \phi(t, 1). \tag{II.57}$$

Besides, we have:

$$\begin{aligned} 0 \leq \lambda(\varepsilon)w_\varepsilon &= (1 - \varepsilon)\mathbb{1}_{g\geq\varepsilon}T_k(w_\varepsilon) - \gamma w_\varepsilon \\ &\leq (1 - \varepsilon)T_k(w_\varepsilon) - \gamma w_\varepsilon \\ &\leq (1 - w_\varepsilon)T_k(w_\varepsilon) - \gamma w_\varepsilon \\ &= F(w_\varepsilon), \end{aligned}$$

where the last inequality follows from the fact that $\|w_\varepsilon\| \leq \varepsilon$. Thus, the path $t \mapsto \phi(t, w_\varepsilon)$ is non-decreasing according to Proposition II.2.10. Hence, it converges pointwise to a limit $h^* \neq 0$ since $w_\varepsilon \in \mathscr{L}^\infty_+ \setminus \{0\}$. This limit has to be an equilibrium by Proposition II.2.13. Since 0 and $\mathfrak{g}$ are the only equilibria by Proposition II.4.12, we have necessarily $h^* = \mathfrak{g}$. We conclude thanks to Equation (II.57). □

II.4.5 Endemic states in the critical regime

Here we show by a counter-example that the integral condition (II.8) is necessary to obtain the convergence towards the disease-free equilibrium in the critical regime. In the following example, the transmission kernel has a bounded density with respect to a finite measure μ and we have $\inf \gamma > 0$ and $\mathfrak{R}_0 = 1$. However, there exists a continuum of distinct equilibria.

Consider the set $\mathbb{N}^*$ equipped with some finite measure μ such that $\mu_n = \mu(\{n\}) > 0$ for all $n \in \mathbb{N}^*$. We choose γ constant equal to 1 and the kernel κ defined for $i, j \in \mathbb{N}^*$ by:

$$\kappa(i, \{j\}) = \begin{cases} \frac{2i+2}{2i-1} & \text{if } j = i+1, \\ 0 & \text{otherwise,} \end{cases} \tag{II.58}$$

Clearly Assumption II.1 is satisfied. Moreover, the kernel κ admits with respect to μ the density k defined by $k(i,j) = \kappa(i,\{j\})/\mu(\{j\})$, for $i, j \in \mathbb{N}^*$. However condition (II.9), and thus (II.8) from Assumption II.2, is not satisfied. Indeed, for all $q > 1$ we have:

$$\sup_{n\in\mathbb{N}^*} \int_{\mathbb{N}^*} k(x,y)^q \, \mu(\mathrm{d}y) = \sup_{n\in\mathbb{N}^*} k(n, n+1)^q \, \mu_{n+1} = \lim_{n\to\infty} \frac{(2n+1)^q}{(2n-1)^q} \, \mu_{n+1}^{1-q} = +\infty,$$

where divergence of the sequence follows from the convergence of μ_{n+1} to 0 (because μ is a finite measure). The following proposition asserts that we are in the critical regime.

Proposition II.4.14. *Let κ be defined by* (II.58)*, k be its density and $\gamma = 1$. We have for the reproduction number: $\mathfrak{R}_0 = \rho(T_{k/\gamma}) = 1$, and for the spectral bound: $s(T_k - \gamma) = 0$.*

Proof. Since γ is the function constant equal to 1, we have $s(T_k - \gamma) = \mathfrak{R}_0 - 1$ and $\mathfrak{R}_0 = \rho(T_k)$. We compute the spectral radius of T_k using Gelfand's formula:

$$\rho(T_k) = \lim_{n\to\infty} \| T_k^n \|^{1/n} = \lim_{n\to\infty} \left(\prod_{i=1}^{n} \frac{2i+2}{2i-1} \right)^{1/n} = 1.$$

The limit is found by applying the logarithm to the sequence and using Cesàro lemma. □

The following result shows that even if we are in the critical regime, the maximal equilibrium $\mathfrak{g}$ is not equal to 0 everywhere, and there exists infinitely many distinct equilibria. For $\alpha \in [0,1]$, we define the function g^*_α on $\mathbb{N}^*$ by $\mathfrak{g}_\alpha(1) = \alpha$ and for $n \in \mathbb{N}^*$:

$$\mathfrak{g}_\alpha(n+1) = \begin{cases} \frac{2n-1}{2n+2} \frac{\mathfrak{g}_\alpha(n)}{1-\mathfrak{g}_\alpha(n)} & \text{if } \mathfrak{g}_\alpha(n) < 1, \\ 0 & \text{if } \mathfrak{g}_\alpha(n) \geq 1. \end{cases}$$

Proposition II.4.15. *Let κ be defined by* (II.58)*, k be its density and $\gamma = 1$.*

(i) The equilibria of Equation (II.3) *are $\{ \mathfrak{g}_\alpha : \alpha \in [0, 1/2] \}$.*

(ii) The function $\alpha \mapsto \mathfrak{g}_\alpha$ defined on $[0,1/2]$ and taking values in $\Delta \subset \mathscr{L}^\infty$ is increasing and continuous (with respect to $\|\cdot\|$). In particular, the set of equilibria is totally ordered, compact and connected.

(iii) The equilibrium $g^*_{1/2}$ is the maximal equilibrium and is given by $g^*_{1/2}(n) = 1/(2n)$ for $n \in \mathbb{N}^*$.

(iv) For every $\alpha \in (0, 1/2)$, there exists a constant c_α such that $g^*_\alpha(n) \sim c_\alpha n^{-3/2}$. Moreover, the map $\alpha \to c_\alpha$ is strictly increasing and continuous on $[0, 1/2)$, with the convention $c_0 = 0$.

Proof. We start by remarking that for $\alpha = 1/2$, the induction may be solved explicitly, so that $\mathfrak{g}_{1/2}(n) = 1/(2n)$. Similarly $\mathfrak{g}_0(n) = 0$.

We prove property (ii) first. Let Γ denote the function $\alpha \mapsto \mathfrak{g}_\alpha$ defined on $[0, 1/2]$ and taking values in $\mathscr{L}^\infty$. Since the function $x \mapsto \lambda x/(1-x)$ is increasing on $[0,1)$ for all $\lambda > 0$, we deduce by induction that $0 \leq g^*_\alpha(n) < g^*_\beta(n) \leq \mathfrak{g}_{1/2}(n)$ for all $0 \leq \alpha < \beta \leq 1/2$ and $n \in \mathbb{N}^*$. As

$\mathfrak{g}_0$ and $\mathfrak{g}_{1/2}$ both belong to Δ, we deduce that Γ takes values in Δ by monotonicity. It is also immediate to check that the function Γ is continuous for the pointwise convergence in Δ. Since $\lim_{n\to\infty} \sup_{\alpha\in[0,1/2]} g^*_\alpha(n) = \lim_{n\to\infty} g^*_{1/2}(n) = 0$, this continuity also holds with respect to the uniform convergence in Δ. This proves property (ii).

We now prove property (i). It is clear that if h^* is an equilibrium, then $h^*(n) < 1$ for all $n \in \mathbb{N}^*$ thanks to Remark II.2.16 and by the definition of the kernel κ we have that:

$$h^*(n+1) = \frac{2n-1}{2n+2}\,\frac{h^*(n)}{1-h^*(n)} \quad \text{for all } n \in \mathbb{N}^*. \tag{II.59}$$

This readily implies that $h^\star = \mathfrak{g}_\alpha$ where $\alpha = h^*(0)$, so that the only possible equilibria are the $\mathfrak{g}_\alpha$ for $\alpha \in [0,1]$.

If $\alpha \in [0,1/2]$, $\mathfrak{g}_\alpha$ is indeed an equilibrium as $\mathfrak{g}_\alpha(n) \leq \mathfrak{g}_{1/2}(n) = 1/(2n)$ and $\mathfrak{g}_\alpha$ solves (II.59). On the contrary, since $\mathfrak{g}_1(1) = 1$, the function $\mathfrak{g}_1$ is not an equilibrium.

Let $\alpha \in (1/2, 1)$. We shall now prove by contradiction that there exists $n \in \mathbb{N}^*$ such that $\mathfrak{g}_\alpha(n) \geq 1$. Let us assume that $\mathfrak{g}_\alpha(n) < 1$ for all $n \in \mathbb{N}^*$. Arguing as in the first part of the proof, we get $\mathfrak{g}_\alpha(n) > \mathfrak{g}_{1/2}(n)$ for all $n \in \mathbb{N}^*$. Thus the sequence $v = (v_n : n \in \mathbb{N}^*)$ with $v_n = 2n\mathfrak{g}_\alpha(n)$ satisfies the following recurrence for $n \in \mathbb{N}^*$:

$$v_{n+1} = v_n \frac{2n-1}{2n-v_n} \quad \text{and} \quad 1 < v_n < 2n.$$

We deduce that the sequence v is increasing, and thus $v_{n+1} \geq v_n \frac{2n-1}{2n-2\alpha}$, as $v_1 = 2\alpha$. We deduce that $v_n \geq c n^{\alpha-1/2}$ for some positive constant c. This in turn implies that $v_{n+1} \geq v_n \frac{2n-1}{2n-cn^{\alpha-1/2}}$ and thus $v_n \geq c' \exp(c'' n^{\alpha-1/2})$ for some positive constants c' and c''. This contradicts the fact that $v_n < 2n$ for $n \in \mathbb{N}^*$. As a conclusion, there exists $n \in \mathbb{N}^*$ such that $\mathfrak{g}_\alpha(n) \geq 1$. This implies that g^*_α can not be an equilibrium. This ends the proof of property (i).

We end the proof of property (iii). We have already computed $g^*_{1/2}$. We deduce from properties (i) and (ii) that $g^*_{1/2}$ is the maximal equilibrium.

Finally, let us prove the asymptotic result (iv). The asymptotics of $\mathfrak{g}_\alpha(n)$ is obtained by starting from an easy bound on its decay, and then refining it by plugging it back into the induction relation (II.59).

Linear decay. Since $\alpha \leq 1/2$, we already know that $\mathfrak{g}_\alpha(n) \leq g^*_{1/2}(n) = 1/(2n)$. We can prove a little bit better. The sequence $w = (w_n : n \in \mathbb{N}^*)$ defined by $w_n = n\mathfrak{g}_\alpha(n)/\alpha$ satisfies

$$w_{n+1} = \frac{2n-1}{2n-2\alpha w_n} w_n.$$

Since $2\alpha w_n = 2n\mathfrak{g}_\alpha(n) \leq 1$, the sequence w is non-increasing. In particular $w_n \leq w_1 = 1$, so we deduce that:

$$\mathfrak{g}_\alpha(n) \leq \frac{\alpha}{n}. \tag{II.60}$$

Sublinear decay. Let q_n be the quotient $q_n = \mathfrak{g}_\alpha(n+1)/\mathfrak{g}_\alpha(n)$. By the recurrence relation and the linear bound (II.60), we get successively

$$q_n \leq \frac{1-1/(2n)}{1+1/n}\frac{1}{1-\alpha/n} \leq 1 - \left(\frac{3}{2}-\alpha\right)\frac{1}{n} + O(n^{-2}) \quad \text{and} \quad \log(q_n) \leq -(\frac{3}{2}-\alpha)\frac{1}{n} + r_n,$$

where $r_n = O(n^{-2})$. Summing the terms from 1 to $n-1$, we get

$$\log(\mathfrak{g}_\alpha(n)) - \log(\mathfrak{g}_\alpha(1)) \leq -(3/2-\alpha)\log(n) + O(1).$$

Therefore, there exists a constant C (that may depend on α) such that

$$\mathfrak{g}_\alpha(n) \leq \frac{C}{n^{3/2-\alpha}}. \tag{II.61}$$

Optimal decay. We come back to the recurrence relation, and use the sublinear bound (II.61) on the term $1 - \mathfrak{g}_\alpha(n)$ that appears in the denominator. This yields successively

$$q_n = \left(1 - \frac{3}{2n} + O(n^{-2})\right)\left(1 + O(n^{-(3/2-\alpha)})\right) = 1 - \frac{3}{2n} + O(n^{-(3/2-\alpha)}),$$

which gives $\log(q_n) = -\frac{3}{2n} + r_n$ where $r_n = O(n^{-(3/2-\alpha)})$. As $\sum_m r_m$ is finite, summing from 1 up to $n-1$ and taking the exponential yields

$$\mathfrak{g}_\alpha(n) = \alpha \exp\left(-(3/2)\ln(n) + O(1)\right).$$

In other words we obtain, as claimed, the existence of a (non-explicit) constant c_α such that $\mathfrak{g}_\alpha(n) \sim c_\alpha n^{-3/2}$.

Properties of c_α. Notice that for $\alpha \in [0, 1/2)$ the limit $d_\alpha = \lim_{n\to\infty} \prod_{k=1}^{n}(1 - \mathfrak{g}_\alpha(k))^{-1}$ exists and is positive. Let $0 \leq \alpha \leq \beta < 1/2$ and set $D_n = \mathfrak{g}_\beta(n) - \mathfrak{g}_\alpha(n)$ for $n \geq 1$. We have $D_{n+1} = \frac{2n-1}{2n+2}(1 - \mathfrak{g}_\beta(n))^{-1}(1 - \mathfrak{g}_\alpha(n))^{-1} D_n$. So we deduce that:

$$D_n = (\beta - \alpha)\left(\prod_{k=1}^{n-1} \frac{2k-1}{2k+2}\right)\left(\prod_{k=1}^{n-1}(1 - \mathfrak{g}_\alpha(k))^{-1}\right)\left(\prod_{k=1}^{n-1}(1 - g^*_\beta(k))^{-1}\right).$$

Using that $\prod_{k=0}^{n-1} \frac{2k-1}{2k+2} \sim C n^{-3/2}$ for some finite constant $C > 0$, we deduce that

$$c_\beta - c_\alpha = C(\beta - \alpha) d_\alpha d_\beta.$$

This gives the strict monotonicity of the map $\alpha \mapsto c_\alpha$. Then, use that $d_\alpha \leq d_\beta \leq d_{\beta'} < +\infty$ for some $\beta' \in (\beta, 1/2)$ to get the continuity. $\square$

We are not able to describe entirely the basins of attraction of each equilibrium. However, the asymptotic behavior in n of the starting point g tells us quite a lot.

Proposition II.4.16. *For all $g \in \Delta$, and for all $\alpha \in (0, 1/2)$, we have:*

$$\limsup_n n^{3/2} g(n) \leq c_\alpha \implies \limsup_{t\to\infty} \phi(t, g) \leq g^*_\alpha,$$
$$\liminf_n n^{3/2} g(n) \geq c_\alpha \implies \liminf_{t\to\infty} \phi(t, g) \geq g^*_\alpha.$$

In particular, we have:

$$\limsup_n n^{3/2} g(n) = 0 \implies \phi(t, g) \to 0,$$
$$\liminf_n n^{3/2} g(n) = \infty \implies \phi(t, g) \to \mathfrak{g}_{1/2}.$$

Proof. Since k is upper-triangular, the long-time behavior of the dynamic does not depend on the first terms of the initial condition. Indeed, for $n \geq 2$, consider the subspace of functions whose first $n-1$ terms are 0:

$$E_n = \{ g \in \mathscr{L}^\infty \ :\ g(p) = 0 \text{ for } 1 \leq p < n \}.$$

Denote by P_n the canonical projection from $\mathscr{L}^\infty$ on E_n. For $n \geq 2$ and $g \in \Delta$, we have:

$$P_n \phi(t, g) = P_n\left(\phi(t, P_n(g))\right). \tag{II.62}$$

Let us denote by $\preceq$ the partial order defined by $g \preceq h$ if there exists $n \geq 2$ such that $P_n(g) \leq P_n(h)$.

Suppose that $\limsup n^{3/2} g(n) \leq c_\alpha$. Since $\alpha \to c_\alpha$ is strictly increasing, for any $\alpha < \beta < 1/2$, the asymptotics of g and $\mathfrak{g}_\beta$ imply that $g \preceq \mathfrak{g}_\beta$. Since the flow is order-preserving, this entails $\limsup \phi(t, g) \leq \mathfrak{g}_\beta$. This inequality holds for all $\beta > \alpha$: we get the conclusion by continuity of the map $\Gamma : \alpha \to \mathfrak{g}_\alpha$. The proof of the other implication is similar. $\square$

II.4.6 Uniform convergence

In the subcritical case, Theorem II.4.6 shows an exponentially fast convergence, in the uniform norm. By contrast, the convergence results in the critical and supercritical case from Sections II.4.3 and II.4.4 only hold pointwise.

In the next result, we show how to recover a form of uniformity; in particular we recover uniform convergence in the particular case where $\inf \gamma > 0$.

Theorem II.4.17. *Suppose that Assumption II.2 and II.3 are in force and let $A \in \mathscr{F}$. If γ is bounded away from 0 on A, that is:*

$$\inf_{x\in A} \gamma(x) > 0,$$

then, for $g \in \Delta$, with positive integral if $\mathfrak{g} \neq 0$, we have:

$$\lim_{t\to\infty} \sup_{x\in A} |\phi(t,g)(x) - \mathfrak{g}(x)| = 0.$$

Proof. Set $m = \inf_{x\in A} \gamma(x)$. Let us first study the convergence of the trajectory starting from 1. For $s \in \mathbb{R}_+$, we have:

$$\begin{aligned}
\partial_t(\phi(s,1) - \mathfrak{g}) &= F(\phi(s,1)) - F(\mathfrak{g}) \\
&\leq (1 - \phi(s,1))T_k(\phi(s,1) - \mathfrak{g}) - \gamma(\phi(s,1) - \mathfrak{g}) \\
&\leq T_k(\phi(s,1) - \mathfrak{g}) - \gamma(\phi(s,1) - \mathfrak{g}) \\
&\leq M \left\| \phi(s,1) - \mathfrak{g} \right\|_p - \gamma(\phi(s,1) - \mathfrak{g}),
\end{aligned}$$

where we used that T_k is positive for the second inequality and Hölder inequality for the last with $M = \sup_{x\in\Omega} \left(\int_\Omega k(x,y)^q \, \mu(\mathrm{d}y)\right)^{1/q} < \infty$. For $s \in \mathbb{R}_+$, set $v_s = \mathrm{e}^{ms}(\phi(s,1) - \mathfrak{g})$. Notice that $v_s \geq 0$ and that $\partial_t v_s(x) \leq M \left\| v_s \right\|_p$ for $x \in A$. Integrating for $s \in [0,t]$, we deduce that for $x \in A$:

$$\begin{aligned}
0 \leq (\phi(t,1) - \mathfrak{g})(x) &\leq \mathrm{e}^{-mt}(1 - \mathfrak{g}) + M \int_0^t \mathrm{e}^{-m(t-s)} \left\| \phi(s,1) - \mathfrak{g} \right\|_p \, \mathrm{d}s \\
&\leq \mathrm{e}^{-mt} + M \int_0^t \mathrm{e}^{-ms} \left\| \phi(t-s,1) - \mathfrak{g} \right\|_p \, \mathrm{d}s.
\end{aligned}$$

Note the right hand-side does not depend on x. As $\phi(s,1)$ converges pointwise to $\mathfrak{g}$ (see Equation (II.31)) and is bounded by 1, using the dominated convergence theorem, we deduce that the right hand-side goes to 0 as t goes to infinity. So, we obtain that:

$$\lim_{t\to+\infty} \sup_{x\in A} |\phi(t,1)(x) - \mathfrak{g}(x)| = 0. \tag{II.63}$$

If $\mathfrak{g} = 0$, use that $0 \leq \phi(t,g) \leq \phi(t,1)$ for all $g \in \Delta$ and $t \in \mathbb{R}_+$ to conclude.

If $\mathfrak{g}$ is non zero (which corresponds to the super-critical case), consider a function $f \leq g$ with positive integral such that $f \leq \mathfrak{g}$. By monotonicity of the flow, this implies that $0 \leq \mathfrak{g} - \phi(s,f)$ for all $s \in \mathbb{R}_+$. Arguing similarly as above, we get for $s \in \mathbb{R}_+$:

$$\partial_t(\mathfrak{g} - \phi(s,f)) \leq M \left\| \mathfrak{g} - \phi(s,f) \right\|_p - \gamma(\mathfrak{g} - \phi(s,f)).$$

Using that $\phi(s,f)$ converges pointwise to $\mathfrak{g}$ (see Theorem II.4.13), we similarly get that

$$\lim_{t\to+\infty} \sup_{x\in A} |\phi(t,f)(x) - \mathfrak{g}(x)| = 0. \tag{II.64}$$

Then, use the monotonicity of the flow which implies that $\phi(t,f) \leq \phi(t,g) \leq \phi(t,1)$ for $f \leq g \leq 1$ as well as (II.63) and (II.64) to conclude. □

II.5 Vaccination model

II.5.1 Infinite-dimensional models

We write an infinite-dimensional model with two goals in mind: take into account the heterogeneity in the transmission of the infectious disease, in the spirit of (II.3), and model the effect of vaccination by generalizing Equations (II.13) and (II.14), allowing even for different types of vaccine. Recall that the measurable space $(\Omega, \mathscr{F})$ represents the features of the individuals in a given population, the finite measure μ describes the size of the population and its sub-groups, and the number $\gamma(x)$ is the recovery rate of individuals with feature $x \in \Omega$. The transmission kernel κ describes the way the disease is spread among the population without vaccination.

Suppose that we have different vaccines or treatments available that we can give to individuals in order to fight the disease upstream. The set of vaccines is represented by a set Σ which is finite in practice. We endow Σ with a σ-field $\mathscr{G}$. We are also given two measurable functions $e, \delta : \Omega \times \Sigma \to [0, 1]$. For both models, the number $\delta(x, \xi)$ is the relative reduction of infectiousness for people with feature x vaccinated by the vaccine ξ. The coefficient $e(x, \xi)$ is the efficacy of vaccine ξ given on individuals with feature x. We encode the absence of vaccination by a particular type of vaccine $\xi_0 \in \Sigma$. This vaccination has no efficacy upon the individuals: $e(x, \xi_0) = 0$ and $\delta(x, \xi_0) = 0$ for all $x \in \Omega$. We define a vaccination policy as a non-negative kernel $\eta : \Omega \times \mathscr{G} \to [0, 1]$. The probability for an individual with feature type x to be vaccinated by a vaccine in the measurable set $A \in \mathscr{G}$ under the policy η is equal to $\eta(x, A)$. The recovery rate can be affected by the vaccine, and in this case γ is then a non-negative measurable function defined on $\Omega \times \Sigma$, with $\gamma(x, \xi)$ the recovery rate of individuals with feature x and vaccine ξ. The number $u(t, x, \xi)$ is the probability for an individual with feature x which has been inoculated by the vaccine ξ to be infected at time t. The total number of infected individuals at time t is therefore given by:

$$\int_{\Omega\times\Sigma} u(t, x, \xi)\, \eta(x, \mathrm{d}\xi)\mu(\mathrm{d}x). \tag{II.65}$$

The leaky vaccination mechanism

In this setting, $e(x, \xi)$ denotes the leaky vaccine efficacy of $\xi \in \Sigma$ on an individual with feature x, *i.e.*, the relative reduction in the transmission rate. We generalize Equation (II.13) to get the following infinite dimensional evolution equation:

$$\begin{aligned}\partial_t u(t, x, \xi) &= -\gamma(x, \xi)\, u(t, x, \xi) \\ &\quad + (1 - u(t, x, \xi))(1 - e(x, \xi)) \int_{\Omega\times\Sigma} (1 - \delta(y, \zeta))u(t, y, \zeta)\kappa(x, \mathrm{d}y)\eta(y, \mathrm{d}\zeta).\end{aligned} \tag{II.66}$$

The evolution Equation (II.66) can be seen as the SIS evolution Equation (II.3) on an extended feature space:

- the feature $\mathbf{x} = (x, \xi)$ lives in $\mathbf{\Omega} = \Omega \times \Sigma$ endowed with the σ-field $\mathscr{F} \otimes \mathscr{G}$,
- the recovery rate is given by $\boldsymbol{\gamma}(\mathbf{x}) = \gamma(x, \xi)$,
- the extended transmission kernel is given by:

$$\boldsymbol{\kappa}^{\ell}(\mathbf{x}, \mathrm{d}\mathbf{y}) = (1 - e(x, \xi))(1 - \delta(y, \zeta))\kappa(x, \mathrm{d}y)\eta(y, \mathrm{d}\zeta). \tag{II.67}$$

Remark II.5.1. In the leaky mechanism, we suppose that the vaccine acts directly on the susceptibility and the infectiousness of the individuals. Protective gears (like respirators or safety glasses) which are designed to protect the wearer from absorbing airborne microbes or transmitting them have a similar effect. Hence, Equation (II.66) is not limited to vaccination and can also be used as a model for distribution of equipment in the population.

The all-or-nothing mechanism

In this setting, $e(x,\xi)$, is defined as the probability to immunize completely the individual with feature x to the disease with vaccine ξ. We generalize Equation (II.14) to get the following infinite dimensional evolution equation:

$$\begin{aligned}\partial_t u(t,x,\xi) &= -\gamma(x,\xi)\,u(t,x,\xi)\\ &\quad + (1-e(x,\xi)-u(t,x,\xi))\int_{\Omega\times\Sigma}(1-\delta(y,\zeta))u(t,y,\zeta)\kappa(x,\mathrm{d}y)\eta(y,\mathrm{d}\zeta).\end{aligned} \tag{II.68}$$

The probability $v(t,x,\xi) = u(t,x,\xi)/(1-e(x,\xi))$ for an individual with feature x which has not been vaccinated by the inoculation of vaccine ξ to be infected at time t satisfies the following equation:

$$\begin{aligned}\partial_t v(t,x,\xi) &= -\gamma(x)v(t,x,\xi)\\ &\quad + (1-v(t,x,\xi))\int_{\Omega\times\Sigma}(1-\delta(y,\zeta))v(t,y,\zeta)(1-e(y,\zeta))\kappa(x,\mathrm{d}y)\eta(y,\mathrm{d}\zeta).\end{aligned} \tag{II.69}$$

As before, the evolution Equation (II.69) can be seen as the SIS evolution Equation (II.3) on the same extended feature space $\boldsymbol{\Omega} = \Omega\times\Sigma$, still endowed with the σ-field $\mathscr{F}\otimes\mathscr{G}$, with the same recovery rate $\boldsymbol{\gamma}(\mathbf{x}) = \gamma(x,\xi)$, but the transmission kernel now reads:

$$\boldsymbol{\kappa}^a(\mathbf{x},\mathrm{d}\mathbf{y}) = (1-e(y,\zeta))(1-\delta(y,\zeta))\kappa(x,\mathrm{d}y)\eta(y,\mathrm{d}\zeta). \tag{II.70}$$

Notice the difference between the evolution Equation (II.66) for leaky mechanism and the evolution Equation (II.69) for the all-or-nothing mechanism is that $e(y,\zeta)$ in (II.69) (or in the kernel $\boldsymbol{\kappa}^a$ from (II.70)) is replaced by $e(x,\xi)$ in (II.66) (or in the kernel $\boldsymbol{\kappa}^\ell$ from (II.67)).

II.5.2 Discussion on the basic reproduction number

Suppose that Assumption II.2 is in force. Then, we can define a new basic reproduction number for the vaccination models. We consider the following bounded operators on $\mathscr{L}^\infty(\Omega\times\Sigma)$:

$$T(g)(x,\xi) = \int_{\Omega\times\Sigma}(1-\delta(y,\zeta))g(y,\zeta)\,\frac{\kappa(x,\mathrm{d}y)}{\gamma(y,\zeta)}\,\eta(y,\mathrm{d}\zeta),$$

$$M(g)(x,\xi) = (1-e(x,\xi))g(x,\xi).$$

Following Section II.4.1, the all-or-nothing vaccination reproduction number $\mathfrak{R}_0^a(\eta)$ associated to Equation (II.69) and vaccine policy η is:

$$\mathfrak{R}_0^a(\eta) = \rho(TM). \tag{II.71}$$

where we recall that r stands for the spectral radius. For the leaky vaccination, the basic reproduction number associated to Equation (II.66) and vaccine policy η is:

$$\mathfrak{R}_0^\ell(\eta) = \rho(MT). \tag{II.72}$$

In [140], the authors already remarked that the two vaccination mechanisms actually leads to the same basic reproduction number for the one-group models. This result also holds in the infinite-dimension SIS model. Notice that Assumption II.2 insures that those two basic reproduction numbers are well defined.

Proposition II.5.2. *We assume Assumption II.2 holds. Let η be a vaccination policy. The basic reproduction number for the leaky vaccination and the for the all-or-nothing vaccination are the same:*

$$\mathfrak{R}_0^\ell(\eta) = \mathfrak{R}_0^a(\eta).$$

Proof. Thanks to the definition of the spectral radius (II.34) and the basic reproduction numbers defined in (II.71) and (II.72), the result is a direct consequence of the following equality on the spectra:

$$\sigma(MT) \cup \{0\} = \sigma(TM) \cup \{0\}.$$

We prove this later equality by following [127, Appendix A1]. Let $\lambda \in \mathbb{C}\backslash(\sigma(MT) \cup \{0\})$. By definition, there exists a bounded operator A on $\mathscr{L}^\infty(\Omega \times \Sigma)$ such that:

$$A(\lambda \mathrm{Id} - MT) = (\lambda \mathrm{Id} - MT)A = \mathrm{Id},$$

where Id is the identity operator. Then, one can check easily that $\lambda^{-1}(\mathrm{Id} + TAM)$ is the inverse of $\lambda\mathrm{Id} - TM$, whence $\lambda \in \mathbb{C}\backslash(\sigma(TM) \cup \{0\})$. This gives that $\sigma(TM) \cup \{0\}) \subset \sigma(MT) \cup \{0\})$. The other inclusion is proved similarly. □

II.5.3 The perfect vaccine

The simplest case is a situation where there is only one vaccine with complete efficacy on every individual: $\Sigma = \{\xi_0, \xi_1\}$ with $e(x, \xi_1) = 1$ and $\delta(x, \xi_1) = 1$ for all $x \in \Omega$. Recall that ξ_0 corresponds to the absence of vaccine. For simplicity, we denote by $\eta^0(x) = \eta(x, \{\xi_0\})$ the probability for (or the proportion of) individuals of type $x \in \Omega$ which are not vaccinated. We assume for simplicity that initially no vaccinated individuals are infected, that is $u(0, x, \xi_1) = 0$. Since individuals that have been vaccinated are fully immunized, we have $u(t, x, \xi_1) = 0$ for all x and t. The only equation that matters is the one on $u^0(t, x) = u(t, x, \xi_0)$ which represents the proportion of unvaccinated individuals that are infected. For both mechanisms (all-or-nothing and leaky vaccination), the evolution equation of u^0 writes:

$$\partial_t u^0(t, x) = (1 - u^0(t, x)) \int_\Omega u^0(t, y)\eta^0(y)\kappa(x, \mathrm{d}y) - \gamma(x)u^0(t, x). \tag{II.73}$$

We shall use this formulation in a future work to find optimal vaccination policies for a given cost.

II.6 Limiting contacts within the population

Motivated by the recent lockdown policies taken by many countries all around the world to slow down the propagation of Covid-19 in 2020, we propose to investigate the possible impact on our SIS model of the limitations of contacts within the population. We consider the case where κ takes the form of Example II.1.3:

$$\kappa_W(x, \mathrm{d}y) = \beta(x)W(x, y)\theta(y)\,\mu(\mathrm{d}y),$$

where β is the susceptibility function, θ is the infectiousness function, μ is a probability measure on the space Ω of features of the individuals and the graphon W represents the initial graph of the contacts between individuals of the population (recall that $W(x, y) = W(y, x) \in [0, 1]$ is the probability that x and y are connected and can be also seen as the density of contact between the individuals with features x and y). In order to stress the dependence in W, we write $\mathfrak{R}_0(W) = \rho(T_{\kappa_W/\gamma})$ the corresponding basic reproduction number and ϕ_W the semi-flow (II.28) associated to $F = F_W$ in (II.24) given by $F_W(g) = (1 - g)T_{\kappa_W}(g) - \gamma g$. We model the impact of a policy which reduces the contacts between the individuals, by a new graph of contact given by a new graphon W'. We say that W' is a perfect lockdown with respect to W if:

$$W'(x, y) \le W(x, y), \qquad \forall x, y \in \Omega. \tag{II.74}$$

Intuitively x and y have a lesser probability to be connected in the graphon W' than in the graphon W. We get the following intuitive result as a direct application of Theorem II.3.5 (i) and Corollary II.2.4.

Proposition II.6.1 (Perfect Lockdown)**.** *Assume that β and θ are bounded and γ is bounded away from 0. If W' is a perfect lockdown with respect to W then $\mathfrak{R}_0(W') \leq \mathfrak{R}_0(W)$ and $\phi_{W'}(t,g) \leq \phi_W(t,g)$ for all initial condition $g \in \Delta$.*

However, assuming that all the contacts within the population are reduced might be unrealistic (*e.g.* people can have stronger contacts with their family in lockdown). Instead, we can suppose as a weaker condition, that each individual reduces the average number of contacts he has. Recall (II.6) for the definition of the degree $\deg_W(x)$ of an individual $x \in \Omega$ (*i.e.* the average number of his contacts) and the mean degree d_W over the population for a graphon W as:

$$\deg_W(x) = \int_\Omega W(x,y)\,\mu(\mathrm{d}y) \quad \text{and} \quad \mathrm{d}_W = \int_\Omega \deg_W(x)\,\mu(\mathrm{d}x) = \int_{\Omega^2} W(x,y)\,\mu(\mathrm{d}y)\,\mu(\mathrm{d}x).$$

Recall that $\|\cdot\|_1$ is the usual $L^1(\mu)$ norm. The following lemma bounds the basic reproduction number with the supremum and the mean degree of the graphon.

Lemma II.6.2. *Let W be a graphon. Assume that β and θ/γ are bounded. We have that:*

$$\frac{1}{\|\gamma/\beta\theta\|_1}\,\mathrm{d}_W \leq \mathfrak{R}_0(W) \leq \|\beta\theta/\gamma\| \sup_{x\in\Omega} \deg_W(x). \tag{II.75}$$

Proof. Recall T_k is the operator defined by (II.22) with $\kappa(x,\mathrm{d}y) = \mathrm{k}(x,y)\,\mu(\mathrm{d}y)$. Let $M(v)$ be the operator corresponding to the multiplication by the function v. We have:

$$\begin{aligned}
\mathfrak{R}_0(W) &= \rho(M(\beta)\,T_W\,M(\theta/\gamma)) \\
&= \rho(M(\beta\theta/\gamma)\,T_W) \\
&\leq \|M(\beta\theta/\gamma)\,T_W\| \\
&= \sup_{x\in\Omega} \frac{\beta(x)\theta(x)}{\gamma(x)} \int_\Omega W(x,y)\,\mu(\mathrm{d}y) \\
&\leq \|\beta\theta/\gamma\| \sup_{x\in\Omega} \deg_W(x),
\end{aligned}$$

where we used the definition of the basic reproduction number (II.12) for the first equality, arguments similar as in the proof of Proposition II.5.2 for the second, and the (third) definition of the spectral radius (II.34) for the first inequality.

Using similar arguments, we have:

$$\mathfrak{R}_0(W) = \rho(M(\beta)\,T_W\,M(\theta/\gamma)) = r(M(v)\,T_W\,M(v)),$$

with $v = \sqrt{\beta\theta/\gamma}$. Recall notations from Lemma II.3.7, and notice that $M(v)\,T_W\,M(v) = T_\mathrm{k}$ is a bounded integral operator on $\mathscr{L}^\infty$ associated to the symmetric kernel $\mathrm{k}(x,y) = v(x)W(x,y)v(y)$. According to Lemma II.3.7 (iv) with $q = p = 1/2$ and $\hat{T}_\mathrm{k}$ the integral operator on $L^2(\mu)$ with the same kernel k, defined in (II.38), we get $\mathfrak{R}_0(W) = \rho(\hat{T}_\mathrm{k})$. The operator $\hat{T}_\mathrm{k}$ is self-adjoint, as k is symmetric, and compact according to (iii). Thanks to the Courant-Fischer-Weyl min-max principle, we obtain:

$$\mathfrak{R}_0(W) = \rho(\hat{T}_\mathrm{k}) = \sup_{g\in L^2(\mu)\setminus\{0\}} \frac{\langle M(v)\,g,\, T_W\,M(v)\,g\rangle}{\langle g, g\rangle}.$$

Taking $g = 1/v$, we get $M(v)g = 1$ and thus:

$$\mathfrak{R}_0(W) \geq \frac{\langle 1, T_W 1\rangle}{\|\gamma/\beta\theta\|_1} = \frac{d_W}{\|\gamma/\beta\theta\|_1}.$$

This ends the proof of Lemma II.6.2. □

We deduce from Lemma II.6.2 the following result for a lockdown policy W' for which the degree of each individuals is less than the average degree of the initial graphon W.

Corollary II.6.3 (Partial Lockdown)**.** *Assume that β and θ/γ are bounded. If W' is a partial lockdown of W, that is:*

$$\sup_{x\in\Omega} \deg_{W'}(x) \leq C \mathrm{d}_W \quad \textit{with} \quad C = \frac{1}{\|\beta\theta/\gamma\| \, \|\gamma/\beta\theta\|_1}, \tag{II.76}$$

then we have $\mathfrak{R}_0(W') \leq \mathfrak{R}_0(W)$.

In the general case, we have $C \leq 1$. But, if the functions β, θ and γ are constants (or simply if $\beta\theta/\gamma$ constant), then we have $C = 1$ since μ is a probability measure.

Remark II.6.4. Suppose that β, θ and γ are constants (or that $\beta\theta/\gamma$ is constant). Inequality (II.75) shows that the graphon W which corresponds to a minimal basic reproduction number $\mathfrak{R}_0(W)$, when the mean degree d_W is fixed, say equal to p, is any graphon with constant degree equal to p, that is $\deg_W(x) = p$ for all $x \in \Omega$. We then deduce from Lemma II.6.2 that $\mathfrak{R}_0(W) = p\beta\theta/\gamma$.

This is in particular the case for the constant graphon $W = p \in [0, 1]$. According to Example II.1.2(i), this corresponds to the one dimensional SIS model (II.1).

This is also the case for the geometric graphon, where the probability of edges between x and y depends only on the distance between x and y. Keeping notations from Example II.1.2(iii), we consider the population uniformly spread on the unit circle: $\Omega = [0, 2\pi]$ and $\mu(\mathrm{d}x) = \mathrm{d}x/2\pi$, and the graphon W_f defined by $W_f(x, y) = f(x - y)$ for $x, y \in \Omega$, where f is a measurable non-negative function defined on $\mathbb{R}$ which is bounded by 1 and 2π-periodic. Let $p = (2\pi)^{-1} \int_{[0,2\pi]} f(y)\, \mathrm{d}y$. We have: $\deg_W(x) = \mathrm{d}_W = p$; the basic reproduction number $\mathfrak{R}_0(W_f) = p\beta\theta/\gamma$ and the maximal equilibrium $\mathfrak{g} = \max(0, 1 - \mathfrak{R}_0^{-1})$. Furthermore, the graphon W_f minimizes the basic reproduction number among all graphons with mean degree p. It is interesting to notice that $\mathfrak{R}_0(W_f)$ does not depend on the support of f or even on $\sup\{|r| \,:\, r \in [-\pi, \pi] \text{ and } f(r) > 0\}$, which can be seen as the maximal contamination distance from an infected individual.

II.7 Proof of Theorem II.2.4

We use notation from Section II.2.1 and let X be a Banach space. Let us first recall a few definitions and classical properties of ODEs. Let $a > 0$. We consider a function $G : [0, a) \times X \to X$. We suppose that G is locally Lipschitz in the second variable, that is: for all $(t, x) \in [0, a) \times X$, there exist $\eta = \eta(t, x) > 0$, $L = L(t, x) > 0$ and a neighborhood U_x of x such that $\| G(s, y) - G(s, z) \| \leq L \| y - z \|$ for all $s \in [0, a) \cap [t, t + \eta]$ and $y, z \in U_x$. With this assumption over G, the Picard–Lindelöf theorem ensures the existence of $0 < b \leq a$ and a continuously differentiable function y from $J = [0, b)$ to X which is the unique solution of the Cauchy problem:

$$\begin{cases} y'(t) = G(t, y(t)) & t \in J, \\ y(0) = y_0, \end{cases} \tag{II.77}$$

where $y_0 \in X$ is the so-called initial condition (see [34, Section 1.1]). A solution y defined on an interval $[0, b)$ is said to be maximal if there is no solution of Equation (II.77) defined on $[0, c)$ with $c > b$. A solution is said to be global if it is defined on $[0, a)$.

Global existence, existence and theorems on differential inequalities are intimately connected with the flow invariance of certain subsets in the domain of G, *i.e.*, the question whether every solution starting in D remains in D as long as it exists. We recall the definition of flow invariance given in [34, Section 5].

Definition II.7.1 (Forward invariance)**.** *A set $D \subset X$ is said to be forward invariant with respect to G if the maximal solution (y, J) of the Cauchy problem* (II.77) *takes values in D for $t \in J$ provided that $y_0 \in D$.*

In most applications, the set D owns a structure which make the forward invariance easier to show. For instance, when D is the translation of a cone, the forward invariance is implied by the following condition:

Theorem II.7.2. *Let $G : [0,a) \times X \to X$ be locally Lipschitz in the second variable. Let K be a proper cone of X with non-empty interior and $y \in X$. If for all $(x,t) \in \partial K \times [0,a)$ and for all $x^\star \in K^\star$ such that $\langle x^\star, x\rangle = 0$, we have: $\langle x^\star, G(t, y+x)\rangle \geq 0$, then $y + K$ is forward invariant with respect to G.*

Before proving this result let us state two lemmas. Let X be a Banach space. For $x \in X$ and $D \subset X$, we denote by $\mathrm{d}(x, D)$ the distance between x and the set D:

$$\mathrm{d}(x, D) = \inf_{y \in D} \| x - y \| . \tag{II.78}$$

Let $a > 0$ and $G : [0,a) \times X \to X$ be a locally Lipschitz function with respect to the second variable. Recall Definition II.7.1 of a forward invariant set with respect to G. The following result appears in [34, Theorem 5.2].

Lemma II.7.3. *Let D be a closed convex set with non-empty interior. Suppose that G satisfies:*

$$\lim_{\lambda \to 0^+} \frac{1}{\lambda} \mathrm{d}(x + \lambda G(t,x), D) = 0, \qquad \forall (t,x) \in (0,a) \times \partial D. \tag{II.79}$$

Then D is forward invariant with respect to G.

If the set D is a proper cone, the following equivalence enables to establish (II.79) more easily. It is a consequence of [34, Lemma 4.1] and [34, Example 4.1.ii]

Lemma II.7.4. *Let K be a proper cone and let $x \in \partial K$ and $z \in X$. The following conditions are equivalent:*

(i) $\lim_{\lambda \to 0^+} \lambda^{-1} \mathrm{d}(x + \lambda z, K) = 0.$

(ii) For all $x^\star \in K^\star$ such that $\langle x^\star, x\rangle = 0$, we have $\langle x^\star, z\rangle \geq 0$.

We have now all the tools to prove Theorem II.7.2.

Proof of Theorem II.7.2. Let $y \in X$. We assume that, for all $(x,t) \in \partial K \times [0,a)$ and for all $x^\star \in K^\star$ such that $\langle x^\star, x\rangle = 0$, we have: $\langle x^\star, G(t, y+x)\rangle \geq 0$. According to Lemma II.7.4, we obtain:

$$\lim_{\lambda \to 0^+} \lambda^{-1} \mathrm{d}(x + \lambda G(t, y+x), K) = 0,$$

for all $(x,t) \in \partial K \times [0,a)$. Since $\mathrm{d}(y+x+\lambda G(t,y+x), y+K) = \mathrm{d}(x+\lambda G(t,y+x), K)$ by Equation (II.78), we can conclude the proof using Lemma II.7.3 with $D = y + K$. ☐

Finally, the main comparison result used in the text, Theorem II.2.4, may be proved as a corollary.

Proof of Theorem II.2.4. We suppose that F is cooperative on $D_1 \times X$ and the inequality (II.20) holds. Let $w = v - u$. The function w is solution of the ODE $w' = G(t,w)$ where:

$$G(t,x) = F(u(t) + x) - F(u(t)) + d(t) \qquad \text{and} \qquad d(t) = v'(t) - F(v(t)) - u'(t) + F(u(t)).$$

First we show that G is locally Lipschitz with respect to the second variable. Let $(t,x) \in [0,a) \times X$. Let U be a neighborhood of $u(t) + x$ such that F is Lipschitz on U with a Lipschitz constant L. By continuity of u, there exist a neighborhood V_x of x and a positive constant η, such that $u(s) + y \in U$, for all $s \in [t, t+\eta] \cap [0,a)$ and $y \in V_x$. Thus, for all $s \in [t, t+\eta] \cap [0,a)$ and $y, z \in V_x$, we have $\| G(s,y) - G(s,z) \| \leq L \| y - z \|$.

Let $t \in [0,a)$, $x \in \partial K$ and let $x^\star \in K^\star$ such that $\langle x^\star, x\rangle = 0$. Let us prove that $\langle x^\star, G(t,x)\rangle \geq 0$. By (II.20), we know that $d(t)$ belongs to K. Furthermore, the inequality $\langle x^\star, F(u(t)+x) - F(u(t))\rangle \geq 0$ holds because the function F is cooperative on $D_1 \times X$. Thus, $\langle x^\star, G(t,x)\rangle$ is non-negative. Hence, we can apply Theorem II.7.2 with $y = 0$ and obtain that K is forward invariant with respect to G. Since $w(0) \in K$, this shows that $w(t) \in K$ for all $t \in [0,a)$, *i.e.*, $u(t) \leq v(t)$ for all $t \in [0,a)$.

When F is cooperative on $X \times D_2$, the proof is similar. ☐

II.8 The Hausdorff distance on the compact sets of $\mathbb{C}$

Let $\mathcal{K}(\mathbb{C})$ be the set of non-empty compact subsets of $\mathbb{C}$. The Hausdorff distance between A and B in $\mathcal{K}(\mathbb{C})$ is defined as:

$$d_{\mathrm{H}}(A, B) = \max\left\{\sup_{z_1 \in A} \inf_{z_2 \in B} |\, z_1 - z_2 \,|, \quad \sup_{z_2 \in B} \inf_{z_1 \in A} |\, z_2 - z_1 \,|\right\}. \tag{II.80}$$

We recall that the space $(\mathcal{K}(\mathbb{C}), d_{\mathrm{H}})$ is a metric space, see [26, Section 7.3.1]. Since:

$$\sup\{\, |z| \;:\; z \in A \,\} = d_{\mathrm{H}}(A, \{\, 0 \,\}),$$

for all $A \in \mathcal{K}(\mathbb{C})$, we deduce the following result.

Lemma II.8.1. *The map $A \mapsto \sup\{\, |z| \;:\; z \in A \,\}$ from $(\mathcal{K}(\mathbb{C}), d_{\mathrm{H}})$ to $\mathbb{R}$ endowed with the usual Euclidean distance is continuous.*

Chapter III

Targeted vaccination strategies for an infinite-dimensional SIS model

Chapter Abstract

We formalize and study the problem of optimal allocation strategies for a (perfect) vaccine in an infinite-dimensional metapopulation SIS model. The question may be viewed as a bi-objective minimization problem, where one tries to minimize simultaneously the cost of the vaccination, and a loss that may be either the effective reproduction number, or the overall proportion of infected individuals in the endemic state. We prove the existence of Pareto optimal strategies for both loss functions.

We also show that vaccinating according to the profile of the endemic state is a critical allocation, in the sense that, if the initial reproduction number is larger than 1, then this vaccination strategy yields an effective reproduction number equal to 1.

The material for this chapter has been released in [39].

Chapter Content:

III.1 Introduction

III.1.1 Motivation

Increasing the prevalence of immunity from contagious disease in a population limits the circulation of the infection among the individuals who lack immunity. This so-called "herd effect" plays a fundamental role in epidemiology as it has had a major impact in the eradication of smallpox and rinderpest or the near eradication of poliomyelitis; see [60]. Targeted vaccination strategies, based on the heterogeneity of the infection spreading in the population, are designed to increase the level of immunity of the population with a limited quantity of vaccine. These strategies rely on identifying groups of individuals that should be vaccinated in priority in order to slow down or eradicate the disease.

In this chapter, we establish a theoretical framework to study targeted vaccination strategies for the deterministic infinite-dimensional SIS model introduced in Chapter II, that encompasses as particular cases the SIS model on graphs or on stochastic block models. In the next chapters, we provide a series of general and specific examples that complete and illustrate the present work: see Section III.1.5 for more detail.

III.1.2 Herd immunity and targeted vaccination strategies

Let us start by recalling a few classical results in mathematical epidemiology; we refer to Keeling and Rohani's monograph [93] for an extensive introduction to this field, including details on the various classical models (SIS, SIR, etc.)

In an homogeneous population, the basic reproduction number of an infection, denoted by $\mathfrak{R}_0$, is defined as the number of secondary cases one individual generates on average over the course of its infectious period, in an otherwise uninfected (susceptible) population. This number plays a fundamental role in epidemiology as it provides a scale to measure how difficult an infectious disease is to control. Intuitively, the disease should die out if $\mathfrak{R}_0 < 1$ and invade the population if $\mathfrak{R}_0 > 1$. For many classical mathematical models of epidemiology, such as SIS or S(E)IR, this intuition can be made rigorous: the quantity $\mathfrak{R}_0$ may be computed from the parameters of the model, and the threshold phenomenon occurs.

Assuming $\mathfrak{R}_0 > 1$ in an homogeneous population, suppose now that only a proportion η^{uni} of the population can catch the disease, the rest being immunized. An infected individual will now only generate $\eta^{\text{uni}}\mathfrak{R}_0$ new cases, since a proportion $(1 - \eta^{\text{uni}})$ of previously successful infections will be prevented. Therefore, the new *effective reproduction number* is equal to $\mathfrak{R}_e(\eta^{\text{uni}}) = \eta^{\text{uni}}\mathfrak{R}_0$. This fact led to the recognition by Smith in 1970 [141] and Dietz in 1975 [43] of a simple threshold theorem: the incidence of an infection declines if the proportion of non-immune individuals is reduced below $\eta^{\text{uni}}_{\text{crit}} = 1/\mathfrak{R}_0$. This effect is called *herd immunity*, and the corresponding percentage $1 - \eta^{\text{uni}}_{\text{crit}}$ of people that have to be vaccinated is called *herd immunity threshold*; see for instance [147, 148].

It is of course unrealistic to depict human populations as homogeneous, and many generalizations of the homogeneous model have been studied; see [93, Chapter 3] for examples and further references. For most of these generalizations, it is still possible to define a meaningful reproduction number $\mathfrak{R}_0$, as the number of secondary cases generated by a *typical* infectious individual when all other individuals are uninfected; see [42]. After a vaccination campaign, let the vaccination strategy η denote the (non necessarily homogeneous) proportion of the **non-vaccinated** population, and let the effective reproduction number $\mathfrak{R}_e(\eta)$ denote the corresponding reproduction number of the non-vaccinated population. The vaccination strategy η is *critical* if $\mathfrak{R}_e(\eta) = 1$. The possible choices of η naturally raises a question that may be expressed as the following informal optimization problem:

$$\begin{cases} \textbf{Minimize:} & \text{the quantity of vaccine to administrate} \\ \textbf{Such that:} & \text{herd immunity is reached, that is, } \mathfrak{R}_e \leq 1. \end{cases} \tag{III.1}$$

If the quantity of available vaccine is limited, then one is also interested in:

$$\begin{cases} \textbf{Minimize:} & \text{the effective reproduction number } \mathfrak{R}_e \\ \textbf{Such that:} & \text{a given quantity of available vaccine.} \end{cases} \tag{III.2}$$

Interestingly enough, the strategy $\eta^{\text{uni}}_{\text{crit}}$, which consists in delivering the vaccine *uniformly* to the population, without taking inhomogeneity into account, leaves a proportion $\eta^{\text{uni}}_{\text{crit}} = 1/\mathfrak{R}_0$ of the population unprotected, and is therefore critical since $\mathfrak{R}_e(\eta^{\text{uni}}_{\text{crit}}) = 1$. In particular it is admissible for the optimization problem (III.1).

However, herd immunity may be achieved even if the proportion of unprotected people is *greater* than $1/\mathfrak{R}_0$, by targeting certain group(s) within the population; see Figure 3.3 in [93]. For example, the discussion of vaccination control of gonorrhea in [80, Section 4.5] suggests that it may be better to prioritize the vaccination of people that have already caught the disease: this leads us to consider a vaccination strategy guided by the equilibrium state. This strategy denoted by η^{equi} will be defined formally below. Let us mention here an observation in the same vein made by Britton, Ball and Trapman in [25]. Recall that in the S(E)IR model, immunity can be obtained through infection. Using parameters from real-world data, these authors noticed that the disease-induced herd immunity level can, for some models, be substantially lower than the classical herd immunity threshold $1 - 1/\mathfrak{R}_0$. This can be reformulated in term of targeted vaccination strategies: prioritizing the individuals that are more likely to get infected in a S(E)IR epidemic may be more efficient than distributing uniformly the vaccine in the population.

The main goal of this chapter is two-fold: formalize the optimization problems (III.1) and (III.2) for a particular infinite dimensional SIS model, recasting them more generally as a bi-objective optimization problem; and give existence and properties of solutions to this bi-objective problem. We will also consider a closely related problem, where one wishes to minimize the size of the epidemic rather than the reproduction number. We will in passing provide insight on the efficiency of classical vaccination strategies such as $\eta^{\text{uni}}_{\text{crit}}$ or η^{equi}.

III.1.3 Literature on targeted vaccination strategies

Targeted vaccination problems have mainly been studied using two different mathematical frameworks.

On meta-populations models

Problems (III.1) and (III.2) have been examined in depth for deterministic *meta-population* models, that is, models in which an heterogeneous population is stratified into a finite number of homogeneous sub-populations (by age group, gender, ...). Such models are specified by choosing the sizes of the subpopulations and quantifying the degree of interactions between them, in terms of various mixing parameters. In this setting, $\mathfrak{R}_0$ can often be identified as the spectral radius of a *next-generation matrix* whose coefficients depend on the subpopulation sizes, and the mixing parameters. It turns out that the next generation matrices take similar forms for many dynamics (SIS, SIR, SEIR,...); see the discussions in [81, Section 10] and Section IV.2. Vaccination strategies are defined as the levels at which each sub-population is immunized. After vaccination, the next-generation matrix is changed and its new spectral radius corresponds to the effective reproduction number $\mathfrak{R}_e$.

Problem (III.1) has been studied in this setting by Hill and Longini [81]. These authors study the geometric properties of the so-called threshold hypersurface, that is, the vaccination allocations for which $\mathfrak{R}_e = 1$. They also compute the vaccination belonging to this surface with minimal cost for an Influenza A model. Making structural assumptions on the mixing parameters, Poghotanyan, Feng, Glasser and Hill derive in [130] an analytical formula for the solutions of Problem (III.2), for populations divided in two groups. Many papers also contain numerical studies of the optimization problems (III.1) and (III.2) on real-world data using gradient techniques or similar methods; see for example [46, 56, 57, 67, 161].

Finally, the effective reproduction number is not the only reasonable way of quantifying a population's vulnerability to an infection. For an SIR infection for example, the proportion of individuals that eventually catch (and recover from) the disease, often referred to as the *attack rate*, is broadly used. We refer to [46, 47] for further discussion on this topic.

On networks

Whereas the previously cited works typically consider a small number of subpopulations, often with a "dense" structure of interaction (every subpopulation may directly infect all the others), other research communities have looked into a similar problem for graphs. Indeed, given a (large), possibly random graph, with epidemic dynamics on it, and supposing that we are able to suppress vertices by vaccinating, one may ask for the best way to choose the vertices to remove.

The importance of the spectral radius of the network has been rapidly identified as its value determines if the epidemic dies out quickly or survives for a long time [66, 132]. Since Van Mieghem *et al.* proved in [154] that the problem of minimizing spectral radius of a graph by removing a given number of vertices is NP-complete (and therefore unfeasible in practice), many computational heuristics have been put forward to give approximate solutions; see for example [135] and references therein.

III.1.4 Main results

The differential equations governing the epidemic dynamics in meta-population SIS models were developed by Lajmanovich and Yorke in their pioneer paper [102]. In Chapter II, we introduced a natural generalization of their equation, to a possibly infinite space Ω, where $x \in \Omega$ represents a feature and the probability measure $\mu(\mathrm{d}x)$ represents the fraction of the population with feature x.

Regularity of the effective reproduction function $\mathfrak{R}_e$

We consider the effective reproduction function in a general operator framework which we call the *kernel model*. This model, which will be defined in detail below in Section III.2, is characterized by a probability space $(\Omega, \mathscr{F}, \mu)$ and a measurable non-negative kernel $\mathbf{k} : \Omega \times \Omega \to \mathbb{R}_+$. Let $T_\mathbf{k}$ be the corresponding integral operator defined by:

$$T_\mathbf{k}(h)(x) = \int_\Omega \mathbf{k}(x, y)h(y)\, \mu(\mathrm{d}y).$$

In the setting of Chapter II (see in particular Equation (11) therein), $T_\mathbf{k}$ is the so-called *next generation operator*, where the kernel $\mathbf{k}$ is defined in terms of a transmission rate kernel $k(x, y)$ and a recovery rate function γ by the product $\mathbf{k}(x, y) = k(x, y)/\gamma(y)$; the reproduction number $\mathfrak{R}_0$ is then the spectral radius $\rho(T_\mathbf{k})$ of $T_\mathbf{k}$.

Following Chapter II Section 5, we represent a vaccination strategy by a function $\eta : \Omega \to [0, 1]$, where $\eta(x)$ represents the fraction of **non-vaccinated** individuals with feature x; the effective reproduction number associated to η is then given by

$$\mathfrak{R}_e(\eta) = \rho(T_{\mathbf{k}\eta}), \tag{III.3}$$

where ρ stands for the spectral radius and $\mathbf{k}\eta$ stands for the kernel $(\mathbf{k}\eta)(x, y) = \mathbf{k}(x, y)\eta(y)$. If $\mathfrak{R}_0 \geq 1$, then a vaccination strategy η is called *critical* if it achieves precisely the herd immunity threshold, that is $\mathfrak{R}_e(\eta) = 1$.

In particular, the "strategy" that consists in vaccinating no one (resp. everybody) corresponds to $\eta = \mathbb{1}$, the constant function equal to 1, (resp. $\eta = \mathbb{0}$, the constant function equal to 0), and of course $\mathfrak{R}_e(\mathbb{1}) = \mathfrak{R}_0$ (resp. $\mathfrak{R}_e(\mathbb{0}) = 0$). As the spectral radius is positively homogeneous, we also get, when $\mathfrak{R}_0 \geq 1$, that the uniform strategy that corresponds to the constant function:

$$\eta^{\mathrm{uni}}_{\mathrm{crit}} = \frac{1}{\mathfrak{R}_0}\mathbb{1}$$

is critical, as $\mathfrak{R}_e(\eta^{\text{uni}}_{\text{crit}}) = 1$. This is consistent with results obtained in the homogeneous model that were recalled in Section III.1.2.

Let Δ be the set of strategies, that is, the set of $[0, 1]$-valued functions defined on Ω. The usual technique to obtain the existence of solutions to optimization problems like (III.1) or (III.2) is to prove that the function $\mathfrak{R}_e$ is continuous with respect to a topology for which the set of strategies Δ is compact. It is natural to try and prove this continuity by writing $\mathfrak{R}_e$ as the composition of the spectral radius ρ and the map $\eta \mapsto T_{\mathbf{k}\eta}$. The spectral radius is indeed continuous at compact operators (and $T_{\mathbf{k}\eta}$ is in fact compact under a technical integrability assumption on the kernel $\mathbf{k}$ formalized on page 75 as Assumption III.1), if we endow the set of bounded operators with the operator norm topology; see [27, 119]. However, this only works if we equip Δ with the uniform topology, for which it is not compact.

We instead endow Δ with the weak topology, see Section III.3.1, for which compactness holds; see Lemma III.3.1. This forces us to equip the space of bounded operators with the strong topology, for which the spectral radius is in general not continuous; see [92, p. 431]. However, the family of operators $(T_{\mathbf{k}\eta}, \eta \in \Delta)$ is *collectively compact* which enables us to recover continuity, using a series of results obtained by Anselone [6]. This leads to the following statement, proved in Theorem III.4.2 below. We recall that Assumption III.1, formulated on page 75, provides an integrability condition on the kernel $\mathbf{k}$. In particular, it is satisfied if $\mathbf{k}$ is bounded or even in $L^2(\Omega \times \Omega, \mu \otimes \mu)$.

Theorem III.1.1 (Continuity of the spectral radius). *Under Assumption III.1 on the kernel* $\mathbf{k}$*, the function* $\mathfrak{R}_e : \Delta \to \mathbb{R}_+$ *is continuous with respect to the weak topology on* Δ.

In fact, we also prove the continuity of the spectrum with respect to the Hausdorff distance on the set of compact subsets of $\mathbb{C}$. We shall write $\mathfrak{R}_e[\mathbf{k}]$ to stress the dependence of the function $\mathfrak{R}_e$ in the kernel $\mathbf{k}$. In Proposition III.4.3, we prove the stability of $\mathfrak{R}_e$, by giving natural sufficient conditions on a sequence of kernels $(\mathbf{k}_n, n \in \mathbb{N})$ converging to $\mathbf{k}$ which imply that $\mathfrak{R}_e[\mathbf{k}_n]$ converges uniformly towards $\mathfrak{R}_e[\mathbf{k}]$. This result has both a theoretical and a practical interest: the next-generation operator is unknown in practice, and has to be estimated from data. Thanks to this result, the value of $\mathfrak{R}_e$ computed from the estimated operator should converge to the true value.

On the maximal endemic equilibrium in the SIS model

We consider the *SIS model* from Chapter II. This model is characterized by a probability space $(\Omega, \mathscr{F}, \mu)$, the transmission kernel $k : \Omega\times\Omega \to \mathbb{R}_+$ and the recovery rate $\gamma : \Omega \to \mathbb{R}_+^*$. We suppose in the following that the technical Assumption III.2, formulated on page 76, holds, so that the SIS dynamical evolution is well defined.

This evolution is encoded as $u = (u_t, t \in \mathbb{R}_+)$, where $u_t \in \Delta$ for all t and $u_t(x)$ represents the probability of an individual with feature $x \in \Omega$ to be infected at time $t \geq 0$, and follows the equation:

$$\partial_t u_t = F(u_t) \quad \text{for } t \in \mathbb{R}_+, \quad \text{where} \quad F(g) = (1 - g)\mathscr{T}_k(g) - \gamma g \quad \text{for } g \in \Delta, \qquad \text{(III.4)}$$

with an initial condition $u_0 \in \Delta$ and with $\mathscr{T}_k$ the integral operator corresponding to the kernel k acting on the set of bounded measurable functions, see (III.16). It is proved in Chapter II that such a solution u exists and is unique under Assumption III.2. An *equilibrium* of (III.4) is a function $g \in \Delta$ such that $F(g) = 0$. According to Chapter II, there exists a maximal equilibrium $\mathfrak{g}$, *i.e.*, an equilibrium such that all other equilibria $h \in \Delta$ are dominated by $\mathfrak{g}$: $h \leq \mathfrak{g}$. Furthermore, we have $\mathfrak{R}_0 \leq 1$ if and only if $\mathfrak{g} = 0$. In the connected case (for example if $k > 0$), then 0 and $\mathfrak{g}$ are the only equilibria. Besides, $\mathfrak{g}$ is the long-time distribution of infected individuals in the population: $\lim_{t\to+\infty} u_t = \mathfrak{g}$ as soon as the initial condition is non-zero; see Theorem II.1.5.

As hinted in [80, Section 4.5] for vaccination control of gonorrhea, it is interesting to consider vaccinating people with feature x with probability $\mathfrak{g}(x)$; this corresponds to the strategy based on the maximal equilibrium:

$$\eta^{\text{equi}} = 1 - \mathfrak{g}.$$

The following result entails that this strategy is critical and thus achieves the herd immunity threshold. Recall that Assumption III.2, formulated page 76, provides technical conditions on the parameters k and γ of the SIS model. It is fulfilled in particular if k/γ and γ are bounded. The effective reproduction number of the SIS model is the function $\mathfrak{R}_e$ defined in (III.3) with the kernel $\mathbf{k} = k/\gamma$.

Theorem III.1.2 (The maximal equilibrium yields a critical vaccination)**.** *Suppose Assumption III.2 holds. If $\mathfrak{R}_0 \geq 1$, then the vaccination strategy η^{equi} is critical, that is, $\mathfrak{R}_e(\eta^{\mathrm{equi}}) = 1$.*

This result will be proved below as a part of Proposition III.8.2. Let us finally describe informally another consequence of this Proposition. We were able to prove in Theorem II.1.5 that, in the connected case, if $\mathfrak{R}_0 > 1$, the disease-free equilibrium $u = 0$ is unstable. Proposition III.8.2 gives spectral information on the formal linearization of the dynamics (III.4) near any equilibrium h; in particular if $h \neq \mathfrak{g}$ then h is linearly unstable.

Regularity of the total proportion of infected population function $\mathfrak{I}$

According to Section II.5.3, the SIS equation with vaccination strategy η is given by (III.4), where F is replaced by F_η defined by:

$$F_\eta(g) = (1 - g)T_{k\eta}(g) - \gamma g.$$

and u_t now describes the proportion of infected *among the non-vaccinated population.* We denote by $\mathfrak{g}_\eta$ the corresponding maximal equilibrium (thus considering $\eta = \mathbb{1}$ gives $\mathfrak{g} = \mathfrak{g}_{\mathbb{1}}$), so that $F_\eta(\mathfrak{g}_\eta) = 0$. Since the probability for an individual x to be infected in the stationary regime is $\mathfrak{g}_\eta(x)\,\eta(x)$, the *fraction of infected individuals at equilibrium*, $\mathfrak{I}(\eta)$, is thus given by:

$$\mathfrak{I}(\eta) = \int_\Omega \mathfrak{g}_\eta \, \eta \, \mathrm{d}\mu = \int_\Omega \mathfrak{g}_\eta(x)\,\eta(x)\,\mu(\mathrm{d}x). \tag{III.5}$$

As mentioned above, for a SIR model, distributing vaccine so as to minimize the attack rate is at least as natural as trying to minimize the reproduction number, and this problem has been studied for example in [46, 47]. In the SIS model the quantity $\mathfrak{I}$ appears as a natural analogue of the attack rate, and is therefore a natural optimization objective.

We obtain results on $\mathfrak{I}$ that are very similar to the ones on $\mathfrak{R}_e$. Recall that Assumption III.2 on page 76 ensures that the infinite-dimensional SIS model, given by equation (III.4), is well defined. The next theorem corresponds to Theorem III.4.6.

Theorem III.1.3 (Continuity of the equilibrium infection size)**.** *Under Assumption III.2, the function $\mathfrak{I} : \Delta \to \mathbb{R}_+$ is continuous with respect to the weak topology on Δ.*

In Proposition III.4.7, we prove the stability of $\mathfrak{I}$, by giving natural sufficient conditions on a sequence of kernels and functions $((k_n, \gamma_n), n \in \mathbb{N})$ converging to (k, γ) which imply that $\mathfrak{I}[k_n, \gamma_n]$ converges uniformly towards $\mathfrak{I}[k, \gamma]$. We also prove that the loss functions $\mathrm{L} = \mathfrak{R}_e$ and $\mathrm{L} = \mathfrak{I}$ are both non-decreasing ($\eta \leq \eta'$ implies $\mathrm{L}(\eta) \leq \mathrm{L}(\eta')$), and sub-homogeneous ($\mathrm{L}(\lambda\eta) \leq \lambda \mathrm{L}(\eta)$ for all $\lambda \in [0, 1]$); see Propositions III.4.1 and III.4.5.

Optimizing the protection of the population

Consider a cost function $C : \Delta \to [0, 1]$ which measures the cost for the society of a vaccination strategy (production and diffusion). Since the vaccination strategy η represents the non-vaccinated population, the cost function C should be decreasing (roughly speaking $\eta < \eta'$ implies $C(\eta) > C(\eta')$; see Definition III.5.1). We shall also assume that C is continuous with respect to the weak topology on Δ, and that doing nothing costs nothing, that is, $C(\mathbb{1}) = 0$. A simple and natural choice is the uniform cost C_{uni} given by the overall proportion of vaccinated individuals:

$$C_{\mathrm{uni}}(\eta) = \int_\Omega (1 - \eta)\, \mathrm{d}\mu = 1 - \int_\Omega \eta \, \mathrm{d}\mu.$$

See Remark III.5.2 for comments on other examples of cost functions.

Our problem may now be seen as a bi-objective minimization problem: we wish to minimize both the loss $\mathrm{L}(\eta)$ and the cost $C(\eta)$, subject to $\eta \in \Delta$, with the loss function L being either $\mathfrak{R}_e$ or $\mathfrak{I}$. Following classical terminology for multi-objective optimisation problems [117], we call a strategy $\eta_\star$ *Pareto optimal* if no other strategy is strictly better:

$$C(\eta) < C(\eta_\star) \implies \mathrm{L}(\eta) > \mathrm{L}(\eta_\star) \quad \text{and} \quad \mathrm{L}(\eta) < \mathrm{L}(\eta_\star) \implies C(\eta) > C(\eta_\star).$$

The set of Pareto optimal strategies will be denoted by $\mathcal{P}_\mathrm{L}$, and we define the *Pareto frontier* as the set of Pareto optimal outcomes:

$$\mathcal{F}_\mathrm{L} = \{(C(\eta_\star), \mathrm{L}(\eta_\star)) \,:\, \eta_\star \in \mathcal{P}_\mathrm{L}\}.$$

Notice that, with this definition, the Pareto frontier is empty when there is no Pareto optimal strategy.

For any strategy η, the cost and loss of η vary between the following bounds:

$$\begin{aligned} 0 = C(\mathbb{1}) \leq C(\eta) \leq C(\mathbb{0}) = c_{\max} &= \text{cost of vaccinating the whole population,} \\ 0 = \mathrm{L}(\mathbb{0}) \leq \mathrm{L}(\eta) \leq \mathrm{L}(\mathbb{1}) = \ell_{\max} &= \text{loss incurred in the absence of vaccination.} \end{aligned}$$

Let $\mathrm{L}_\star$ be the *optimal loss* function and $C_{\star,\mathrm{L}}$ the *optimal* cost function defined by:

$$\begin{aligned} \mathrm{L}_\star(c) &= \inf\{\mathrm{L}(\eta) \,:\, \eta \in \Delta,\, C(\eta) \leq c\} \quad \text{for } c \in [0, c_{\max}], \\ C_{\star,\mathrm{L}}(\ell) &= \inf\{C(\eta) \,:\, \eta \in \Delta,\, \mathrm{L}(\eta) \leq \ell\} \quad \text{for } \ell \in [0, \ell_{\max}]. \end{aligned}$$

When there is no confusion on the loss function, we simply write $C_\star$ for $C_{\star,\mathrm{L}}$. Proposition III.5.5 (in a more general framework in particular for the cost function) and Lemma III.5.6 state that the Pareto frontier is non empty and has a continuous parametrization for the cost $C = C_{\mathrm{uni}}$ and the loss $\mathrm{L} = \mathfrak{R}_e$ or $\mathrm{L} = \mathfrak{I}$. More formally, we prove the following result, illustrated in Figure III.1(b) below.

Theorem III.1.4 (Properties of the Pareto frontier)**.** *For the kernel model with loss function* $\mathrm{L} = \mathfrak{R}_e$ *or the SIS model with* $\mathrm{L} \in \{\mathfrak{R}_e, \mathfrak{I}\}$*, and the uniform cost function* $C = C_{\mathrm{uni}}$*, the function* $C_{\star,\mathrm{L}}$ *is continuous and decreasing on* $[0, \ell_{\max}]$*, the function* $\mathrm{L}_\star$ *is continuous on* $[0, c_{\max}]$ *decreasing on* $[0, C_{\star,\mathrm{L}}(0)]$ *and zero on* $[C_{\star,\mathrm{L}}(0), c_{\max}]$*; furthermore the Pareto frontier is connected and:*

$$\mathcal{F}_\mathrm{L} = \{(c, \mathrm{L}_\star(c)) \,:\, c \in [0, C_{\star,\mathrm{L}}(0)]\} = \{(C_{\star,\mathrm{L}}(\ell), \ell) \,:\, \ell \in [0, \ell_{\max}]\}.$$

We also establish that $\mathcal{P}_\mathrm{L}$ is compact in Δ for the weak topology in Corollary III.5.7; that the set of outcomes or feasible region $\mathbf{F} = \{(C(\eta), \mathrm{L}(\eta)),\, \eta \in \Delta\}$ has no holes in Proposition III.6.1; and that the Pareto frontier is convex if C and L are convex in Proposition III.6.6. We study in Proposition III.6.2 the stability of the Pareto frontier and the set of Pareto optima when the parameters vary.

In a sense the Pareto optimal strategies are intuitively the "best" strategies. Similarly, we also study the "worst" strategies, which we call anti-Pareto optimal strategies, and describe the corresponding anti-Pareto frontier. Understanding the "worst strategies" also helps to avoid pitfalls when one has to consider sub-optimal strategies: for example, we prove in Chapter IV that disconnecting strategies are not the "worst" strategies, and we provide in Chapter V Section 4 an elementary example where the same strategies can be "best" or "worst" according to model parameters values. Surprisingly, proving properties of the anti-Pareto frontier sometimes necessitates stronger assumptions than in the Pareto case. For example, under some *irreducibility* assumption on the kernel developed in Section III.5.4, we establish the connectedness of the anti-Pareto frontier when the loss is given by $\mathrm{L} = \mathfrak{R}_e$; and a slightly different behavior when the loss is given by $\mathrm{L} = \mathfrak{I}$.

Remark III.1.5 (Eradication strategies do not depend on the loss). In Chapter II, we proved that, for all $\eta \in \Delta$, the equilibrium infection size $\mathfrak{I}(\eta)$ is non zero if and only if $\mathfrak{R}_e(\eta) > 1$. Consider the uniform cost $C = C_{\text{uni}}$. First, this implies that $\mathcal{P}_{\mathfrak{I}}$ is a subset of $\{\eta \in \Delta \ : \ \mathfrak{R}_e(\eta) \geq 1\}$. econdly, a vaccination strategy $\eta \in \Delta$ is Pareto optimal for the objectives $(\mathfrak{R}_e, C)$ and satisfies $\mathfrak{R}_e(\eta) = 1$ if and only if η is Pareto optimal for the objectives $(\mathfrak{I}, C)$ and satisfies $\mathfrak{I}(\eta) = 0$:

$$\eta \in \mathcal{P}_{\mathfrak{R}_e} \text{ and } \mathfrak{R}_e(\eta) = 1 \quad \iff \quad \eta \in \mathcal{P}_{\mathfrak{I}} \text{ and } \mathfrak{I}(\eta) = 0. \tag{III.6}$$

Remark III.1.6 (Minimal cost of eradication). Assume $\mathfrak{R}_0 > 1$ and consider the uniform cost $C = C_{\text{uni}}$. The equivalence (III.6) implies directly that:

$$C_{\star,\mathfrak{R}_e}(1) = C_{\star,\mathfrak{I}}(0).$$

thus, this latter quantity can be seen as the minimal cost (or minimum percentage of people that have to be vaccinated) required to eradicate the infection. Recall that the vaccination strategies $\eta^{\text{uni}}_{\text{crit}} = \mathfrak{R}_0^{-1}\mathbb{1}$ and $\eta^{\text{equi}} = 1 - \mathfrak{g}$ are critical (as $\mathfrak{R}_e(\eta^{\text{uni}}_{\text{crit}}) = \mathfrak{R}_e(\eta^{\text{equi}}) = 1$). Since $C(\eta^{\text{uni}}_{\text{crit}}) = 1 - 1/\mathfrak{R}_0$ and $C(\eta^{\text{equi}}) = \int_\Omega \mathfrak{g} \, \mathrm{d}\mu = \mathfrak{I}(\mathbb{1})$, we obtain the following upper bounds of the minimal cost required to eradicate the infection:

$$C_{\star,\mathfrak{R}_e}(1) = C_{\star,\mathfrak{I}}(0) \leq \min\left(1 - \frac{1}{\mathfrak{R}_0}\, , \int_\Omega \mathfrak{g} \, \mathrm{d}\mu\right).$$

Equivalence of models

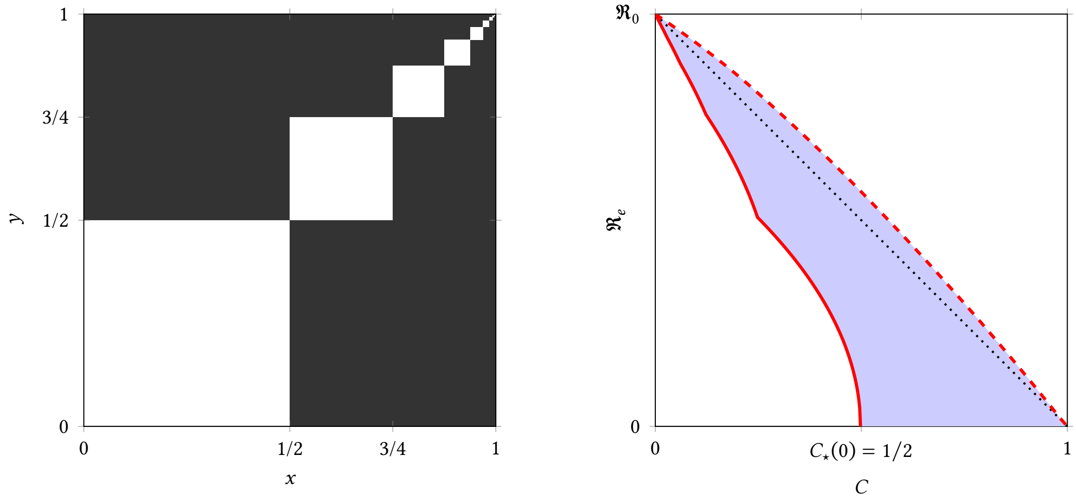
Our last results address a natural question stemming from our choice of a very general framework to modelize the infection. Since our models are infinite dimensional and depend on the choices of the probability space $(\Omega, \mathscr{F}, \mu)$, the kernel $\mathbf{k}$ (for the kernel model) or the kernel k and recovery rate γ (for the SIS model), there are different, equivalent ways to model the same situation. We study in Section III.7 a way to ensure that, even if the parameters are different, we end up with the same Pareto frontiers. This situation is similar to random variables having the same law in probability theory, or to equivalent graphons in graphon theory. In particular it allows us to treat the same meta-population model in either a discrete or a continuous setting, see Example III.1.7 and Figure III.5 for an illustration.

An illustrative example: the multipartite graphon

Let us illustrate some of our results on an example, which will be discussed in details in Chapter V.

Example III.1.7 (Multipartite graphon). Graphs whose vertices can be colored with ℓ colors, so that the endpoints of every edge are colored differently, are known as ℓ-partite graphs. In a biological setting, this corresponds to a population of ℓ groups, such that individuals in a group only contaminate individuals of the other groups. Let us generalize and assume there is an infinity of groups, $\ell = \infty$, of respective size $(2^{-n},\ n \in \mathbb{N}^*)$, and that the next generation kernel $\mathbf{k}$ is equal to the constant $\kappa > 0$ between individuals of different groups and equal to 0 between individuals of the same group (so there is no intra-group contamination). Using the equivalence of models from Section III.7, we can represent this model by using a continuous state space $\Omega = [0, 1]$, endowed with μ the Lebesgue measure on Ω, the group n being represented by the interval $I_n = [1 - 2^{-n+1}, 1 - 2^{-n})$ for $n \in \mathbb{N}^*$. The kernel $\mathbf{k}$ is then given by $\mathbf{k} = \kappa(1 - \sum_{n\in\mathbb{N}^*} \mathbb{1}_{I_n \times I_n})$; it is represented in Figure III.1(A).

Consider the loss $\mathrm{L} = \mathfrak{R}_e$ and the cost $C = C_{\text{uni}}$ giving the overall proportion of vaccinated individuals. Building on the results of [54, 149], we prove in Chapter V that the vaccination strategies $\mathbb{1}_{[0,1-c]}$, with cost $C(\mathbb{1}_{[0,1-c]}) = c \in [0, 1/2]$, are Pareto optimal. Remembering that the natural definition of the degree in a continuous graph is given by $\deg(x) = \int_\Omega \mathbf{k}(x, y)\, \mu(\mathrm{d}y)$, we note that the vaccination strategy $\mathbb{1}_{[0,1-c]}$ corresponds to vaccinating individuals with feature $x \in (1-c, 1]$, that is, the individuals with the highest degree. In Figure III.1(B), the corresponding Pareto frontier (*i.e.*, the outcome of the "best" vaccination strategies) is drawn as the solid red line; the



(A) Grayplot of the kernel $\mathbf{k}$, with $\Omega = [0, 1]$ and μ the Lebesgue measure ($\mathbf{k}$ is equal to the constant $\kappa > 0$ on the black zone and to 0 on the white zone).

(B) The Pareto frontier in solid red line compared to the cost and loss of the uniform vaccinations in dotted line and the worst vaccination strategy in red dashed line.

Figure III.1: Example of optimization with $\mathrm{L} = \mathfrak{R}_e$.

blue-colored zone corresponds to the feasible region that is, all the possible values of $(C(\eta), \mathfrak{R}_e(\eta))$, where η ranges over Δ; the dotted line corresponds to the outcome of the uniform vaccination strategy $\eta = c\mathbb{1}$, that is $(C(\eta), \mathfrak{R}_e(\eta)) = (c, (1 - c)\mathfrak{R}_0)$ where c ranges over $[0, 1]$; and the red dashed curve corresponds to the anti-Pareto frontier (*i.e.*, the outcome of the "worst" vaccination strategies), which for this model correspond to the uniform vaccination of the nodes with the updated lower degree; see Chapter V. Notice that the path $(\mathbb{1}_{[0,1-c]}, c \in [0, 1/2])$ is an increasing continuous (for the topology of the simple convergence and thus the $L^1(\mu)$ topology) path of Pareto optima which gives a complete parametrization of the Pareto frontier. The latter has been computed numerically using the power iteration method. In particular, we obtained the following value: $\mathfrak{R}_0 \simeq 0.697\kappa$.

III.1.5 On the next chapters and the companion papers

Let us discuss briefly the results that can be found in Chapter IV and V and in the companion papers [37, 40].

In Chapter IV, motivated by a conjecture formulated by Hill and Longini in finite dimension (see [81, Conjecture 8.1]), we investigate the convexity and concavity of the effective reproduction function $\mathfrak{R}_e$. We also prove that a disconnecting strategy is better than the worst, *i.e.*, is not anti-Pareto optimal.

In [40], under monotonicity properties of the kernel, satisfied for example by the configuration model, we prove that vaccinating the individuals with the highest (resp. lowest) number of contacts is Pareto (resp. anti-Pareto) optimal. In this case the greedy algorithm, which performs infinitesimal locally optimal steps, is optimal as it browses continuously the set of Pareto (resp. anti-Pareto) optimal strategies, providing an increasing parametrization of the Pareto (resp. anti-Pareto) frontier. In this setting, we provide some examples of SIS models where the set of Pareto optimal strategies coincide for the losses $\mathfrak{R}_e$ and $\mathfrak{I}$:

$$\mathcal{P}_{\mathfrak{I}} = \mathcal{P}_{\mathfrak{R}_e} \cap \{\eta \in \Delta \ : \ \mathfrak{R}_e(\eta) \geq 1\}. \tag{III.7}$$

In Chapter V, we study the Pareto and anti-Pareto frontier for a number of examples. In some of them, like the multipartite kernel of Example III.1.7, the optimal vaccinations target the individuals with the highest number of contacts. When the individuals have the same number of contacts, this heuristic cannot be used, and the situations may be extremely varied: for example, we give models for which the uniform vaccination is Pareto optimal, or anti-Pareto optimal, or not optimal for either problem. We also provide an example where the set $\mathcal{P}_{\mathfrak{R}_e}$ has a countable number of connected components (and is thus not connected). This implies in particular that the greedy algorithm is not optimal in this case.

In [37], we give a comprehensive treatment of the two groups model, $\Omega = \{1, 2\}$, for $\mathrm{L} = \mathfrak{R}_e$, and some partial results for $\mathrm{L} = \mathfrak{I}$. Despite its apparent simplicity, the derivation of formulae for the Pareto optimal strategies is non trivial, see also [130]. In addition, this model is rich enough to give examples of various interesting behaviours:

- *On the critical strategies $\eta^{\text{uni}}_{\text{crit}}$ and η^{equi}.* Depending on the parameters, the strategies $\eta^{\text{uni}}_{\text{crit}}$ and/or η^{equi} may or may not be Pareto optimal, and the cost $C(\eta^{\text{uni}}_{\text{crit}})$ may be larger than, smaller than or equal to $C(\eta^{\text{equi}})$.

- *Vaccinating the most connected individuals.* The intuitive idea of vaccinating the individuals with the highest number of contacts may or may not provide the optimal strategies, depending on the parameters.

- *Dependence on the choice of the loss function.* For examples where $\mathfrak{R}_0 > 1$, the optimal strategies for the losses $\mathfrak{I}$ and $\mathfrak{R}_e$ may coincide, so that (III.7) holds, or not at all, so that $\mathcal{P}_{\mathfrak{I}} \cap \mathcal{P}_{\mathfrak{R}_e} \cap \{\eta \in \Delta \,:\, 1 < \mathfrak{R}_e(\eta) < \mathfrak{R}_0\} = \emptyset$, depending on the parameters.

III.1.6 Structure of the chapter

Section III.2 is dedicated to the presentation of the vaccination model and the various assumptions on the parameters. We also define properly the so-called loss functions $\mathfrak{R}_e$ and $\mathfrak{I}$. After recalling a few topological facts in Section III.3, we study the regularity properties of $\mathfrak{R}_e$ and $\mathfrak{I}$ in Section III.4. We present the multi-objective optimization problem in Section III.5 under general conditions on the loss function L and cost function C and prove the first results on the Pareto frontier. This is completed in Section III.6 with miscellaneous properties of the Pareto frontier. In Section III.7, we discuss the equivalent representation of models with different parameters. The proofs of a few technical results are gathered in Section III.8.

III.2 Setting and notation

III.2.1 Spaces, operators, spectra

All metric spaces (S, d) are endowed with their Borel σ-field denoted by $\mathscr{B}(S)$. The set $\mathscr{K}$ of compact subsets of $\mathbb{C}$ endowed with the Hausdorff distance d_{H} is a metric space, and the function rad from $\mathscr{K}$ to $\mathbb{R}_+$ defined by $\mathrm{rad}(K) = \max\{|\lambda|\,,\, \lambda \in K\}$ is Lipschitz continuous from $(\mathscr{K}, d_{\mathrm{H}})$ to $\mathbb{R}$ endowed with its usual Euclidean distance.

Let $(\Omega, \mathscr{F}, \mu)$ be a probability space. We denote by $\mathscr{L}^\infty$, the Banach spaces of bounded real-valued measurable functions defined on Ω equipped with the sup-norm, $\mathscr{L}^\infty_+$ the subset of $\mathscr{L}^\infty$ of non-negative function, and $\Delta = \{f \in \mathscr{L}^\infty \,:\, 0 \leq f \leq 1\}$ the subset of non-negative functions bounded by 1. For f and g real-valued functions defined on Ω, we may write $\langle f, g\rangle$ or $\int_\Omega f g \,\mathrm{d}\mu$ for $\int_\Omega f(x) g(x)\, \mu(\mathrm{d}x)$ whenever the latter is meaningful. For $p \in [1, +\infty]$, we denote by $L^p = L^p(\mu) = L^p(\Omega, \mu)$ the space of real-valued measurable functions g defined Ω such that $\| g \|_p = \left(\int |g|^p \,\mathrm{d}\mu\right)^{1/p}$ (with the convention that $\| g \|_\infty$ is the μ-essential supremum of $|g|$) is finite, where functions which agree μ-almost surely are identified. We denote by L^p_+ the subset of L^p of non-negative functions.

Let $(E, \|\cdot\|)$ be a Banach space. We denote by $\|\cdot\|_E$ the operator norm on $\mathcal{L}(E)$ the Banach algebra of bounded operators. The spectrum $\mathrm{Spec}(T)$ of $T \in \mathcal{L}(E)$ is the set of $\lambda \in \mathbb{C}$ such that $T - \lambda \mathrm{Id}$ does not have a bounded inverse operator, where Id is the identity operator on E. Recall that $\mathrm{Spec}(T)$ is a compact subset of $\mathbb{C}$, and that the spectral radius of T is given by:

$$\rho(T) = \mathrm{rad}(\mathrm{Spec}(T)) = \lim_{n\to\infty} \| T^n \|_E^{1/n}. \tag{III.8}$$

The element $\lambda \in \mathrm{Spec}(T)$ is an eigenvalue if there exists $x \in E$ such that $Tx = \lambda x$ and $x \neq 0$.

If E is also a functional space, for $g \in E$, we denote by M_g the (possibly unbounded) multiplication operator defined by $M_g(h) = gh$ for all $h \in E$.

III.2.2 Kernel operators

We define a *kernel* (resp. *signed kernel*) on Ω as a $\mathbb{R}_+$-valued (resp. $\mathbb{R}$-valued) measurable function defined on $(\Omega^2, \mathscr{F}^{\otimes 2})$. For f, g two non-negative measurable functions defined on Ω and $\mathbf{k}$ a kernel on Ω, we denote by $f\mathbf{k}g$ the kernel defined by:

$$f\mathbf{k}g \,:\, (x,y) \mapsto f(x)\,\mathbf{k}(x,y)g(y). \tag{III.9}$$

When γ is a positive measurable function defined on Ω, we write $\mathbf{k}/\gamma$ for $\mathbf{k}\gamma^{-1}$, and remark that it may differ from $\gamma^{-1}\mathbf{k}$.

For $p \in (1, +\infty)$, we define the double norm of a signed kernel $\mathbf{k}$ by:

$$\|\mathbf{k}\|_{p,q} = \left(\int_\Omega \left(\int_\Omega |\mathbf{k}(x,y)|^q \,\mu(\mathrm{d}y)\right)^{p/q} \mu(\mathrm{d}x)\right)^{1/p} \quad \text{with } q \text{ given by } \quad \frac{1}{p} + \frac{1}{q} = 1. \tag{III.10}$$

Assumption III.1 (On the kernel model $[(\Omega, \mathscr{F}, \mu), \mathbf{k}]$)**.** *The kernel* $\mathbf{k}$ *defined on the probability space* $(\Omega, \mathscr{F}, \mu)$ *has a* finite double-norm, *that is,* $\|\mathbf{k}\|_{p,q} < +\infty$ *for some* $p \in (1, +\infty)$.

To a kernel k such that $\|\mathbf{k}\|_{p,q} < +\infty$, we associate the integral operator $T_\mathbf{k}$ on L^p defined by:

$$T_\mathbf{k}(g)(x) = \int_\Omega \mathbf{k}(x,y)g(y)\,\mu(\mathrm{d}y) \quad \text{for } g \in L^p \text{ and } x \in \Omega. \tag{III.11}$$

This operator is positive (in the sense that $T_\mathbf{k}(L^p_+) \subset L^p_+$), and compact (see [68, p. 293]). It is well known and easy to check that:

$$\|T_\mathbf{k}\|_{L^p} \leq \|\mathbf{k}\|_{p,q}. \tag{III.12}$$

For $\eta \in \Delta$, the kernel $\mathbf{k}\eta$ has also a finite double norm on L^p and the operator M_η is bounded, so that the operator $T_{\mathbf{k}\eta} = T_\mathbf{k} M_\eta$ is compact. We can define the *effective spectrum* function $\mathrm{Spec}[\mathbf{k}]$ from Δ to $\mathscr{K}$ by:

$$\mathrm{Spec}[\mathbf{k}](\eta) = \mathrm{Spec}(T_{\mathbf{k}\eta}), \tag{III.13}$$

the *effective reproduction number* function $\mathfrak{R}_e[\mathbf{k}] = \mathrm{rad} \circ \mathrm{Spec}[\mathbf{k}]$ from Δ to $\mathbb{R}_+$ by:

$$\mathfrak{R}_e[\mathbf{k}](\eta) = \mathrm{rad}(\mathrm{Spec}(T_{\mathbf{k}\eta})) = \rho(T_{\mathbf{k}\eta}), \tag{III.14}$$

and the corresponding *reproduction number*:

$$\mathfrak{R}_0[\mathbf{k}] = \mathfrak{R}_e[\mathbf{k}](\mathbb{1}) = \rho(T_\mathbf{k}). \tag{III.15}$$

When there is no ambiguity, we simply write $\mathfrak{R}_e$ for the function $\mathfrak{R}_e[\mathbf{k}]$, and $\mathfrak{R}_0$ for the number $\mathfrak{R}_0[\mathbf{k}]$. We say a vaccination strategy $\eta \in \Delta$ is *critical* if $\mathfrak{R}_e(\eta) = 1$.

Following the framework of Chapter II, for $q \in (1, +\infty)$, we also consider the following norm for the kernel $\mathbf{k}$:

$$\|\mathbf{k}\|_{\infty,q} = \sup_{x\in\Omega} \left(\int_\Omega \mathbf{k}(x,y)^q \,\mu(\mathrm{d}y)\right)^{1/q}.$$

If this norm $\|\mathbf{k}\|_{\infty,q}$ is finite, then for p such that $1/p + 1/q = 1$, the norm $\|\mathbf{k}\|_{p,q}$ is also finite. When $\|\mathbf{k}\|_{\infty,q} < +\infty$, the corresponding positive bounded linear integral operator $\mathscr{T}_\mathbf{k}$ on $\mathscr{L}^\infty$ is similarly defined by:

$$\mathscr{T}_\mathbf{k}(g)(x) = \int_\Omega \mathbf{k}(x,y)g(y)\,\mu(\mathrm{d}y) \quad \text{for } g \in \mathscr{L}^\infty \text{ and } x \in \Omega. \tag{III.16}$$

Notice that the integral operators $\mathscr{T}_\mathbf{k}$ and $T_\mathbf{k}$ corresponds respectively to the operators $T_\mathbf{k}$ and $\hat{T}_\mathbf{k}$ in Chapter II. According to Lemma II.3.7, the operator $\mathscr{T}_\mathbf{k}^2$ on $\mathscr{L}^\infty$ is compact and $\mathscr{T}_\mathbf{k}$ has the same spectral radius as $T_\mathbf{k}$:

$$\rho(\mathscr{T}_\mathbf{k}) = \rho(T_\mathbf{k}). \tag{III.17}$$

III.2.3 Dynamics for the SIS model and equilibria

In accordance with Chapter II, we consider the following assumption. Recall that $k/\gamma = k\gamma^{-1}$.

Assumption III.2 (On the SIS model $[(\Omega, \mathscr{F}, \mu), k, \gamma]$). *The recovery rate function γ, defined on a probability space $(\Omega, \mathscr{F}, \mu)$, is bounded and non negative.*

The transmission rate kernel k on Ω^2 is such that $\| k/\gamma \|_{\infty,q} < +\infty$ for some $q \in (1, +\infty)$.

If k and γ satisfy Assumption III.2, then $\mathbf{k} = k/\gamma$ clearly satisfies Assumption III.1. Under Assumption III.2, we also consider the bounded operators $\mathscr{T}_{k/\gamma}$ on $\mathscr{L}^\infty$, as well as $T_{k/\gamma}$ on L^p, which are the so called *next-generation operator.* The SIS dynamics considered in Chapter II (under Assumption III.2) follows the vector field F defined on $\mathscr{L}^\infty$ by:

$$F(g) = (1 - g)\mathscr{T}_k(g) - \gamma g. \tag{III.18}$$

More precisely, we consider $u = (u_t, t \in \mathbb{R})$, where $u_t \in \Delta$ for all $t \in \mathbb{R}_+$ such that:

$$\partial_t u_t = F(u_t) \quad \text{for } t \in \mathbb{R}_+, \tag{III.19}$$

with initial condition $u_0 \in \Delta$. The value $u_t(x)$ models the probability that an individual of feature x is infected at time t; it is proved in Chapter II that such a solution u exists and is unique.

An *equilibrium* of (III.19) is a function $g \in \Delta$ such that $F(g) = 0$. According to Chapter II, there exists a maximal equilibrium $\mathfrak{g}$, *i.e.*, an equilibrium such that all other equilibria $h \in \Delta$ are dominated by $\mathfrak{g}$: $h \leq \mathfrak{g}$. The *reproduction number* $\mathfrak{R}_0$ associated to the SIS model given by (III.19) is the spectral radius of the next-generation operator, so that using the definition of the effective reproduction number (III.14), (III.15) and (III.17), this amounts to:

$$\mathfrak{R}_0 = \rho(\mathscr{T}_{k/\gamma}) = \mathfrak{R}_0[k/\gamma] = \mathfrak{R}_e[k/\gamma](\mathbb{1}). \tag{III.20}$$

If $\mathfrak{R}_0 \leq 1$ (sub-critical and critical case), then u_t converges pointwise to 0 when $t \to \infty$. In particular, the maximal equilibrium $\mathfrak{g}$ is equal to 0 everywhere. If $\mathfrak{R}_0 > 1$ (super-critical case), then 0 is still an equilibrium but different from the maximal equilibrium $\mathfrak{g}$, as $\int_\Omega \mathfrak{g} \, \mathrm{d}\mu > 0$.

III.2.4 Vaccination strategies

A *vaccination strategy* η of a vaccine with perfect efficiency is an element of Δ, where $\eta(x)$ represents the proportion of ***non-vaccinated*** individuals with feature x. Notice that $\eta \, \mathrm{d}\mu$ corresponds in a sense to the effective population.

Recall the definition of the kernel $f\mathbf{k}g$ from (III.9). For $\eta \in \Delta$, the kernels $k\eta/\gamma$ and $k\eta$ have finite norm $\| \cdot \|_{\infty,q}$ under Assumption III.2, so we can consider the bounded positive operators $\mathscr{T}_{k\eta/\gamma}$ and $\mathscr{T}_{k\eta}$ on $\mathscr{L}^\infty$. According to Section 5.3.] of Chapter II, the SIS equation with vaccination strategy η is given by (III.19), where F is replaced by F_η defined by:

$$F_\eta(g) = (1 - g)\mathscr{T}_{k\eta}(g) - \gamma g. \tag{III.21}$$

We denote by $u^\eta = (u_t^\eta, t \geq 0)$ the corresponding solution with initial condition $u_0^\eta \in \Delta$. We recall that $u_t^\eta(x)$ represents the probability for an non-vaccinated individual of feature x to be infected at time t. Since the effective reproduction number is the spectral radius of $\mathscr{T}_{k\eta/\gamma}$, we recover (III.14) as $\rho(\mathscr{T}_{k\eta/\gamma}) = \rho(T_{k\eta/\gamma}) = \mathfrak{R}_e[k/\gamma](\eta)$ with $\mathbf{k} = k/\gamma$. We denote by $\mathfrak{g}_\eta$ the corresponding maximal equilibrium (so that $\mathfrak{g} = \mathfrak{g}_{\mathbb{1}}$). In particular, we have:

$$F_\eta(\mathfrak{g}_\eta) = 0. \tag{III.22}$$

We will denote by $\mathfrak{I}$ the *fraction of infected individuals at equilibrium.* Since the probability for an individual with feature x to be infected in the stationary regime is $\mathfrak{g}_\eta(x)\,\eta(x)$, this fraction is given by the following formula:

$$\mathfrak{I}(\eta) = \int_\Omega \mathfrak{g}_\eta \, \eta \, \mathrm{d}\mu = \int_\Omega \mathfrak{g}_\eta(x)\, \eta(x)\, \mu(\mathrm{d}x). \tag{III.23}$$

We deduce from (III.21) and (III.22) that $\mathfrak{g}_\eta \eta = 0$ μ-almost surely is equivalent to $\mathfrak{g}_\eta = 0$. Applying the results of Chapter II to the kernel $k\eta$, we deduce that:

$$\mathfrak{I}(\eta) > 0 \iff \mathfrak{R}_e[k/\gamma](\eta) > 1. \tag{III.24}$$

III.3 Preliminary topological results

III.3.1 On the weak topology

We first recall briefly some properties we shall use frequently. We can see Δ as a subset of L^1, and consider the corresponding *weak topology*: a sequence $(g_n,\ n \in \mathbb{N})$ of elements of Δ converges weakly to g if for all $h \in L^\infty$ we have:

$$\lim_{n\to\infty} \int_\Omega h g_n \, \mathrm{d}\mu = \int_\Omega hg \, \mathrm{d}\mu. \tag{III.25}$$

Notice that (III.25) can easily be extended to any function $h \in L^q$ for any $q \in (1, +\infty)$; so that the weak-topology on Δ, seen as a subset of L^p with $1/p + 1/q = 1$, can be seen as the trace on Δ of the weak topology on L^p. The main advantage of this topology is the following compactness result.

Lemma III.3.1 (Topological properties of Δ)**.** *We have that:*

(i) The set Δ endowed with the weak topology is compact and sequentially compact.

(ii) A function from Δ (endowed with the weak topology) to a metric space (endowed with its metric topology) is continuous if and only if it is sequentially continuous.

Proof. Let $p \in (1, +\infty)$, and consider the weak topology on Δ as the trace on Δ of the weak topology on L^p. We first prove (i). Since L^p is reflexive, by the Banach-Alaoglu theorem [30, Theorem V.4.2], its unit ball is weakly compact. The set Δ is closed and convex, therefore it is weakly closed; see [30, Corollary V.1.5]. Thus, Δ is weakly compact as a weakly closed subset of the weakly compact unit ball. By the Eberlein–Šmulian theorem [30, Theorem V.13.1], Δ is also weakly sequentially compact.

We now prove (ii). A continuous function is sequentially continuous. Conversely, the inverse image of a closed set by a sequentially continuous function is sequentially closed. Besides, a sequentially closed subset of a sequentially compact set is sequentially compact. Using the Eberlein–Šmulian theorem, we deduce that the inverse images of closed sets are compact. In particular they are closed, which proves that a sequentially continuous function is continuous. □

III.3.2 Invariance and continuity of the spectrum for compact operators

We recall a few facts on operators. Let $(E, \|\cdot\|)$ be a Banach space. Let $A \in \mathcal{L}(E)$. We denote by $A^\top$ the adjoint of A. A sequence $(A_n, n \in \mathbb{N})$ of elements of $\mathcal{L}(E)$ converges strongly to $A \in \mathcal{L}(E)$ if $\lim_{n\to\infty} \| A_n x - Ax\| = 0$ for all $x \in E$. Following [6], a set of operators $\mathscr{A} \subset \mathcal{L}(E)$ is *collectively compact* if the set $\{Ax : \ A \in \mathscr{A},\ \|x\| \leq 1\}$ is relatively compact.

We collect some known results on the spectrum of to compact operators. Recall that the spectrum of a compact operator is finite or countable and has at most one accumulation point, which is 0. Furthermore, 0 belongs to the spectrum of compact operators in infinite dimension. We refer to [136] for an introduction to Banach lattices; we shall only consider the Banach lattices $L^p(\Omega, \mu)$ for $p \geq 1$ on a probability space $(\Omega, \mathscr{F}, \mu)$ and a bounded operator A is positive if $A(L^p_+) \subset L^p_+$.

Lemma III.3.2. *Let A, B be elements of $\mathcal{L}(E)$.*

(i) If E is a Banach lattice, and if A, B and $A - B$ are positive operators, then we have:

$$\rho(A) \geq \rho(B). \tag{III.26}$$

(ii) If A is compact, then we have AB and BA compact and:

$$\mathrm{Spec}(A) = \mathrm{Spec}(A^\top) \tag{III.27}$$
$$\mathrm{Spec}(AB) = \mathrm{Spec}(BA) \tag{III.28}$$

and in particular:

$$\rho(AB) = \rho(BA). \tag{III.29}$$

(iii) Let $(E', \|\cdot\|')$ be a Banach space such that E' is continuously and densely embedded in E. Assume that $A(E') \subset E'$, and denote by A' the restriction of A to E' seen as an operator on E'. If A and A' are compact, then we have:

$$\mathrm{Spec}(A) = \mathrm{Spec}(A'). \tag{III.30}$$

(iv) Let $(A_n, n \in \mathbb{N})$ be a collectively compact sequence which converges strongly to A. Then, we have $\lim_{n\to\infty} \mathrm{Spec}(A_n) = \mathrm{Spec}(A)$ in $(\mathscr{K}, d_{\mathrm{H}})$, and $\lim_{n\to} \rho(T_n) = \rho(T)$.

Proof. Property (i) can be found in [114, Theorem 4.2]. Equation (III.27) from Property (ii) can be deduced from the [99, Theorem page 20]. Using the [99, Proposition page 25], we get that $\mathrm{Spec}(AB) \cap \mathbb{C}^* = \mathrm{Spec}(BA) \cap \mathbb{C}^*$, and thus (III.29). As A is compact we get that AB and BA are compact, thus 0 belongs to their spectrum in infinite dimension. Whereas in finite dimension, as $\det(AB) = \det(A)\det(B) = \det(BA)$ (where A and B denote also the matrix of the corresponding operator in a given base), we get that 0 belongs to the spectrum of AB if and only if it belongs to the spectrum of BA. This gives (III.28).

Property (iii) follows from [72, Corollary 1 and Section 6]. We eventually check Property (iv). We deduce from [6, Theorems 4.8 and 4.16] (see also (d) and (e) in [7, Section 3]) that

$$\lim_{n\to\infty} \mathrm{Spec}(T_n) = \mathrm{Spec}(T).$$

Then use that the function rad is continuous to deduce the convergence of the spectral radius from the convergence of the spectra (see also (f) in [7, Section 3]). □

III.4 First properties of the functions $\mathfrak{R}_e$ and $\mathfrak{I}$

III.4.1 The effective reproduction number $\mathfrak{R}_e$

We consider the kernel model $[(\Omega, \mathscr{F}, \mu), \mathbf{k}]$ under Assumption III.1, so that $\mathbf{k}$ is a kernel on Ω with finite double norm. Recall the effective reproduction number function $\mathfrak{R}_e[\mathbf{k}]$ defined on Δ by (III.14): $\mathfrak{R}_e[\mathbf{k}](\eta) = \rho(T_\mathbf{k} M_\eta)$, and the reproduction number $\mathfrak{R}_0[\mathbf{k}] = \rho(T_\mathbf{k})$. When there is no risk of confusion on the kernel $\mathbf{k}$, we simply write $\mathfrak{R}_e$ and $\mathfrak{R}_0$ for $\mathfrak{R}_e[\mathbf{k}]$ and $\mathfrak{R}_0[\mathbf{k}]$.

Proposition III.4.1 (Basic properties of $\mathfrak{R}_e$)**.** *Suppose Assumption III.1 holds. Let $\eta, \eta_1, \eta_2 \in \Delta$. The function $\mathfrak{R}_e = \mathfrak{R}_e[\mathbf{k}]$ satisfies the following properties:*

(i) $\mathfrak{R}_e(\eta_1) = \mathfrak{R}_e(\eta_2)$ if $\eta_1 = \eta_2$ μ-almost surely.

(ii) $\mathfrak{R}_e(\mathbb{0}) = 0$ and $\mathfrak{R}_e(\mathbb{1}) = \mathfrak{R}_0$.

(iii) $\mathfrak{R}_e(\eta_1) \leq \mathfrak{R}_e(\eta_2)$ if $\eta_1 \leq \eta_2$ μ-almost surely.

(iv) $\mathfrak{R}_e(\lambda\eta) = \lambda\mathfrak{R}_e(\eta)$ for all $\lambda \in [0, 1]$.

Proof. If $\eta_1 = \eta_2$ μ-almost surely, then we have that $T_{\mathbf{k}\eta_1} = T_{\mathbf{k}\eta_2}$, and thus $\mathfrak{R}_e(\eta_1) = \mathfrak{R}_e(\eta_2)$. This gives Point (i). Point (ii) is a direct consequence of the definition of $\mathfrak{R}_e$. Since for any fixed $\lambda \in \mathbb{C}$ and any operator A, the spectrum of λA is equal to $\{\lambda s,\ s \in \mathrm{Spec}(A)\}$, Point (iv) is clear. Finally, note that if $\eta_1 \leq \eta_2$ μ-almost everywhere, then the operator $T_{\mathbf{k}\eta_2} - T_{\mathbf{k}\eta_1}$ is positive. According to (III.26), we get that $\rho(T_{\mathbf{k}\eta_1}) \leq \rho(T_{k\eta_2})$. This concludes the proof of Point (iii). □

We generalize a continuity property on the spectral radius originally stated in Chapter II by weakening the topology.

Theorem III.4.2 (Continuity of $\mathfrak{R}_e[\mathbf{k}]$ and $\mathrm{Spec}[\mathbf{k}]$)**.** *Suppose Assumption III.1 holds. Then, the functions* $\mathrm{Spec}[\mathbf{k}]$ *and* $\mathfrak{R}_e[\mathbf{k}]$ *are continuous functions from* Δ *(endowed with the weak-topology) respectively to* $\mathscr{K}$ *(endowed with the Hausdorff distance) and to* $\mathbb{R}_+$ *(endowed with the usual Euclidean distance).*

Let us remark the proof holds even if $\mathbf{k}$ takes negative values.

Proof. Let B denote the unit ball in L^p, with $p \in (1, +\infty)$ from Assumption III.1. Since the operator $T_\mathbf{k}$ is compact, the set $T_\mathbf{k}(B)$ is relatively compact. For all $\eta \in \Delta$, set $\eta B = \{\eta g : g \in B\}$. As $\eta B \subset B$, we deduce that $T_{\mathbf{k}\eta}(B) = T_\mathbf{k}(\eta B) \subset T_\mathbf{k}(B)$. This implies that the family $(T_{\mathbf{k}\eta}, \eta \in \Delta)$ is collectively compact.

Let $(\eta_n, n \in \mathbb{N})$ be a sequence in Δ converging weakly to some $\eta \in \Delta$. Let $g \in L^p$. The weak convergence of η_n to η implies that $(T_{\mathbf{k}\eta_n}(g), n \in \mathbb{N})$ converges μ-almost surely to $T_{\mathbf{k}\eta}(g)$. Consider the function:

$$K(x) = \left(\int_\Omega \mathbf{k}(x, y)^q \, \mu(\mathrm{d}y) \right)^{1/q},$$

which belongs to L^p, thanks to (III.10). Since for all x,

$$\left| T_{\mathbf{k}\eta_n}(g)(x) \right| \leq T_\mathbf{k}(|\eta_n g|)(x) \leq K(x) \, \|\eta_n g\|_p \leq K(x) \, \| g \|_p ,$$

we deduce, by dominated convergence, that the convergence holds also in L^p:

$$\lim_{n\to\infty} \left\| T_{\mathbf{k}\eta_n}(g) - T_{\mathbf{k}\eta}(g) \right\|_p = 0, \tag{III.31}$$

so that $T_{\mathbf{k}\eta_n}$ converges strongly to $T_{\mathbf{k}\eta}$. Using Lemma III.3.2 (iv) (with $T_n = T_{\mathbf{k}\eta_n}$ and $T = T_{\mathbf{k}\eta}$) on the continuity of the spectrum, we get that $\lim_{n\to\infty} \mathrm{Spec}[\mathbf{k}](\eta_n) = \mathrm{Spec}[\mathbf{k}](\eta)$. The function $\mathrm{Spec}[\mathbf{k}]$ is thus sequentially continuous, and, thanks to Lemma III.3.1, it is continuous from Δ endowed with the weak topology to the metric space $\mathscr{K}$ endowed with the Hausdorff distance. The continuity of $\mathfrak{R}_e[\mathbf{k}]$ then follows from its definition (III.8) as the composition of the continuous functions rad and $\mathrm{Spec}[\mathbf{k}]$. □

We now give a stability property of the spectrum and spectral radius with respect to the kernel $\mathbf{k}$.

Proposition III.4.3 (Stability of $\mathfrak{R}_e[\mathbf{k}]$ and $\mathrm{Spec}[\mathbf{k}]$)**.** *Let* $p \in (1, +\infty)$*. Let* $(\mathbf{k}_n, n \in \mathbb{N})$ *and* $\mathbf{k}$ *be kernels on* Ω *with finite double norms on* L^p*. If* $\lim_{n\to\infty} \| \mathbf{k}_n - \mathbf{k} \|_{p,q} = 0$*, then we have:*

$$\lim_{n\to\infty} \sup_{\eta\in\Delta} \left| \mathfrak{R}_e[\mathbf{k}_n](\eta) - \mathfrak{R}_e[\mathbf{k}](\eta) \right| = 0 \quad \text{and} \quad \lim_{n\to\infty} \sup_{\eta\in\Delta} d_\mathrm{H}\Big(\mathrm{Spec}[\mathbf{k}_n](\eta), \mathrm{Spec}[\mathbf{k}](\eta) \Big) = 0. \tag{III.32}$$

Proof. Let us first prove that, if $(\eta_n, n \in \mathbb{N})$ is a sequence in Δ which converges weakly to $\eta \in \Delta$, then $\mathrm{Spec}[\mathbf{k}_n](\eta_n)$ converges to $\mathrm{Spec}[\mathbf{k}](\eta)$ in Haussdorff distance.

All the operators in $\mathscr{A} = \{T_\mathbf{k}\} \cup \{T_{\mathbf{k}_n} : n \in \mathbb{N}\}$ are compact, and we deduce from (III.12) that:

$$\lim_{n\to\infty} \left\| T_{\mathbf{k}_n} - T_\mathbf{k} \right\|_{L^p} = 0.$$

Therefore $\mathscr{A}$ is a compact set (in the uniform topology) of compact operators: by [5, Theorem 2.4], $\mathscr{A}$ is collectively compact. This implies, see [6, Proposition 4.1(2)] for details, that the family $\mathscr{A}' = \{T' M_\eta : , T' \in \mathscr{A} \text{ and } \eta \in \Delta\}$ is collectively compact. A fortiori the sequence $(T_n = T_{\mathbf{k}_n\eta_n} = T_{\mathbf{k}_n} M_{\eta_n}, n \in \mathbb{N})$ of elements of $\mathscr{A}'$ is collectively compact, and $T = T_{\mathbf{k}\eta} = T_\mathbf{k} M_\eta$ is compact.

Let $g \in L^p$. We have:

$$\| T_n(g) - T(g) \|_p \leq \left\| T_{\mathbf{k}_n} - T_\mathbf{k} \right\|_{L^p} \| g \|_p + \left\| T_{\mathbf{k}\eta_n}(g) - T_{\mathbf{k}\eta}(g) \right\|_p .$$

Using $\lim_{n\to\infty} \left\| T_{\mathbf{k}_n} - T_{\mathbf{k}} \right\|_{L^p} = 0$ and (III.31), we get that $\lim_{n\to\infty} \| T_n(g) - T(g) \|_p$, thus $(T_n, n \in \mathbb{N})$ converges strongly to T. With Lemma III.3.2 (iv), we get that $\lim_{n\to\infty} \mathrm{Spec}(T_n) = \mathrm{Spec}(T)$, that is $\lim_{n\to\infty} \mathrm{Spec}[\mathbf{k}_n](\eta_n) = \mathrm{Spec}[\mathbf{k}](\eta)$.

Then, as the function $\eta \mapsto d_{\mathrm{H}}\Big(\mathrm{Spec}[\mathbf{k}_n](\eta), \mathrm{Spec}[\mathbf{k}](\eta)\Big)$ is continuous on the compact set Δ, thanks to Theorem III.4.2, it reaches its maximum say at $\eta_n \in \Delta$ for $n \in \mathbb{N}$. As Δ is compact, consider a sub-sequence which converges weakly to a limit say η. Since

$$\begin{aligned}\sup_{\eta\in\Delta} d_{\mathrm{H}}\Big(\mathrm{Spec}[\mathbf{k}_n](\eta), \mathrm{Spec}[\mathbf{k}](\eta)\Big)& \\ &= d_{\mathrm{H}}\Big(\mathrm{Spec}[\mathbf{k}_n](\eta_n), \mathrm{Spec}[\mathbf{k}](\eta_n)\Big) \\ &\leq d_{\mathrm{H}}\Big(\mathrm{Spec}[\mathbf{k}_n](\eta_n), \mathrm{Spec}[\mathbf{k}](\eta)\Big) + d_{\mathrm{H}}\Big(\mathrm{Spec}[\mathbf{k}](\eta_n), \mathrm{Spec}[\mathbf{k}](\eta)\Big),\end{aligned}$$

using the continuity of $\mathrm{Spec}[\mathbf{k}]$, we deduce that along this sub-sequence the right hand side converges to 0. Since this result holds for any converging sub-sequence, we get the second part of (III.32). The first part then follows from the definition (III.8) of $\mathfrak{R}_e$ as a composition, and the Lipschitz continuity of the function rad. □

III.4.2 The asymptotic proportion of infected individuals $\mathfrak{I}$

We consider the SIS model $[(\Omega, \mathscr{F}, \mu), k, \gamma]$ under Assumption III.2. Recall from (III.23) that the asymptotic proportion of infected individuals $\mathfrak{I}$ is given on Δ by $\mathfrak{I}(\eta) = \int_\Omega \mathfrak{g}_\eta \, \eta \, \mathrm{d}\mu$, where $\mathfrak{g}_\eta$ is the maximal solution in Δ of the equation $F_\eta(h) = 0$. We first give preliminary results on the maximal equilibrium, which complete what is known from Chapter II. Notice that, if $\mathfrak{R}_0 > 1$, then Property (iii) implies that the strategy $1 - \mathfrak{g}$ is critical.

Lemma III.4.4 (Properties of the maximal equilibrium). *Suppose Assumption III.2 holds and write $\mathfrak{R}_e$ for $\mathfrak{R}_e[k/\gamma]$.*

(i) Let $\eta, g \in \Delta$. If $F_\eta(g) \geq 0$, then we have $g \leq \mathfrak{g}_\eta$.

(ii) For any $h \in \Delta$, $h = \mathfrak{g}$ if and only if $F(h) = 0$ and $\mathfrak{R}_e(1 - h) \leq 1$.

(iii) If $\mathfrak{g} \neq 0$, then $\mathfrak{R}_e(1 - \mathfrak{g}) = 1$.

Proof. For the first point, consider the solution u_t of the SIS model with vaccination $\partial_t u_t = F_\eta(u_t)$ and initial condition $u_0 = g$. According to Proposition II.2.10, this solution is non-decreasing since $F_\eta(g) \geq 0$. According to Proposition II.2.13, the pointwise limit of u_t is an equilibrium. As this limit is dominated by the maximal equilibrium $\mathfrak{g}_\eta$ and since u_t is non-decreasing, this proves that $g \leq \mathfrak{g}_\eta$.

The other two points follow from Proposition III.8.2 and will be proved in Section III.8.1. □

We may now state the main properties of the function $\mathfrak{I}$.

Proposition III.4.5 (Basic properties of $\mathfrak{I}$). *Suppose that Assumption III.2 holds. Let $\eta, \eta_1, \eta_2 \in \Delta$. The function $\mathfrak{I}$ has the following properties:*

(i) $\mathfrak{I}(\eta_1) = \mathfrak{I}(\eta_2)$ if $\eta_1 = \eta_2$ μ-almost surely.

(ii) $\mathfrak{I}(\eta) = 0$ if and only if $\mathfrak{R}_e[k/\gamma](\eta) \leq 1$.

(iii) $\mathfrak{I}(\eta_1) \leq \mathfrak{I}(\eta_2)$ if $\eta_1 \leq \eta_2$ μ-almost surely.

(iv) $\mathfrak{I}(\lambda\eta) \leq \lambda\mathfrak{I}(\eta)$ for all $\lambda \in [0, 1]$.

Proof. If $\eta_1 = \eta_2$ μ-almost surely, then the operators $\mathcal{T}_{k\eta_1}$ and $\mathcal{T}_{k\eta_2}$ are equal. Thus, the equilibria $\mathfrak{g}_{\eta_1}$ and $\mathfrak{g}_{\eta_2}$ are also equal, which in turns implies that $\mathfrak{I}(\eta_1) = \mathfrak{I}(\eta_2)$. Point (ii) is already stated in Equation (III.24).

To prove the monotonicity (Point (iii)), consider $\eta_1 \leq \eta_2$. Since $\mathcal{T}_{k\eta_1} \leq \mathcal{T}_{k\eta_2}$, we get $F_{\eta_1}(g) \leq F_{\eta_2}(g)$ for all $g \in \Delta$. In particular, taking $g = \mathfrak{g}_{\eta_1}$ and using (III.22), we get $F_{\eta_2}(\mathfrak{g}_{\eta_1}) \geq 0$. By Lemma III.4.4 this implies $\mathfrak{g}_{\eta_1} \leq \mathfrak{g}_{\eta_2}$. To sum up, we get:

$$\eta_1 \leq \eta_2 \quad \Longrightarrow \quad \mathfrak{g}_{\eta_1} \leq \mathfrak{g}_{\eta_2}. \tag{III.33}$$

This readily implies that $\mathfrak{I}(\eta_1) = \int_\Omega \mathfrak{g}_{\eta_1} \eta_1 \,\mathrm{d}\mu \leq \int_\Omega \mathfrak{g}_{\eta_2} \eta_2 \,\mathrm{d}\mu = \mathfrak{I}(\eta_2)$. We conclude using Point (i).

We now consider Point (iv). Since $\lambda \in [0,1]$, we deduce from (III.33) that $\mathfrak{g}_{\lambda\eta} \leq \mathfrak{g}_\eta$. This implies that $\mathfrak{I}(\lambda\eta) = \int_\Omega \mathfrak{g}_{\lambda\eta} \lambda\eta \,\mathrm{d}\mu \leq \lambda \int_\Omega \mathfrak{g}_\eta \eta \,\mathrm{d}\mu = \lambda\mathfrak{I}(\eta)$. □

The proof of the following continuity results are both postponed to Section III.8.1.

Theorem III.4.6 (Continuity of $\mathfrak{I}$). *Suppose that Assumption III.2 holds. The function $\mathfrak{I}$ defined on Δ is continuous with respect to the weak topology.*

We write $\mathfrak{I}[k,\gamma]$ for $\mathfrak{I}$ to stress the dependence on the parameters k, γ of the SIS model.

Proposition III.4.7 (Stability of $\mathfrak{I}$). *Let $((k_n, \gamma_n), n \in \mathbb{N})$ and (k, γ) be a sequence of kernels and functions satisfying Assumption III.2. Assume furthermore that there exists $p' \in (1, +\infty)$ such that $\mathbf{k} = \gamma^{-1}k$ and $(\mathbf{k}_n = \gamma_n^{-1}k_n, n \in \mathbb{N})$ have finite double norm in $L^{p'}$ and that $\lim_{n\to\infty} \|\mathbf{k}_n - \mathbf{k}\|_{p',q'} = 0$. Then we have:*

$$\lim_{n\to\infty} \sup_{\eta\in\Delta} \left|\mathfrak{I}[k_n,\gamma_n](\eta) - \mathfrak{I}[k,\gamma](\eta)\right| = 0. \tag{III.34}$$

III.5 Pareto and anti-Pareto frontiers

III.5.1 The setting

To any vaccination strategy $\eta \in \Delta$, we associate a cost and a loss.

- ***The cost function***. The *cost* $C(\eta)$ measures all the costs of the vaccination strategy (production and diffusion). The cost is expected to be a decreasing function of η, since η encodes the non-vaccinated population. Since doing nothing costs nothing, we also expect $C(\mathbb{1}) = 0$, see Assumptions III.3 below. We shall also consider natural hypothesis on C, see Assumptions III.4 (p. 86) and III.6 (p. 87). A simple cost model is the affine cost given by:

  $$C_{\text{aff}}(\eta) = \int_\Omega (1 - \eta(x))\, c_{\text{aff}}(x)\, \mu(\mathrm{d}x), \tag{III.35}$$

  where $c_{\text{aff}}(x)$ is the cost of vaccination of population of feature x, with $c_{\text{aff}} \in L^1$ positive. The particular case $c_{\text{aff}} = 1$ is the uniform cost $C = C_{\text{uni}}$:

  $$C_{\text{uni}}(\eta) = \int_\Omega (1 - \eta)\,\mathrm{d}\mu. \tag{III.36}$$

  The real cost of the vaccination may be a more complicated function $\psi(C_{\text{aff}}(\eta))$ of the affine cost, for example if the marginal cost of producing a vaccine depends on the quantity already produced. However, as long as ψ is strictly increasing, this will not affect the optimal strategies.

- ***The loss function***. The *loss* $\mathrm{L}(\eta)$ measures the (non)-efficiency of the vaccination strategy η. Different choices are possible here. We prove in this section general results that only depend on a few natural hypothesis for L; see Assumptions III.3 (p. 82), III.5 (p. 86) and III.7 (p. 88). These hypothesis are in particular satisfied if the loss is the effective reproduction number $\mathfrak{R}_e$ (kernel and SIS models), or the asymptotic proportion of infected individuals $\mathfrak{I}$ (SIS model); more precisely see Lemmas III.5.6, III.5.14 and III.5.16.

We shall consider cost and loss functions with some regularities.

Definition III.5.1. *We say that a real-valued function H defined on Δ endowed with the weak topology is:*

- **Continuous**: *if H is continuous with respect to the weak topology on Δ.*
- **Non-decreasing**: *if for any $\eta_1, \eta_2 \in \Delta$ such that $\eta_1 \leq \eta_2$, we have $H(\eta_1) \leq H(\eta_2)$.*
- **Decreasing**: *if for any $\eta_1, \eta_2 \in \Delta$ such that $\eta_1 \leq \eta_2$ and $\int_\Omega \eta_1 \, \mathrm{d}\mu < \int_\Omega \eta_2 \, \mathrm{d}\mu$, we have $H(\eta_1) > H(\eta_2)$.*
- **Sub-homogeneous**: *if $H(\lambda\eta) \leq \lambda H(\eta)$ for all $\eta \in \Delta$ and $\lambda \in [0, 1]$.*

The definition of non-increasing function and increasing function are similar.

Assumption III.3 (On the cost function and loss function)**.** *The loss function $\mathrm{L} : \Delta \to \mathbb{R}$ is non-decreasing and continuous with $\mathrm{L}(\mathbb{0}) = 0$. The cost function $C : \Delta \to \mathbb{R}$ is non-increasing and continuous with $C(\mathbb{1}) = 0$. We also have:*

$$\ell_{\max} := \max_\Delta \mathrm{L} > 0 \quad \textit{and} \quad c_{\max} := \max_\Delta C > 0.$$

Assumption III.3 will always hold. In particular, the loss and the cost functions are non-negative and non-constant.

We will consider the multi-objective minimization and maximization problems:

$$\begin{cases} \textbf{Minimize:} & (C(\eta), \mathrm{L}(\eta)) \\ \textbf{subject to:} & \eta \in \Delta \end{cases} \quad \text{and} \quad \begin{cases} \textbf{Maximize:} & (C(\eta), \mathrm{L}(\eta)) \\ \textbf{subject to:} & \eta \in \Delta \end{cases} \tag{III.37}$$

Remark III.5.2 (On the generality of the uniform cost). For the reproduction number optimization in the vaccination context, one can without loss of generality consider the uniform cost instead of the affine cost. Indeed, consider the kernel model $\mathrm{Param} = [(\Omega, \mathscr{F}, \mu), \mathbf{k}]$ with the affine cost function C_{aff} and the loss $\mathfrak{R}_e$. If we assume furthermore that c_{aff} is bounded and bounded away from 0 (that is c_{aff} and $1/c_{\mathrm{aff}}$ belongs to $\mathscr{L}^\infty_+$), and, without loss of generality, that $\int c_{\mathrm{aff}} \, \mathrm{d}\mu = 1$, then we can consider the weighted kernel model $\mathrm{Param}_0 = [(\Omega, \mathscr{F}, \mu_0), \mathbf{k}_0]$ with measure $\mu_0(\mathrm{d}x) = c_{\mathrm{aff}}(x) \, \mu(\mathrm{d}x)$ and kernel $\mathbf{k}_0 = \mathbf{k}/c_{\mathrm{aff}}$. (Notice that if Assumption III.2 holds for the model Param, then it also holds for the model Param_0.) Consider the loss $\mathrm{L} = \mathfrak{R}_e$. Then for a strategy $\eta \in \Delta$, we get that $(C_{\mathrm{aff}}(\eta), \mathrm{L}(\eta))$ for the model Param is equal to $(C_{\mathrm{uni}}(\eta), \mathrm{L}(\eta))$ for the model Param_0. Therefore, for the loss function $\mathrm{L} = \mathfrak{R}_e$, instead of the affine cost C_{aff}, one can consider without any real loss of generality the uniform cost. (This holds also for the SIS model.) However, this is no longer the case for the loss function $\mathrm{L} = \mathfrak{I}$ in the SIS model.

Multi-objective problems are in a sense ill-defined because in most cases, it is impossible to find a single solution that would be optimal to all objectives simultaneously. Hence, we recall the concept of Pareto optimality. Since the minimization problem is crucial for vaccination, we shall define Pareto optimality for the bi-objective minimization problem. A strategy $\eta_\star \in \Delta$ is said to be *Pareto optimal* for the minimization problem in (III.37) if any improvement of one objective leads to a deterioration of the other, for $\eta \in \Delta$:

$$C(\eta) < C(\eta_\star) \implies \mathrm{L}(\eta) > \mathrm{L}(\eta_\star) \quad \text{and} \quad \mathrm{L}(\eta) < \mathrm{L}(\eta_\star) \implies C(\eta) > C(\eta_\star). \tag{III.38}$$

The set of Pareto optimal strategies will be denoted by $\mathcal{P}_\mathrm{L}$, and the Pareto frontier is defined as the set of Pareto optimal outcomes:

$$\mathcal{F}_\mathrm{L} = \{(C(\eta), \mathrm{L}(\eta)) \, : \, \eta \text{ Pareto optimal}\}.$$

Similarly, a strategy $\eta^\star \in \Delta$ is *anti-Pareto optimal* if it is Pareto optimal for the bi-objective maximization problem in (III.37). Intuitively, the "best" vaccination strategies are the Pareto optima and the "worst" vaccination strategies are the anti-Pareto optima. We denote similarly by $\mathcal{P}_{\mathrm{L}}^{\mathrm{Anti}}$ the set of anti-Pareto optimal strategies, and by $\mathcal{F}_{\mathrm{L}}^{\mathrm{Anti}}$ its frontier:

$$\mathcal{F}_{\mathrm{L}}^{\mathrm{Anti}} = \{(C(\eta), \mathrm{L}(\eta)) \;:\; \eta \text{ anti-Pareto optimal}\}.$$

Finally, we define the *feasible region* as the set of all possible outcomes:

$$\mathbf{F} = \{(C(\eta), \mathrm{L}(\eta)),\, \eta \in \Delta\}.$$

We now consider the classical, single-objective minimization problems related to the "best" vaccination strategies, with a fixed loss $\ell \in [0, \ell_{\max}]$ or a fixed cost $c \in [0, c_{\max}]$:

$$\textbf{Minimize:} \quad \mathrm{L}(\eta) \tag{III.39a}$$
$$\textbf{subject to:} \quad \eta \in \Delta,\, C(\eta) \leq c, \tag{III.39b}$$

as well as

$$\textbf{Minimize:} \quad C(\eta) \tag{III.40a}$$
$$\textbf{subject to:} \quad \eta \in \Delta,\, \mathrm{L}(\eta) \leq \ell. \tag{III.40b}$$

We denote the values of Problems (III.39) and (III.40) by:

$$\begin{aligned} \mathrm{L}_\star(c) &= \inf\{\mathrm{L}(\eta) \;:\; \eta \in \Delta \text{ and } C(\eta) \leq c\} \quad \text{for } c \in [0, c_{\max}], \\ C_\star(\ell) &= \inf\{C(\eta) \;:\; \eta \in \Delta \text{ and } \mathrm{L}(\eta) \leq \ell\} \quad \text{for } \ell \in [0, \ell_{\max}]. \end{aligned}$$

Similarly, the maximization problem related to the "worst" vaccination strategies for a fixed loss $\ell \in [0, \ell_{\max}]$ or a fixed cost $c \in [0, c_{\max}]$ are defined by:

$$\textbf{Maximize:} \quad \mathrm{L}(\eta) \tag{III.41a}$$
$$\textbf{subject to:} \quad \eta \in \Delta,\, C(\eta) \geq c, \tag{III.41b}$$

as well as

$$\textbf{Maximize:} \quad C(\eta) \tag{III.42a}$$
$$\textbf{subject to:} \quad \eta \in \Delta,\, \mathrm{L}(\eta) \geq \ell. \tag{III.42b}$$

We denote the values of Problems (III.41) and (III.42) by:

$$\begin{aligned} \mathrm{L}^\star(c) &= \sup\{\mathrm{L}(\eta) \;:\; \eta \in \Delta \text{ and } C(\eta) \geq c\} \quad \text{for } c \in [0, c_{\max}], \\ C^\star(\ell) &= \sup\{C(\eta) \;:\; \eta \in \Delta \text{ and } \mathrm{L}(\eta) \geq \ell\} \quad \text{for } \ell \in [0, \ell_{\max}]. \end{aligned}$$

If necessary, we may write $C_{\star,\mathrm{L}}$ and $C^{\star,\mathrm{L}}$ to stress the dependence of the function $C_\star$ and $C^\star$ in the loss function L.

Under Assumption III.3, as the loss and the cost functions are continuous on the compact set Δ, the infima in the definitions of the value functions $C_\star$ and $\mathrm{L}_\star$ are minima; and the suprema in the definition of the value functions $C^\star$ and $\mathrm{L}^\star$ are maxima. Since Δ in endowed with the weak topology, we will consider the set of Pareto and anti-Pareto optimal vaccination modulo μ-almost sure equality.

See Figure III.2 for a typical representation of the possible aspects of the feasible region $\mathbf{F}$ (in light blue), the value functions and the Pareto and anti-Pareto frontiers under the general Assumption III.3, and the connected Pareto and anti-Pareto frontiers under further regularity on the cost and loss functions (see Assumption III.4-III.7 below) in Figure III.2(D). In Figure III.1(B), we have plotted in solid red line the Pareto frontier and in dashed red line the anti-Pareto frontier from Example III.1.7.

Outline of the section

It turns out that the anti-Pareto optimization problem can be recast as a Pareto optimization problem by changing signs and exchanging the cost and loss functions. In order to make use of this property for the kernel and SIS models, we study the Pareto problem under assumptions on the cost that are general enough to cover the choices C_{uni} and $-\mathrm{L}$, and assumptions on the loss that cover the choices $\mathfrak{R}_e$, $\mathfrak{I}$ and $-C_{\text{uni}}$.

The main result of this section states that all the solutions of the optimization Problems (III.39) or (III.40) are Pareto optimal, and gives a description of the Pareto frontier $\mathcal{F}_{\mathrm{L}}$ as a graph in Section III.5.2, and similarly for the anti-Pareto frontier in Section III.5.3. Surprisingly, those two problems are not completely symmetric for the loss functions $\mathfrak{R}_e$ and $\mathfrak{I}$ considered in Section III.5.4, see Lemmas III.5.14 and III.5.16, where one uses some irreducibility condition on the kernels to study the anti-Pareto frontier.

III.5.2 On the Pareto frontier

We first check that Problems (III.39) and (III.40) have solutions.

Proposition III.5.3 (Optimal solutions for fixed cost or fixed loss). *Suppose that Assumption III.3 holds. For any cost $c \in [0, c_{\max}]$, there exists a minimizer of the loss under the cost constraint $C(\cdot) \leq c$, that is, a solution to Problem* (III.39). *Similarly, for any loss $\ell \in [0, \ell_{\max}]$, there exists a minimizer of the cost under the loss constraint $\mathrm{L}(\cdot) \leq \ell$, that is a solution to Problem* (III.40).

Proof. Let $c \in [0, c_{\max}]$. Since $C(\mathbb{1}) = 0$, the set $\{\eta \in \Delta \,:\, C(\eta) \leq c\}$ is non-empty. It is also compact as C is continuous on the compact set Δ (for the weak topology). Therefore, since the loss function L is continuous (for the weak topology), we get that L restricted to this compact set reaches its minimum. Thus, Problem (III.39) has a solution. The proof is similar for the existence of a solution to Problem (III.40). ☐

We start by a general result concerning the links between the three problems.

Proposition III.5.4 (Single-objective and bi-objective problems). *Suppose Assumption III.3 holds.*

(i) If $\eta_\star$ is Pareto optimal, then $\eta_\star$ is a solution of (III.39) *for the cost $c = C(\eta_\star)$, and a solution of* (III.40) *for the loss $\ell = \mathrm{L}(\eta_\star)$. Conversely, if $\eta_\star$ is a solution to both problems* (III.39) *and* (III.40) *for some values c and ℓ, then $\eta_\star$ is Pareto optimal.*

(ii) The Pareto frontier is the intersection of the graphs of $C_\star$ and $\mathrm{L}_\star$:

$$\mathcal{F}_{\mathrm{L}} = \{(c, \ell) \in [0, c_{\max}] \times [0, \ell_{\max}] \,:\, c = C_\star(\ell) \text{ and } \ell = \mathrm{L}_\star(c)\}.$$

(iii) The points $(0, \mathrm{L}_\star(0))$ and $(C_\star(0), 0)$ both belong to the Pareto frontier, and:

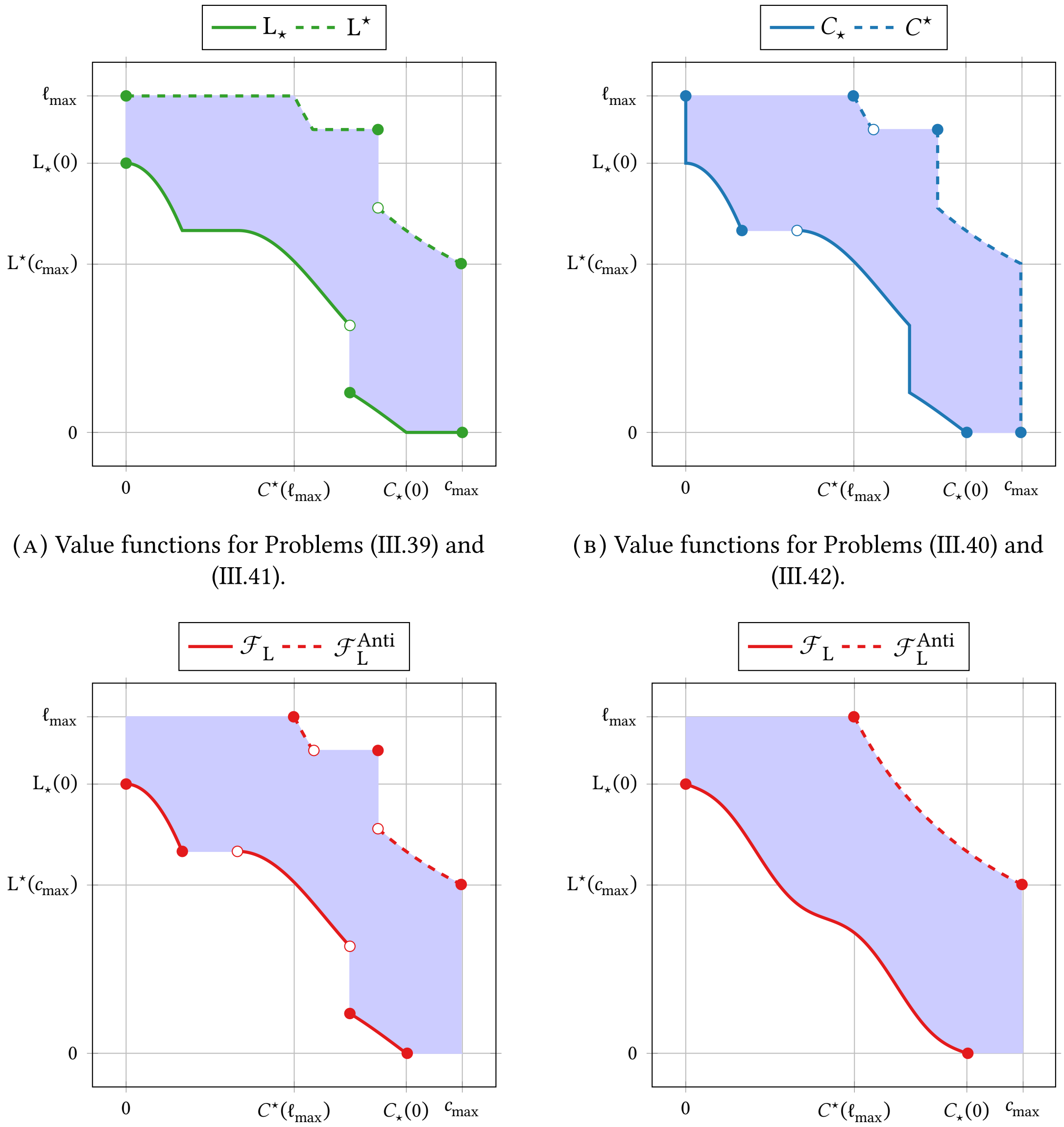
$$C_\star(\mathrm{L}_\star(0)) = \mathrm{L}_\star(C_\star(0)) = 0.$$

Moreover, we also have $C_\star(\ell) = 0$ for $\ell \in [\mathrm{L}_\star(0), \ell_{\max}]$, and $\mathrm{L}_\star(c) = 0$ for $c \in [C_\star(0), c_{\max}]$.

Proof. Let us prove (i). If $\eta_\star$ is Pareto optimal, then for any strategy η, if $C(\eta) \leq C(\eta_\star)$ then $\mathrm{L}(\eta) \geq \mathrm{L}(\eta_\star)$ by taking the contraposition in (III.38), and $\eta_\star$ is indeed a solution of Problem (III.39) with $c = C(\eta_\star)$. Similarly $\eta_\star$ is a solution of Problem (III.40).

For the converse statement, let $\eta_\star$ be a solution of (III.39) for some c and of (III.40) for some ℓ. It is also a solution of (III.39) with $c = C(\eta_\star)$. In particular, we get that for $\eta \in \Delta$, $\mathrm{L}(\eta) < \mathrm{L}(\eta_\star)$ implies that $C(\eta) > c = C(\eta_\star)$, which is the second part of (III.38). Similarly, use that $\eta_\star$ is a solution to (III.40), to get that the first part of (III.38) also holds. Thus the strategy $\eta_\star$ is Pareto optimal.

To prove Point (ii), we first prove that $\mathcal{F}_{\mathrm{L}}$ is a subset of $\{(c, \ell) \,:\, c = C_\star(\ell) \text{ and } \ell = \mathrm{L}_\star(c)\}$. A point in $\mathcal{F}_{\mathrm{L}}$ may be written as $(C(\eta_\star), \mathrm{L}(\eta_\star))$ for some Pareto optimal strategy $\eta_\star$. By Point (i), $\eta_\star$



(A) Value functions for Problems (III.39) and (III.41).

(B) Value functions for Problems (III.40) and (III.42).

(C) Pareto and anti-Pareto frontier.

(D) Pareto and anti-Pareto frontier under additional regularity Assumptions III.4-III.7.

Figure III.2: An example of the possible aspects of the feasible region **F** (in light blue), the value functions $L_\star, L^\star, C_\star, C^\star$, and the Pareto and anti-Pareto frontier (in red) under Assumption III.3.

solves Problem (III.39) for the cost $C(\eta_\star)$, so $L_\star(C(\eta_\star)) = L(\eta_\star)$. Similarly, we have $C_\star(L(\eta_\star)) = C(\eta_\star)$, as claimed.

We now prove the reverse inclusion. Assume that $c = C_\star(\ell)$ and $\ell = L_\star(c)$, and consider η a solution of Problem (III.40) for the loss ℓ: $L(\eta) \leq \ell$ and $C(\eta) = C_\star(\ell) = c$. Then η is admissible for Problem (III.39) with cost $c = C_\star(\ell)$, so $L(\eta) \geq L_\star(C_\star(\ell)) = L_\star(c) = \ell$. Therefore, we get $L(\eta) = L^\star(c)$, and η is also a solution of Problem (III.39). By Point (i), η is Pareto optimal, so $(C(\eta), L(\eta)) = (c, \ell) \in \mathcal{F}_L$, and the reverse inclusion is proved.

Finally we prove Point (iii). We have $C_\star(0) = \min\{C(\eta) : \eta \in \Delta \text{ and } L(\eta) = 0\} \in [0, c_{\max}]$. Let $\eta \in \Delta$ such that $L(\eta) = 0$ and $C(\eta) = C_\star(0)$. We deduce that $L_\star(C_\star(0)) \leq L(\eta) = 0$ and thus $L_\star(C_\star(0)) = 0$ as L is non-negative. We deduce from (ii) that $(C_\star(0), 0)$ belongs to $\mathcal{F}_L$. Since $C_\star$ is non-increasing, we also get that $C_\star = 0$ on $[C_\star(0), c_{\max}]$. The other properties of (iii) are proved similarly. □

The next two hypotheses on C and L will imply that the Pareto frontier is connected.

Assumption III.4. *If the cost C has a local minimum (for the weak topology) at η, then $C(\eta) = 0$, that is, η is a global minimum of C.*

Assumption III.5. *If the loss L has a local minimum (for the weak topology) at η, then $\mathrm{L}(\eta) = 0$, that is, η is a global minimum of L.*

Under these hypotheses, the picture becomes much nicer, see Figure III.2(D), where the only flat parts of the graphs of $C_\star$ and $\mathrm{L}_\star$ occur at zero cost or zero loss.

Proposition III.5.5. *Under Assumption III.3 and III.4 the following properties hold:*

(i) The optimal cost $C_\star$ is decreasing on $[0, \mathrm{L}_\star(0)]$.

(ii) If η solves Problem (III.40) for the loss $\ell \in [0, \mathrm{L}_\star(0)]$, then $\mathrm{L}(\eta) = \ell$ (that is, the constraint is binding). Moreover η is Pareto optimal, and:

$$\mathrm{L}_\star(C_\star(\ell)) = \ell. \tag{III.43}$$

(iii) The Pareto frontier is the graph of $C_\star$:

$$\mathcal{F}_{\mathrm{L}} = \{(C_\star(\ell), \ell) \,:\, \ell \in [0, \mathrm{L}_\star(0)]\}. \tag{III.44}$$

Similarly, under Assumptions III.3 and III.5, the following properties hold:

(iv) The optimal loss $\mathrm{L}_\star$ is decreasing on $[0, C_\star(0)]$.

(v) If η solves Problem (III.39) for the cost $c \in [0, C_\star(0)]$, then $C(\eta) = c$. Moreover η is Pareto optimal, and $C_\star(\mathrm{L}_\star(c)) = c$.

(vi) The Pareto frontier is the graph of $\mathrm{L}_\star$:

$$\mathcal{F}_{\mathrm{L}} = \{(c, \mathrm{L}_\star(c)) \,:\, c \in [0, C_\star(0)]\}. \tag{III.45}$$

Finally, if Assumptions III.4 and III.5 hold, then $\mathrm{L}_\star$ is a continuous decreasing bijection of $[0, C_\star(0)]$ onto $[0, L^\star(0)]$ and $C_\star$ is the inverse bijection, and the Pareto frontier is compact and connected.

Proof. We prove (i). Let $0 \leq \ell < \ell' \leq \mathrm{L}_\star(0)$, and let $\eta_\star$ be a solution of Problem (III.40):

$$C(\eta_\star) = C_\star(\ell) \quad \text{and} \quad \mathrm{L}(\eta_\star) \leq \ell. \tag{III.46}$$

The set $\mathcal{O} = \{\eta \,:\, \mathrm{L}(\eta) < \ell'\}$ is open and contains $\eta_\star$. Since $\mathrm{L}(\eta_\star) < \mathrm{L}_\star(0)$, we get $C(\eta_\star) > 0$, so $\eta_\star$ is not a global minimum for C. By Assumption III.4, it cannot be a local minimum for C, so $\mathcal{O}$ contains at least one point η' for which $C(\eta') < C(\eta_\star)$. Since $\eta' \in \mathcal{O}$, we get $\mathrm{L}(\eta') \leq \ell'$, so that $C_\star(\ell') \leq C(\eta') < C(\eta_\star) = C_\star(\ell)$. Since $\ell < \ell'$ are arbitrary, $C_\star$ is decreasing on $[0, \mathrm{L}_\star(0)]$.

We now prove (ii). If the inequality in (III.46) was strict, that is $\mathrm{L}(\eta_\star) < \ell$, then we would get a contradiction as $C(\eta_\star) \geq C_\star(\mathrm{L}(\eta_\star)) > C_\star(\ell) = C(\eta_\star)$. Therefore any solution $\eta_\star$ of (III.40) satisfies $\mathrm{L}(\eta_\star) = \ell$, and in particular $C_\star(\mathrm{L}(\eta_\star)) = C_\star(\ell) = C(\eta_\star)$. This implies in turn that $\eta_\star$ also solves (III.39): if η satisfies $\mathrm{L}(\eta) < \mathrm{L}(\eta_\star)$, then using the definition of $C_\star$, the fact that it decreases, and the definition of $\eta_\star$, we get:

$$C(\eta) \geq C_\star(\mathrm{L}(\eta)) > C_\star(\mathrm{L}(\eta_\star)) = C(\eta_\star).$$

By contraposition, we have $\mathrm{L}(\eta) \geq \mathrm{L}(\eta_\star)$ for any η such that $C(\eta) \leq C(\eta_\star)$, proving that $\eta_\star$ is also a solution of (III.39) with $c = C(\eta_*)$. By Point (i) of Proposition III.5.4, $\eta_\star$ is Pareto optimal. Therefore $(C(\eta_\star), \mathrm{L}(\eta_\star)) = (C_\star(\ell), \ell)$ belongs to the Pareto frontier. Using Point (ii) of Proposition III.5.4, we deduce that $\ell = \mathrm{L}_\star(C_\star(\ell))$.

To prove Point (iii), note that Equation (III.43) shows that, if $c = C_\star(\ell)$ for $\ell \in [0, \mathrm{L}_\star(0)]$, then $\ell = \mathrm{L}_\star(c)$. Use Point (ii) and (iii) of Proposition III.5.4, to get that $\mathcal{F}_{\mathrm{L}} = \{(c, \ell) \,:\, c = C_\star(\ell), \ell \in [0, \mathrm{L}_\star(0)]\}$.

The claims (iv), (v) and (vi) are proved in the same way, exchanging the roles of L and C.

To conclude the proof, it remains to check that $C_\star$ and $\mathrm{L}_\star$ are continuous under Assumptions III.3, III.4 and III.5. We deduce from Point (ii) and Proposition III.5.3 that $[0, \mathrm{L}_\star(0)]$ is in the range of $\mathrm{L}_\star$. Since $\mathrm{L}_\star$ is decreasing, thanks to Point (iv) and $\mathrm{L}_\star(C_\star(0)) = 0$, see Proposition III.5.4 (iii), the function $\mathrm{L}_\star$ is continuous and decreasing on $[0, \mathrm{L}_\star(0)]$, and thus one-to-one from $[0, C_\star(0)]$ onto $[0, \mathrm{L}^\star(0)]$. Thanks to (III.43), its inverse bijection is the function $C_\star$. Since the frontier $\mathcal{F}_\mathrm{L}$ is given by (III.45), and $\mathrm{L}_\star$ is continuous, $\mathcal{F}_\mathrm{L}$ is compact and connected. □

Finally, let us check that Assumptions III.4 and III.5 hold under very simple assumptions, which are in particular satisfied by the cost functions C_{uni} and C_{aff} and the loss functions $\mathfrak{R}_e$ and $\mathfrak{I}$ (recall from Propositions III.4.1 and III.4.5 that $\mathfrak{R}_e$ and $\mathfrak{I}$ are sub-homogeneous).

Lemma III.5.6. *Suppose Assumption III.3 holds. If the cost function C is decreasing, then Assumption III.4 holds and $\mathrm{L}_\star(0) = \ell_{\max}$. If the loss function L is sub-homogeneous, then Assumption III.5 holds.*

Proof. Let $\eta \in \Delta$. If C has a local minimum at η, then, as C is non-increasing, for $\varepsilon > 0$ small enough, we get that $C(\eta) \geq C(\eta + \varepsilon(1 - \eta)) \geq C(\eta)$. If C is decreasing, this is only possible if $\eta = \mathbb{1}$, so that η is a global minimum of C. This also gives $\mathrm{L}_\star(0) = \ell_{\max}$. Similarly if L has a local minimum at η, then for $\varepsilon > 0$ small enough $\mathrm{L}(\eta) \leq \mathrm{L}((1-\varepsilon)\eta) \leq (1-\varepsilon)\mathrm{L}(\eta)$, so $\mathrm{L}(\eta) = 0$ and η is a global minimum of L. □

Corollary III.5.7. *Suppose that Assumptions III.3, III.4 and III.5 hold. The set of Pareto optimal strategies $\mathcal{P}_\mathrm{L}$ is compact (for the weak topology).*

Proof. Since $\mathrm{L}_\star$ is continuous thanks to Proposition III.5.5, we deduce that $\mathcal{F}_\mathrm{L}$, which is given by (III.45), is compact and thus closed. Since $\mathcal{P}_\mathrm{L} = f^{-1}(\mathcal{F}_\mathrm{L})$, where the function $f = (C, \mathrm{L})$ defined on Δ is continuous, we deduce that $\mathcal{P}_\mathrm{L}$ is closed and thus compact as Δ is compact. □

III.5.3 On the anti-Pareto frontier

Letting $C'(\eta) = \ell_{\max} - \mathrm{L}(\eta)$ and $\mathrm{L}'(\eta) = c_{\max} - C(\eta)$, it is easy to see that:

$$C'_\star(c) = \ell_{\max} - \mathrm{L}^\star(c_{\max} - c) \quad \text{and} \quad \mathrm{L}'_\star(\ell) = c_{\max} - C^\star(\ell_{\max} - \ell),$$

so that Proposition III.5.5 may be applied to the cost function C' and the loss function L' to yield the following result.

Proposition III.5.8 (Single-objective and bi-objective problems for the anti Pareto)**.** *Suppose Assumption III.3 holds.*

(i) If $\eta^\star$ is anti-Pareto optimal, then $\eta^\star$ is a solution of (III.41) *for the cost $c = C(\eta^\star)$, and a solution of* (III.42) *for the loss $\ell = \mathrm{L}(\eta^\star)$. Conversely, if $\eta^\star$ is a solution to both problems* (III.41) *and* (III.42) *for some values c and ℓ, then $\eta^\star$ is anti-Pareto optimal.*

(ii) The anti-Pareto frontier is the intersection of the graphs of $C^\star$ and $\mathrm{L}^\star$:

$$\mathcal{F}_\mathrm{L}^{\mathrm{Anti}} = \{(c, \ell) \in [0, c_{\max}] \times [0, \ell_{\max}] \ : \ c = C^\star(\ell) \text{ and } \ell = \mathrm{L}^\star(c)\}.$$

(iii) The points $(C^\star(\ell_{\max}), \ell_{\max})$ and $(c_{\max}, \mathrm{L}^\star(c_{\max}))$ both belong to the anti-Pareto frontier, and we have $C^\star(\mathrm{L}^\star(c_{\max})) = c_{\max}$ and $\mathrm{L}^\star(C^\star(\ell_{\max})) = \ell_{\max}$.

Moreover, we also have $C^\star(\ell) = c_{\max}$ for $\ell \in [0, \mathrm{L}^\star(c_{\max})]$, and $\mathrm{L}^\star(c) = \ell_{\max}$ for $c \in [0, C^\star(\ell_{\max})]$.

The following additional hypotheses rule out the occurrence of flat parts in the anti-Pareto frontier.

Assumption III.6. *If the cost C has a local maximum at η (for the weak topology), then $C(\eta) = c_{\max}$ and η is a global maximum of C.*

Assumption III.7. *If the loss* L *has a local maximum at* η *(for the weak topology), then* $\mathrm{L}(\eta) = \ell_{\max}$ *and* η *is a global maximum of* L.

The following result is now a consequence of Proposition III.5.5 and Corollary III.5.7 applied to the loss function L' and cost function C'.

Proposition III.5.9. *Under Assumption III.3 and III.6 the following properties hold:*

(i) The optimal cost $C^\star$ *is decreasing on* $[\mathrm{L}^\star(c_{\max}), \ell_{\max}]$.

(ii) If η *solves Problem* (III.42) *for the loss* $\ell \in [\mathrm{L}^\star(c_{\max}), \ell_{\max}]$*, then* $\mathrm{L}(\eta) = \ell$ *(that is, the constraint is binding). Moreover* η *is anti-Pareto optimal, and* $\mathrm{L}^\star(C^\star(\ell)) = \ell$.

(iii) The anti-Pareto frontier is the graph of $C^\star$*:*

$$\mathcal{F}_{\mathrm{L}}^{\mathrm{Anti}} = \{(C^\star(\ell), \ell) \;:\; \ell \in [\mathrm{L}^\star(c_{\max}), \ell_{\max}]\}. \tag{III.47}$$

Similarly, under Assumptions III.3 and III.7, the following properties hold:

(iv) The optimal loss $\mathrm{L}^\star$ *is decreasing on* $[C^\star(\ell_{\max}), c_{\max}]$.

(v) If η *solves Problem* (III.41) *for the cost* $c \in [C^\star(\ell_{\max}), c_{\max}]$*, then* $C(\eta) = c$*. Moreover* η *is anti-Pareto optimal, and* $C^\star(\mathrm{L}^\star(c)) = c$.

(vi) The anti-Pareto frontier is the graph of $\mathrm{L}^\star$*:*

$$\mathcal{F}_{\mathrm{L}}^{\mathrm{Anti}} = \{(c, \mathrm{L}^\star(c)) \;:\; c \in [C^\star(\ell_{\max}), c_{\max}]\}. \tag{III.48}$$

Finally, suppose that Assumptions III.3, III.6 and III.7 hold. Then, $\mathrm{L}^\star$ *is a continuous decreasing bijection of* $[C^\star(\ell_{\max}), c_{\max}]$ *onto* $[\mathrm{L}^\star(c_{\max}), \ell_{\max}]$*,* $C^\star$ *is the inverse bijection, and the anti-Pareto frontier is compact and connected. Furthermore, the set of anti-Pareto optimal strategies* $\mathcal{P}_{\mathrm{L}}^{\mathrm{Anti}}$ *is compact (for the weak topology).*

The following result is similar to the first part of Lemma III.5.6.

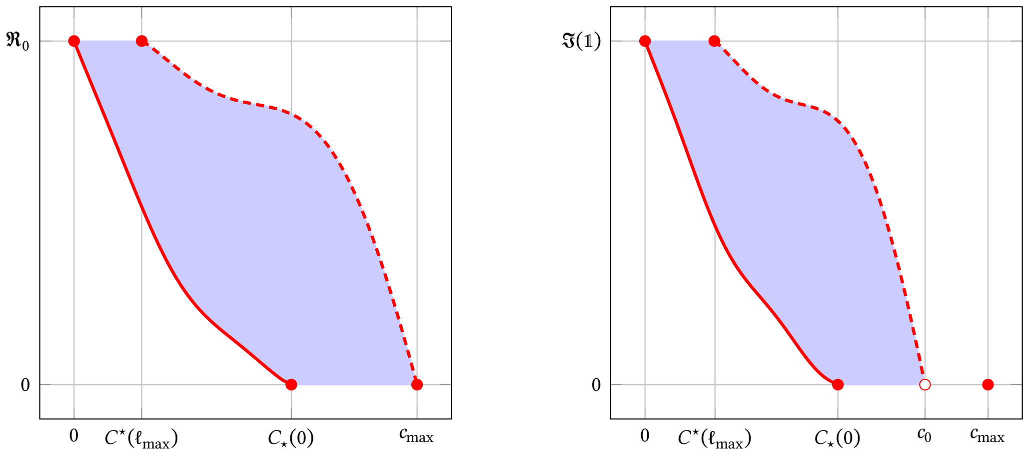
Lemma III.5.10. *Suppose Assumption III.3 holds. If the cost function* C *is decreasing, then Assumption III.6 holds and* $\mathrm{L}^\star(c_{\max}) = 0$.

Proof. Let $\eta \in \Delta$ and $\varepsilon \in (0,1)$. Since C is decreasing, $C((1-\varepsilon)\eta) \geq C(\eta)$, with equality if and only if $\eta = \mathbb{0}$ μ-almost surely. Therefore the only local maximum of C is $\eta = \mathbb{0}$, and it is a global maximum. Since $C(\eta) = c_{\max}$ implies that $\eta = \mathbb{0}$ μ-almost surely, we also get that $\mathrm{L}^\star(c_{\max}) = \mathrm{L}(\mathbb{0}) = 0$. □

III.5.4 The particular case of the kernel and SIS models

We check in this section that the loss functions $\mathfrak{R}_e$ and $\mathfrak{I}$ satisfy Assumption III.5. We also prove under some irreducibility condition on the kernel that $\mathfrak{R}_e$ satisfies Assumption III.7. The situation is a bit more complicated fo the loss $\mathfrak{I}$, for which Assumption III.7 does not hold. However, $\mathfrak{I}$ satisfies a weakened version; see Assumption III.8 below and its consequences in Proposition III.5.15 (to be compared with items (iv)-(vi) from Proposition III.5.9). The reducible case is more delicate and it is studied in more details in Section 7 of Chapter IV for the loss function $\mathrm{L} = \mathfrak{R}_e$; in particular Assumption III.7 may not hold and the anti-Pareto frontier may not be connected.

We represent in Figure III.3 the typical Pareto and anti-Pareto frontiers for the loss $\mathfrak{R}_e[\mathbf{k}]$ and $\mathfrak{I}[k,\gamma]$ with a continuous decreasing cost function, when the kernel $\mathbf{k}$ and k have only one irreducible component, that is are monatomic (see definition in the next section).



(A) The loss function $\mathrm{L} = \mathfrak{R}_e$ for the kernel model with $\mathbf{k}$ monatomic.

(B) The loss function $\mathrm{L} = \mathfrak{I}$ for the SIS model with k (or k/γ) monatomic and $c_0 = C^\star(0+)$.

Figure III.3: Typical shape of the Pareto frontier $\mathcal{F}$ in solid red line, the anti-Pareto frontier $\mathcal{F}^{\text{Anti}}$ in red dashed line and the feasible region $\mathbf{F}$ in light blue for monatomic cases with continuous decreasing cost function C and loss functions $\mathfrak{R}_e$ or $\mathfrak{I}$.

Irreducible, quasi-irreducible and monatomic kernels

We first define *irreducible* and *monatomic* kernels. For $A, B \in \mathscr{F}$, we write $A \subset B$ a.s. if $\mu(B^\complement \cap A) = 0$ and $A = B$ a.s. if a.s. $A \subset B$ and $B \subset A$. For $A, B \in \mathscr{F}$, $x \in \Omega$ and an integrable kernel $\mathbf{k}$, we simply write $\mathbf{k}(x, A) = \int_A \mathbf{k}(x, y)\, \mu(\mathrm{d}y)$, $\mathbf{k}(B, x) = \int_B \mathbf{k}(z, x)\, \mu(\mathrm{d}z)$ and:

$$\mathbf{k}(B, A) = \int_{B\times A} \mathbf{k}(z, y)\, \mu(\mathrm{d}z)\mu(\mathrm{d}y).$$

A set $A \subset \mathscr{F}$ is $\mathbf{k}$-*invariant*, or simply *invariant* when there is no ambiguity on the kernel $\mathbf{k}$, if $\mathbf{k}(A^\complement, A) = 0$. If $\mathbf{k}$ is symmetric, then A is invariant if and only if $A^\complement$ is invariant. In the epidemiological setting, the set A is invariant if the subpopulation A does not infect the subpopulation $A^\complement$.

A kernel $\mathbf{k}$ is *irreducible* or *connected* if any $\mathbf{k}$-invariant set A is such that a.s. $A = \emptyset$ or a.s. $A = \Omega$. According to [136, Theorem V.6.6], if $\mathbf{k}$ is an irreducible kernel with finite double norm, then we have $\mathfrak{R}_0[\mathbf{k}] > 0$. If the kernel is a.s. positive, then it is irreducible. Following [17, Definition 2.11], we say that a kernel is *quasi-irreducible* if $\mathbf{k}$ restricted to $\{\mathbf{k} \equiv 0\}^\complement$, with $\{\mathbf{k} \equiv 0\} = \{x \in \Omega \,:\, \mathbf{k}(x, \Omega) + \mathbf{k}(\Omega, x) = 0\}$, is irreducible. The quasi-irreducible property was introduced for symmetric kernel; for general kernel one can consider the following weaker property. A kernel $\mathbf{k}$ is *monatomic* if the operator $T_\mathbf{k}$ has a unique (up to a multiplicative constant) non-negative eigenfunction. Intuitively, this corresponds to have only one irreducible component. Formally, this is also equivalent to the following two properties:

(i) There exists a measurable subset $\Omega_\mathrm{a} \subset \Omega$, the irreducible component or *atom* such that:

- $\mu(\Omega_\mathrm{a}) > 0$ and the kernel $\mathbf{k}$ restricted to Ω_a is irreducible.
- If a.s. $\Omega_\mathrm{a}^\complement \neq \emptyset$ then the restriction of $T_\mathbf{k}$ to $\Omega_\mathrm{a}^\complement$ is quasi-nilpotent, that is, $\mathfrak{R}_e[\mathbf{k}](\mathbb{1}_{\Omega_\mathrm{a}^\complement}) = 0$.

(ii) There exists a measurable subset $\Omega_\mathrm{i} \subset \Omega_\mathrm{a}^\complement$, “the subpopulation infected by” Ω_a such that:

- The sets $\Omega_\mathrm{a} \cup \Omega_\mathrm{i}$ and Ω_i are invariant.
- The set Ω_i is the minimal set such that $\Omega_\mathrm{a} \cup \Omega_\mathrm{i}$ is invariant: if A is invariant and $\Omega_\mathrm{a} \subset A$ then a.s. $\Omega_\mathrm{i} \subset A$.

In the epidemiological setting, the subpopulation Ω_i can only infect itself, and the subpopulation Ω_a infects only itself and Ω_i. We refer to [138] for further details on the decomposition of a kernel on its irreducible components; in particular the sets Ω_a and Ω_i are unique up to the a.s. equivalence.

The following result explains the usefulness of monatomicity: it implies in particular that the effective spectral radius of a vaccination only depends on its value on the atom.

Lemma III.5.11 (Reduction to the atom)**.** *Assume that the kernel* $\mathbf{k}$ *is monoatomic with atom* Ω_a*. Let* $\mathbf{k}_a$ *be the restriction of* $\mathbf{k}$ *to* $\Omega_\mathrm{a} \times \Omega_\mathrm{a}$*. Then, for any* $\eta \in \Delta$*,*

$$\mathfrak{R}_e[\mathbf{k}](\eta) = \mathfrak{R}_e[\mathbf{k}_\mathrm{a}](\eta_\mathrm{a}),$$

where η_a *is the restriction of* η *to* Ω_a*.*

Proof. We deduce from [138, Lemma 11] that if $A, B \in \mathscr{F}$ are such that $A \cap B = \emptyset$ a.s. and $\mathbf{k}'(B, A) = 0$ for a kernel $\mathbf{k}'$ with finite double norm, then the spectral radius of the operator $T_{\mathbf{k}'}$ restricted to $A \cup B$ is the maximum of the spectral radii of the operator $T_{\mathbf{k}'}$ restricted to A and to B:

$$\mathfrak{R}_e[\mathbf{k}'](\mathbb{1}_{A\cup B}) = \max\left(\mathfrak{R}_e[\mathbf{k}'](\mathbb{1}_A), \mathfrak{R}_e[\mathbf{k}'](\mathbb{1}_B)\right). \tag{III.49}$$

Set $\Omega' = (\Omega_\mathrm{a} \cup \Omega_\mathrm{i})^\mathsf{c}$. We deduce that:

$$\mathfrak{R}_e[\mathbf{k}](\eta) = \max\left(\mathfrak{R}_e[\mathbf{k}\eta](\mathbb{1}_{\Omega_\mathrm{a}}), \mathfrak{R}_e[\mathbf{k}\eta](\mathbb{1}_{\Omega_\mathrm{a}}), \mathfrak{R}_e[\mathbf{k}\eta](\mathbb{1}_{\Omega'})\right) = \mathfrak{R}_e[\mathbf{k}\eta](\mathbb{1}_{\Omega_\mathrm{a}}),$$

where we used (III.49) once with $A = B^\mathsf{c} = \Omega_\mathrm{a} \cup \Omega_\mathrm{i}$, which is invariant by definition, and once with $A = \Omega_\mathrm{i}$, which is invariant by definition, and $B = \Omega_\mathrm{a}$ for the first equality, and that the restriction of $\mathbf{k}$ to $\Omega_\mathrm{a}^\mathsf{c} = \Omega_\mathrm{i} \cup \Omega'$ is quasi-nilpotent. We deduce that:

$$\mathfrak{R}_e[\mathbf{k}](\eta) = \mathfrak{R}_e[\mathbf{k}_\mathrm{a}](\eta_\mathrm{a}),$$

where $\mathbf{k}_\mathrm{a}$ is the kernel $\mathbf{k}$ restricted to Ω_a and η_a is the restriction of η to Ω_a. □

Remark III.5.12. Irreducible and quasi-irreducible kernels are also monatomic (take $\Omega_\mathrm{a} = \{\mathbf{k} \equiv 0\}^c$ and $\Omega_\mathrm{i} = \emptyset$). If the kernel $\mathbf{k}$ is monatomic and symmetric, then we get $\mathbf{k} = \mathbb{1}_{\Omega_\mathrm{a}}\, \mathbf{k}\, \mathbb{1}_{\Omega_\mathrm{a}}$ and thus the kernel $\mathbf{k}$ is quasi-irreducible.

The notion of irreducibility of a kernel depends only on its support: the kernel $\mathbf{k}$ is irreducible (resp. quasi-irreducible, resp. monatomic) if and only if the kernel $\mathbb{1}_{\{\mathbf{k}>0\}}$ is irreducible (resp. quasi-irreducible, resp. monatomic). Furthermore, if $\mathbf{k}$ is monatomic, then the kernels $\mathbf{k}$ and $\mathbb{1}_{\{\mathbf{k}>0\}}$ have the same atom Ω_a and the same set Ω_i infected by Ω_a.

In the epidemiological setting, the support of the endemic equilibrium in the supercritical regime ($\mathfrak{R}_0 > 1$) for a monatomic kernel is given by the atom and the subpopulation it infects, that is, $\Omega_\mathrm{a} \cup \Omega_\mathrm{i}$. A very similar proof yields that $\Omega_\mathrm{a} \cup \Omega_\mathrm{i}$ is also the support of the non-negative right Perron eigenfunction of the operator $T_{k/\gamma}$ (or of the operator $T_{\gamma^{-1}k}$).

Lemma III.5.13 (Support of the equilibrium in the monatomic case)**.** *Consider the supercritical SIS model* $\mathrm{Param} = [(\Omega, \mathscr{F}, \mu), k, \gamma]$ *under Assumption III.2 (that is,* $\mathfrak{R}_0[k/\gamma] > 1$*). If* k *is monatomic with atom* Ω_a *and* Ω_i *the smallest subpopulation infected by* Ω_a*, then the maximal equilibrium* $\mathfrak{g}$ *is the only non-zero equilibrium, and the following equality is satisfied up to a zero-measure set:*

$$\{\mathfrak{g} > 0\} = \Omega_\mathrm{a} \cup \Omega_\mathrm{i}.$$

Proof. Let h be a (non-zero) equilibrium. We recall from (III.18) with $F(h) = 0$ that:

$$\frac{h}{1-h} = \gamma^{-1} \int_\Omega k(\cdot, y)\, h(y)\, \mu(\mathrm{d}y). \tag{III.50}$$

In a first step, we prove that $\{h > 0\} \subset \Omega_\mathrm{a} \cup \Omega_\mathrm{i}$. Let A be an invariant subset. Since h is an equilibrium, we deduce using (III.50) that h restricted to A^c is also an equilibrium of the SIS model restricted to A^c. As k restricted to $\Omega' = (\Omega_a \cup \Omega_i)^c$ is quasi-nilpotent, we get that the only equilibrium of the SIS model restricted to Ω' is zero. Since $\Omega_a \cup \Omega_i$ is invariant, we deduce that $h = 0$ on Ω' and thus $\{h > 0\} \subset \Omega_a \cup \Omega_i$.

Let us prove the reverse inclusion. Since $k(\Omega_a, \Omega_i) = 0$ as Ω_i is invariant, we deduce from (III.50) that h restricted to the atom Ω_a, say h_a, is an equilibrium of the SIS model restricted to Ω_a. Since k restricted to Ω_a is irreducible, the function h_a is either identically zero, or positive everywhere. By definition of Ω_i and using (III.50), we deduce that if $h_a = 0$, then the restriction of h to Ω_i, say h_i, is an equilibrium of the SIS model restricted to Ω_i. Since k is quasi-nilpotent on Ω_i, we deduce that $h_i = 0$, and thus $h = 0$, which is absurd by hypothesis. Therefore h_a is positive, thus $\Omega_a \subset \{h > 0\}$.

Let us now check that $\Omega_i \subset \{h > 0\}$. We deduce from (III.50) that for $x \in \{h = 0\}$ we have $\int_\Omega k(x, y)\, h(y)\, \mu(\mathrm{d}y) = 0$ and thus $k(x, \{h > 0\}) = 0$. This implies that $\{h > 0\}$ is invariant. Since $\Omega_a \cup \Omega_i$ is by definition the smallest invariant set containing Ω_a and since $\Omega_a \subset \{h > 0\}$, we must have $\Omega_i \subset \{h > 0\}$. In conclusion, we have obtained that the support of any non-zero equilibrium h is almost surely equal to $\Omega_a \cup \Omega_i$.

Finally let us prove that $\mathfrak{g}$ is the only non-zero equilibrium. Let h be another non-zero equilibrium. The previous points show that h_a and $\mathfrak{g}_a$, the restrictions of h and $\mathfrak{g}$ to the atom Ω_a, are positive equilibria for the SIS model restricted to Ω_a; since this model is irreducible, they are equal, so $h = \mathfrak{g}$ on Ω_a. Now, using Lemma III.5.11, for the first and third equalities, and Lemma III.4.4, point (iii) for the last one, we get:

$$\mathfrak{R}_e[\mathbf{k}](1 - h) = \mathfrak{R}_e[\mathbf{k}_a](1 - h_a) = \mathfrak{R}_e[\mathbf{k}_a](1 - \mathfrak{g}_a) = \mathfrak{R}_e[\mathbf{k}](1 - \mathfrak{g}) = 1.$$

By Lemma III.4.4, point (ii), this implies that h is the maximal equilibrium $\mathfrak{g}$. □

The kernel model

We now check Assumptions III.5 and III.7 for the loss $\mathrm{L} = \mathfrak{R}_e$.

Lemma III.5.14. *Consider the kernel model* $\mathrm{Param} = [(\Omega, \mathscr{F}, \mu), \mathbf{k}]$ *under Assumption III.1 with the loss* $\mathrm{L} = \mathfrak{R}_e[\mathbf{k}]$ *and* $\ell_{\max} = \mathfrak{R}_0[\mathbf{k}]$.

(i) Assumption III.5 holds, and if $\mathfrak{R}_0[\mathbf{k}] > 0$*, then the part of Assumption III.3 on the loss holds.*

(ii) If $\mathbf{k}$ *is monatomic with atom* Ω_a*, then Assumption III.7 holds, and we have* $\mathfrak{R}_0[\mathbf{k}] > 0$ *and* $C^\star(\ell_{\max}) = C(\mathbb{1}_{\Omega_a})$ *(which is 0 if* $\mathbf{k}$ *is irreducible).*

Proof. Since the loss $\mathrm{L} = \mathfrak{R}_e[\mathbf{k}]$ is homogeneous according to Proposition III.4.1, we deduce from Lemma III.5.6 that Assumption III.5 holds. Using Theorem III.4.2, for the continuity, and Proposition III.4.1, for the monotonicity of the function $\mathfrak{R}_e[\mathbf{k}]$, and the fact that $\mathfrak{R}_0[\mathbf{k}] > 0$, the hypotheses on the loss in Assumption III.3 hold.

To prove that Assumption III.7 holds, we first assume that the kernel $\mathbf{k}$ is irreducible. In particular, we have $\mathbf{k}(\Omega, y) > 0$ almost surely. Let $\eta \in \Delta$ be a local maximum; we want to show that it is also a global maximum.

Suppose first that $\inf \eta > 0$. Then $\mathbf{k}\eta$ is irreducible with finite double norm. According [136, Theorem V.6.6 and Example V.6.5.b], the eigenspace of $T_{\mathbf{k}\eta}$ associated to $\mathfrak{R}_e(\eta)$ is one-dimensional and it is spanned by a vector v_d such that $v_d > 0$ almost surely, and the corresponding left eigenvector associated to $\mathfrak{R}_e(\eta)$, say v_g, can be chosen such that $\langle v_g, v_d \rangle = 1$ and $v_g > 0$ almost surely. According to [98, Theorem 2.6], applied to $L_0 = T_{\mathbf{k}\eta}$ and $L = T_{\mathbf{k}(\eta + \varepsilon(1-\eta))}$ with $\varepsilon \in (0, 1)$, we have, using that $\| L_0 - L \| = O(\varepsilon)$ thanks to (III.12):

$$\mathfrak{R}_e(\eta + \varepsilon(1 - \eta)) = \mathfrak{R}_e(\eta) + \varepsilon \langle v_g, T_{\mathbf{k}(1-\eta)} v_d \rangle + O(\varepsilon^2).$$

Since $\mathfrak{R}_e$ has a local maximum at η, the first order term on the right hand side vanishes, so $v_g(x)\mathbf{k}(x, y)(1 - \eta(y))v_d(y) = 0$ for μ almost all x and y. Since v_g and v_d are positive almost surely and $\mathbf{k}$ is irreducible, we get that $\mathbf{k}(\Omega, y)(1 - \eta(y)) = 0$ almost surely and thus $\eta(y) = 1$ almost surely. Therefore $\eta = \mathbb{1}$, which is a global maximum for $\mathfrak{R}_e$.

Finally, suppose that $\inf \eta = 0$. Let $\mathcal{O}$ be an open subset of Δ on which $\mathfrak{R}_e \leq \mathfrak{R}_e(\eta)$ and with $\eta \in \mathcal{O}$. For $\varepsilon > 0$ small enough, the strategy $\eta_\varepsilon = \eta + \varepsilon(1-\eta)$ belongs to $\mathcal{O}$ and satisfies $\mathfrak{R}_e(\eta) \leq \mathfrak{R}_e(\eta_\varepsilon) \leq \mathfrak{R}_e(\eta)$ (where the first inequality comes from the fact that $\mathfrak{R}_e$ is non-decreasing). Therefore η_ε is a local maximum with $\inf \eta_\varepsilon \geq \varepsilon$, and thus, thanks to the first part of the proof, $\eta_\varepsilon = \mathbb{1}$. This readily implies that $\eta = \mathbb{1}$.

We deduce that if η is a local maximum, then $\eta = \mathbb{1}$. Thus η is a global maximum and $C^\star(\ell_{\max}) = C(\mathbb{1}) = 0$. This ends the proof for the irreducible case.

To treat the monatomic case, recall that for any η, we know by Lemma III.5.11 that

$$\mathfrak{R}_e[\mathbf{k}](\eta) = \mathfrak{R}_e[\mathbf{k}_\mathrm{a}](\eta_\mathrm{a}),$$

where η_a is the restriction of η to the atom Ω_a. Since by hypothesis $\mathbf{k}_\mathrm{a}$ is irreducible, we deduce from [136, Theorem V.6.6] that $\mathfrak{R}_0[\mathbf{k}_\mathrm{a}] > 0$ and thus $\mathfrak{R}_0[\mathbf{k}] > 0$. Furthermore, if η is a local maximum for $\mathfrak{R}_e[\mathbf{k}]$, then η_a is a local maximum for $\mathfrak{R}_e[\mathbf{k}_\mathrm{a}]$. Since by hypothesis $\mathbf{k}_\mathrm{a}$ is irreducible, we deduce from the first part of the proof that $\eta_\mathrm{a} = \mathbb{1}_\mathrm{a}$, and thus $\eta \geq \mathbb{1}_{\Omega_\mathrm{a}}$ as well as $\mathfrak{R}_e[\mathbf{k}](\eta) \geq \mathfrak{R}_e[\mathbf{k}](\mathbb{1}_{\Omega_\mathrm{a}}) = \mathfrak{R}_e[\mathbf{k}](\mathbb{1})$. Thus, the strategy η is a global maximum. This implies that Assumption III.7 holds.

Use that $\mathbb{1}_\mathrm{a}$, the unity function defined on Ω_a, is the only global maximum of $\mathfrak{R}_e[\mathbf{k}_\mathrm{a}]$ thanks to the first part of the proof, to deduce that η is a global maximum of $\mathfrak{R}_e[\mathbf{k}]$ if and only if a.s. $\eta \geq \mathbb{1}_{\Omega_\mathrm{a}}$. We deduce that $C^\star(\ell_{\max}) = C(\mathbb{1}_{\Omega_\mathrm{a}})$. □

The SIS model

The loss $\mathrm{L} = \mathfrak{I}$ does not satisfies Assumption III.7 in general even when the kernel $\mathbf{k}$ is irreducible. Indeed, by continuity of $\mathfrak{R}_e$, there exists a (weakly) open neighborhood $\mathcal{O}$ of $\mathbb{0}$ such that $\mathfrak{R}_e(\eta) < 1$ for all $\eta \in \mathcal{O}$: consequently $\mathfrak{I}$ is identically zero on $\mathcal{O}$, and any $\eta \in \mathcal{O}$ is a local maximum of $\mathrm{L} = \mathfrak{I}$. However, these maxima are not global in the supercritical case where $\mathfrak{I}(\mathbb{1}) > 0$. For this reason, we shall consider the following variant of Assumption III.7, where one does not consider the zeros of the loss.

Assumption III.8. *If the loss* L *has a local maximum at* η *(for the weak topology) and* $\mathrm{L}(\eta) > 0$, *then* $\mathrm{L}(\eta) = \ell_{\max}$ *and* η *is a global maximum of* L.

We set:

$$c_0 = C^\star(0+) = \lim_{\substack{\ell \to 0 \\ \ell > 0}} C^\star(\ell).$$

Under Assumption III.7, we have $c_0 = c_{\max}$. If Assumption III.8 holds and $c_0 = c_{\max}$, then Assumption III.7 holds; so we only need to consider Assumption III.8 with $c_0 < c_{\max}$. Eventually, if $c_0 < c_{\max}$, then we have $\mathrm{L}^\star(c_{\max}) = 0$.

Considering Assumption III.8 with $c_0 < c_{\max}$ instead of Assumption III.7 impacts only items (iv), (v), (vi) and the conclusion of Proposition III.5.9 as follows; we refer to Figure IV.3(A) for an illustration, and leave the proof to the reader.

Proposition III.5.15. *Under Assumptions III.3 and III.8 and* $c_0 < c_{\max}$, *the following properties hold:*

(iv) The optimal loss $\mathrm{L}^\star$ *is decreasing on* $[C^\star(\ell_{\max}), c_0]$ *and zero on* $(c_0, c_{\max}]$.

(v) If η *solves Problem* (III.41) *for the cost* $c \in [C^\star(\ell_{\max}), c_0)$, *then* $C(\eta) = c$. *Moreover* η *is anti-Pareto optimal, and* $C^\star(\mathrm{L}^\star(c)) = c$.

(vi) The anti-Pareto frontier is also given by:

$$\mathcal{F}_{\mathrm{L}}^{\mathrm{Anti}} = \{(c, \mathrm{L}^\star(c)) : c \in [C^\star(\ell_{\max}), c_0)\} \cup \{(c_{\max}, 0)\}. \tag{III.51}$$

If furthermore Assumption III.6 holds, then the function $\mathrm{L}^\star$ *is a continuous decreasing bijection of* $[C^\star(\ell_{\max}), c_0)$ *onto* $(0, \ell_{\max}]$, $C^\star$ *is the inverse bijection, and the union of the anti-Pareto frontier with its limit point* $\{(c_0, 0)\}$ *is compact but not connected.*

We are now ready to check that Assumption III.8 holds for the loss $\mathrm{L} = \mathfrak{I}$ when the kernel k is monatomic.

Lemma III.5.16. *Consider the SIS model* $\mathrm{Param} = [(\Omega, \mathscr{F}, \mu), k, \gamma]$ *under Assumption III.2 with the loss* $\mathrm{L} = \mathfrak{I}$ *and* $\ell_{\max} = \mathfrak{I}(\mathbb{1})$.

(i) Assumption III.5 holds, and if $\mathfrak{I}(\mathbb{1}) > 0$, *then the part of Assumption III.3 on the loss holds.*

(ii) If k *is monatomic and* $\mathfrak{I}(\mathbb{1}) > 0$, *then Assumption III.8 holds for* $\mathrm{L} = \mathfrak{I}$. *Moreover,* $C^\star(\ell_{\max}) = C(\mathbb{1}_{\{\mathfrak{g}>0\}})$, *where* $\mathfrak{g}$ *is the maximal equilibrium, and* $\ell_{\max} = \mathfrak{I}(\mathbb{1})$ *(we also have that* $C^\star(\ell_{\max}) = 0$ *if* k *is irreducible). Furthermore, if the cost function* C *is decreasing, then we have* $c_0 < c_{\max}$.

Proof. We prove Point (i). Since the loss $\mathrm{L} = \mathfrak{I}$ is sub-homogeneous, see Proposition III.4.5, we deduce from Lemma III.5.6 that Assumption III.5 holds. Using Theorem III.4.6 (for the continuity), Proposition III.4.5 (for the monotonicity of the function $\mathfrak{I}$) and the fact that $\mathfrak{I}(\mathbb{1}) > 0$, the hypotheses on the loss in Assumption III.3 hold.

We now prove Point (ii). Assume that $\mathfrak{I}(\mathbb{1}) > 0$ and set $\mathbf{k} = k/\gamma$. Since $\{\mathbf{k} > 0\} = \{k > 0\}$, and k is monatomic, we deduce that $\mathbf{k}$ is monatomic with the same atom Ω_a and same smallest subpopulation Ω_i infected by Ω_a. Let $\mathfrak{g}$ be the maximal equilibrium. It is non-zero as $\mathfrak{I}(\mathbb{1}) > 0$. Suppose that $\mathfrak{I}$ has a local maximum at some $\eta \in \Delta$ and $\mathfrak{I}(\eta) > 0$. For $\varepsilon \in (0, 1)$, set $\eta_\varepsilon = \eta + \varepsilon(1-\eta)$. We have that for $\varepsilon > 0$ small enough:

$$\mathfrak{I}(\eta) \geq \mathfrak{I}(\eta_\varepsilon) = \int_\Omega \mathfrak{g}_{\eta_\varepsilon}\, \eta_\varepsilon \,\mathrm{d}\mu \geq \int_\Omega \mathfrak{g}_{\eta_\varepsilon}\, \eta \,\mathrm{d}\mu \geq \int_\Omega \mathfrak{g}_{\eta}\, \eta \,\mathrm{d}\mu = \mathfrak{I}(\eta), \tag{III.52}$$

where we used that $\eta \leq \eta_\varepsilon$ and $0 \leq \mathfrak{g}_\eta \leq \mathfrak{g}_{\eta_\varepsilon}$, see (III.33). Therefore all these quantities are equal. Since $\{k\eta_\varepsilon > 0\} = \{k > 0\}$, we deduce that $\mathbf{k}\eta_\varepsilon$ is monatomic with atom Ω_a and the smallest subpopulation Ω_i infected by Ω_a. We deduce from Lemma III.5.13, that $\{\mathfrak{g}_{\eta_\varepsilon} > 0\} = \Omega_\mathrm{a} \cup \Omega_\mathrm{i}$. We deduce from (III.52), as all the inequalities are equalities, that $\eta_\varepsilon = \eta$ on $\Omega_\mathrm{a} \cup \Omega_\mathrm{i}$, and thus $\eta \geq \mathbb{1}_{\Omega_\mathrm{a}\cup\Omega_\mathrm{i}}$.

Recall from (III.33) that $\mathfrak{g}_\eta \leq \mathfrak{g}$. and from Lemma III.5.13 that $\{\mathfrak{g} > 0\} = \Omega_\mathrm{a} \cup \Omega_\mathrm{i}$. So $\mathfrak{g}_\eta$ is zero outside $\Omega_\mathrm{a} \cup \Omega_\mathrm{i}$, and we deduce that changing the value of η outside $\Omega_\mathrm{a} \cup \Omega_\mathrm{i}$ does not affect the value of $\mathfrak{I}(\eta)$. In conclusion, $\eta \in \Delta$ is a local maximum such that $\mathfrak{I}(\eta) > 0$ if and only if $\eta \geq \mathbb{1}_{\Omega_\mathrm{a}\cup\Omega_\mathrm{i}}$, and thus is a global maximum. We deduce that $C^\star(\ell_{\max}) = C(\mathbb{1}_{\{\mathfrak{g}>0\}})$.

Eventually, notice that the set $\{\eta \in \Delta : \mathfrak{R}_e[\mathbf{k}](\eta) < 1\}$ is an open neighborhood of $\mathbb{0}$ on which $\mathfrak{I}$ is zero. If the cost function C is decreasing, this implies there exists $\eta \in \Delta$ such that $C(\eta) < c_{\max}$ and $\mathfrak{I}(\eta) = 0$ and thus $c_0 < c_{\max}$. □

III.6 Miscellaneous properties of the feasible region and the Pareto frontier

We prove results concerning the feasible region, the stability of the Pareto frontier and its geometry.

III.6.1 The feasible region

In the following proposition we check a number of topological properties of the set of outcomes $\mathbf{F} = \{(C(\eta), \mathrm{L}(\eta)), \eta \in \Delta\}$.

Proposition III.6.1 (No hole in the feasible region)**.** *Suppose that Assumption III.3 holds. The feasible region* $\mathbf{F}$ *is compact, path connected, and its complement is connected in* $\mathbb{R}^2$. *It is the whole region between the graphs of the one-dimensional value functions:*

$$\begin{aligned}\mathbf{F} &= \{(c, \ell) \in \mathbb{R}^2 \ :\ 0 \leq c \leq c_{\max},\ \mathrm{L}_\star(c) \leq \ell \leq \mathrm{L}^\star(c)\}\\ &= \{(c, \ell) \in \mathbb{R}^2 \ :\ 0 \leq \ell \leq \ell_{\max},\ C_\star(\ell) \leq c \leq C^\star(\ell)\}.\end{aligned}$$

Proof. The region $\mathbf{F}$ is compact and path-connected as a continuous image by (C, L) of the compact, path-connected set Δ.

By symmetry, it is enough to prove that $\mathbf{F}$ is equal to $F_1 = \{(c, \ell) \in \mathbb{R}^2 \; : \; 0 \le c \le c_{\max}, \, \mathrm{L}_\star(c) \le \ell \le \mathrm{L}^\star(c)\}$. Let $(c, \ell) \in \mathbf{F}$ and $\eta \in \Delta$ be such that $(c, \ell) = (C(\eta), \mathrm{L}(\eta))$. By definition of $\mathrm{L}_\star$ and $\mathrm{L}^\star$, we have: $\mathrm{L}_\star(c) = \mathrm{L}_\star(C(\eta)) \le \mathrm{L}(\eta) \le \mathrm{L}^\star(C(\eta)) = \mathrm{L}^\star(c)$. We deduce that $(c, \ell) \in F_1$.

Let us now prove that $F_1 \subset \mathbf{F}$. Let us first consider a point of the form $(c, \mathrm{L}_\star(c))$, where $0 \le c \le c_{\max}$. By definition, there exists η such that $C(\eta) \le c$ and $\mathrm{L}(\eta) = \mathrm{L}_\star(c)$. Let $\eta_t = t\eta$. The map $t \mapsto C(\eta_t)$ is continuous from $[0, 1]$ to $[C(\eta), c_{\max}]$, and $c \in [C(\eta), c_{\max}]$, so there exists s such that $C(\eta_s) = c$. Since L is non-decreasing, $\mathrm{L}(\eta_s) \le \mathrm{L}(\eta)$. By definition of $\mathrm{L}_\star(c)$, $\mathrm{L}(\eta_s) \ge \mathrm{L}_\star(c)$. Therefore $(c, \mathrm{L}_\star(c)) = (C(\eta_s), \mathrm{L}(\eta_s))$ belongs to $\mathbf{F}$. Similarly the graphs of $C_\star$, $C^\star$ and $\mathrm{L}^\star$ are also included in $\mathbf{F}$.

So, it is enough to check that, if $A = (c, \ell)$ is in F_1, with $c \in (0, c_{\max})$ and $\ell \in (\mathrm{L}_\star(c), \mathrm{L}^\star(c))$, then A belongs to $\mathbf{F}$. We shall assume that $A \notin \mathbf{F}$ and derive a contradiction by building a loop in $\mathbf{F}$ that encloses A and which can be continuously contracted into a point in $\mathbf{F}$.

Since $\mathrm{L}_\star(c) < \ell < \mathrm{L}^\star(c)$, there exist η_{SO} and η_{NE} such that:

$$C(\eta_{\mathrm{SO}}) \le c, \quad \mathrm{L}(\eta_{\mathrm{SO}}) < \ell, \quad C(\eta_{\mathrm{NE}}) \ge c \quad \text{and} \quad \mathrm{L}(\eta_{\mathrm{NE}}) > \ell.$$

We concatenate the four paths defined for $u \in [0, 1]$:

$$u \mapsto u\eta_{\mathrm{SO}}, \quad u \mapsto (1-u)\eta_{\mathrm{SO}} + u, \quad u \mapsto (1-u) + u\eta_{\mathrm{NE}} \quad \text{and} \quad u \mapsto (1-u)\eta_{\mathrm{NE}},$$

to obtain a continuous loop $(\eta_t, t \in [0, 4])$ from $[0, 4]$ to Δ, such that:

$$(\eta_0, \eta_1, \eta_2, \eta_3, \eta_4) = (\mathbb{0}, \eta_{\mathrm{SO}}, \mathbb{1}, \eta_{\mathrm{NE}}, \mathbb{0}).$$

We now define a continuous family of loops $(\gamma_s, s \in [0, 1])$ in $\mathbb{R}^2$ by

$$\gamma_s(t) = (C(s\eta_t), \mathrm{L}(s\eta_t), t \in [0, 4]).$$

By definition, for all $s \in [0, 1]$, γ_s is a continuous loop in F. Since $A = (c, \ell) \notin \mathbf{F}$, the loops γ_s do not contain A, so the winding number $W(\gamma_s, A)$ is well-defined (see for example [88, Definition 6.1]). As $A \notin \mathbf{F}$, we get that γ_s is a continuous deformation in $\mathbb{R}^2 \setminus \{A\}$ from γ_1 to γ_0. Thanks to [88, Theorem 6.5], this implies that $W(\gamma_s, A)$ does not depend on $s \in [0, 1]$.

For $s = 0$, the loop degenerates to the single point $(C(0), 0)$ so the winding number is 0. For $s = 1$, let us check that the winding number is 1, which will provide the contradiction. To do this, we compare γ_1 with a simpler loop δ defined by:

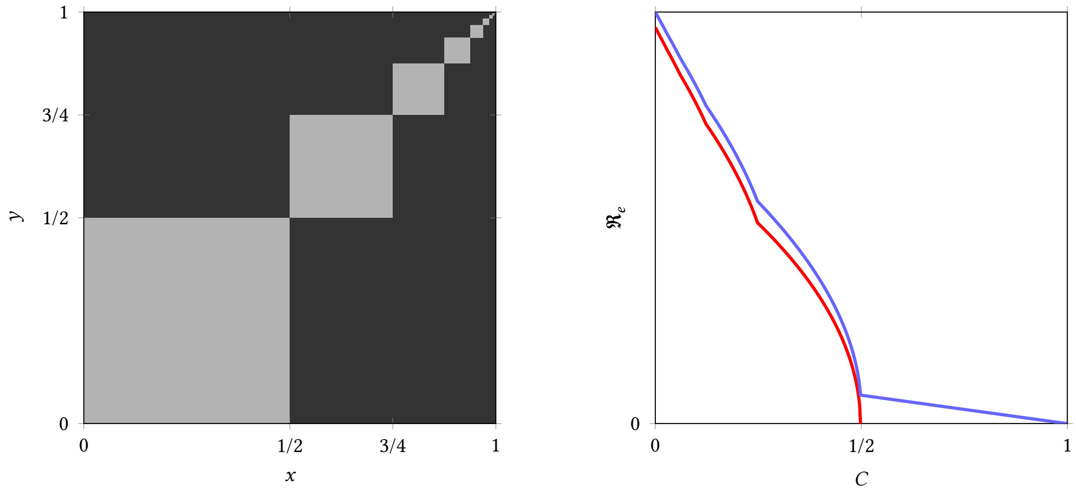
$$\delta(0) = \delta(4) = (c_{\max}, 0), \quad \delta(1) = (0, 0), \quad \delta(2) = (0, \ell_{\max}) \quad \text{and} \quad \delta(3) = (c_{\max}, \ell_{\max}),$$

and by linear interpolation for non integer values of t: in other words, δ runs around the perimeter of the axis-aligned rectangle with corners $(0, 0)$ and $(c_{\max}, \ell_{\max})$. Clearly, we have $W(\delta, A) = 1$.

Let M_t, N_t denote $\gamma_1(t)$ and $\delta(t)$ respectively. For $t \in [0, 1]$, we have $N_t = ((1-t)c_{\max}, 0)$, so the second coordinate of $\overrightarrow{AN_t}$ is non-positive. On the other hand $\mathrm{L}(t\eta_{\mathrm{SO}}) \le \mathrm{L}(\eta_{\mathrm{SO}}) < \ell$, so the second coordinate of $\overrightarrow{AM_t}$ is negative. Therefore the two vectors $\overrightarrow{AN_t}$ and $\overrightarrow{AM_t}$ cannot point in opposite directions. Similar considerations for the other values of $t \in [1, 4]$ show that $\overrightarrow{AN_t}$ and $\overrightarrow{AM_t}$ never point in opposite directions. By [88, Theorem 6.1], the winding numbers $W(\gamma_1, A)$ and $W(\delta, A)$ are equal, and thus $W(\gamma_1, A) = 1$.

This gives that $A \in \mathbf{F}$ by contradiction, and thus $F_1 \subset \mathbf{F}$.

Finally, it is easy to check that F_1 has a connected complement, because F_1 is bounded, and all the points in F_1^{c} can reach infinity by a straight line: for example, if $\ell > \mathrm{L}^\star(c)$, then the half-line $\{(c, \ell'), \ell' \ge \ell\}$ is in F_1^{c}. □



(A) Grayplot of the kernel $\mathbf{k}_\varepsilon$, with $\Omega = [0,1]$ and μ the Lebesgue measure ($\mathbf{k}_\varepsilon$ is equal to the constant $\kappa > 0$ on the black zone and to ε on the gray zone, with $\varepsilon > 0$ small).

(B) In red, the Pareto frontier of the kernel $\mathbf{k}$ represented in Figure III.1(A) compared to the Pareto frontier of the kernel $\mathbf{k}_\varepsilon$ in blue.

Figure III.4: On the stability of the Pareto frontier

III.6.2 Stability

We can consider the stability of the Pareto frontier and the set of Pareto optima. Recall that, thanks to (III.45), the graph $\{(c, \mathrm{L}_\star(c)) : c \in [0, c_{\max}]\}$ of $\mathrm{L}_\star$ is the union of the Pareto frontier and the straight line joining $(0, C_\star(0))$ to $(0, c_{\max})$ and can thus be seen as an extended Pareto frontier. The proof of the following proposition is immediate. It implies in particular the convergence of the extended Pareto frontier. This result can also easily be adapted to the anti-Pareto frontier.

Proposition III.6.2. *Let C be a cost function and $(\mathrm{L}^{(n)}, n \in \mathbb{N})$ a sequence of loss functions converging uniformly on Δ to a loss function L. Assume that Assumptions III.3, III.4 and III.5 hold for the cost C and the loss functions $\mathrm{L}^{(n)}$, $n \in \mathbb{N}$, and L. Then $\mathrm{L}_\star^{(n)}$ converges uniformly to $\mathrm{L}_\star$. Let $\eta \in \Delta$ be the weak limit of a sequence $(\eta_n, n \in \mathbb{N})$ of Pareto optima, that is $\eta_n \in \mathcal{P}_{\mathrm{L}^{(n)}}$ for all $n \in \mathbb{N}$. If $C(\eta) \leq C_\star(0)$, then we have $\eta \in \mathcal{P}_{\mathrm{L}}$.*

Remark III.6.3 (On the continuity of the Pareto Frontier). It might happen that some elements of $\mathcal{P}_{\mathrm{L}}$ are not weak limits of any sequences of elements of $\mathcal{P}_{\mathrm{L}^{(n)}}$; see [37] for an example of such a discontinuity. It might also happen that a sequence $(\eta_n, n \in \mathbb{N})$ such that $\eta_n \in \mathcal{P}_{\mathrm{L}^{(n)}}$ and $\mathrm{L}^{(n)}(\eta_n) > 0$ converges to some η that does not belong to $\mathcal{P}_{\mathrm{L}}$ if $\mathrm{L}(\eta) = 0$. In particular, in this case, $C_{\star,\mathrm{L}^{(n)}}(0)$ does not converge to $C_{\star,\mathrm{L}}(0)$, where $C_{\star,\mathrm{L}'}$ is the value function $C_\star$ associated to the loss L'. This situation is represented in Figure III.4. In Figure III.4(A), we have plotted a perturbation $\mathbf{k}_\varepsilon = \mathbf{k} + \varepsilon \sum_{n \in \mathbb{N}^\star} \mathbb{1}_{I_n \times I_n}$ of the multipartite kernel $\mathbf{k}$ defined in Example III.1.7 for $\varepsilon > 0$ small. According to Proposition III.4.3, $\mathfrak{R}_e[\mathbf{k}_\varepsilon]$ converges uniformly to $\mathfrak{R}_e[\mathbf{k}]$ when ε vanishes. However, the Pareto optimal strategies for $\mathbf{k}_\varepsilon$ that cost more than $1/2$ do not converge to some Pareto optimal strategies for $\mathbf{k}$. This can be seen in Figure III.4(B), where the Pareto frontier of $\mathbf{k}_\varepsilon$ (in blue) corresponding to costs larger than 1/2 does not have a counterpart in the Pareto frontier of $\mathbf{k}$ (in red).

III.6.3 Geometric properties

If the cost function is affine, then there is a nice geometric property of the Pareto frontier.

Lemma III.6.4. *Suppose that Assumption III.3 holds, the cost function $C = C_{\mathrm{aff}}$ is given by* (III.35), *with $c_{\mathrm{aff}} \in L^1$ positive, and the loss function L is sub-homogeneous. Then, we have $\mathrm{L}_\star(\theta c + (1 - \theta) c_{\max}) \leq \theta \mathrm{L}_\star(c)$ for all $c \in [0, c_{\max}]$ and $\theta \in [0, 1]$.*

Remark III.6.5. Geometrically, Lemma III.6.4 means that the graph of the loss $\mathrm{L}_\star : [0, c_{\max}] \to [0, \ell_{\max}]$ is below its chords with end point $(1, \mathrm{L}_\star(\mathbb{1})) = (1, 0)$. See Figures III.1(b) for a typical representation of the Pareto frontier (red solid line).

Proof. Let $c \in [0, c_{\max}]$ and $\theta \in [0, 1]$. Thanks to Lemma III.5.6, Assumption III.5 holds. Thus, thanks to Proposition III.5.5 (iv), there exists $\eta \in \mathcal{P}_\mathrm{L}$ with cost $C(\eta) = c$ and thus $\mathrm{L}(\eta) = \mathrm{L}_\star(c)$. Since C is affine, we have:

$$C(\theta\eta) = \theta C(\eta) + (1 - \theta)c_{\max} \leq \theta c + (1 - \theta)c_{\max}.$$

Therefore, $\theta\eta$ is admissible for Problem (III.39) with cost constraint $C(\cdot) \leq \theta c + (1 - \theta)c_{\max}$. This implies that $\mathrm{L}_\star(\theta c + (1 - \theta)c_{\max}) \leq \mathrm{L}(\theta\eta) \leq \theta \mathrm{L}_\star(c)$, thanks to the sub-homogeneity of the loss function L. □

In some cases, see for example Section 5 in Chapter IV, it is possible to prove that the loss function is a convex function of η (which in turn implies Assumption III.5). In this case, choosing a convex cost function implies that Assumption III.4 holds and the Pareto frontier is convex. A similar result holds in the concave case. We provide a short proof of this result.

Proposition III.6.6. *Suppose that Assumption III.3 holds. If the cost function C and the loss function* L *are convex, then the functions $C_\star$ and $\mathrm{L}_\star$ are convex. If the cost function C and the loss function* L *are concave, then the functions $C^\star$ and $\mathrm{L}^\star$ are concave.*

Proof. Let $\ell_0, \ell_1 \in [0, \ell_{\max}]$. By Proposition III.5.3, there exist η_0, η_1 such that $\mathrm{L}(\eta_i) \leq \ell_i$ and $C(\eta_i) = C_\star(\ell_i)$ for $i \in \{0, 1\}$. For $\theta \in [0, 1]$, let $\ell = (1 - \theta)\ell_0 + \theta\ell_1$. Since C and L are assumed to be convex, $\eta = (1 - \theta)\eta_0 + \theta\eta_1$ satisfies:

$$C(\eta) \leq (1 - \theta)C_\star(\ell_0) + \theta C_\star(\ell_1) \quad \text{and} \quad \mathrm{L}(\eta) \leq (1 - \theta)\ell_0 + \theta\ell_1.$$

Therefore, we get that $C_\star((1 - \theta)\ell_0 + \theta\ell_1) \leq C(\eta) \leq (1 - \theta)C_\star(\ell_0) + \theta C_\star(\ell_1)$, and $C_\star$ is convex. The proof of the convexity of $\mathrm{L}_\star$ is similar. The concave case is also similar. □

III.7 Equivalence of models by coupling

III.7.1 Motivation

The aim of this section is to provide examples of different set of parameters for which two kernel or SIS models are "equivalent", in the intuitive sense that their Pareto frontiers are the same (as subsets of $\mathbb{R}_+^2$), and it is possible to map nicely the Pareto optima from one model to the another. In Section III.7.4, we present an example where discrete models can be represented as a continuous models and an example based on measure preserving transformation in the spirit of the graphon theory. We shall consider the two families of models:

- ***the kernel model*** characterized by $\mathrm{Param} = [(\Omega, \mathscr{F}, \mu), \mathbf{k}]$, with Assumption III.1 fulfilled, and loss function $\mathrm{L} = \mathfrak{R}_e$;
- ***the SIS model*** characterized by $\mathrm{Param} = [(\Omega, \mathscr{F}, \mu), k, \gamma]$, with Assumption III.2 fulfilled, and loss function $\mathrm{L} \in \{\mathfrak{R}_e, \mathfrak{I}\}$;

where $(\Omega, \mathscr{F}, \mu)$ is a probability space, $\mathbf{k}$ and k are non-negative kernels on Ω and γ is a non-negative function on Ω.

In order to emphasize the dependence of a quantity H on the parameters Param of the model, we shall write $H[\mathrm{Param}]$ for H. For example we write: $\Delta[\mathrm{Param}]$ for the set of functions $\{\eta \in \mathscr{L}^\infty(\Omega, \mathscr{F}) : 1 \geq \eta \geq 0\}$, which clearly depends on the parameters Param; and the effective reproduction function $\mathfrak{R}_e[\mathrm{Param}]$. For example, under Assumption III.2, we have the equality of the following functions: $\mathfrak{R}_e[(\Omega, \mathscr{F}, \mu), k, \gamma] = \mathfrak{R}_e[(\Omega, \mathscr{F}, \mu), k/\gamma, 1] = \mathfrak{R}_e[(\Omega, \mathscr{F}, \mu), k/\gamma]$, where for the last equality the left hand-side refers to the SIS model and the right hand-side refers to the kernel model (where Assumption III.1 holds as a consequence of Assumption III.2).

If $\inf \gamma > 0$, then, using (III.29) (see Section IV.3 for details and more general results), we also have $\mathfrak{R}_e[(\Omega, \mathscr{F}, \mu), k/\gamma] = \mathfrak{R}_e[(\Omega, \mathscr{F}, \mu), \gamma^{-1}k]$.

Remark III.7.1 (On the cost function). Even if, in full generality, the cost function could also be treated as a parameter, we shall for simplicity consider only the uniform cost C_{uni} given by (III.36) in this section. The interested reader can use Remark III.5.2 for a first generalization to the affine cost function given by (III.35).

III.7.2 On measurability

Let us recall some well-known facts on measurability. Let $(E, \mathscr{E})$ and $(E', \mathscr{E}')$ be two measurable spaces. If $E' = \mathbb{R}$, then we take $\mathscr{E}' = \mathcal{B}(\mathbb{R})$ the Borel σ-field. Let f be a function from E to E'. We denote by $\sigma(f) = \{f^{-1}(A) : A \in \mathscr{E}'\}$ the σ-field generated by f. In particular f is measurable from $(E, \mathscr{E})$ to $(E', \mathscr{E}')$ if and only if $\sigma(f) \subset \mathscr{E}$. Let φ be a measurable function from $(E, \mathscr{E})$ to $(E', \mathscr{E}')$. For ν a measure on $(E, \mathscr{E})$, we write $\varphi_\# \nu$ for the for the push-forward measure on $(E', \mathscr{E}')$ of the measure ν by the function φ (that is $\varphi_\# \nu(A) = \nu(\varphi^{-1}(A))$ for all $A \in \mathscr{E}'$). By definition of $\varphi_\# \nu$, for a non-negative measurable function g defined from $(E', \mathscr{E}')$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, we have:

$$\int_{E'} g \, \mathrm{d}\varphi_\# \nu = \int_E g \circ \varphi \, \mathrm{d}\nu. \tag{III.53}$$

Let f be a measurable function from $(E, \mathscr{E})$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. We recall that:

$$\sigma(f) \subset \sigma(\varphi) \implies f = g \circ \varphi, \tag{III.54}$$

for some measurable function g from $(E', \mathscr{E}')$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

The random variables we consider are defined on a probability space, say $(\Omega_0, \mathscr{F}_0, \mathbb{P})$.

III.7.3 Coupled models

We refer the reader to [90] for a similar development in the graphon setting. We first define coupled models in the next definition and state in Proposition III.7.4 that coupled models have related (anti-)Pareto optima and the same (anti-)Pareto frontiers.

In the kernel model, we consider the models $\text{Param}_i = [(\Omega_i, \mathscr{F}_i, \mu_i), \mathbf{k}_i]$ for $i \in \{1, 2\}$, where Assumption III.1 holds for each model; in the SIS model, we consider the models

$$\text{Param}_i = [(\Omega_i, \mathscr{F}_i, \mu_i), k_i, \gamma_i],$$

for $i \in \{1, 2\}$, where Assumption III.2 holds for each model. In what follows, we simply write Δ_i the set of functions Δ for the model Param_i.

A measure π on $(\Omega_1 \times \Omega_2, \mathscr{F}_1 \otimes \mathscr{F}_2)$ is a *coupling* if its marginals are μ_1 and μ_2.

Definition III.7.2 (Coupled models). *The models* Param_1 *and* Param_2 *are* coupled *if there exists two independent* $\Omega_1 \times \Omega_2$*-valued random vectors* (X_1, X_2) *and* (Y_1, Y_2) *(defined on a probability space* $(\Omega_0, \mathscr{F}_0, \mathbb{P})$*) with the same distribution given by a coupling (*i.e. X_i *and* Y_i *have distribution* μ_i*) such that,* $\mathbb{P}$*-almost surely:*

$$\begin{aligned} \textit{Kernel model:} \quad & \mathbf{k}_1(X_1, Y_1) = \mathbf{k}_2(X_2, Y_2), \\ \textit{SIS model:} \quad & \gamma_1(X_1) = \gamma_2(X_2) \quad \textit{and} \quad k_1(X_1, Y_1) = k_2(X_2, Y_2). \end{aligned}$$

In this case, two real-valued measurable functions v_1 *and* v_2 *defined respectively on* Ω_1 *and* Ω_2 *are* coupled *(through* V*) if there exists a real-valued* $\sigma(X_1, X_2)$*-measurable integrable random variable* V *such that* $\mathbb{P}$*-almost surely:*

$$\mathbb{E}\left[V | X_i\right] = v_i(X_i) \quad \textit{for } i \in \{1, 2\}.$$

Remark III.7.3. We keep notation from Definition III.7.2

(i) Since V is real-valued and $\sigma(X_1, X_2)$-measurable, we deduce from (III.54) that there exits a measurable function v defined on $\Omega_1 \times \Omega_2$ such that $V = v(X_1, X_2)$, thus the following equality holds $\mathbb{P}$-almost surely:

$$\mathbb{E}\left[v(Y_1, Y_2)|\, Y_i\right] = v_i(Y_i) \quad \text{for } i \in \{1, 2\}.$$

(ii) If W is a real-valued integrable $\sigma(X_1) \cap \sigma(X_2)$-measurable random variable, then setting $v_i(X_i) = \mathbb{E}\left[W|X_i\right] = W$, the equality $v_1(X_1) = v_2(X_2)$ holds almost surely, and we get that v_1 and v_2 are coupled (through W).

(iii) Let $\eta_1 \in \Delta_1$. According to (III.54), there exists $\eta_2 \in \Delta_2$ such that $\mathbb{E}\left[\eta_1(X_1)|X_2\right] = \eta_2(X_2)$. Thus, by definition η_1 and η_2 are coupled (through $V = \eta_1(X_1)$).

The main result of this section, whose proof is given in Section III.8.2, states that coupled models have coupled Pareto optimal strategies, and thus the same (anti-)Pareto frontier.

Proposition III.7.4 (Coupling and Pareto optimality)**.** *Let* Param_1 *and* Param_2 *be two coupled (kernel or SIS) models with the uniform cost function* $C = C_{\mathrm{uni}}$ *and loss function* L *(with* $\mathrm{L} = \mathfrak{R}_e$ *in the kernel model and* $\mathrm{L} \in \{\mathfrak{R}_e, \mathfrak{I}\}$ *in the SIS model). If the functions* $\eta_1 \in \Delta_1$ *and* $\eta_2 \in \Delta_2$ *are coupled, then:*

$$\eta_1 \textit{ is Pareto optimal (for } \mathrm{Param}_1\textit{)} \iff \eta_2 \textit{ is Pareto optimal (for } \mathrm{Param}_2\textit{)}.$$

Furthermore, if $\eta_1 \in \Delta_1$ *is Pareto optimal (for* Param_1*), then there exists a Pareto optimal (for* Param_2*) strategy* $\eta_2 \in \Delta_2$ *such that* η_1 *and* η_2 *are coupled. In particular, the (anti-)Pareto frontiers are the same for the two models* Param_1 *and* Param_2*.*

The next Corollary is useful for model reduction, which corresponds to merging individuals with identical behavior, see the examples in Sections III.7.4 and III.7.4. Equation (III.55) below could also be stated for anti-Pareto optima; and the adaptation to the kernel model is immediate.

Corollary III.7.5. *Let* $\mathrm{Param} = [(\Omega, \mathscr{F}, \mathbb{P}), k, \gamma]$ *be a SIS model with the uniform cost function* $C = C_{\mathrm{uni}}$ *and loss function* $\mathrm{L} \in \{\mathfrak{R}_e, \mathfrak{I}\}$*. Let* $\mathscr{G} \subset \mathscr{F}$ *be a* σ*-field such that* γ *is* $\mathscr{G}$*-measurable and* k *is* $\mathscr{G} \otimes \mathscr{G}$*-measurable. Then, for any* $\eta \in \Delta[\mathrm{Param}]$*, we have:*

$$\eta \textit{ is Pareto optimal} \iff \mathbb{E}\left[\eta|\, \mathscr{G}\right] \textit{ is Pareto optimal.} \tag{III.55}$$

Proof. Let $\Omega_0 = \Omega^2$ endowed with the product σ-field and the product probability measure $\mathbb{P}_0$, and X (resp. Y) be the projection on the first (resp. second) coordinate. Thus the random variables X and Y are independent, $(\Omega, \mathscr{F})$-valued with distribution $\mathbb{P}$. Write (X', Y') for (X, Y) when considered as $(\Omega, \mathscr{G})$-valued random variables. Notice that X' and Y' are by construction independent with distribution $\mathbb{P}'$, where $\mathbb{P}'$ is the restriction of $\mathbb{P}$ to $\mathscr{G}$. As γ is $\mathscr{G}$-measurable and k is $\mathscr{G} \otimes \mathscr{G}$-measurable, we can consider the model $\mathrm{Param}' = [(\Omega, \mathscr{G}, \mathbb{P}'), k, \gamma]$. Then (X, X') and (Y, Y') are two trivial couplings such that $k(X, Y) = k(X', Y')$ and $\gamma(X) = \gamma(X')$. Thus the models Param and Param$'$ are coupled. We have that $\eta \in \Delta$ and $\eta' = \mathbb{E}[\eta|\, \mathscr{G}] \in \Delta'$ are coupled through $\eta \circ X$ since $\mathbb{E}_0[\eta \circ X|\, \sigma(X)] = \eta \circ X$ and $\mathbb{E}_0[\eta \circ X|\, \sigma(X')] = \eta' \circ X'$ as $\sigma(X') = X^{-1}(\mathscr{G})$ and $X = X'$ can be seen as the identity map on Ω. The conclusion then follows from Proposition III.7.4. □

III.7.4 Examples of couplings

In this section, we consider the SIS model as the kernel model can be handled in the same way. We denote by Leb the Lebesgue measure.

Discrete and continuous models

We now formalize how finite population models can be seen as particular cases of models with a continuous population. Let $\Omega_{\mathrm{d}} \subset \mathbb{N}$, $\mathscr{F}_{\mathrm{d}}$ the set of subsets of Ω_{d} and μ_{d} a probability measure on Ω_{d}. Without loss of generality, we can assume that $\mu_{\mathrm{d}}(\{\ell\}) > 0$ for all $\ell \in \Omega_{\mathrm{d}}$. We set $\Omega_{\mathrm{c}} = [0, 1)$, with $\mathscr{F}_c$ its Borel σ-field and $\mu_c = \mathrm{Leb}$. Let $(B_\ell, \ell \in \Omega_d)$ be a partition of $[0,1)$ in measurable sets such that $\mathrm{Leb}(B_\ell) = \mu_d(\{\ell\})$ for all $\ell \in \Omega_d$. The measure π on $\Omega_d \times \Omega_c$ uniquely defined by:

$$\pi(\{\ell\} \times A) = \mathrm{Leb}(B_\ell \cap A)$$

for all measurable $A \subset [0,1)$ and $\ell \in \Omega_d$ is clearly a coupling of μ_d and μ_c. If the kernels k_d on Ω_d and k_c on Ω_c and the functions γ_d and γ_c are related through the formula:

$$\gamma_c(x) = \gamma_d(\ell) \quad \text{and} \quad k_c(x,y) = k_d(\ell, j), \quad \text{for } x \in B_\ell,\ y \in B_j \text{ and } \ell, j \in \Omega_d,$$

then the discrete model $\mathrm{Param}_d = [(\Omega_d, \mathscr{F}_d, \mu_d), k_d, \gamma_d]$ and the continuous model $\mathrm{Param}_c = [([0,1), \mathscr{F}_c, \mathrm{Leb}), k_c, \gamma_c]$ are coupled. Roughly speaking, we can blow up the atomic part of the measure μ_d into a continuous part, or, conversely, merge all points that behave similarly for k_c and γ_c into an atom, without altering the Pareto frontier.

Example III.7.6. We consider the so called stochastic block model, with 2 populations for simplicity, in the setting of the SIS model, and give in this elementary case the corresponding discrete and continuous models. Then, we explicit the relation with the formalism of the same model developed in [102] by Lajmanovich and Yorke.

The discrete SIS model is defined on $\Omega_d = \{1,2\}$ with the probability measure μ_d defined by $\mu_d(\{1\}) = 1 - \mu_d(\{2\}) = p$ with $p \in (0,1)$, and a kernel k_d and recovery function γ_d given by the matrix and the vector:

$$k_d = \begin{pmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{pmatrix} \quad \text{and} \quad \gamma_d = \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix}.$$

Notice p is the relative size of population 1. The corresponding discrete model is $\mathrm{Param}_d = [(\{1,2\}, \mathscr{F}_d, \mu_d), k_d, \gamma_d]$; see Figure III.5(b).

The continuous model is defined on the state space $\Omega_c = [0,1)$ is endowed with its Borel σ-field, $\mathscr{F}_c$, and the Lebesgue measure $\mu_c = \mathrm{Leb}$. The segment $[0,1)$ is partitioned into two intervals $B_1 = [0,p)$ and $B_2 = [p,1)$, the transmission kernel k_c and recovery rate γ_c are given by:

$$k_c(x,y) = k_{ij} \quad \text{and} \quad \gamma_c(x) = \gamma_i \quad \text{for } x \in B_i,\ y \in B_j, \text{ and } i,j \in \{1,2\}.$$

The corresponding continuous model is $\mathrm{Param}_c = [([0,1), \mathscr{F}_c, \mathrm{Leb}), k_c, \gamma_c]$; see Figure III.5(a). By the general discussion above, the discrete and continuous models are coupled, and in particular they have the same Pareto and anti-Pareto frontiers.

Furthermore, in this simple example, it is easily checked that a discrete vaccination $\eta_d = (\eta_1, \eta_2)$ and a continuous vaccination $\eta_c = (\eta_c(x), x \in [0,1))$ are coupled if and only if there exists a function η defined on $\Omega_c \times \Omega_d = [0,1) \times \{1,2\}$ such that:

$$\begin{cases} \eta_i = \dfrac{1}{\mathrm{Leb}(B_i)} \displaystyle\int_{B_i} \eta(x,i)\,\mathrm{d}x, \quad i \in \{1,2\}, \\ \eta_c(x) = \eta(x,1)\mathbb{1}_{B_1}(x) + \eta(x,2)\mathbb{1}_{B_2}(x), \quad \text{Leb-a.s.,} \end{cases}$$

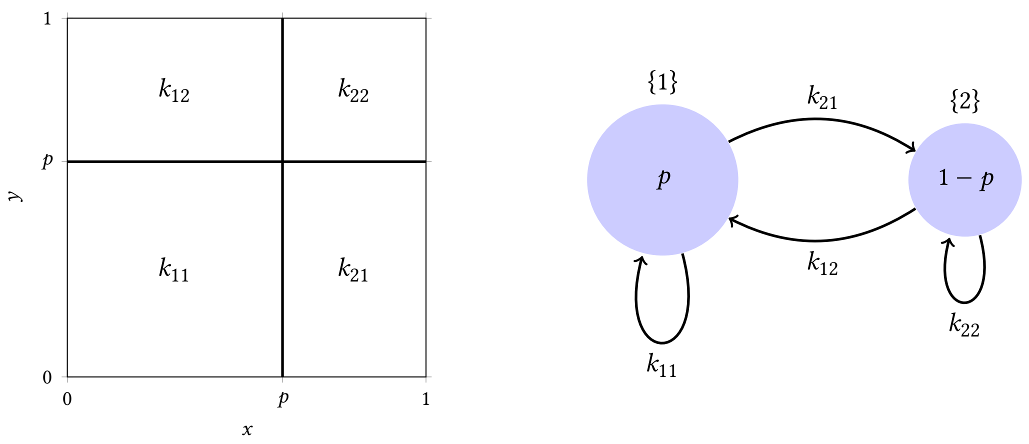
which occurs if and only if:

$$\eta_i = \frac{1}{\mathrm{Leb}(B_i)} \int \eta_c(x)\mathbb{1}_{B_i}(x)\,\mathrm{d}x, \quad i \in \{1,2\}.$$

Therefore, in this case, the optimal strategies of the continuous model are easily deduced from the optimal strategies of the discrete model.

To conclude this example, we rewrite, using the formalism of the discrete model Param_d, the next-generation matrix K in the setting of [102], and the effective next-generation matrix $K_e(\eta)$ when the vaccination strategy η is in force (recall η_i is the proportion of population with feature i which is not vaccinated):

$$K = \begin{pmatrix} \mathbf{k}_{11}\, p & \mathbf{k}_{12}\,(1-p) \\ \mathbf{k}_{21}\, p & \mathbf{k}_{22}\,(1-p) \end{pmatrix} \quad \text{and} \quad K_e(\eta) = \begin{pmatrix} \mathbf{k}_{11}\, p\eta_1 & \mathbf{k}_{12}\,(1-p)\,\eta_2 \\ \mathbf{k}_{21}\, p\eta_1 & \mathbf{k}_{22}\,(1-p)\,\eta_2 \end{pmatrix},$$



(A) Continuous model: kernel k_c on $\Omega_c = [0, 1)$ with the Lebesgue measure.

(B) Discrete model: kernel k_d on $\Omega_d = \{1, 2\}$ with the measure $p\delta_1 + (1-p)\delta_2$.

Figure III.5: Coupled continuous model (left) and discrete model (right).

with $p = \mu_d(\{1\})$, $1 - p = \mu_d(\{2\})$ and $\mathbf{k}_d = k_d/\gamma_d$, that is:

$$\mathbf{k}_d = \begin{pmatrix} \mathbf{k}_{11} & \mathbf{k}_{12} \\ \mathbf{k}_{21} & \mathbf{k}_{22} \end{pmatrix} = \begin{pmatrix} k_{11}/\gamma_1 & k_{12}/\gamma_2 \\ k_{21}/\gamma_1 & k_{22}/\gamma_2 \end{pmatrix}.$$

Measure preserving function

This section is motivated by the theory of graphons, which are indistinguishable by measure preserving transformation, see [111, Sections 7.3 and 10.7]. Let $(\Omega, \mathscr{F}, \mu)$ be a measurable space. We say a measurable function φ from $(\Omega, \mathscr{F})$ to itself is *measure preserving* if $\mu = \varphi_\# \mu$. For example the function $\varphi : x \mapsto 2x \mod (1)$ defined on the probability space $([0, 1], \mathscr{B}([0, 1], \text{Leb})$ is measure preserving.

Let φ be measure preserving function on Ω. Let k_1 be a kernel and γ_1 a function on Ω such that the model $\text{Param}_1 = [(\Omega, \mathscr{F}, \mu), k_1, \gamma_1]$ satisfies Assumption III.2. Let X_1 be a random variable with probability distribution μ and let $X_2 = \varphi(X_1)$, so that (X_1, X_2) is a coupling of $(\Omega, \mathscr{F}, \mu)$ with itself. Then for the kernel k_2 and the function γ_2 defined by:

$$k_2(x, y) = k_1(\varphi(x), \varphi(y)) \quad \text{and} \quad \gamma_2(x) = \gamma_1(\varphi(x)),$$

the models Param_1 and $\text{Param}_2 = [(\Omega, \mathscr{F}, \mu), k_2, \gamma_2]$ are coupled. Roughly speaking, we can give different labels to the features of the population without altering the Pareto and anti-Pareto frontiers.

Model reduction using deterministic coupling

This example is in the spirit of Section III.7.4, where one merges individual with identical behavior. We consider a SIS model $\text{Param}_1 = [(\Omega_1, \mathscr{F}_1, \mu_1), k_1, \gamma_1]$. Let φ be a measurable function from $(\Omega_1, \mathscr{F}_1)$ to $(\Omega_2, \mathscr{F}_2)$. Assume that:

$$\sigma(\gamma_1) \subset \sigma(\varphi) \quad \text{and} \quad \sigma(k_1) \subset \sigma(\varphi) \otimes \sigma(\varphi).$$

We can then build an elementary coupling. Let X_1 and Y_1 be independent μ_1 distributed random elements of Ω_1, and set $(X_2, Y_2) = (\varphi(X_1), \varphi(Y_1))$. Since $\sigma(\gamma_1) \subset \sigma(\varphi)$ and $\sigma(k_1) \subset \sigma(\varphi) \otimes \sigma(\varphi)$, we get that $\gamma_1(X_1)$ is $\sigma(X_2)$-measurable and $k_1(X_1, Y_1)$ is $\sigma(X_2, Y_2)$-measurable. According to (III.54), there exists two measurable functions $\gamma_2 : \Omega_2 \to \mathbb{R}$ and $k_2 : \Omega_2 \times \Omega_2 \to \mathbb{R}$ such that $\gamma_1 = \gamma_2 \circ \varphi$ and $k_1 = k_2(\varphi \otimes \varphi)$ that is almost surely:

$$\gamma_1(X_1) = \gamma_2(X_2) \quad \text{and} \quad k_1(X_1, Y_1) = k_2(X_2, Y_2).$$

Let $\mu_2 = \varphi_\# \mu_1$ be the push-forward measure of μ_1 by φ. Using (III.53) it is easy to check that the integrability condition from Assumption III.2 is fulfilled, so we can consider the reduced model $\mathrm{Param}_2 = [(\Omega_2, \mathscr{F}_2, \mu_2), k_2, \gamma_2]$. By Definition III.7.2, Param_1 is coupled with Param_2 through the (deterministic) coupling π given by the distribution of $(X_1, \varphi(X_1))$.

Eventually, we get from Corollary III.7.5 with $\mathscr{G} = \sigma(\varphi)$, that $\eta_1 \in \Delta_1$ is Pareto optimal if and only if $\mathbb{E}_1[\eta_1 | \varphi]$ is Pareto optimal (for the model Param_1), where $\mathbb{E}_1$ correspond to the expectation with respect to the probability measure μ_1 on $(\Omega_1, \mathscr{F}_1)$.

III.8 Technical proofs

III.8.1 The SIS model: properties of $\mathfrak{I}$ and of the maximal equilibrium

We prove here Theorem III.4.6 and Proposition III.4.7, and properties of the maximal equilibrium. For the convenience of the reader, we only use references to the results recalled in Chapter II for positive operators on Banach spaces. For an operator A, we denote by $A^\top$ its adjoint. We first give a preliminary lemma.

Lemma III.8.1. *Suppose Assumption III.2 holds, and consider the positive bounded linear integral operator $\mathscr{T}_{k/\gamma}$ on $\mathscr{L}^\infty$. If there exists $g \in \mathscr{L}^\infty_+$, with $\int_\Omega g \,\mathrm{d}\mu > 0$ and $\lambda > 0$ satisfying:*

$$\mathscr{T}_{k/\gamma}(g)(x) > \lambda g(x), \quad \textit{for all} \quad x \textit{ such that } g(x) > 0,$$

then we have $\rho(\mathscr{T}_{k/\gamma}) > \lambda$.

Proof. Set $\mathscr{T} = \mathscr{T}_{k/\gamma}$. Let $A = \{ g > 0 \}$ be the support of the function g. Let $\mathscr{T}'$ be the bounded operator defined by $\mathscr{T}'(f) = \mathbb{1}_A \mathscr{T}(\mathbb{1}_A f)$. Since $\mathscr{T}'(g) = \mathbb{1}_A \mathscr{T}(\mathbb{1}_A g) = \mathbb{1}_A \mathscr{T}(g) > \lambda g$, we deduce from the Collatz-Wielandt formula, see Proposition II.3.6, that $\rho(\mathscr{T}') \geq \lambda > 0$. According to Lemma II.3.7 (v), there exists $v \in L^q_+ \setminus \{0\}$, seen as an element of the topological dual of $\mathscr{L}^\infty$, a left Perron eigenfunction of $\mathscr{T}'$, that is such that $(\mathscr{T}')^\top(v) = \rho(\mathscr{T}')v$. In particular, we have $v = \mathbb{1}_A v$ and thus $\int_A v \,\mathrm{d}\mu > 0$ and $\int_\Omega v g \,\mathrm{d}\mu > 0$. We obtain:

$$(\rho(\mathscr{T}') - \lambda) \langle v, g \rangle = \langle v, \mathscr{T}'(g) - \lambda g \rangle > 0.$$

This implies that $\rho(\mathscr{T}') > \lambda$. Since $\mathscr{T} - \mathscr{T}'$ is a positive operator, we deduce from (III.26) that $\rho(\mathscr{T}) \geq \rho(\mathscr{T}') > \lambda$. □

We now state an interesting result on the characterization of the maximal equilibrium $\mathfrak{g}$. We keep notations from Sections III.2.3 and III.2.4 and write $\mathfrak{R}_e$ for $\mathfrak{R}_e[k/\gamma]$. Recall that $\mathfrak{R}_0 = \mathfrak{R}_e(\mathbb{1})$ and F defined by (III.18). Let $DF[h]$ denote the bounded linear operator on $\mathscr{L}^\infty$ of the derivative of the map $f \mapsto F(f)$ defined on $\mathscr{L}^\infty$ at point h:

$$DF[h](g) = (1-h)\mathscr{T}_k(g) - (\gamma + \mathscr{T}_k(h))g \quad \text{for } h, g \in \mathscr{L}^\infty.$$

Let $s(A)$ denote the spectral bound of the bounded operator A, see (33) in Chapter II.

Proposition III.8.2. *Suppose Assumption III.2 holds and write $\mathfrak{R}_e$ for $\mathfrak{R}_e[k/\gamma]$. Let h in Δ be an equilibrium, that is $F(h) = 0$. The following properties are equivalent:*

(i) $h = \mathfrak{g}$,

(ii) $s(DF[h]) \leq 0$,

(iii) $\mathfrak{R}_e((1-h)^2) \leq 1$.

(iv) $\mathfrak{R}_e(1-h) \leq 1$.

We also have: $\mathfrak{g} = 0 \Longleftrightarrow \mathfrak{R}_0 \leq 1$; as well as: $\mathfrak{g} \neq 0 \Longrightarrow \mathfrak{R}_e(1-\mathfrak{g}) = 1$.

Proof. Let $h \in \Delta$ be an equilibrium, that is $F(h) = 0$.

Let us show the equivalence between (ii) and (iii). According to Proposition II.4.2, we have $s(DF[h]) \leq 0$ if and only if:

$$\rho(\mathscr{T}_{\mathbf{k}}) \leq 1 \quad \text{with} \quad \mathbf{k}(x,y) = (1-h(x))\frac{k(x,y)}{\gamma(y) + \mathscr{T}_k(h)(y)}.$$

Since $F(h) = 0$, we have $(1-h)/\gamma = 1/(\gamma + \mathscr{T}_k(h))$. This gives:

$$\mathsf{k}(x,y) = (1-h(x))\frac{k(x,y)(1-h(y))}{\gamma(y)} \tag{III.56}$$

and thus $\mathscr{T}_{\mathbf{k}} = M_{1-h}\,\mathscr{T}_{k/\gamma}M_{1-h}$, where M_f is the multiplication operator by f. Recall the definition (III.14) of $\mathfrak{R}_e$. According to (III.29), we have:

$$\rho(\mathscr{T}_{\mathbf{k}}) = \rho\left(\mathscr{T}_{k/\gamma}M_{(1-h)^2}\right) = \mathfrak{R}_e((1-h)^2). \tag{III.57}$$

This gives the equivalence between (ii) and (iii).

We prove that (i) implies (iv). Suppose that $\mathfrak{R}_e(1-h) > 1$. Thanks to (III.29), we have $\rho(M_{1-h}\mathscr{T}_{k/\gamma}) = \rho(\mathscr{T}_{k/\gamma}M_{1-h}) = \mathfrak{R}_e(1-h) > 1$. According to Lemma II.3.7 (v), there exists $v \in L^q_+ \setminus \{0\}$ a left Perron eigenfunction of $\mathscr{T}_{(1-h)k/\gamma}$, that is $\mathscr{T}^\top_{(1-h)k/\gamma}(v) = \mathfrak{R}_e(1-h)v$. Using $F(h) = 0$, and thus $(1-h)\mathscr{T}_k(h) = \gamma h$, for the last equality, we have:

$$\mathfrak{R}_e(1-h)\,\langle v, \gamma h\rangle = \langle v, (1-h)\mathscr{T}_{k/\gamma}(\gamma h)\rangle = \langle v, \gamma h\rangle\,.$$

We get $\langle v, \gamma h\rangle = 0$ and thus $\langle v, \mathbb{1}_A\rangle = 0$, where $A = \{\,h > 0\,\}$ denote the support of the function h. Since $\mathscr{T}^\top_{(1-h)k/\gamma}(v) = \mathfrak{R}_e(1-h)v$ and setting $v' = (1-h)v$ (so that $v' = v$ μ-almost surely on A^{c}), we deduce that:

$$\mathscr{T}^\top_{k'/\gamma}(v') = \mathfrak{R}_e(1-h)v',$$

where $k' = \mathbb{1}_{A^{\mathsf{c}}}\,k\,\mathbb{1}_{A^{\mathsf{c}}}$. This implies that $\rho(\mathscr{T}_{k'/\gamma}) \geq \mathfrak{R}_e(1-h)$. Since $k' = (1-h)k'$ and $\mathscr{T}_{k/\gamma} - \mathscr{T}_{k'/\gamma}$ is a positive operator as $k - k' \geq 0$, we get, using (III.26) for the inequality, that $\rho(\mathscr{T}_{k'/\gamma}) = \rho(M_{1-h}\mathscr{T}_{k'/\gamma}) \leq \rho(M_{1-h}\mathscr{T}_{k/\gamma}) = \mathfrak{R}_e(1-h)$. Thus, the spectral radius of $\mathscr{T}_{k'/\gamma}$ is equal to $\mathfrak{R}_e(1-h)$. According to Proposition II.4.2, since $\rho(\mathscr{T}_{k'/\gamma}) > 1$, there exists $w \in \mathscr{L}^\infty_+ \setminus \{0\}$ and $\lambda > 0$ such that:

$$\mathscr{T}_{k'}(w) - \gamma w = \lambda w.$$

This also implies that $w = 0$ on $A = \{\,h > 0\,\}$, that is $wh = 0$ and thus $w\mathscr{T}_k(h) = 0$ as $\mathscr{T}_k(h) = \gamma h/(1-h)$. Using that $F(h) = 0$, $\mathscr{T}_k(w) = \mathscr{T}_{k'}(w) = (\gamma + \lambda)w$ and $h\mathscr{T}_k(w) = 0$, we obtain:

$$F(h + w) = w(\lambda - \mathscr{T}_k(w)).$$

Taking $\varepsilon > 0$ small enough so that $\varepsilon\mathscr{T}_k(w) \leq \lambda/2$ and $\varepsilon w \leq 1$, we get $h + \varepsilon w \in \Delta$ and $F(h + \varepsilon w) \geq 0$. Then use Lemma III.4.4 to deduce that $h + \varepsilon w \leq \mathfrak{g}$ and thus $h \neq \mathfrak{g}$.

To see that (iv) implies (iii), notice that $(1-h)^2 \leq (1-h)$, and then deduce from Proposition III.4.1 (iii) that $\mathfrak{R}_e((1-h)^2) \leq \mathfrak{R}_e(1-h)$.

We prove that (iii) implies (i). Notice that $F(g) = 0$ and $g \in \Delta$ implies that $g < 1$. Assume that $h \neq \mathfrak{g}$. Notice that $\gamma/(1-h) = \gamma + \mathscr{T}_k(h)$, so that $\gamma(\mathfrak{g} - h)/(1-h) \in \mathscr{L}^\infty_+$. An elementary computation, using $F(h) = F(\mathfrak{g}) = 0$ and $\mathbf{k}$ defined in (III.56), gives:

$$\mathscr{T}_{\mathbf{k}}\left(\gamma \times \frac{\mathfrak{g} - h}{1-h}\right) = (1-h)\mathscr{T}_k(\mathfrak{g} - h) = \gamma \times \frac{\mathfrak{g} - h}{1-\mathfrak{g}} = \frac{1-h}{1-\mathfrak{g}} \times \gamma \times \frac{\mathfrak{g} - h}{1-h}.$$

Since $h \neq \mathfrak{g}$ and $h \leq \mathfrak{g}$, we deduce that $(1-h)/(1-\mathfrak{g}) \geq 1$, with strict inequality on $\{\,\mathfrak{g} - h > 0\,\}$ which is a set of positive measure. We deduce from Lemma III.8.1 (with k replaced by $\mathbf{k}\gamma$) that $\rho(\mathscr{T}_{\mathbf{k}}) > 1$. Then use (III.57) to conclude.

To conclude notice that $\mathfrak{g} = 0 \iff \mathfrak{R}_0 \leq 1$ is a consequence of the equivalence between (i) and (iv) with $h = 0$ and $\mathfrak{R}_0 = \mathfrak{R}_e(\mathbb{1})$.

Using that $F(\mathfrak{g}) = 0$, we get $\mathcal{T}_k(\mathfrak{g}) = \gamma\mathfrak{g}/(1-\mathfrak{g})$. We deduce that $\mathcal{T}_{k(1-\mathfrak{g})/\gamma}(\mathcal{T}_k(\mathfrak{g})) = \mathcal{T}_k(\mathfrak{g})$. If $\mathfrak{g} \neq 0$, we get $\mathcal{T}_k(\mathfrak{g}) \neq 0$ (on a set of positive μ-measure). This implies that $\mathfrak{R}_e(1-\mathfrak{g}) \geq 1$. Then use (iv) to deduce that $\mathfrak{R}_e(1-\mathfrak{g}) = 1$ if $\mathfrak{g} \neq 0$. □

In the SIS model, in order to stress, if necessary, the dependence of a quantity H, such as F_η, $\mathfrak{R}_e$ or $\mathfrak{g}_\eta$, in the parameters k and γ (which satisfy Assumption III.2) of the model, we shall write $H[k,\gamma]$. Recall that if k and γ satisfy Assumption III.2, then the kernel k/γ has a finite double norm on L^p for some $p \in (1,+\infty)$. We now consider the continuity property of the maps $\eta \mapsto \mathfrak{g}_\eta[k,\gamma]$ and $(k,\gamma,\eta) \mapsto \mathfrak{g}_\eta[k,\gamma]$.

Lemma III.8.3. *Let $((k_n,\gamma_n), n \in \mathbb{N})$ and (k,γ) be kernels and functions satisfying Assumption III.2 and $(\eta_n,\, n \in \mathbb{N})$ be a sequence of elements of Δ converging weakly to η.*

(i) We have $\lim_{n\to\infty} \mathfrak{g}_{\eta_n}[k,\gamma] = \mathfrak{g}_\eta[k,\gamma]$ μ-almost surely.

(ii) Assume furthermore there exists $p' \in (1,+\infty)$ such that $\mathbf{k} = \gamma^{-1}k$ and $(\mathbf{k}_n = \gamma_n^{-1}k_n, n \in \mathbb{N})$ have finite double norm on $L^{p'}$ and that $\lim_{n\to\infty} \|\mathbf{k}_n - \mathbf{k}\|_{p',q'} = 0$. Then, we have:

$$\lim_{n\to\infty} \mathfrak{g}_{\eta_n}[k_n,\gamma_n] = \mathfrak{g}_\eta[k,\gamma] \ \mu - a.s.$$

Proof. The proof of (i) and (ii) being rather similar, we only provide the latter and indicate the difference when necessary. To simplify, we write $g_n = \mathfrak{g}_{\eta_n}[k_n,\gamma_n]$. We set $h_n = \eta_n g_n \in \Delta$ for $n \in \mathbb{N}$. Since Δ is sequentially weakly compact, up to extracting a subsequence, we can assume that h_n converges weakly to a limit $h \in \Delta$. Since $F_{\eta_n}[k_n,\gamma_n](g_n) = 0$ for all $n \in \mathbb{N}$, see (III.22), we have:

$$g_n = \frac{\mathcal{T}_{\mathbf{k}_n}(\eta_n g_n)}{1+\mathcal{T}_{\mathbf{k}_n}(\eta_n g_n)} = \frac{\mathcal{T}_{\mathbf{k}_n}(h_n)}{1+\mathcal{T}_{\mathbf{k}_n}(h_n)}. \tag{III.58}$$

We set $g = \mathcal{T}_{\mathbf{k}}(h)/(1+\mathcal{T}_{\mathbf{k}}(h))$. Notice that $\mathcal{T}_{\mathbf{k}_n}(h_n) = (\mathcal{T}_{\mathbf{k}_n} - \mathcal{T}_{\mathbf{k}})(h_n) + \mathcal{T}_{\mathbf{k}}(h_n)$. We have $\lim_{n\to\infty} \mathcal{T}_{\mathbf{k}}(h_n) = \mathcal{T}_{\mathbf{k}}(h)$ pointwise. Since $\left\|(\mathcal{T}_{\mathbf{k}_n} - \mathcal{T}_{\mathbf{k}})(h_n)\right\|_{p'} \leq \|\mathbf{k}_n - \mathbf{k}\|_{p',q'}$, up to taking a subsequence, we deduce that $\lim_{n\to\infty}(\mathcal{T}_{\mathbf{k}_n} - \mathcal{T}_{\mathbf{k}})(h_n) = 0$ almost surely. (Notice the previous step is not used in the proof of (i) as $\mathbf{k}_n = \mathbf{k}$ and $\lim_{n\to\infty} \mathcal{T}_k(h_n) = \mathcal{T}_k(h)$ pointwise.) This implies that g_n converges almost surely to g. By the dominated convergence theorem, we deduce that g_n converges also in L^p to g. This proves that $h = \eta g$ almost surely. We get $g = \mathcal{T}_{\mathbf{k}}(\eta g)/(1+\mathcal{T}_{\mathbf{k}}(\eta g))$ and thus $F_\eta[k,\gamma](g) = 0$: g is an equilibrium for $F_\eta[k,\gamma]$. We recall from Section IV.3 the functional equality $\mathfrak{R}_e[k'h] = \mathfrak{R}_e[hk']$, where k' is a kernel, h a non-negative functions such that the kernels $k'h$ and hk' have some finite double norm. We deduce from the weak-continuity and the stability of $\mathfrak{R}_e$, see Theorem III.4.2 and Proposition III.4.3, that:

$$\begin{aligned}\mathfrak{R}_e[k/\gamma](\eta(1-g)) = \mathfrak{R}_e[\mathbf{k}](\eta(1-g)) &= \lim_{n\to\infty} \mathfrak{R}_e[\mathbf{k}_n](\eta_n(1-g_n)) \\ &= \lim_{n\to\infty} \mathfrak{R}_e[k_n/\gamma_n](\eta_n(1-g_n)) \\ &\leq 1.\end{aligned}$$

(Only the weak-continuity of $\eta' \mapsto \mathfrak{R}_e[k/\gamma](\eta')$ is used in the proof of (i) to get $\mathfrak{R}_e[k/\gamma](\eta(1-g)) \leq 1$.) We deduce that property (iv) of Proposition III.8.2 holds with k replaced by $k\eta$, and thus property (i) therein implies that $g = \mathfrak{g}_\eta[k,\gamma]$. □

Proofs of Theorem III.4.6 and Proposition III.4.7. Under the assumptions of Lemma III.8.3, taking the pair (k_n,γ_n) equal to (k,γ) in the case (i) therein, we deduce that $(\eta_n\, \mathfrak{g}_{\eta_n}[k_n,\gamma_n], n \in \mathbb{N})$ converges weakly to $\eta\, \mathfrak{g}_\eta[k,\gamma]$. This implies that:

$$\lim_{n\to\infty} \mathfrak{I}[k_n,\gamma_n](\eta_n) = \lim_{n\to\infty} \int_\Omega \eta_n\, \mathfrak{g}_\eta[k_n,\gamma_n]\, \mathrm{d}\mu = \int_\Omega \eta\, \mathfrak{g}_\eta[k,\gamma]\, \mathrm{d}\mu = \mathfrak{I}[k,\gamma](\eta).$$

Taking $(k_n,\gamma_n) = (k,\gamma)$ provides the continuity of $\mathfrak{I}[k,\gamma]$ and thus Theorem III.4.6. Then, arguing as in the end of the proof of Proposition III.4.3, we get Proposition III.4.7. □

III.8.2 Coupling and Pareto optimality

We prove here Proposition III.7.4. We only consider the SIS model $\mathrm{Param} = [(\Omega, \mathscr{F}, \mu), k, \gamma]$, as the kernel model can be handled similarly. We suppose throughout this section that Assumption III.2 holds.

The random variables we consider, are defined on a probability space, say $(\Omega_0, \mathscr{F}_0, \mathbb{P})$. We recall an elementary result on conditional independence. Let $\mathscr{A}$, $\mathscr{B}$ and $\mathscr{H}$ be σ-fields subsets of $\mathscr{F}_0$, such that $\mathscr{H} \subset \mathscr{A} \cap \mathscr{B}$. Then, according to [91, Theorem 8.9], we have that for any integrable real-valued random variable X which is $\mathscr{B}$-measurable:

$$\mathscr{A} \text{ and } \mathscr{B} \text{ are conditionally independent given } \mathscr{H} \implies \mathbb{E}\left[X|\,\mathscr{A}\right] = \mathbb{E}\left[X|\,\mathscr{H}\right]. \tag{III.59}$$

We now state two technical lemmas.

Lemma III.8.4 (Measurability). *Let* Param_1 *and* Param_2 *be coupled models with independent coupling* (X_1, X_2) *and* (Y_1, Y_2)*. Then the random variable* $\gamma_1(X_1)$ *is* $\sigma(X_1) \cap \sigma(X_2)$*-measurable. For any measurable function* $v : \Omega_1 \times \Omega_2 \to \mathbb{R}$*, such that* $k_1(X_1, Y_1)v(Y_1, Y_2)$ *is integrable, the random variable* $\mathbb{E}\left[k_1(X_1, Y_1)v(Y_1, Y_2)|X_1\right]$ *is also* $\sigma(X_1) \cap \sigma(X_2)$*-measurable.*

Proof. The $\sigma(X_1) \cap \sigma(X_2)$-measurability of $\gamma_1(X_1)$ is an immediate consequence of the almost-sure equality $\gamma_1(X_1) = \gamma_2(X_2)$. Since $\mathbb{E}\left[k(X_1, Y_1)v(Y_1, Y_2)|X_1\right]$ is $\sigma(X_1)$-measurable, it remains to prove that it is also $\sigma(X_2)$-measurable. Since (X_1, X_2) is independent from (Y_1, Y_2), the σ-fields $\mathscr{A} = \sigma(X_1, X_2)$ and $\mathscr{B} = \sigma(X_1, Y_1, Y_2)$ are conditionally independent given $\mathscr{H} = \sigma(X_1)$. Using (III.59), we deduce that:

$$\mathbb{E}\left[k_1(X_1, Y_1)v(Y_1, Y_2)|X_1\right] = \mathbb{E}\left[k_1(X_1, Y_1)v(Y_1, Y_2)|X_1, X_2\right].$$

Since $k_1(X_1, Y_1) = k_2(X_2, Y_2)$ $\mathbb{P}$-almost surely, we get:

$$\begin{aligned}\mathbb{E}\left[k_1(X_1, Y_1)v(Y_1, Y_2)|X_1\right] &= \mathbb{E}\left[k_2(X_2, Y_2)v(Y_1, Y_2)|X_1, X_2\right] \\ &= \mathbb{E}\left[k_2(X_2, Y_2)v(Y_1, Y_2)|X_2\right],\end{aligned}$$

where the last equality follows from another application of (III.59) with $\mathscr{A} = \sigma(X_1, X_2)$, $\mathscr{B} = \sigma(X_2, Y_1, Y_2)$ which are conditionally independent given $\mathscr{H} = \sigma(X_2)$. The last expression is $\sigma(X_2)$ measurable, so the proof is complete. ☐

In the following key lemma, we simply write H_i for $H[\mathrm{Param}_i]$ for H the loss functions $\mathfrak{R}_e$ and $\mathfrak{I}$, the cost function $C = C_{\mathrm{uni}}$ and the spectrum Spec.

Lemma III.8.5. *If* Param_1 *and* Param_2 *are coupled models, and if the functions* $\eta_1 \in \Delta_1$ *and* $\eta_2 \in \Delta_2$ *are coupled, then* $\mathrm{Spec}_1(\eta_1) \cup \{0\} = \mathrm{Spec}_2(\eta_2) \cup \{0\}$ *and for* H *any one of the mappings* C_{uni}, $\mathfrak{R}_e$ *or* $\mathfrak{I}$*:*

$$H_1(\eta_1) = H_2(\eta_2). \tag{III.60}$$

Proof. Let (X_1, X_2) and (Y_1, Y_2) be two independent couplings, and assume that η_1 and η_2 are coupled through the function η, see Remark III.7.3 (i):

$$\mathbb{E}\left[\eta(X_1, X_2)|\, X_i\right] = \eta_i(X_i) \quad \text{for } i \in \{1, 2\}. \tag{III.61}$$

Step 1: The cost function ($H = C_{\mathrm{uni}}$). We directly have:

$$C_1(\eta_1) = 1 - \mathbb{E}\left[\eta_1(X_1)\right] = 1 - \mathbb{E}\left[\eta(X_1, X_2)\right] = 1 - \mathbb{E}\left[\eta_2(X_2)\right] = C_2(\eta_2).$$

Step 2: The spectrum and the effective reproduction function ($H = \mathfrak{R}_e$). Set $\mathbf{k}_i = k_i/\gamma_i$ for $i \in \{1, 2\}$. Let λ be a non-zero eigenvalue of $T_{\mathbf{k}_1\eta_1}$ associated with an eigenvector v_1. Notice that $\mathbf{k}(X_1, Y_1)\eta_1(Y_1)v(Y_1)$ is integrable thanks to the integrability condition from Assumption III.2. By definition of eigenvectors, $v_1(X_1)$ is a version of the conditional expectation:

$$\lambda^{-1}\mathbb{E}\left[\mathbf{k}_1(X_1, Y_1)\,\eta_1(Y_1)v_1(Y_1)|X_1\right].$$

By Lemma III.8.4 applied to the function $v(y_1, y_2) = (v_1\eta_1/\gamma_1)(y_1)$, the real-valued random variable $v_1(X_1)$ is $\sigma(X_1) \cap \sigma(X_2)$-measurable and thus $\sigma(X_2)$-measurable. Thanks to (III.54), there exists v_2 such that $v_2(X_2) = v_1(X_1)$ almost surely. Since (Y_1, Y_2) is distributed as (X_1, X_2), we deduce that (III.61) holds also with (X_1, X_2) replaced by (Y_1, Y_2) and that $v_2(Y_2) = v_1(Y_1)$ almost surely. Recall that $\mathbf{k}_i = k_i/\gamma_i$, so that $\mathbf{k}_1(X_1, Y_1) = \mathbf{k}_2(X_2, Y_2)$ almost surely. We may now compute:

$$\begin{aligned}
\lambda v_2(X_2) &= \lambda v_1(X_1) \\
&= \mathbb{E}\left[k_1(X_1, Y_1)\, \eta_1(Y_1) v_1(Y_1) | X_1\right] \\
&= \mathbb{E}\left[\mathbf{k}_1(X_1, Y_1)\, \eta(Y_1, Y_2) v_1(Y_1) | X_1\right] && \text{(de-conditioning on } (Y_1, X_1)) \\
&= \mathbb{E}\left[\mathbf{k}_1(X_1, Y_1)\, \eta(Y_1, Y_2) v_1(Y_1) | X_2\right] && \text{(Lemma III.8.4)} \\
&= \mathbb{E}\left[\mathbf{k}_2(X_2, Y_2)\, \eta(Y_1, Y_2) v_2(Y_2) | X_2\right] && \text{(a.s. equality)} \\
&= \mathbb{E}\left[\mathbf{k}_2(X_2, Y_2)\, \eta_2(Y_2) v_2(Y_2) | X_2\right] && \text{(conditioning on } (Y_2, X_2)) \\
&= T_{\mathbf{k}_2\eta_2} v_2(X_2).
\end{aligned} \tag{III.62}$$

Since the distribution of X_2 is μ_2, we have $\lambda v_2 = T_{\mathbf{k}_2\eta_2} v_2$ μ_2-almost surely. Therefore λ is also an eigenvalue for $T_{\mathbf{k}_2\eta_2}$. By symmetry we deduce that the spectrum up to $\{0\}$ of $T_{\mathbf{k}_1\eta_1}$ and $T_{\mathbf{k}_2\eta_2}$ coincide, that is $\mathrm{Spec}_1(\eta_1) \cup \{0\} = \mathrm{Spec}_2(\eta_2) \cup \{0\}$, and in particular the spectral radius coincide.

Step 3: The total proportion of infected population function ($H = \mathfrak{I}$). We assume without loss of generality that $\rho(\mathcal{T}_{k_1/\gamma_1}) > 1$, which is equivalent to $\rho(\mathcal{T}_{k_2/\gamma_2}) > 1$, thanks to (III.60) with $H = \mathfrak{R}_e$ and $\eta_1 = \eta_2 = \mathbb{1}$. Let $g_1 = \mathfrak{g}_{\eta_1}$ be the maximal equilibrium for the model Param_1. Since $F_{\eta_1}(g_1) = 0$, see (III.22), we have:

$$g_1 = \frac{\mathcal{T}_{k_1}(\eta_1 g_1)}{\gamma_1 + \mathcal{T}_{k_1}(\eta_1 g_1)}. \tag{III.63}$$

By Lemma III.8.4, this implies that $g_1(X_1)$ is $\sigma(X_1) \cap \sigma(X_2)$ measurable. Thus, there exists g_2' such that $g_2'(X_2) = g_1(X_1)$ $\mathbb{P}$-almost surely.. Therefore, by the same computation as in (III.62):

$$\mathcal{T}_{k_1}(\eta_1 g_1)(X_1) = \mathcal{T}_{k_2}(\eta_2 g_2')(X_2) \quad \mathbb{P}-\text{a.s.}$$

We set:

$$g_2 = \frac{\mathcal{T}_{k_2}(\eta_2 g_2')}{\gamma_2 + \mathcal{T}_{k_2}(\eta_2 g_2')}. \tag{III.64}$$

Then, we deduce from (III.63) that $g_2(X_2) = g_2'(X_2)$ $\mathbb{P}$-almost surely, that is $g_2 = g_2'$ μ_2-almost surely. Thus (III.64) holds with g_2' replaced by g_2. In other words, g_2 satisfies (III.22): it is an equilibrium for the model given by Param_2.

Let us now prove that g_2 is in fact the maximal equilibrium. Since $g_2(X_2) = g_1(X_1)$ $\mathbb{P}$-almost surely and $g_1(X_1)$ is $\sigma(X_1) \cap \sigma(X_2)$-measurable, we deduce from Remark III.7.3 (ii), that $(1 - g_1)$ and $(1 - g_2)$ are coupled, so $\mathfrak{R}_e[\mathrm{Param}_1](1 - g_1) = \mathfrak{R}_e[\mathrm{Param}_2](1 - g_2)$, by Property (III.60) applied to $H = \mathfrak{R}_e$. Since $\mathfrak{R}_0 > 1$ and g_1 is the maximal equilibrium for Param_1, we deduce from Proposition III.8.2 that $\mathfrak{R}_e[\mathrm{Param}_1](1 - g_1) = 1$. Using again Proposition III.8.2, this gives that g_2 is the maximal equilibrium for Param_2.

We may now compute:

$$\begin{aligned}
\mathfrak{I}_1(\eta_1) &= \mathbb{E}\left[\eta_1(X_1)\, g_1(X_1)\right] \\
&= \mathbb{E}\left[\eta(X_1, X_2)\, g_1(X_1)\right] && \text{(deconditioning on } X_1) \\
&= \mathbb{E}\left[\eta(X_1, X_2)\, g_2(X_2)\right] && \text{(a.s. equality)} \\
&= \mathbb{E}\left[\eta_2(X_2)\, g_2(X_2)\right] && \text{(conditioning on } X_2) \\
&= \mathfrak{I}_2(\eta_2),
\end{aligned}$$

thus (III.60) holds for $H = \mathfrak{I}$, and the proof is complete. □

We now give the proof of Proposition III.7.4. Its first part is an elementary consequence of the Lemma III.8.5; and the second part is a direct consequence of Remark III.7.3 (iii).

Chapter IV

Effective reproduction number: convexity, invariance and cordons sanitaires

Chapter Abstract

We consider the problem of optimal allocation strategies for a (perfect) vaccine in an infinite-metapopulation model (including SIS, SIR, SEIR, ...), when the loss function is given by the effective reproduction number $\mathfrak{R}_e$, which is defined as the spectral radius of the effective next generation matrix (in finite dimension) or more generally of the effective next generation operator (in infinite dimension). We give sufficient conditions for $\mathfrak{R}_e$ to be a convex or a concave function of the vaccination strategy. Then, following a previous work, we consider the bi-objective problem of minimizing simultaneously the cost and the loss of the vaccination strategies. In particular, we prove that a cordon sanitaire might not be optimal, but it is still better than the "worst" vaccination strategies. Inspired by the graph theory, we compute the minimal cost which ensures that no infection occurs using independent sets. Using Frobenius decomposition of the whole population into irreducible sub-populations, we give some explicit formulae for optimal ("best" and "worst") vaccinations strategies. Eventually, we provide equivalence properties on models which ensure that the function $\mathfrak{R}_e$ is unchanged; in the matrix setting this corresponds to identify the preservers for the spectral radius of matrices.

The material for this chapter has been released in [36].

Chapter Content:

IV.1 Introduction

IV.1.1 Vaccination in metapopulation models

The study of vaccination strategies for metapopulation models with $N \geq 2$ sub-populations, naturally leads to an easily stated linear algebra problem:c3- given a matrix K, of size $N \times N$, with non-negative entries, what can be said about the function

$$\mathfrak{R}_e : \begin{cases} \Delta & \to \mathbb{R}, \\ \eta & \mapsto \text{spectral radius of } K \cdot \mathrm{Diag}(\eta), \end{cases} \tag{IV.1}$$

where $\Delta = [0,1]^N$, $\mathrm{Diag}(\eta)$ denotes the $N \times N$ matrix with diagonal elements $\eta = (\eta_1, \ldots, \eta_N)$, and the spectral radius is the largest modulus of the eigenvalues. In this form, the problem appears for instance, with a mathematical point of view, in Elsner and Hadeler [51], see also Friedland [64] and Nussbaum [122].

In metapopulation epidemiological models, the indices $i = 1, \ldots, N$ correspond to various sub-populations with respective proportional size $\mu_1, \ldots, \mu_N$. Following [81], the entry K_{ij} of the so-called next-generation matrix K is equal to the expected number of secondary infections for people in subgroup i resulting from a single randomly selected non-vaccinated infectious person in subgroup j. Finally, η represents a vaccination strategy, that is, η_i is the fraction of non-vaccinated individuals in the i^{th} sub-population; thus $\eta_i = 0$ when the i^{th} sub-population is fully vaccinated, and 1 when it is not vaccinated at all. (This seemingly unnatural convention is in particular motivated by the simple form of Equation IV.1). So, the strategy $\mathbb{1} \in \Delta$, with all its entries equal to 1, corresponds to an entirely non-vaccinated population. The quantity $\mathfrak{R}_e$, referred to as the *effective reproduction number*, may then be interpreted as the mean number of infections coming from a typical case. In particular, we denote by $\mathfrak{R}_0 = \mathfrak{R}_e(\mathbb{1})$ the so-called *basic reproduction number* associated to the metapopulation epidemiological model. With the interpretation of the function $\mathfrak{R}_e$ in mind, it is then very natural to minimize it under a constraint on the cost $C(\eta)$ of the vaccination strategies η. A natural choice for the cost function is given by the uniform cost $C(\eta) = 1 - \sum_i \eta_i \mu_i$, which corresponds to the fraction of vaccinated individuals in the population. This constrained optimization problem appears in most of the literature for designing efficient vaccination strategies for multiple epidemic situation (SIR/SEIR), see [29, 46, 52, 81, 116, 130, 161]. Note that in some of these references, the effective reproduction is defined as the spectral radius of the matrix $\mathrm{Diag}(\eta) \cdot K$. Since the eigenvalues of $\mathrm{Diag}(\eta) \cdot K$ are exactly the eigenvalues of the matrix $K \cdot \mathrm{Diag}(\eta)$, this actually defines the same function $\mathfrak{R}_e$. In Section IV.2, we discuss the generalization of the effective reproduction number to the kernel model that offers a finer description of the contacts within the population.

The goal of this chapter is to prove a number of properties of $\mathfrak{R}_e$, that shed a light on how to vaccinate in the best possible way. In previous chapters, we introduced a general infinite-dimensional kernel framework in which the matrix formulation appears as a special finite-dimensional case. We state our results in this general framework, but for ease of presentation, we shall stick to the matrix formulation in this introduction. Finally, many results of this chapter are applied and illustrated in detail on various examples in the next chapters and in future papers [37, 40].

IV.1.2 Convexity properties of the effective reproduction number

Given the importance of convexity to solve optimization problems efficiently, it is natural to look for conditions on the matrix K that imply convexity or concavity for the map $\mathfrak{R}_e$ defined by (IV.1). In their investigation of the behavior of this map in the finite dimensional matrix setting, Hill and Longini conjecture in [81] sufficient spectral conditions to get either concavity or convexity. More precisely, guided by explicit examples, they state that $\mathfrak{R}_e$ should be convex if all the eigenvalues of K are non negative real numbers, and that it should be concave if all eigenvalues are real, with only one positive eigenvalue.

Our first series of results show that, while this conjecture cannot hold in full generality, see Section IV.5.1, it is true under an additional symmetry hypothesis. Recall that a matrix K is called diagonally symmetrizable if there exist positive numbers $(d_1, \dots d_N)$ such that for all i, j, $d_i K_{ij} = d_j K_{ji}$. Such a matrix is necessarily diagonalizable with real eigenvalues. The following result, which appears below in the text as Theorem IV.5.1, settles the conjecture for diagonally symmetrizable matrices. It is a special case of the more general Theorem IV.5.5, which holds in the infinite dimensional kernel setting, and for which the symmetry assumption has to be carefully worded. Let us mention that the eigenvalue λ_1 in the theorem below is non-negative and is equal to the spectral radius of K, that is, $\lambda_1 = \mathfrak{R}_e(\mathbb{1}) = \mathfrak{R}_0$, thanks to the Perron-Frobenius theory.

Theorem IV.1.1. *Let K be an $N \times N$ matrix with non-negative entries. Suppose that K is diagonally symmetrizable with eigenvalues $\lambda_1 \geq \lambda_2 \cdots \geq \lambda_N$.*

(i) If $\lambda_N \geq 0$, then the function $\mathfrak{R}_e$ is convex.

(ii) If $\lambda_2 \leq 0$, then the function $\mathfrak{R}_e$ is concave.

Note that the case (i) appears already in Cairns [29]; see also [57, 64] and Section IV.5.1 below for a detailed comparison with existing results.

IV.1.3 Equivalence properties

When studying the effective reproduction number $\mathfrak{R}_e$, it is natural to ask what kind of transformations may be done on the matrix K that leave the function $\mathfrak{R}_e$ unchanged. It is easy to see that if K and K' are diagonally similar up to transposition, they define the same function $\mathfrak{R}_e$; we check in Section IV.4.1 that this is essentially still true in the generalized kernel setting. We also investigate, in the matrix case, whether diagonal similarity up to transposition is necessary for defining the same $\mathfrak{R}_e$, and give partial results in this case, in the spirit of Hartfiel and Loewy [77]. In the terminology of linear algebra, this corresponds to identify all the preservers of the map $K \mapsto \mathfrak{R}_e[K]$, where we write $\mathfrak{R}_e[K]$ to stress the dependence of the function $\mathfrak{R}_e$ in (IV.1) on the parameter K.

IV.1.4 Properties of Pareto and anti-Pareto optima, cordons sanitaires

Let us now come back to the problem of finding optimal vaccination strategies. In contrast with the previous chapter, where we put minimal assumptions on the loss function which measures the efficiency of the vaccination strategies, we consider here that the loss of a strategy η is given by its effective reproduction number $\mathfrak{R}_e(\eta)$. This focus and the fact that we consider strictly decreasing cost functions (because vaccinating more costs more, see Section IV.6.1), allow us to simplify some of the statements of Chapter III and to give additional specific results.

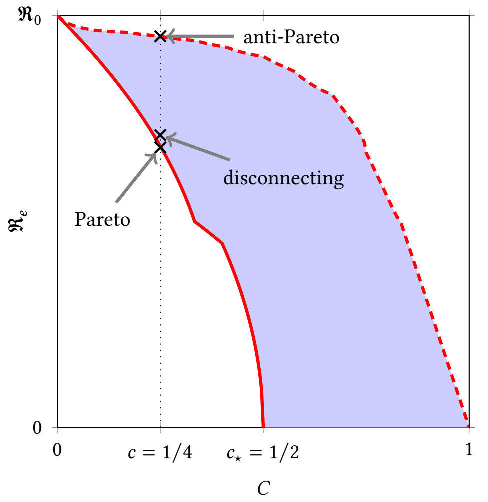
The problem of minimizing the effective reproduction number while keeping the cost of the vaccination low leads to a bi-objective optimization problem. We recall in Section IV.6.1 the setting introduced in detail in Chapter III for a general framework. One can identify Pareto optimal and anti-Pareto optimal vaccinations strategies, informally “best” and “worst” vaccination strategies, and consider the Pareto frontier $\mathcal{F}$ (resp. anti-Pareto frontier $\mathcal{F}^{\text{Anti}}$) as the outcomes $(C(\eta), \mathfrak{R}_e(\eta))$ of the Pareto (resp. anti-Pareto) optimal strategies η.

In Figure IV.1(a), we have plotted in red the Pareto frontier and in a dashed line the anti-Pareto frontier when the next-generation matrix is the adjacency matrix of the non-oriented cycle graph with $N = 12$ nodes from Figure IV.2(a) and Example IV.1.2.

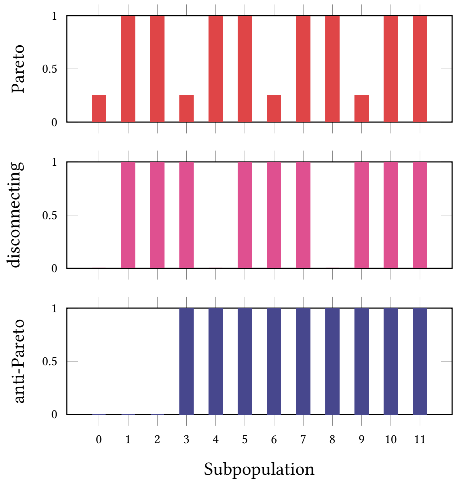
A cordon sanitaire is not the worst vaccination strategy

Recall that a matrix K is reducible if there exists a permutation σ such that $(K_{\sigma(i)\sigma(j)})_{i,j}$ is block upper triangular, and irreducible otherwise. A *cordon sanitaire* is a vaccination strategy η such that the infection matrix between non-vaccinated people, $K \cdot \text{Diag}(\eta)$, is reducible: informally, such a vaccination cuts the effective population in two or more groups that do not infect one another.

Disconnecting the population by creating a cordon sanitaire is not always the “best” choice, that is, it may not be Pareto optimal. However, we prove in Proposition IV.6.5 that a cordon



(A) Thick red line: Pareto frontier; Dashed line: anti-Pareto frontier; x marker: outcomes of various strategies.



(B) Profile of various strategies.

Figure IV.1: Efficacy of the disconnecting vaccination strategy "one in 4" for the non-oriented cycle graph with 12 nodes and uniform cost $1/4$.

sanitaire can never be anti-Pareto optimal; this result still holds in the general kernel framework, provided that the definition of cordon sanitaires is generalized in an appropriate way.

Example IV.1.2 (Non-oriented cycle graph). Suppose that the matrix K is given by the adjacency matrix (see Figure IV.2(B) for a grayplot representation) of the non-oriented cycle graph with $N = 12$ nodes; see Figure IV.2(A). For a cost $C_{\text{uni}} = 1/4$, there is a disconnecting strategy η that consists in vaccinating one sub-population in four; see Figure IV.2(C) (and Figure IV.2(D) for a grayplot representation of the corresponding adjacency matrix). The effective reproduction number associated is equal to $\sqrt{2}$. This strategies performs better than the anti-Pareto optimal strategy and is out-performed by the Pareto optimal one as we can see in Figure IV.1. This example is discussed in detail in Section 2.4 of Chapter V.

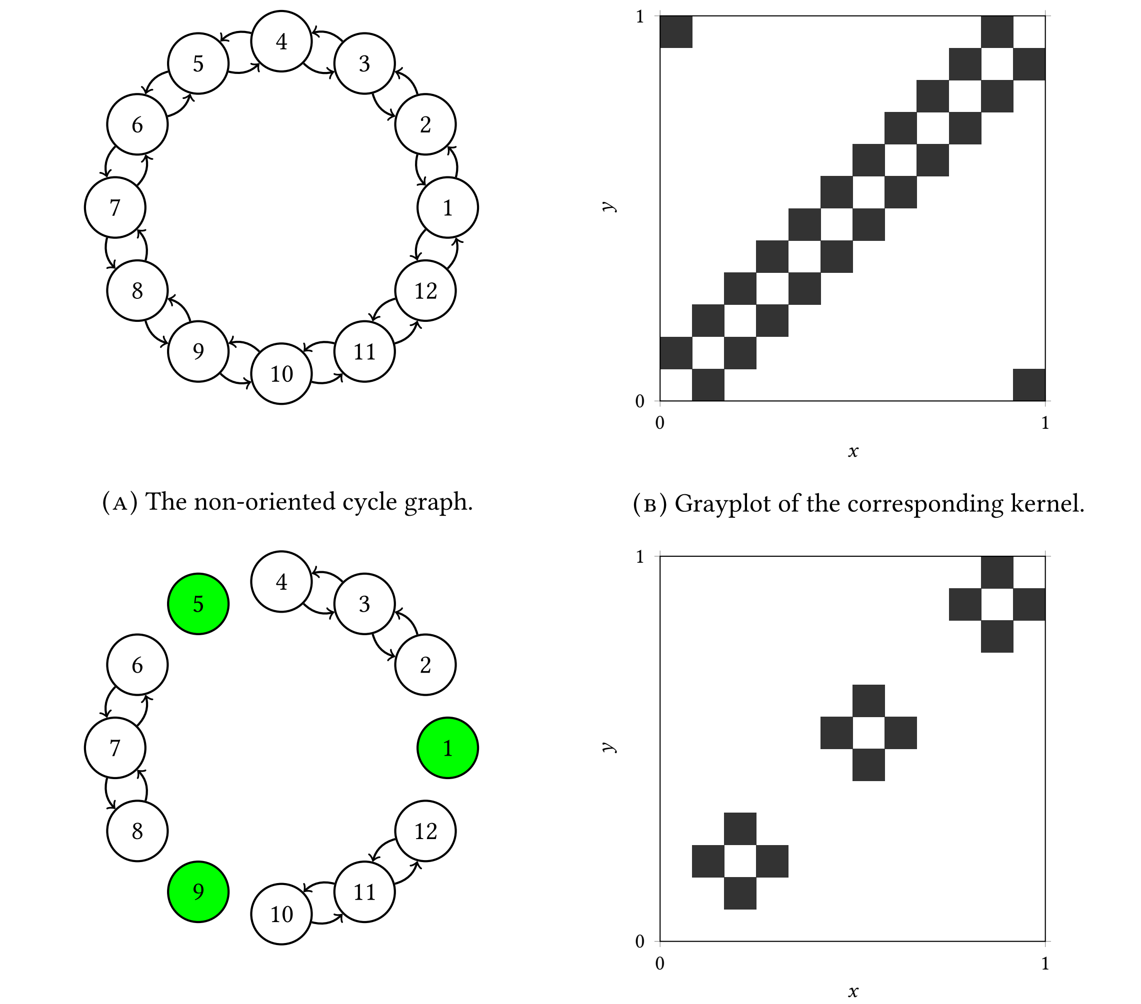
Minimal cost required to completely stop the transmission of the disease

A vaccination strategy η such that $\mathfrak{R}_e(\eta) = 0$ completely eradicates the epidemic. Section IV.6.4 is devoted to the characterization of the minimal cost of such vaccinations, which is denoted by $c_\star$. This quantity is introduced and discussed in Chapter III under general assumption for the loss function. Since we consider here the special case of measuring the loss by the effective reproduction number $\mathfrak{R}_e$, we are able to give in Proposition IV.6.9 an explicit expression of this quantity in the kernel model. In the symmetric matrix case, when the cost is uniform (the cost is proportional to the number of vaccinated individuals), this expression is proportional to the size of maximal independent sets of the non-oriented graph with vertices $\{1, \dots, N\}$, where there is an edge between i and j if and only if $K_{ij} > 0$.

We can observe this property in Figure IV.1(A) as the size of the maximal independent set of the non-oriented cycle graph of size N from Example IV.1.2 is equal to $\lfloor N/2 \rfloor$.

Reducible case

When the matrix K happens to be reducible, up to a relabeling, we may assume that it is block upper triangular. Denoting by m the number of blocks and $I_1, \dots, I_m$ the sets of indices describing the blocks, this means that for all $\ell > k$ and $(i, j) \in I_\ell \times I_k$, we have $K_{ij} = 0$. In the epidemiological interpretation, this means that the populations with indices in I_k never infect the ones with indices in I_ℓ. One may then hope that the study of $\mathfrak{R}_e$ can be effectively reduced to the study of the effective



(A) The non-oriented cycle graph.

(B) Grayplot of the corresponding kernel.

(C) Cordon sanitaire corresponding to the "one in 4" vaccination strategy (in green the vaccinated groups).

(D) Grayplot of the corresponding kernel.

Figure IV.2: An example of disconnecting vaccination strategy on the non-oriented cycle graph.

radius of the square sub-matrices $(K_{ij})_{i,j\in I_k}$ describing the infections within block I_k. This is indeed the case, and we give in Section IV.7 a complete picture of the Pareto and anti-Pareto frontiers of $\mathfrak{R}_e$, in terms of the effective reproduction numbers restricted to each irreducible component of the infection kernel or matrix. In particular, this allows a better understanding of the possible disconnection of the anti-Pareto frontier, whereas the Pareto frontier is always connected. Once more, special care has to be taken with the definitions when handling the infinite dimensional kernel case.

Optimal ray

It is observed by Poghotanyan, Feng, Glasser and Hill in [130], that if there exists a Pareto optimal strategy η with all its entries strictly less than 1, then all the strategies $\lambda\eta$, with $\lambda \geq 0$ such that $\lambda\eta \in \Delta$, are Pareto optimal. We give a short proof on the existence of such optimal rays in Section IV.6.2, when one assumes that the cost function C is affine on Δ.

IV.1.5 Structure of the chapter

We discuss in Section IV.2 the generality of the setting, showing that studying vaccination strategies in many different epidemic models gives rise to the same optimization problem. After recalling formally our infinite dimensional kernel setting in Section IV.3, we discuss invariance properties of $\mathfrak{R}_e$ in Section IV.4. The convexity properties of $\mathfrak{R}_e$ and the related conjecture of Hill and Longini are discussed in Section IV.5. Various properties of the Pareto and anti-Pareto frontiers, and in particular the fact that establishing a *cordon sanitaire* by disconnecting the population is never the worst solution, are discussed in Section IV.6. Finally, the case of reducible kernels is treated in Section IV.7.

IV.2 Discussion on the next-generation operator

In Chapters II and III, we developed a framework that we call the kernel model where the population is represented as an abstract probability space $(\Omega, \mathscr{F}, \mu)$. Individuals are characterized by a feature $x \in \Omega$, and the relative size of the sub-population with feature is given by $\mu(\mathrm{d}x)$. The underlying structure described by this trait can be very varied, typical examples being one or several of the following characteristics: spatial position, social contacts, susceptibility, infectivity, characteristics of the immunological response...The analogue of the next-generation matrix K is then the kernel operator defined formally by:

$$T_{\mathrm{k}}(g)(x) = \int_\Omega \mathrm{k}(x, y)\, g(y)\, \mathrm{d}\mu(y);$$

where the nonnegative kernel k is defined on $\Omega \times \Omega$ and $\mathrm{k}(x, y)$ still represents a strength of infection from y to x. Vaccination strategies $\eta : \Omega \to [0, 1]$ encode the density of non-vaccinated individuals with respect to the measure μ. The (sub-probability) measure $\eta(y)\, \mu(\mathrm{d}y)$ may then be understood as an effective population, giving rise to an effective next-generation operator:

$$T_{\mathrm{k}\eta}(g)(x) = \int_\Omega \mathrm{k}(x, y)\, g(y)\, \eta(y)\, \mu(\mathrm{d}y).$$

The effective reproduction number is then defined by $\mathfrak{R}_e(\eta) = \rho(T_{\mathrm{k}\eta})$, where ρ stands for the spectral radius of the operator and $\mathrm{k}\eta$ for the kernel $(\mathrm{k}\eta)(x, y) = \mathrm{k}(x, y)\eta(y)$.

Most of the results mentioned in the introduction will be given in this general framework as we argue that the latter is sufficiently flexible to describe a wide range of epidemic models from the literature including the metapopulation models. We give in the following a few examples to support this claim: in each of them, the spectral radius of a particular, explicit kernel operator appears as a threshold parameter, and the epidemic either "invades/survives" or "dies out" depending on the value of this parameter. Classical notations are used: S denotes the proportion of susceptible individuals, E the proportion of those who have been exposed to the disease, I the proportion of infected individuals, R the proportion of removed individuals in the population.

Example IV.2.1 (Meta-population models). Recall that in metapopulation models, the population is divided into $N \geq 2$ different sub-populations of respective proportional size $\mu_1, \dots, \mu_N$, and the reproduction number is given by $\mathfrak{R}_e(\eta) = \rho(K \cdot \mathrm{Diag}(\eta))$, where K is the next generation matrix and η belongs to $[0, 1]^N$ and gives the proportion of non-vaccinated individuals in each sub-population. To express the function $\mathfrak{R}_e$ as the effective reproduction number of a kernel model, consider the discrete state space $\Omega_{\mathrm{d}} = \{1, \dots, N\}$ equipped with the probability measure μ_{d} defined by $\mu_{\mathrm{d}}(\{i\}) = \mu_i$, and let $\mathrm{k_d}$ denote the discrete kernel on Ω_{d} defined by:

$$\mathrm{k_d}(i, j) = K_{ij}/\mu_j. \tag{IV.2}$$

For all $\eta \in \Delta = [0, 1]^N$, the matrix $K \cdot \mathrm{Diag}(\eta)$ is the matrix representation of the endomorphism $T_{\mathrm{k_d}\eta}$ in the canonical basis of $\mathbb{R}^N$. In particular, we have: $\mathfrak{R}_e(\eta) = \rho(T_{\mathrm{k}\eta}) = \rho(K \cdot \mathrm{Diag}(\eta))$.

In Figure IV.2(B), we have plotted the kernel on $[0, 1]$ associated to $\mathrm{k_d}$ for the non-oriented cycle graph when the sub-populations have the same size.

Example IV.2.2 (An SIR model with nonlinear incidence rate and vital dynamics). In [150], Thieme proposed an SIR model in an infinite-dimensional population structure with a nonlinear incidence rate. The structure space is given by Ω a compact subset of $\mathbb{R}^N$ and it is equipped with the normalized Lebesgue measure. We restrict slightly his assumption so that the incidence rate is a linear function of the number of susceptible. The dynamic of the epidemic then writes:

$$\text{For } t \geq 0,\ x \in \Omega, \qquad \begin{cases} \partial_t S(t,x) = \Lambda(x) - \nu(x)S(t,x) - S(t,x)\int_\Omega f(I(t,y),x,y)\,\mathrm{d}y, \\[2ex] \partial_t I(t,x) = S(t,x)\int_\Omega f(I(t,y),x,y)\,\mathrm{d}y - (\gamma(x)+\nu(x))I(t,x), \\[2ex] \partial_t R(t,x) = \gamma(x)I(t,x). \end{cases} \tag{IV.3}$$

Here $\Lambda(x)$ is the rate at which fresh susceptibles are recruited into the population at location x, $\nu(x)$ is the *per capita* death rate of the individuals, and $\gamma(x)$ is the *per capita* recovery rate of infectious individuals The integral term describes the incidence at x at time t, *i.e.*, the rate of new infections. Thieme identified a threshold parameter that plays the role of the reproduction number, and is given by the spectral radius of the operator T_{k} with the kernel given by:

$$\mathrm{k}(x,y) = \frac{\Lambda(x)}{\gamma(x)+\nu(x)}\partial_I f(0,x,y), \quad x,y \in \Omega, \tag{IV.4}$$

where $\partial_I f(0,x,y)$, the derivative of f with respect to I, is supposed to be non-negative for all $x,y \in \Omega$.

Suppose that individuals at location x are vaccinated with probability $1-\eta(x)$ at birth so that the susceptible individuals with feature x are recruited at rate $\eta(x)\Lambda(x)$ and recovered/immunized individuals are also recruited at rate $(1-\eta(x))\Lambda(x)$ at location x. The threshold parameter $\mathfrak{R}_e(\eta)$ is then given by the spectral radius of the integral operator $T_{\eta\mathrm{k}}$ with kernel $\eta\mathrm{k}$ given by $(\eta\mathrm{k})(x,y) = \eta(x)\mathrm{k}(x,y)$. According to Lemma IV.3.1 (ii), we have $\rho(T_{\eta\mathrm{k}}) = \rho(T_{\mathrm{k}\eta})$, and our framework can be used for this model.

Under regularity assumptions on the parameters of the model, Thieme proved that if $\mathfrak{R}_e(\eta)$ is greater than 1, then there exists an endemic equilibrium that attracts all the solutions while if $\mathfrak{R}_e(\eta)$ is smaller than 1, then $I(t,x)$ converges to 0 for all $x \in \Omega$ as t goes to infinity.

Example IV.2.3 (An SEIR model without vital dynamics). In [3], Almeida, Bliman, Nadin and Perthame studied an heterogeneous SEIR model where the population is again structured with a bounded subset $\Omega \subset \mathbb{R}^N$ equipped with the normalized Lebesgue measure. The dynamic of the susceptible, exposed, infected and recovered individuals writes:

$$\text{For } t \geq 0,\ x \in \Omega, \qquad \begin{cases} \partial_t S(t,x) = -S(t,x)\int_\Omega k(x,y)I(t,y)\,\mathrm{d}y, \\[2ex] \partial_t E(t,x) = S(t,x)\int_\Omega k(x,y)I(t,y)\,\mathrm{d}y - \alpha(x)E(t,x), \\[2ex] \partial_t I(t,x) = \alpha(x)E(t,x) - \gamma(x)I(t,x), \\[2ex] \partial_t R(t,x) = \gamma(x)I(t,x). \end{cases} \tag{IV.5}$$

Here, the average incubation rate is denoted by $\alpha(x)$ and the average recovery rate by $\gamma(x)$; both quantities may depend upon the trait x. The function k is the transmission kernel of the disease. In this model, the basic reproduction number is given by the spectral radius of the integral operator T_{k} with kernel $\mathrm{k} = k/\gamma$:

$$\mathrm{k}(x,y) = k(x,y)/\gamma(y). \tag{IV.6}$$

Suppose that, prior to the beginning of the epidemic, the decision maker immunizes a density $1-\eta$ of individuals. Then the effective reproduction is given by $\rho(T_{\eta\mathrm{k}})$ which is also equal to $\rho(T_{\mathrm{k}\eta})$, see Lemma IV.3.1 (ii) below, and our model is indeed suitable for studying the vaccination strategies in this context.

Example IV.2.4 (An SIS model without vital dynamic). In Chapter II, we introduced the following heterogeneous SIS model where the population is structured with an abstract probability space $(\Omega, \mathscr{F}, \mu)$:

$$\text{For } t \geq 0,\, x \in \Omega, \qquad \begin{cases} \partial_t S(t,x) = -S(t,x) \int_\Omega k(x,y) I(t,y)\, \mathrm{d}y + \gamma(x) I(t,x), \\ \\ \partial_t I(t,x) = S(t,x) \int_\Omega k(x,y) I(t,y)\, \mathrm{d}y - \gamma(x) I(t,x). \end{cases} \tag{IV.7}$$

The function γ is the *per-capita* recovery rate and k is the transmission kernel. For this model, $\mathfrak{R}_e(\eta) = \rho(T_{\mathrm{k}\eta})$ where $\mathrm{k} = k/\gamma$ is defined by (IV.6).

Suppose that, prior to the beginning of the epidemic, a density $1-\eta$ of individuals is vaccinated with a perfect vaccine. In the same way as for the SEIR model, we proved, as t goes to infinity, that if $\mathfrak{R}_e(\eta)$ is smaller than or equal to 1, then $I(t,\cdot)$ converges to 0, and, under a connectivity assumption on the kernel k, that if $\mathfrak{R}_e(\eta)$ is greater than 1, then $I(t,\cdot)$ converges to an endemic equilibrium. This highlights the importance of $\mathfrak{R}_e$ in the design of vaccination strategies.

IV.3 Setting, notations and previous results

IV.3.1 Spaces, operators, spectra

All metric spaces (S,d) are endowed with their Borel σ-field denoted by $\mathscr{B}(S)$. The set $\mathscr{K}$ of compact subsets of $\mathbb{C}$ endowed with the Hausdorff distance d_{H} is a metric space, and the function rad from $\mathscr{K}$ to $\mathbb{R}_+$ defined by $\mathrm{rad}(K) = \max\{|\lambda|,\, \lambda \in K\}$ is Lipschitz continuous from $(\mathscr{K}, d_{\mathrm{H}})$ to $\mathbb{R}$ endowed with its usual Euclidean distance.

Let $(\Omega, \mathscr{F}, \mu)$ be a probability space. We denote by Δ the set of $[0,1]$-valued measurable functions defined on Ω. For f and g real-valued functions defined on Ω, we may write $\langle f, g\rangle$ or $\int_\Omega f g\, \mathrm{d}\mu$ for $\int_\Omega f(x) g(x)\, \mu(\mathrm{d}x)$ whenever the latter is meaningful. For $p \in [1, +\infty]$, we denote by $L^p = L^p(\mu) = L^p(\Omega, \mu)$ the space of real-valued measurable functions g defined Ω such that $\| g \|_p = \left(\int |g|^p \,\mathrm{d}\mu\right)^{1/p}$ (with the convention that $\| g \|_\infty$ is the μ-essential supremum of $|g|$) is finite, where functions which agree μ-a.s. are identified. We denote by L^p_+ the subset of L^p of non-negative functions.

Let $(E, \|\cdot\|)$ be a Banach space. We denote by $\|\cdot\|_E$ the operator norm on $\mathcal{L}(E)$ the Banach algebra of bounded operators. The spectrum $\mathrm{Spec}(T)$ of $T \in \mathcal{L}(E)$ is the set of $\lambda \in \mathbb{C}$ such that $T - \lambda \mathrm{Id}$ does not have a bounded inverse operator, where Id is the identity operator on E. Recall that $\mathrm{Spec}(T)$ is a compact subset of $\mathbb{C}$, and that the spectral radius of T is given by:

$$\rho(T) = \mathrm{rad}(\mathrm{Spec}(T)) = \lim_{n\to\infty} \| T^n \|_E^{1/n}. \tag{IV.8}$$

The element $\lambda \in \mathrm{Spec}(T)$ is an eigenvalue if there exists $x \in E$ such that $Tx = \lambda x$ and $x \neq 0$. Following [99], we define the (algebraic) multiplicity of $\lambda \in \mathbb{C}$ by:

$$\mathrm{m}(\lambda, T) = \dim \left(\bigcup_{k \in \mathbb{N}^*} \ker(T - \lambda \mathrm{Id})^k \right),$$

so that λ is an eigenvalue if $\mathrm{m}(\lambda, T) \geq 1$. We say the eigenvalue λ of T is *simple* if $\mathrm{m}(\lambda, T) = 1$.

If E is also an algebra, for $g \in E$, we denote by M_g the multiplication (possibly unbounded) operator defined by $M_g(h) = gh$ for all $h \in E$.

IV.3.2 Invariance and continuity of the spectrum for compact operators

We collect some known results on the spectrum and multiplicity of eigenvalues related to compact operators. Let $(E, \|\cdot\|)$ be a Banach space. Let $A \in \mathcal{L}(E)$. We denote by $A^\top$ the adjoint of A. A sequence $(A_n, n \in \mathbb{N})$ of elements of $\mathcal{L}(E)$ converges strongly to $A \in \mathcal{L}(E)$ if $\lim_{n\to\infty} \| A_n x - Ax \| = 0$ for all $x \in E$. Following [6], a set of operators $\mathscr{A} \subset \mathcal{L}(E)$ is *collectively compact* if the set $\{Ax :\ A \in \mathscr{A},\, \|x\| \leq 1\}$ is relatively compact. Recall that the spectrum of a compact operator is finite or countable and has at most one accumulation point, which is 0. Furthermore, 0 belongs to the spectrum of compact operators in infinite dimension. We refer to [136] for an introduction to Banach lattices and positive operators; we shall only consider the Banach lattices $L^p(\Omega, \mu)$ for $p \geq 1$ on a probability space $(\Omega, \mathscr{F}, \mu)$ and a bounded operator A is positive if $A(L^p_+) \subset L^p_+$.

Lemma IV.3.1. *Let A, B be elements of $\mathcal{L}(E)$.*

(i) If E is a Banach lattice, and if A, B and $A - B$ are positive operators, then we have:

$$\rho(A) \geq \rho(B). \tag{IV.9}$$

(ii) If A is compact, then we have AB and BA compact and:

$$\mathrm{Spec}(A) = \mathrm{Spec}(A^\top) \quad \text{and} \quad \mathrm{m}(\lambda, A) = \mathrm{m}(\lambda, A^\top) \quad \text{for } \lambda \in \mathbb{C}^*, \tag{IV.10}$$

$$\mathrm{Spec}(AB) = \mathrm{Spec}(BA) \quad \text{and} \quad \mathrm{m}(\lambda, AB) = \mathrm{m}(\lambda, BA) \quad \text{for } \lambda \in \mathbb{C}^*, \tag{IV.11}$$

and in particular:

$$\rho(AB) = \rho(BA). \tag{IV.12}$$

(iii) Let $(E', \|\cdot\|')$ be a Banach space such that E' is continuously and densely embedded in E. Assume that $A(E') \subset E'$, and denote by A' the restriction of A to E' seen as an operator on E'. If A and A' are compact, then we have:

$$\mathrm{Spec}(A) = \mathrm{Spec}(A') \quad \text{and} \quad \mathrm{m}(\lambda, A) = \mathrm{m}(\lambda, A') \quad \text{for } \lambda \in \mathbb{C}^*. \tag{IV.13}$$

(iv) Let $(A_n, n \in \mathbb{N})$ be a collectively compact sequence which converges strongly to A. Then, we have $\lim_{n\to\infty} \mathrm{Spec}(A_n) = \mathrm{Spec}(A)$ in $(\mathscr{K}, d_{\mathrm{H}})$, $\lim_{n\to} \rho(T_n) = \rho(T)$ and for $\lambda \in \mathrm{Spec}(A) \cap \mathbb{C}^*$, $r > 0$ such that $\lambda' \in \mathrm{Spec}(A)$ and $|\lambda - \lambda'| \leq r$ implies $\lambda = \lambda'$, and all n large enough:

$$\mathrm{m}(\lambda, A) = \sum_{\lambda' \in \mathrm{Spec}(A_n),\, |\lambda - \lambda'| \leq r} \mathrm{m}(\lambda', A_n). \tag{IV.14}$$

Proof. Property (i) can be found in [114, Theorem 4.2]. Equation (IV.10) from Property (ii) can be deduced from from [99, Theorem p. 20]. Using [99, Proposition p. 25], we get the second part of (IV.11) and $\mathrm{Spec}(AB) \cap \mathbb{C}^* = \mathrm{Spec}(BA) \cap \mathbb{C}^*$, and thus (IV.12) holds. To get the first part of (IV.11), see Lemma III.3.2.

We now provide a short proof for Property (iii). According to [72, Corollary 1 and Section 6], we have $\mathrm{Spec}(A) = \mathrm{Spec}(A')$. Let $\lambda \in \mathrm{Spec}(A) \cap \mathbb{C}^*$. Since the multiplicity of λ for A is finite, we get that $\mathrm{m}(\lambda, A) = \dim(\ker(A - \lambda \mathrm{Id})^n)$ for n large enough, and similarly for $\mathrm{m}(\lambda, A')$. Clearly, we have $\ker(A' - \lambda\mathrm{Id})^n \subset \ker(A - \lambda\mathrm{Id})^n$. Let us prove that $\ker(A - \lambda\mathrm{Id})^n \subset \ker(A' - \lambda\mathrm{Id})^n$. Let $x \in \ker(A - \lambda\mathrm{Id})^n$ and $(x_\ell, \ell \in \mathbb{N})$ be a sequence of elements of E' which converges (in E) towards x. Up to taking a sub-sequence, since A' is compact, we can assume that $A' x_\ell$ converges in E', say towards $y \in E'$. We deduce that:

$$\begin{aligned}
\lambda^n x &= \sum_{k=1}^n \binom{n}{k} (-\lambda)^{n-k+1} A^k x \\
&= \lim_{\ell\to\infty} \sum_{k=1}^n \binom{n}{k} (-\lambda)^{n-k+1} A^k x_\ell \\
&= \lim_{\ell\to\infty} \sum_{k=1}^n \binom{n}{k} (-\lambda)^{n-k+1} (A')^{k-1} (A' x_\ell) \\
&= \sum_{k=1}^n \binom{n}{k} (-\lambda)^{n-k+1} (A')^{k-1} y.
\end{aligned}$$

Since $\lambda \neq 0$, we get that x belongs to E' and thus $(A' - \lambda\mathrm{Id})^n x = (A - \lambda\mathrm{Id})^n x = 0$, that is $\ker(A - \lambda\mathrm{Id})^n \subset \ker(A' - \lambda\mathrm{Id})^n$. Then use the definition of the multiplicity to conclude.

We eventually check Point (iv). We deduce from [6, Theorems 4.8 and 4.16] (see also (d), (g) [take care that $d(\lambda, K)$ therein is the algebraic multiplicity of φ for the compact operator K and not the geometric multiplicity] and (e) in [7, Section 3]) that $\lim_{n\to\infty} \mathrm{Spec}(T_n) = \mathrm{Spec}(T)$ and (IV.14). Then use that the function rad is continuous to deduce the convergence of the spectral radius from the convergence of the spectra. □

IV.3.3 Kernel operators

We define a *kernel* (resp. *signed kernel*) on Ω as a $\mathbb{R}_+$-valued (resp. $\mathbb{R}$-valued) measurable function defined on $(\Omega^2, \mathscr{F}^{\otimes 2})$. For f, g two non-negative measurable functions defined on Ω and k a kernel on Ω, we denote by $f\mathrm{k}g$ the kernel defined by:

$$f\mathrm{k}g \,:\, (x,y) \mapsto f(x)\,\mathrm{k}(x,y)g(y). \tag{IV.15}$$

For $p \in (1, +\infty)$, we define the double norm of a signed kernel k on L^p by:

$$\|\mathrm{k}\|_{p,q} = \left(\int_\Omega \left(\int_\Omega |\mathrm{k}(x,y)|^q \, \mu(\mathrm{d}y) \right)^{p/q} \mu(\mathrm{d}x) \right)^{1/p} \quad \text{with } q \text{ given by } \quad \frac{1}{p} + \frac{1}{q} = 1. \tag{IV.16}$$

We say that k has a finite double norm, if there exists $p \in (1, +\infty)$ such that $\|\mathrm{k}\|_{p,q} < +\infty$. To such a kernel k, we then associate the positive integral operator T_k on L^p defined by:

$$T_\mathrm{k}(g)(x) = \int_\Omega \mathrm{k}(x,y)\, g(y)\, \mu(\mathrm{d}y) \quad \text{for } g \in L^p \text{ and } x \in \Omega. \tag{IV.17}$$

According to [68, p. 293], T_k is compact. It is well known and easy to check that:

$$\|T_\mathrm{k}\|_{L^p} \leq \|\mathrm{k}\|_{p,q}\,. \tag{IV.18}$$

We define the *reproduction number* associated to the operator T_k as:

$$\mathfrak{R}_0[\mathrm{k}] = \rho(T_\mathrm{k}). \tag{IV.19}$$

The proof of the next stability result appears already in Chapter III (but for (IV.20) whose proof relies on (IV.14) and is left to the reader).

Corollary IV.3.2. *Let $p \in (1, +\infty)$. Let $(\mathrm{k}_n, n \in \mathbb{N})$ and k be kernels on Ω with finite double norms on L^p such that $\lim_{n\to\infty} \|\mathrm{k}_n - \mathrm{k}\|_{p,q} = 0$. Then, we have $\lim_{n\to\infty} \mathrm{Spec}(T_{\mathrm{k}_n}) = \mathrm{Spec}(T_\mathrm{k})$ in $(\mathscr{K}, d_\mathrm{H})$, $\lim_{n\to} \rho(T_{\mathrm{k}_n}) = \rho(T_\mathrm{k})$ and for $\lambda \in \mathrm{Spec}(T_\mathrm{k}) \cap \mathbb{C}^*$, $r > 0$ such that $\lambda' \in \mathrm{Spec}(T_\mathrm{k})$ and $|\lambda - \lambda'| \leq r$ implies $\lambda = \lambda'$, and all n large enough:*

$$\mathrm{m}(\lambda, T_\mathrm{k}) = \sum_{\lambda' \in \mathrm{Spec}(T_{\mathrm{k}_n}),\, |\lambda - \lambda'| \leq r} \mathrm{m}(\lambda', T_{\mathrm{k}_n}). \tag{IV.20}$$

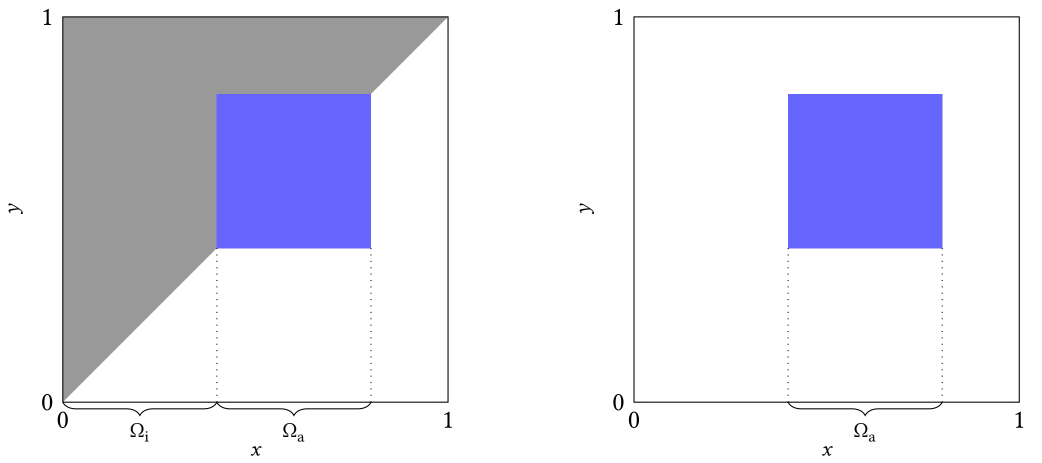
IV.3.4 Irreducibility, quasi-irreducibility and monatomic kernel

We first define *irreducible* and *monatomic* kernels. For $A, B \in \mathscr{F}$, we write $A \subset B$ a.s. if $\mu(B^\mathsf{c} \cap A) = 0$ and $A = B$ a.s. if $A \subset B$ a.s. and $B \subset A$ a.s. For $A, B \in \mathscr{F}$, $x \in \Omega$ and an integrable kernel k, we simply write $\mathrm{k}(x, A) = \int_A \mathrm{k}(x,y)\, \mu(\mathrm{d}y)$, $\mathrm{k}(B, x) = \int_B \mathrm{k}(z,x)\, \mu(\mathrm{d}z)$ and:

$$\mathrm{k}(B, A) = \int_{B\times A} \mathrm{k}(z,y)\, \mu(\mathrm{d}z)\mu(\mathrm{d}y).$$

A set $A \subset \mathscr{F}$ is k-*invariant*, or simply *invariant* when there is no ambiguity on the kernel k, if $\mathrm{k}(A^\mathsf{c}, A) = 0$. In the epidemiological setting, the set A is invariant if the sub-population A does not infect the sub-population A^c. If k is symmetric, then A is invariant if and only if A^c is invariant.

A kernel k is said to be *irreducible* or *connected* if any k-invariant set A is such that $A = \emptyset$ a.s. or $A = \Omega$ a.s. According to [136, Theorem V.6.6], if k is an irreducible kernel with finite double norm, then we have $\mathfrak{R}_0[\mathrm{k}] > 0$. If the kernel is positive a.s., then it is irreducible. Following [17, Definition 2.11], we say that a kernel is *quasi-irreducible* if k restricted to $\{\mathrm{k} \equiv 0\}^\mathsf{c}$, with $\{\mathrm{k} \equiv 0\} = \{x \in \Omega \,:\, \mathrm{k}(x,\Omega) + \mathrm{k}(\Omega, x) = 0\}$, is irreducible. The quasi-irreducible property was introduced for symmetric kernel; for general kernel one can consider the following weaker property. A kernel k is said to be *monatomic* if the operator T_k has a unique (up to a multiplicative constant) non-negative eigenfunction. Intuitively, this corresponds to have only one irreducible component. Formally, this is also equivalent to the following two properties:



(A) A representation of a monatomic kernel. (B) A representation of a quasi-irreducible kernel.

Figure IV.3: Example of monatomic and quasi-irreducible kernels $(x, y) \mapsto \mathrm{k}(x, y)$, where $\mathrm{k}(x, y) = 0$ on the white zone and k reduced to the blue zone is irreducible.

(i) There exists a measurable subset $\Omega_a \subset \Omega$, the irreducible component or *atom* such that:

- $\mu(\Omega_a) > 0$ and the kernel k restricted to Ω_a is irreducible.
- If $\Omega_a^c \neq \emptyset$ a.s. then the restriction of T_k to Ω_a^c is quasi-nilpotent, that is, $\mathfrak{R}_e[\mathrm{k}](\mathbb{1}_{\Omega_a^c}) = 0$.

(ii) There exists a measurable subset $\Omega_i \subset \Omega_a^c$, "the sub-population infected by" Ω_a such that:

- The sets $\Omega_a \cup \Omega_i$ and Ω_i are invariant.
- The set Ω_i is the minimal set such that $\Omega_a \cup \Omega_i$ is invariant: if A is invariant and $\Omega_a \subset A$ then $\Omega_i \subset A$ a.s.

In the epidemiological setting, the sub-population Ω_i can only infect itself, and the sub-population Ω_a infects only itself and Ω_i. The set $\Omega_a \cup \Omega_i$ corresponds to the support of the endemic equilibrium in the supercritical regime; see Lemma III.5.12. We refer to [138] for further details on the decomposition of a kernel on its irreducible components; in particular the sets Ω_a and Ω_i are unique up to the a.s. equivalence. We represented in Figure IV.3(A) a monatomic kernel and in Figure IV.3(B) a quasi-irreducible kernel; the set Ω being "nicely ordered" so that the representation of the kernels are upper triangular.

Remark IV.3.3. Irreducible and quasi-irreducible kernels are also monatomic (take $\Omega_a = \{\mathrm{k} \equiv 0\}^c$ and $\Omega_i = \emptyset$). If the kernel k is monatomic and symmetric, then we get $\mathrm{k} = \mathbb{1}_{\Omega_a} \mathrm{k} \, \mathbb{1}_{\Omega_a}$ and thus the kernel k is quasi-irreducible.

The notion of irreducibility of a kernel depends only on its support: the kernel k is irreducible (resp. quasi-irreducible, resp. monatomic) if and only if the kernel $\mathbb{1}_{\{\mathrm{k}>0\}}$ is irreducible (resp. quasi-irreducible, resp. monatomic). Furthermore, if k is monatomic, then the kernels k and $\mathbb{1}_{\{\mathrm{k}>0\}}$ have the same atom Ω_a and the same set Ω_i infected by Ω_a.

The introduction of monatomic kernel is also motivated by the following result which can be deduced from [136, Theorem V.6.6] and [138, Theorem 8], see also Section IV.7.

Lemma IV.3.4. *Let* k *be a kernel with finite double norm and set* $\mathfrak{R}_0 = \mathfrak{R}_0[\mathrm{k}]$. *If the kernel* k *is monatomic then* $\mathfrak{R}_0 > 0$ *and* $\mathfrak{R}_0$ *is simple* (i.e. $\mathrm{m}(\mathfrak{R}_0, T_k) = 1$)*. If* $\mathfrak{R}_0$ *is simple and the only eigenvalue in* $(0, +\infty)$*, then the kernel* k *is monatomic.*

IV.3.5 The effective reproduction number $\mathfrak{R}_e$

A *vaccination strategy* η of a vaccine with perfect efficiency is an element of Δ, where $\eta(x)$ represents the proportion of non-vaccinated individuals with feature x. In particular $\eta = \mathbb{1}$ (the constant function equal to 1) corresponds to no vaccination and $\eta = \mathbb{0}$ (the constant function equal to 0) corresponds to the whole population vaccinated. Notice that $\eta \, d\mu$ corresponds in a sense to the effective population. Let k be a kernel on Ω with finite double norm on L^p. For $\eta \in \Delta$, the kernel $\mathrm{k}\eta$ has also a finite double norm on L^p and the operator M_η is bounded, so that the operator $T_{\mathrm{k}\eta} = T_\mathrm{k} M_\eta$ is compact. We can define the *effective spectrum* function Spec[k] from Δ to $\mathscr{K}$ by:

$$\mathrm{Spec}[\mathrm{k}](\eta) = \mathrm{Spec}(T_{\mathrm{k}\eta}), \tag{IV.21}$$

the *effective reproduction number* function $\mathfrak{R}_e[\mathrm{k}] = \mathrm{rad} \circ \mathrm{Spec}[\mathrm{k}]$ from Δ to $\mathbb{R}_+$ by:

$$\mathfrak{R}_e[\mathrm{k}](\eta) = \mathrm{rad}(\mathrm{Spec}(T_{\mathrm{k}\eta})) = \rho(T_{\mathrm{k}\eta}), \tag{IV.22}$$

and the corresponding reproduction number is then given by $\mathfrak{R}_0[\mathrm{k}] = \mathfrak{R}_e[\mathrm{k}](\mathbb{1})$. When there is no risk of confusion on the kernel k, we simply write $\mathfrak{R}_e$ and $\mathfrak{R}_0$ for the function $\mathfrak{R}_e[\mathrm{k}]$ and the number $\mathfrak{R}_0[\mathrm{k}]$.

We can see Δ as a subset of L^1, and consider the corresponding *weak topology*: we say that a sequence $(g_n, n \in \mathbb{N})$ of functions in Δ converges weakly to g if for all $h \in L^\infty$ we have:

$$\lim_{n\to\infty} \int_\Omega h g_n \, d\mu = \int_\Omega h g \, d\mu. \tag{IV.23}$$

Notice that (IV.23) can easily be extended to any function $h \in L^q$ for any $q \in (1, +\infty)$; so that the weak-topology on Δ, seen as a subset of L^p with $1/p + 1/q = 1$, can be seen as the trace on Δ of the weak topology on L^p. We recall the following property Lemma III.3.1 that follows from the Banach-Alaoglu theorem.

Lemma IV.3.5 (Topological properties of Δ)**.** *The set Δ endowed with the weak topology is compact and sequentially compact.*

We also recall the properties of the effective reproduction number given in Proposition III.4.1 and Theorem III.4.2.

Proposition IV.3.6. *Let* k *be a kernel on a probability space $(\Omega, \mathscr{F} \mu)$ with finite double norm. Then, the functions* Spec[k] *and $\mathfrak{R}_e = \mathfrak{R}_e[\mathrm{k}]$ are continuous functions from Δ respectively to $\mathscr{K}$ (endowed with the Hausdorff distance) and to $\mathbb{R}_+$. Furthermore, the function $\mathfrak{R}_e = \mathfrak{R}_e[\mathrm{k}]$ satisfies the following properties:*

(i) $\mathfrak{R}_e(\eta_1) = \mathfrak{R}_e(\eta_2)$ *if* $\eta_1 = \eta_2$, μ *a.s., and* $\eta_1, \eta_2 \in \Delta$,

(ii) $\mathfrak{R}_e(\mathbb{0}) = 0$ *and* $\mathfrak{R}_e(\mathbb{1}) = \mathfrak{R}_0$,

(iii) $\mathfrak{R}_e(\eta_1) \leq \mathfrak{R}_e(\eta_2)$ *for all* $\eta_1, \eta_2 \in \Delta$ *such that* $\eta_1 \leq \eta_2$,

(iv) $\mathfrak{R}_e(\lambda\eta) = \lambda\mathfrak{R}_e(\eta)$, *for all* $\eta \in \Delta$ *and* $\lambda \in [0, 1]$.

We complete Corollary IV.3.2 on the stability property of the spectrum and spectral radius with respect to the kernel k; see Proposition III.4.3.

Proposition IV.3.7 (Stability of $\mathfrak{R}_e[\mathrm{k}]$ and Spec[k])**.** *Let $p \in (1, +\infty)$. Let $(\mathrm{k}_n, n \in \mathbb{N})$ and* k *be kernels on Ω with finite double norms on L^p. If $\lim_{n\to\infty} \| \mathrm{k}_n - \mathrm{k} \|_{p,q} = 0$, then we have:*

$$\lim_{n\to\infty} \sup_{\eta\in\Delta} \left| \mathfrak{R}_e[\mathrm{k}_n](\eta) - \mathfrak{R}_e[\mathrm{k}](\eta) \right| = 0 \quad \textit{and} \quad \lim_{n\to\infty} \sup_{\eta\in\Delta} d_\mathrm{H}\Big(\mathrm{Spec}[\mathrm{k}_n](\eta), \mathrm{Spec}[\mathrm{k}](\eta) \Big) = 0. \tag{IV.24}$$

IV.4 Spectrum-preserving transformations

IV.4.1 Diagonal similarity — the operator case

In this section, we consider a given probability state space $(\Omega, \mathscr{F}, \mu)$, and we discuss two operations on the kernel k that leave the functions $\mathrm{Spec}[\mathrm{k}]$ and $\mathfrak{R}_e[\mathrm{k}]$ defined on Δ. Recall the convention (IV.15) for the kernel $f\mathrm{k}g$ defined from the kernel k and the non-negative functions f and g.

Lemma IV.4.1. *Let* k *be a kernel on* Ω *and* h *be a non-negative measurable function on* Ω.

(i) If $h\mathrm{k}$ *and* $\mathrm{k}h$ *have finite double norms (with possibly different* p*), then we have:*

$$\mathrm{Spec}[h\mathrm{k}] = \mathrm{Spec}[h\mathrm{k}\mathbb{1}_{\{h>0\}}] = \mathrm{Spec}[\mathbb{1}_{\{h>0\}}\mathrm{k}h] = \mathrm{Spec}[\mathrm{k}h],$$
$$\mathfrak{R}_e[h\mathrm{k}] = \mathfrak{R}_e[h\mathrm{k}\mathbb{1}_{\{h>0\}}] = \mathfrak{R}_e[\mathbb{1}_{\{h>0\}}\mathrm{k}h] = \mathfrak{R}_e[\mathrm{k}h].$$

(ii) If h *is positive and if* k *and* $h\mathrm{k}/h$ *have finite double norms (with possibly different* p*), then we have:*

$$\mathrm{Spec}[\mathrm{k}] = \mathrm{Spec}[h\mathrm{k}/h] \quad \textit{and} \quad \mathfrak{R}_e[\mathrm{k}] = \mathfrak{R}_e[h\mathrm{k}/h].$$

(iii) If k *and its transpose* $\mathrm{k}^\top : (x, y) \mapsto \mathrm{k}(y, x)$ *have finite double norms (with possibly different* p*), then we have:*

$$\mathrm{Spec}[\mathrm{k}] = \mathrm{Spec}[\mathrm{k}^\top] \quad \textit{and} \quad \mathfrak{R}_e[\mathrm{k}] = \mathfrak{R}_e[\mathrm{k}^\top].$$

Even if (ii) is a consequence of (i), we state it separately since (ii) and (iii) describe two modifications of k that leave the functions $\mathfrak{R}_e$ and Spec invariant.

Proof. Since $\mathfrak{R}_e = \mathrm{rad} \circ \mathrm{Spec}$, we only need to prove (i)-(iii) for the function Spec. We give the detailed proof of (ii) and leave the proof of (i), which is very similar, to the reader. We first assume that k, h and $1/h$ are bounded. The operators $T_{\mathrm{k}\eta}$ and $T_{h\mathrm{k}\eta/h}$ and the multiplication operators M_h and $M_{1/h}$ are bounded operators on L^p for $p \in (1, +\infty)$. We have, using that $T_{\mathrm{k}\eta/h} = T_\mathrm{k}\, M_{\eta/h}$ is compact and (IV.11) for the second equality:

$$\mathrm{Spec}(T_{\mathrm{k}\eta}) = \mathrm{Spec}(T_{\mathrm{k}\eta/h}\, M_h) = \mathrm{Spec}(M_h T_{\mathrm{k}\eta/h}) = \mathrm{Spec}(T_{h\mathrm{k}\eta/h}).$$

Since $\eta \in \Delta$ is arbitrary, this gives that $\mathrm{Spec}[\mathrm{k}] = \mathrm{Spec}[h\mathrm{k}/h]$.

In the general case, we use an approximation scheme. Define the kernel $\mathrm{k}_n = (v_n \mathrm{k} v_n) \wedge n$ with $v_n = \mathbb{1}_{\{n \geq h \geq 1/n\}}$ and the function $h_n = n^{-1} \vee (h \wedge n)$ for $n \in \mathbb{N}^*$. From the first part of the proof, we get $\mathrm{Spec}[\mathrm{k}_n] = \mathrm{Spec}[\mathrm{k}_n']$, with $\mathrm{k}_n' = h_n \mathrm{k}_n / h_n$. Since $\|\mathrm{k}\|_{p,q}$ is finite for some $p \in (1, +\infty)$, we get by dominated convergence that $\lim_{n\to\infty} \|\mathrm{k} - \mathrm{k}_n\|_{p,q} = 0$, and we deduce from Proposition IV.3.7 that $\lim_{n\to\infty} \mathrm{Spec}[\mathrm{k}_n] = \mathrm{Spec}[\mathrm{k}]$. Similarly, setting $\mathrm{k}' = h\mathrm{k}/h$, the norm $\|\mathrm{k}'\|_{p',q'}$ is finite for some $p' \in (1, +\infty)$, and thus $\lim_{n\to\infty} \|\mathrm{k}' - \mathrm{k}_n'\|_{p',q'} = 0$, so that $\lim_{n\to\infty} \mathrm{Spec}[\mathrm{k}_n'] = \mathrm{Spec}[\mathrm{k}']$. This proves that $\mathrm{Spec}[\mathrm{k}] = \mathrm{Spec}[\mathrm{k}']$, and thus (ii).

We now prove (iii). For any $\eta \in \Delta$, the kernel $\mathrm{k}^\top\eta$ defines a bounded integral operator in L^q, whose adjoint is $T_{\eta\mathrm{k}}$. Since the spectrum of an operator and its adjoint are the same, we get $\mathrm{Spec}[\mathrm{k}^\top](\eta) = \mathrm{Spec}(T_{\mathrm{k}^\top\eta}) = \mathrm{Spec}(T_{\eta\mathrm{k}}) = \mathrm{Spec}(M_\eta T_\mathrm{k}) = \mathrm{Spec}(T_\mathrm{k} M_\eta) = \mathrm{Spec}[\mathrm{k}](\eta)$, where the fourth equality follows once more from (IV.11). Since this is true for any $\eta \in \Delta$, this gives $\mathrm{Spec}[\mathrm{k}^\top] = \mathrm{Spec}[\mathrm{k}]$. □

Remark IV.4.2. In the infinite dimensional SIS model developed in Chapter II, the next generation operator is given by the integral operator T_k, where the kernel $\mathrm{k} = k/\gamma$ is defined in terms of a transmission rate kernel $k(x, y)$ and a recovery rate function γ by the product $\mathrm{k}(x, y) = k(x, y)/\gamma(y)$; and the reproduction number $\mathfrak{R}_0$ is then the spectral radius $\rho(T_\mathrm{k})$ of T_k. Furthermore the operator $T_{\gamma^{-1}k}$ appears very naturally in the definition of the maximal equilibrium $\mathfrak{g}$ which is solution to Equation (II.24), that is $T_{\gamma^{-1}k}(\mathfrak{g}) = \mathfrak{g}/(1 - \mathfrak{g})$. According to Lemma IV.4.1 (i), provided that k/γ and $\gamma^{-1}k$ have finite double norms, the next generation operator and $T_{\gamma^{-1}k}$ have the same effective spectrum function.

We shall use the following extension in the proof of Lemma IV.5.12.

Remark IV.4.3. Following closely the proof of Lemma IV.4.1 (ii) and using Corollary IV.3.2, we also get that if h is positive and if k and hk$/h$ have finite double norms (with possibly different p), then we have:

$$\mathrm{m}(\lambda, T_{\mathrm{k}}) = \mathrm{m}(\lambda, T_{h\mathrm{k}/h}) \quad \text{for all } \lambda \in \mathbb{C}^*. \tag{IV.25}$$

IV.4.2 The matrix case

It is then natural to ask if the invariance properties stated in Lemma IV.4.1 describe all possible cases. In other words, does $\mathrm{Spec}[\mathrm{k}] = \mathrm{Spec}[\tilde{\mathrm{k}}]$ or even the weaker condition $\mathfrak{R}_e[\mathrm{k}] = \mathfrak{R}_e[\tilde{\mathrm{k}}]$ imply that k and $\tilde{\mathrm{k}}$, or k and $\tilde{\mathrm{k}}^\top$, are diagonally similar? Building on results from [77, 109], we give a partial answer in the matrix case.

For clarity's sake let us describe how our general notation adapts to the matrix case. Let K be an $n \times n$ matrix with $n \in \mathbb{N}^*$, and $\Delta = [0,1]^n$. For $\eta \in \Delta$, let $K\eta$ denote the square matrix $K \cdot \mathrm{Diag}(\eta)$, defined by:

$$(K\eta)_{ij} = K_{ij}\,\eta_j \quad \text{for} \quad 1 \le i,j \le n.$$

We define two maps:

$$\mathrm{Spec}[K] : \Delta \to \mathscr{K} \quad \text{and} \quad \mathfrak{R}_e[K] : \Delta \to \mathbb{R}_+,$$

where for all $\eta \in \Delta$, $\mathrm{Spec}[K](\eta)$ (resp. $\mathfrak{R}_e[K](\eta)$) is the spectrum (resp. the spectral radius) of the square matrix $K\eta$. We denote by $\mathcal{E}(\Delta) = \{0,1\}^n$ the extreme points of Δ.

For α and β non-empty subsets of $\{1, \dots, n\}$ we denote by $K[\alpha, \beta]$ the sub-matrix of K obtained by keeping the lines in α and the columns in β, and let $K[\alpha] = K[\alpha, \alpha]$. The determinant of $K[\alpha]$ is called a *principal minor* of K, whose size is the cardinal of α. It is elementary to check that the characteristic polynomial of K may be written as:

$$\chi_K(t) = \sum_{k=0}^{n} (-1)^k c_{n-k} t^k, \tag{IV.26}$$

where $c_0 = 1$ and, for $j \ge 1$, c_j is the sum of all principal minors of size j of K.

Definition IV.4.4. *Let K be a square matrix of size n. A non-empty subset α of $\{1, \dots, n\}$ is a* clan *if it satisfies $2 \le \mathrm{Card}(\alpha) \le n-2$ and the submatrices $K[\alpha, \alpha^c]$ and $K[\alpha^c, \alpha]$ have rank at most 1. The matrix K is* clan-free *if there exists no clan.*

Remark IV.4.5. A square matrix of size $n \le 3$ is automatically clan-free.

Assume that $\alpha = \{1, \dots, m\}$ is a clan for K. Then, there exists vectors v, w of size m, and b, c of size $n-m$ such that K may be written in block form as:

$$K = \begin{pmatrix} A & vb^\top \\ cw^\top & B \end{pmatrix}. \tag{IV.27}$$

Let us then say that:

$$\tilde{K} = \begin{pmatrix} A^\top & wb^\top \\ cv^\top & B \end{pmatrix} \tag{IV.28}$$

is a *partial transpose* of K (note that the partial transpose is not unique in general).

Remark IV.4.6. Such transformations have been considered in the special case where $v = w$ in [109, Lemma 5]; see also [22] where a similar transformation called *clan reversal* is introduced for skew symmetric matrices.

Our main result in this direction is summarized in the following proposition. Recall the matrix K is *diagonally similar* to a matrix $\tilde{K}$ if there exists a non singular real diagonal matrix D such that $K = D \cdot \tilde{K} \cdot D^{-1}$. The matrix K is *irreducible* if $K[\alpha, \alpha^c] \ne 0$ for all subsets α such that α and α^c are non-empty. The matrix K is *completely reducible* if $K[\alpha, \alpha^c] = 0$ implies $K[\alpha^c, \alpha] = 0$ whenever

α and α^{C} are non-empty. We have the following graph interpretation when K has non-negative entries: consider the oriented graph $G = (V, E)$ with $V = \{1, \dots, n\}$ and $ij \in E$, that is ij is an oriented edge of G, if and only if $K_{ij} > 0$. Then the matrix K is irreducible if for any choice of vertices $i, j \in V$ there is an oriented path from i to j; the matrix K is completely reducible if for any vertices $i, j \in V$ there is an oriented path from i to j if and only if there is an oriented path from j to i.

Proposition IV.4.7 (Matrix case)**.** *Let K and $\tilde{K}$ be square matrices of the same size with non-negative entries.*

(i) Assume K is symmetric and $\mathfrak{R}_e[K] = \mathfrak{R}_e[\tilde{K}]$. If $\tilde{K}$ is symmetric then $\tilde{K} = K$; if $\tilde{K}$ is completely reducible then $\tilde{K}$ is diagonally similar to K.

(ii) Assume that K is irreducible, clan-free and $\mathfrak{R}_e[K] = \mathfrak{R}_e[\tilde{K}]$. Then $\tilde{K}$ is diagonally similar to K or to $K^\top$.

(iii) If K is not clan-free, then we have $\mathfrak{R}_e[K] = \mathfrak{R}_e[\tilde{K}]$ for any partial transpose $\tilde{K}$ of K.

The proof of this proposition, which is postponed to the end of this section, hinges on the following characterization of matrices whose functions $\mathfrak{R}_e$ coincide.

Lemma IV.4.8. *Let K and $\tilde{K}$ be square matrices of the same size with non-negative entries. The following are equivalent:*

(i) The functions $\mathfrak{R}_e[K]$ and $\mathfrak{R}_e[\tilde{K}]$ coincide on Δ.

(ii) The functions $\mathfrak{R}_e[K]$ and $\mathfrak{R}_e[\tilde{K}]$ coincide on $\mathcal{E}(\Delta)$.

(iii) The functions $\mathrm{Spec}[K]$ *and* $\mathrm{Spec}[\tilde{K}]$ *coincide on Δ.*

(iv) The functions $\mathrm{Spec}[K]$ *and* $\mathrm{Spec}[\tilde{K}]$ *coincide on $\mathcal{E}(\Delta)$.*

(v) All principal minors of K and $\tilde{K}$ coincide.

Before giving the proof of this result, we comment on the non-negativeness condition for the entries of the matrices.

Remark IV.4.9 (When the matrices K and $\tilde{K}$ have signed entries). If two effective reproduction functions coincide on $\mathcal{E}(\Delta)$, they may not coincide on Δ nor does the principal minors coincide in general if the entries of the matrices are signed (that is Property (ii) from Lemma IV.4.8 does not imply (i) nor (v)). Indeed, consider the following two matrices:

$$K = \begin{pmatrix} 1 & \beta \\ \beta & 1 \end{pmatrix} \quad \text{and} \quad \tilde{K} = \begin{pmatrix} 1 & -\gamma \\ \gamma & 1 \end{pmatrix},$$

where $\gamma > 0$ and $\beta = \sqrt{1+\gamma^2} - 1$. We have $\det(K) \neq \det(\tilde{K})$, so that all the principal minors of size 1 coincide but the principal minor of size two is different. The eigenvalues of K are $\sqrt{1+\gamma^2}$ and $2 - \sqrt{1+\gamma^2}$; the eigenvalues of $\tilde{K}$ are $1 \pm \gamma i$. In particular, the two matrices have the same spectral radius $\sqrt{1+\gamma^2}$. The functions $\mathfrak{R}_e[K]$ and $\mathfrak{R}_e[\tilde{K}]$ clearly coincide on $\mathcal{E}(\Delta) = \{(1,1),\ (1,0),\ (0,1)\,(0,0)\}$ even if $\mathfrak{R}_e[K] \neq \mathfrak{R}_e[\tilde{K}]$.

Proof of Lemma IV.4.8. Clearly (iii) $\Longrightarrow$ (i) $\Longrightarrow$ (ii), and (iii) $\Longrightarrow$ (iv) $\Longrightarrow$ (ii).

Let us check that (v) implies (iii). Assume that all principal minors of K and $\tilde{K}$ coincide. Recall that any vector η, $K\eta$ denotes the square matrix $K \cdot \mathrm{Diag}(\eta)$. For any vector η and any set of indices α, by multi-linearity of the determinant,

$$\det\Big((K\eta)[\alpha]\Big) = \Big(\prod_{i\in\alpha} \eta_i\Big) \det\left(K[\alpha]\right).$$

Consequently, all principal minors of $(K\eta)$ and $(\tilde{K}\eta)$ coincide. By (IV.26) this implies that $K\eta$ and $\tilde{K}\eta$ have the same spectrum. Thus, Point (iii) holds.

Therefore, it is enough to prove that (ii) implies (v). The proof is an induction on the dimension. The result is clear in dimension 1. Assume that it holds for any matrix of dimension smaller than or equal to n. Let K and $\tilde{K}$ be two matrices of dimension $n+1$, and assume that $\mathfrak{R}_e[K]$ and $\mathfrak{R}_e[\tilde{K}]$ coincide on $\mathcal{E}(\Delta)$. For any non-empty $\alpha \subset \{1, ..., n+1\}$, let η_α be the column vector $(\mathbb{1}_\alpha(i), 1 \leq i \leq n+1)$. Recall that $K\eta = K \cdot \mathrm{Diag}(\eta)$ for $\eta \in \Delta$. Notice that for any matrix K':

$$\mathfrak{R}_e[K'](\eta_\alpha) = \rho(K'\eta_\alpha) = \rho(K'[\alpha]).$$

Fix $\alpha \subset \{1, ..., n+1\}$ nonempty, with $\alpha \neq \{1, ..., n+1\}$. Let $\beta \subset \alpha$ and set $\tilde{\eta}_\beta = (\mathbb{1}_\beta(i), i \in \alpha)$. We have:

$$\mathfrak{R}_e[K'[\alpha]](\tilde{\eta}_\beta) = \rho(K'[\alpha]\tilde{\eta}_\beta) = \rho(K'\eta_\alpha\eta_\beta) = \rho(K'\eta_\beta) = \mathfrak{R}_e[K'](\eta_\beta). \tag{IV.29}$$

Since the vector η_β is extremal in Δ, we get $\mathfrak{R}_e[K](\eta_\beta) = \mathfrak{R}_e[\tilde{K}](\eta_\beta)$ for all $\beta \subset \alpha$. We deduce from (IV.29) that $\mathfrak{R}_e[K[\alpha]] = \mathfrak{R}_e[\tilde{K}[\alpha]]$ on the extremal points. By the induction hypothesis the principal minors of $K[\alpha]$ and $\tilde{K}[\alpha]$ are equal, that is all principal minors of size less than or equal to n of K and $\tilde{K}$ coincide. It remains to check that the determinants are the same. Since all principal minors of size less than or equal to n coincide, we deduce from (IV.26) that:

$$\chi_K(t) - \det(K) = \chi_{\tilde{K}}(t) - \det(\tilde{K}). \tag{IV.30}$$

Since K and K' have non-negative entries, by Perron-Frobenius theorem, their spectral radius $\mathfrak{R}_e[K](\mathbb{1})$ and $\mathfrak{R}_e[K'](\mathbb{1})$ is also an eigenvalue, and thus a root of their characteristic polynomial. As $\mathfrak{R}_e[K](\mathbb{1}) = \mathfrak{R}_e[K'](\mathbb{1})$, we deduce from (IV.30) that $\det(K) = \det(\tilde{K})$. This ends the proof of the induction step. □

Proof of Proposition IV.4.7. To prove the first two points (i) and (ii), we use that the principal minors of K and $\tilde{K}$ coincide thanks to Lemma IV.4.8. The results then follow directly from [53, Theorem 3.5], for the symmetric case, [77, Theorem 3] for the irreducible case when $n \leq 3$ (by Remark IV.4.5, there can be no clan in this case), and [109, Theorem 1] for the clan-free case when $n \geq 4$.

To prove Point (iii), suppose that K has a clan α, and let $\tilde{K}$ be a partial transpose of K, so that K and $\tilde{K}$ may be given by (IV.27) and (IV.28). For any $\lambda \notin \mathrm{Spec}(B)$, using a classical formula for determinants of block matrices, we get:

$$\begin{aligned}
\det(K - \lambda I) &= \det(A - \lambda I - vb^\top(B - \lambda I)^{-1}cw^\top)\det(B - \lambda I), \\
\det(\tilde{K} - \lambda I) &= \det(A^\top - \lambda I - wb^\top(B - \lambda I)^{-1}cv^\top)\det(B - \lambda I) \\
&= \det(A - \lambda I - vc^\top((B - \lambda I)^{-1})^\top bw^\top)\det(B - \lambda I).
\end{aligned}$$

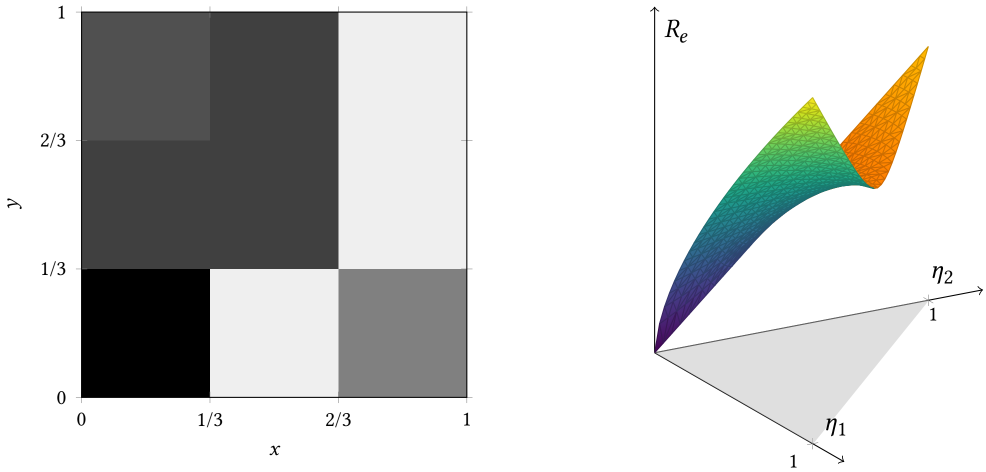
Since $b^\top(B - \lambda I)^{-1}c$ is a one-dimensional matrix, it is equal to its transpose, so that $\det(K - \lambda I) = \det(\tilde{K} - \lambda I)$ are equal for all $\lambda \notin \mathrm{Spec}(B)$, and thus for all $\lambda \in \mathbb{C}$ by continuity. Consequently, the matrices K and $\tilde{K}$ have the same spectrum. For any β, it is easily seen that $K[\beta]$ and $\tilde{K}[\beta]$ are partial transposes of each other, so that $K[\beta]$ and $\tilde{K}[\beta]$ also have the same spectrum, and in particular the same spectral radius. Therefore $\mathfrak{R}_e[K]$ and $\mathfrak{R}_e[\tilde{K}]$ coincide as (i) and (ii) are equivalent in Lemma IV.4.8. □

IV.5 Sufficient conditions for convexity or concavity of $\mathfrak{R}_e$

IV.5.1 A conjecture from Hill and Longini

Recall that, in the metapopulation framework, the effective reproduction number is equal to the spectral radius of the matrix $K \cdot \mathrm{Diag}(\eta)$, where K has non-negative entries and is the next-generation matrix and η is the vaccination strategy giving the proportion of non-vaccinated people in each groups. The Hill-Longini conjecture appears in [81] and gives conditions on the spectrum of the next-generation matrix that implies the convexity or the concavity of the effective reproduction number. It states that the function $\mathfrak{R}_e[K]$ is:

(i) convex when $\mathrm{Spec}(K) \subset \mathbb{R}_+$,



(A) Grayplot of the kernel corresponding to the next-generation matrix (IV.31) when the sub-populations have same size.

(B) The plan of strategies η with cost $C(\eta) = 2/3$ is represented as a gray surface. The triangled surface corresponds to the graph of $\mathfrak{R}_e$ restricted to this set.

Figure IV.4: Counter-example of the Hill-Longini conjecture (convex case).

(ii) concave when $\mathrm{Spec}(K)\backslash\{\mathfrak{R}_0\} \subset \mathbb{R}_-$.

It turns out that the conjecture cannot be true without additional assumption on the matrix K. Indeed, consider the following next-generation matrix:

$$K = \begin{pmatrix} 16 & 12 & 11 \\ 1 & 12 & 12 \\ 8 & 1 & 1 \end{pmatrix} \tag{IV.31}$$

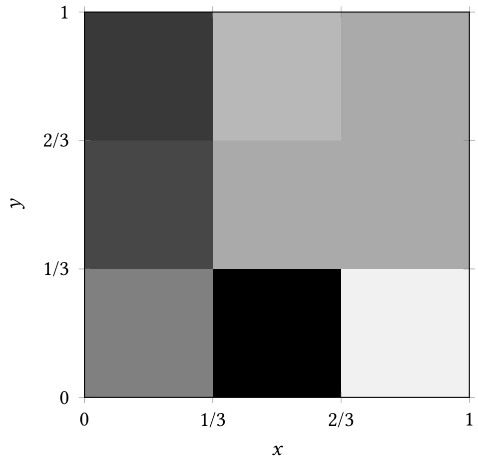
Its eigenvalues are approximately equal to 24.8, 2.9 and 1.3. Since $\mathfrak{R}_e$ is homogeneous, the function is entirely determined by the value it takes on the plane $\{\eta\ :\ \eta_1 + \eta_2 + \eta_3 = 1/3\}$. The graph of the function $\mathfrak{R}_e$ restricted to this set has been represented in Figure IV.4(B). The view clearly shows the saddle nature of the surface. Hence, the Hill-Longini conjecture (i) is contradicted in its original formulation. In Figure IV.4(A), we have represented the corresponding kernel model when the population is splitted equally into three groups, *i.e.*, $\mu_1 = \mu_2 = \mu_3 = 1/3$.

In the same manner, the eigenvalues of the following next-generation matrix:

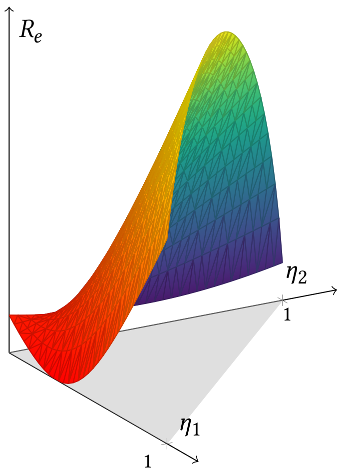
$$K = \begin{pmatrix} 9 & 13 & 14 \\ 18 & 6 & 5 \\ 1 & 6 & 6 \end{pmatrix} \tag{IV.32}$$

are approximately equal to 26.3, −1.4 and −3.9. Thus, K satisfies the condition that should imply the concavity of the effective reproduction number in the Hill-Longini conjecture (ii). However, as we can see in Figure IV.5(B), the function $\mathfrak{R}_e$ is neither convex nor concave. In Figure IV.5(A), we have represented the corresponding kernel model when the population is splitted equally into three groups, *i.e.*, $\mu_1 = \mu_2 = \mu_3 = 1/3$.

Despite these counter-examples, the Hill-Longini conjecture is indeed true when making further assumption on the next-generation matrix. Let M be a square real matrix. The matrix M is *diagonally similar* to a matrix M' if there exists a non singular real diagonal matrix D such that $M = D \cdot M' \cdot D^{-1}$. The matrix M is said to be *diagonally symmetrizable* or simply *symmetrizable* if it is diagonally similar to a symmetric matrix, or, equivalently, if M admits a decomposition $M = D \cdot S$ (or $M = S \cdot D$), where D is a diagonal matrix with positive diagonal entries and S is a symmetric matrix. If a matrix M is diagonally symmetrizable, then its eigenvalues are real since similar matrices share the same spectrum. We obtained the following result when the next-generation matrix is symmetrizable.



(A) Grayplot of the kernel corresponding to the next-generation matrix (IV.32) when the sub-populations have same size.



(B) The plan of strategies η with cost $C(\eta) = 2/3$ is represented as a gray surface. The triangled surface corresponds to the graph of $\mathfrak{R}_e$ restricted to this set.

Figure IV.5: Counter-example of the Hill-Longini conjecture (concave case).

Theorem IV.5.1. *Suppose the non-negative matrix K is diagonally symmetrizable.*

(i) If $\operatorname{Spec}(K) \subset \mathbb{R}_+$, then the function $\mathfrak{R}_e[K]$ is convex.

(ii) If $\mathfrak{R}_0$ is a simple eigenvalue of K and $\operatorname{Spec}(K) \backslash \{\mathfrak{R}_0\} \subset \mathbb{R}_-$, then the function $\mathfrak{R}_e[K]$ is concave.

This result is a particular case of Theorem IV.5.5 below. The first point (i) has been proved by Cairns in [29]. In [64], Friedland obtained that, if the next-generation matrix K is not singular and if its inverse is an M-matrix (*i.e.*, its non-diagonal coefficients are non-positive), then $\mathfrak{R}_e$ is convex. Friedland's condition does not imply that K is symmetrizable nor that $\operatorname{Spec}(K) \subset \mathbb{R}_+$. On the other hand, the following matrix is symmetric definite positive (and thus $\mathfrak{R}_e$ is convex) but its inverse is not an M-matrix.

$$K = \begin{pmatrix} 3 & 2 & 0 \\ 2 & 2 & 1 \\ 0 & 1 & 4 \end{pmatrix} \qquad \text{with inverse} \quad K^{-1} = \begin{pmatrix} 1.4 & -1.6 & 0.4 \\ -1.6 & 2.4 & -0.6 \\ 0.4 & -0.6 & 0.4 \end{pmatrix}.$$

Thus Friedland's condition and Property (i) in Theorem IV.5.1 are not comparable. Note that if K is symmetrizable and its inverse is an M-matrix, then the eigenvalues of K are actually non-negative thanks to [15, Chapter 6 Theorem 2.3] and one can apply Theorem IV.5.1 (i).

IV.5.2 Generalization for the kernel model

In this section, we give the analogue of Theorem IV.5.1 for kernels instead of matrices. First, we proceed with some definitions.

We say that a kernel k' is an Hilbert-Schmidt non-negative symmetric kernel if $\mathrm{k}' \geq 0$, $\|\mathrm{k}'\|_{2,2} < +\infty$ and $\mu(\mathrm{d}x) \otimes \mu(\mathrm{d}y)$-a.e. $\mathrm{k}'(x,y) = \mathrm{k}'(y,x)$. By analogy with the matrix case and following [160, Example A, p252] we introduce the notion of symmetrizability in the context of kernels.

Definition IV.5.2 (Diagonally HS kernel)**.** *A kernel k on Ω is diagonally HS if there exists an Hilbert-Schmidt symmetric non-negative kernel k' on Ω and two positive measurable functions f, g defined on Ω such that $\mathrm{k} = f\mathrm{k}'g$ a.s., that is $\mu(\mathrm{d}x) \otimes \mu(\mathrm{d}y)$-a.s.:*

$$\mathrm{k}(x,y) = f(x)\,\mathrm{k}'(x,y)\,g(y). \tag{IV.33}$$

If furthermore f and g are bounded and bounded away from 0, then we say that the kernel k is strongly diagonally HS.

The notion of diagonally HS kernel appears naturally when considering the SIS model on graphons, see Example II.1.3, where the kernel k is written as $\mathrm{k} = \beta W \theta$, where $\beta(x)$ represents the susceptibility and $\theta(x)$ the infectiousness of the individuals with feature x, and W models the graph of the contacts within the population with the quantity $W(x, y) = W(y, x) \in [0, 1]$ representing the density of contacts between individuals with features x and y.

Remark IV.5.3. We complete the notion of diagonally HS kernel with three comments.

(i) In finite dimension (*i.e.* Ω finite), a kernel diagonally HS is strongly diagonally HS.

(ii) Notice that a strongly diagonally HS kernel has finite double norm in L^2.

(iii) Consider the decomposition (IV.33), where f and g are assumed to be non-negative instead of positive, with the other assumptions unchanged. Then using Lemma IV.4.1 (i), we get that $\mathfrak{R}_e[\mathrm{k}] = \mathfrak{R}_e[\mathrm{k}_0]$ coincide on Δ, where $\mathrm{k}_0 = \mathbb{1}_{\{fg>0\}} \,\mathrm{k}\, \mathbb{1}_{\{fg>0\}}$. As $\mathrm{k}_0 = f' \,\mathrm{k}_0'\, g'$ with $\mathrm{k}_0' = \mathbb{1}_{\{fg>0\}} \,\mathrm{k}'\, \mathbb{1}_{\{fg>0\}}$ and $h' = h + \mathbb{1}_{\{fg=0\}}$ for $h \in \{f, g\}$, we get that the kernel k_0 is diagonally HS (indeed f' and g' are positive, and the other assumptions hold). So, as far as the study of $\mathfrak{R}_e[\mathrm{k}]$ is concerned, without loss of generality one can indeed assume that the functions f and g which appear in the decomposition of a diagonally HS kernel are positive instead of non-negative.

The following elementary lemma states that the integral operator of a diagonally HS kernel has real eigenvalues.

Lemma IV.5.4. *Let* k *be a diagonally HS kernel with finite double norm. The spectrum of* T_k *is real:* $\mathrm{Spec}(T_\mathrm{k}) \subset \mathbb{R}$.

Proof. Let k', f and g as in (IV.33) and for $n \in \mathbb{N}^*$ set:

$$v_n = \mathbb{1}_{\{n \geq f \geq 1/n \text{ and } n \geq g \geq 1/n\}}. \tag{IV.34}$$

Let $p \in (1, +\infty)$ be such that $\|\mathrm{k}\|_{p,q}$ is finite. By the monotone convergence theorem, we have

$$\lim_{n\to\infty} \|\mathrm{k} - f v_n \mathrm{k}' v_n g\|_{p,q} = 0.$$

We deduce that:

$$\mathrm{Spec}(T_\mathrm{k}) = \mathrm{Spec}(T_{f\mathrm{k}'g}) = \lim_{n\to\infty} \mathrm{Spec}(T_{f v_n \mathrm{k}' v_n g}) = \lim_{n\to\infty} \mathrm{Spec}(T_{\sqrt{fg}\, v_n \mathrm{k}' v_n \sqrt{fg}}),$$

where we used (IV.33) for the first equality, Corollary IV.3.2 for the second, Lemma IV.4.1 (ii) with $h = v_n \sqrt{g/f} + (1 - v_n)$ for the last. Since the kernel $\sqrt{fg}\, v_n \mathrm{k}' v_n \sqrt{fg}$ is symmetric with finite double norm in L^2, we deduce that the associated compact integral operator is self-adjoint, and thus $\mathrm{Spec}(T_{\sqrt{fg}\, v_n \mathrm{k}' v_n \sqrt{fg}}) \subset \mathbb{R}$. Then, use that $\mathbb{R}$ is closed for the Hausdorff distance to deduce that $\mathrm{Spec}(T_\mathrm{k}) \subset \mathbb{R}$. □

For a compact operator T, we denote by $\mathrm{p}(T)$ and $\mathrm{n}(T)$ the number of its positive and negative eigenvalues with their multiplicity:

$$\mathrm{p}(T) = \sum_{\lambda > 0} \mathrm{m}(\lambda, T) \quad \text{and} \quad \mathrm{n}(T) = \sum_{\lambda < 0} \mathrm{m}(\lambda, T).$$

Note that $\mathfrak{R}_0[\mathrm{k}] > 0$ implies that $\mathrm{p}(T_\mathrm{k}) \geq 1$.

The following result is the analogue of Theorem IV.5.1 for the kernel model.

Theorem IV.5.5 (Convexity/Concavity of $\mathfrak{R}_e$). *Let* k *be a strongly diagonally HS kernel. We consider the function* $\mathfrak{R}_e = \mathfrak{R}_e[\mathrm{k}]$ *defined on* Δ.

(i) If $\mathrm{n}(T_\mathrm{k}) = 0$*, then the function* $\mathfrak{R}_e$ *is convex.*

(ii) If $\mathrm{p}(T_\mathrm{k}) = 1$*, then the function* $\mathfrak{R}_e$ *is concave.*

In the case of diagonally HS kernels, we have the following partial result.

Proposition IV.5.6. *Let* k *be a diagonally HS kernel of finite double norm, with the HS kernel* k′ *from* (IV.33). *We consider the function* $\mathfrak{R}_e = \mathfrak{R}_e[\mathrm{k}]$ *defined on* Δ.

(i) If $\mathrm{n}(T_{\mathrm{k}'}) = 0$, *then* $\mathrm{n}(T_{\mathrm{k}}) = 0$ *and the function* $\mathfrak{R}_e$ *is convex.*

(ii) If $\mathrm{p}(T_{\mathrm{k}'}) = 1$, *then* $\mathrm{p}(T_{\mathrm{k}}) = 1$ *and the function* $\mathfrak{R}_e$ *is concave.*

The proof for HS kernels is given in Section IV.5.4 for the convex case and in Section IV.5.4 for the concave case. The extension to (strongly) diagonally HS kernel follows from Sections IV.5.5.

Remark IV.5.7. The fact that $\mathfrak{R}_0 > 0$, where we write $\mathfrak{R}_0 = \mathfrak{R}_0[\mathrm{k}]$, and $\mathrm{p}(T_{\mathrm{k}}) = 1$ in Theorem IV.5.5 (ii) implies that k is monatomic, see Lemma IV.3.4. Using the decomposition of a reducible kernel from Lemma IV.7.2, we get that if $\mathrm{Spec}(T_{\mathrm{k}}) \in \mathbb{R}_- \cup \{\mathfrak{R}_0\}$, then the function $\mathfrak{R}_e$ is the maximum of $m = \mathrm{m}(\mathfrak{R}_0, T_{\mathrm{k}})$ concave functions which are non-zero on m pairwise disjoint subsets of Δ.

Remark IV.5.8. It is unclear whether or not $\mathrm{p}(T_{\mathrm{k}}) = 1$ (resp. $\mathrm{n}(T_{\mathrm{k}}) = 0$) in Proposition IV.5.6 implies that $\mathrm{p}(T_{\mathrm{k}'}) = 1$ (resp. $\mathrm{n}(T_{\mathrm{k}'}) = 0$).

Remark IV.5.9. A *configuration model* corresponds in finite dimension to the next generation matrix having rank one, this is the so-called *proportionate mixing* model in the metapopulation literature; see Cairns [29] for optimal vaccinations strategies in this setting.

Motivated by the finite dimensional case, we say that a kernel k is a *configuration* kernel if there exist $p \in (1, +\infty)$, $f \in L^p$ and $g \in L^q$ where $q = p/(p-1)$ such that $\mathrm{k}(x, y) = f(x)g(y)$, $\mu\otimes\mu$-almost surely. We also suppose that $\mu(fg > 0) > 0$. Such a kernel has finite double norm and, following Remark IV.5.3 (iii), we have $\mathfrak{R}_e[\mathrm{k}] = \mathfrak{R}_e[\mathbb{1}_{fg>0}\,\mathrm{k}\,\mathbb{1}_{fg>0}]$ with $\mathbb{1}_{fg>0}\,\mathrm{k}\,\mathbb{1}_{fg>0}$ diagonally HS. Indeed, we have $\mathbb{1}_{fg>0}\,\mathrm{k}\,\mathbb{1}_{fg>0} = (f+\mathbb{1}_{f=0})\mathbb{1}_{fg>0}\mathbb{1}_{fg>0}(g+\mathbb{1}_{g=0})$. Besides, the only eigenvalue of the kernel $\mathbb{1}_{fg>0}(x)\mathbb{1}_{fg>0}(y)$ different from 0 is its spectral radius equal to $\mu(fg > 0)$ and it has multiplicity 1. Applying Proposition IV.5.6, we obtain that $\mathfrak{R}_e$ is convex and concave and thus linear. This can be checked directly as $\mathfrak{R}_e$ has the following expression:

$$\mathfrak{R}_e[\mathrm{k}](\eta) = \int_\Omega fg\eta\,\mathrm{d}\mu. \tag{IV.35}$$

We shall provide a deeper study of configuration kernels in the context of epidemiology, in a future paper [40].

IV.5.3 Sylvester's inertia theorem

Following [128, Section 4.1.2], we state and provide a short proof for the Sylvester's inertia theorem in our context; see also [87, Theorem 4.5.8] in finite dimension. This result will be used to prove the concavity of $\mathfrak{R}_e$.

Theorem IV.5.10 (Sylvester's inertia theorem). *Let* $(\Omega, \mathscr{F}, \mu)$ *be a probability space. Let* T' *be a self-adjoint compact operator on* $L^2(\mu)$, *and two non-negative measurable functions* f, g *defined on* Ω *which are bounded and bounded away from* 0. *Set* $T = M_f T' M_g$. *Then, we have* $\mathrm{Spec}(T) \subset \mathbb{R}$ *and:*

$$\mathrm{p}(T) = \mathrm{p}(T') \quad \textit{as well as} \quad \mathrm{n}(T) = \mathrm{n}(T'). \tag{IV.36}$$

Proof. Set $h = \sqrt{f/g}$, $M = M_{\sqrt{fg}}$ and

$$T'' = MT'M,$$

so that $T = M_h T'' M_{1/h}$. Thanks to (IV.11), we get that $\mathrm{m}(\lambda, T) = \mathrm{m}(\lambda, T'')$ for all $\lambda \in \mathbb{C}^*$. So, we need to prove (IV.36) with T replaced by T''. We only consider the number of positive eigenvalues as the number of negative eigenvalues can be handled similarly.

We introduce some general notations. For a self-adjoint compact operator S on L^2, let $(u_i, i \in I)$, with I at most countable and $\sharp I = \mathrm{p}(S)$, be a sequence of orthogonal eigenvectors associated to the positive eigenvalues $(\lambda_i, i \in I)$ of S. Let $U \subset L^2$ be the (closed) vector sub-space spanned by $(u_i, i \in I)$. The orthogonal complement of U, say $U^\top$ is the (closed) vector space spanned by the kernel of I and the eigenvectors associated to the negative eigenvalues. We consider the quadratic form Q_S on L^2 defined by:

$$Q_S(u) = \langle Su, u\rangle.$$

Let P_S be the orthogonal projection on $U^\top$. By decomposing u on $U \oplus U^\top$, we get:

$$Q_S(u) = \sum_{i\in I} \lambda_i \langle u, u_i\rangle^2 + Q_S(P_S(u)),$$

and the quadratic form $Q_S \circ P_S$ is negative semi-definite.

We shall now prove that $\mathrm{p}(T'') = \mathrm{p}(T')$ by contradiction. First assume that $\mathrm{p}(T') < \mathrm{p}(T'')$, so in particular $\mathrm{p}(T')$ is finite. Let $(u_i'', i \in I'')$ be a sequence of orthogonal eigenvectors associated to the positive eigenvalues $(\lambda_i'', i \in I'')$ of T''. Set $v_i = Mu_i''$ for $i \in I''$. In particular, the dimension of the space spanned by $(v_i, i \in I'')$, which is equal to $\mathrm{p}(T'')$, is larger than the finite dimension of the space U spanned by the orthogonal eigenvectors $(u_i', i \in I')$ associated to the positive eigenvalues of T'. Thus, solving a linear system, we get there exists $(c_i, i \in I'')$ such that $c_i \neq 0$ for at most $\mathrm{p}(T') + 1$ indices, $u = \sum_{i\in I''} c_i v_i \neq 0$, and $u \in U^\top$. On one hand, since $Q_{T'}$ is negative semi-definite on $U^\top$, we get $Q_{T'}(u) \leq 0$. On the other hand, we have:

$$Q_{T'}(u) = \langle u, T'u\rangle = \sum_{i,j\in I''} c_i c_j \langle v_i, T' v_j\rangle = \sum_{i,j\in I''} c_i c_j \langle u_i'', T'' u_j''\rangle = \sum_{i\in I''} c_i^2 \lambda_i'' > 0.$$

By contradiction, we deduce that $\mathrm{p}(T') \geq \mathrm{p}(T'')$, and by symmetry $\mathrm{p}(T') = \mathrm{p}(T'')$. □

IV.5.4 The symmetric case

Let k be an Hilbert-Schmidt non-negative symmetric kernel. As $\mathfrak{R}_0[\mathrm{k}] = 0$ implies $\mathfrak{R}_e[\mathrm{k}] = 0$ by (IV.9), we shall only consider the case $\mathfrak{R}_0[\mathrm{k}] > 0$. We now prove Theorem IV.5.5 when k is symmetric with finite double norm in L^2.

The convex case

The proof relies on an idea from [64] (see therein just before Theorem 4.3). Let k be an Hilbert-Schmidt non-negative symmetric kernel such that $\mathrm{Spec}(T_\mathrm{k}) \subset \mathbb{R}_+$, where T_k is the corresponding integral operator on L^2. Since T_k is a self-adjoint positive semi-definite operator on L^2, there exists a self-adjoint positive semi-definite operator Q on L^2 such that $Q^2 = T$. Recall that for a real-valued function u defined on Ω, M_u denotes the multiplication by u operator. Thanks to (IV.12), we have for $\eta \in \Delta$:

$$\mathfrak{R}_e[\mathrm{k}](\eta) = \rho(T_\mathrm{k}\, M_\eta) = \rho(Q^2\, M_\eta) = \rho(Q\, M_\eta\, Q).$$

Since the self-adjoint operator $Q\, M_\eta\, Q$ (on L^2) is also positive semi-definite, we deduce from the Courant-Fischer-Weyl min-max principle that:

$$\mathfrak{R}_e[\mathrm{k}](\eta) = \rho\left(Q\, M_\eta\, Q\right) = \sup_{u\in L^2\setminus\{0\}} \frac{\langle u, Q\, M_\eta\, Qu\rangle}{\langle u, u\rangle}.$$

Since the map $\eta \mapsto \langle u, Q\, M_\eta\, Qu\rangle$ defined on Δ is linear, we deduce that $\eta \mapsto \mathfrak{R}_e[\mathrm{k}](\eta)$ is convex as a supremum of linear functions.

The concave case

Let k be an Hilbert-Schmidt non-negative symmetric kernel such that $\mathrm{p}(T_\mathrm{k}) = 1$. In particular k is monatomic, see Lemma IV.3.4. Let Δ^* be the subset of Δ of the functions which are bounded away from 0. The set Δ^* is a dense convex subset of Δ. So its suffice to prove that $\mathfrak{R}_e = \mathfrak{R}_e[\mathrm{k}]$ is concave on Δ^*. Let η_0, η_1 be elements of Δ^*, and set $\eta_\alpha = (1-\alpha)\eta_0 + \alpha\eta_1$ for $\alpha \in [0, 1]$ (which is also an element of Δ^*). We write $T_\alpha = T_{\mathrm{k}\eta_\alpha}$, so that $T_\alpha = T_0 + \alpha T_{\mathrm{k}} M$, where M is the multiplication by $(\eta_1 - \eta_0)$ operator, and:

$$R(\alpha) = \mathfrak{R}_e(\eta_\alpha) = \rho(T_\alpha) = \rho(T_0 + \alpha T_{\mathrm{k}} M).$$

So, to prove that $\mathfrak{R}_e$ is concave on Δ^* (and thus on Δ), it is enough to prove that $\alpha \mapsto R(\alpha)$ is concave on $(0, 1)$. As η_α is also bounded away from 0, we get that $\mathrm{k}\eta_\alpha$ is monatomic and its spectral radius $R(\alpha)$ is positive and a simple eigenvalue, thanks to Lemma IV.3.4. Thanks to Sylvester's inertia theorem, see Theorem IV.5.10 (with $f = 1$ and $g = \eta_\alpha$), we also get that $\mathrm{p}(T_\alpha) = 1$.

We consider the following scalar product on L^2 defined by $\langle u, v\rangle_\alpha = \langle u, \eta_\alpha v\rangle$. The operator T_α is self-adjoint and compact on $L^2(\eta_\alpha \mathrm{d}\mu)$ with spectrum $\mathrm{Spec}(T_\alpha)$ thanks to Lemma IV.3.1 (iii). Let $(\lambda_n, n \in I = [\![0, N[\![)$, with $N \in \mathbb{N} \cup \{\infty\}$ be an enumeration of the non-zero eigenvalues of T_α with their multiplicity so that $\lambda_0 = R(\alpha) > 0$ and thus $\lambda_n < 0$ for $n \in I^* = I \setminus \{0\}$; and denote by $(u_n, n \in I)$ a corresponding sequence of orthogonal eigenvectors. The functions $v_\alpha = u_0$ and $\phi_\alpha = \eta_\alpha u_0$ are the right and left-eigenvectors for T_α (seen as an operator on $L^2(\mu)$) associated to $R(\alpha)$.

We now follow [98] to get that $\alpha \mapsto R(\alpha) = \rho(T_0 + \alpha T_{\mathrm{k}} M)$ is analytic and compute its second derivative. Let π_α be the projection on the $(\langle\cdot,\cdot\rangle_\alpha)$-orthogonal of v_α, and define:

$$S_\alpha = (T_\alpha - R(\alpha))^{-1} \pi_\alpha.$$

In other words, S_α maps u_0 to 0 and u_i to $(\lambda_i - R(\alpha))^{-1}\, u_i$. Let $\alpha \in (0, 1)$ and ε small enough so that $\alpha + \varepsilon \in [0, 1]$. We have:

$$T_{\alpha+\varepsilon} = T_\alpha + \varepsilon T_{\mathrm{k}} M,$$

and thus $\| T_{\alpha+\varepsilon} - T_\alpha \|_{L^2(\mu)} = O(\varepsilon)$. Using [98, Theorem 2.6] on the Banach space $L^2(\eta_\alpha\, \mathrm{d}\mu)$, we get that:

$$R(\alpha + \varepsilon) = R(\alpha) + \varepsilon \langle v_\alpha, T_{\mathrm{k}} M v_\alpha\rangle_\alpha - \varepsilon^2 \langle v_\alpha, T_{\mathrm{k}} M S_\alpha T_{\mathrm{k}} M v_\alpha\rangle_\alpha + O(\varepsilon^3).$$

Let $N_\alpha = M_{1/\eta_\alpha} M = M M_{1/\eta_\alpha}$ be the multiplication by $(\eta_1 - \eta_0)/\eta_\alpha$ bounded operator. Since $\alpha \mapsto R(\alpha)$ is analytic and T_{k} self-adjoint (with respect to $\langle\cdot,\cdot\rangle$), we get that:

$$\begin{aligned} R''(\alpha) &= -2 \langle v_\alpha, T_{\mathrm{k}} M S_\alpha T_{\mathrm{k}} M v_\alpha\rangle_\alpha \\ &= -2 \langle M T_\alpha v_\alpha, S_\alpha T_{\mathrm{k}} M v_\alpha\rangle \\ &= -2R(\alpha) \langle M v_\alpha, S_\alpha T_{\mathrm{k}} M v_\alpha\rangle \\ &= -2R(\alpha) \langle N_\alpha v_\alpha, S_\alpha T_\alpha N_\alpha v_\alpha\rangle_\alpha . \end{aligned}$$

Since the kernel and the image of T_α are orthogonal (in $L^2(\eta_\alpha \mathrm{d}\mu)$), and the latter is generated by $(u_n, n \in I)$, we have the decomposition $N_\alpha v_\alpha = g + \sum_{n\in I} a_n u_n$ with $g \in \mathrm{Ker}(T_\alpha)$ and $a_n = \langle N_\alpha v_\alpha, u_n\rangle_\alpha$. This gives, with $I^* = I \setminus \{0\}$:

$$R''(\alpha) = 2R(\alpha) \sum_{n\in I^*} \frac{\lambda_n}{R(\alpha) - \lambda_n}\, a_n^2\, \langle u_n, u_n\rangle_\alpha . \tag{IV.37}$$

Since $\lambda_n < 0$ for all $n \in I^*$, we deduce that R is concave on $[0, 1]$. This implies that $\mathfrak{R}_e[\mathrm{k}]$ is concave.

Remark IV.5.11. The same proof with obvious changes gives that if k is an Hilbert-Schmidt non-negative symmetric monatomic (and thus quasi-irreducible) kernel such that $\mathrm{n}(T_{\mathrm{k}}) = 0$, then $\mathfrak{R}_e[\mathrm{k}]$ is convex on Δ.

IV.5.5 Proof of Theorem IV.5.5 and Proposition IV.5.6

We first consider the following technical Lemma.

Lemma IV.5.12. *Let* k *be a diagonally HS kernel, with the HS kernel* k′ *from* (IV.33). *We have:*

$$\mathrm{p}(T_{\mathrm{k}}) \leq \mathrm{p}(T_{\mathrm{k}'}) \quad \textit{and} \quad \mathrm{n}(T_{\mathrm{k}}) \leq \mathrm{n}(T_{\mathrm{k}'}).$$

If furthermore k *is strongly diagonally HS, then the previous inequalities are equalities.*

Proof. We only consider the number of positive eigenvalues as the number of negative eigenvalues can be handled similarly. Let f, g be the functions from (IV.33) and v_n defined in (IV.34) for $n \in \mathbb{N}^*$. For simplicity, we write $\mathrm{p}(\mathrm{k}'')$ for $\mathrm{p}(T_{\mathrm{k}''})$ when k″ is a kernel with finite double norm. Let $m \in \mathbb{N}^*$. As the function $w_{n,m} = \sqrt{fg}\, v_n + m^{-1}(1 - v_n)$ is bounded and bounded away from 0, we deduce from the Sylvester's inertia Theorem IV.5.10 that:

$$\mathrm{p}(\mathrm{k}') = \mathrm{p}\left(w_{n,m}\, \mathrm{k}'\, w_{n,m}\right). \tag{IV.38}$$

Notice that $\lim_{m\to\infty} \left\| \sqrt{fg}\, v_n\, \mathrm{k}'\, v_n \sqrt{fg} - w_{n,m}\, \mathrm{k}'\, w_{n,m} \right\|_{2,2} = 0$. Letting m goes to infinity, we deduce from (IV.20) in Corollary IV.3.2 and the fact that the spectrum is real that:

$$\mathrm{p}(\mathrm{k}') \geq \mathrm{p}\left(\sqrt{fg}\, v_n\, \mathrm{k}'\, v_n \sqrt{fg}\right). \tag{IV.39}$$

We also deduce from Remark IV.4.3, with $h = \sqrt{f/g}\, v_n + (1 - v_n)$ that:

$$\mathrm{p}\left(\sqrt{fg}\, v_n\, \mathrm{k}'\, v_n \sqrt{fg}\right) = \mathrm{p}\left(f v_n\, \mathrm{k}'\, v_n g\right).$$

Recall that k has a finite double norm on some L^p space. By the monotone convergence theorem, we get that:

$$\lim_{m\to\infty} \| f\mathrm{k}'\, g - f v_n\, \mathrm{k}'\, v_n g \|_{p,q} = 0.$$

Letting n goes to infinity, we also deduce from (IV.20) in Corollary IV.3.2 and the fact that the spectrum is real according to Lemma IV.5.4, that:

$$\liminf_{n\to\infty} p\left(f v_n\, \mathrm{k}'\, v_n g\right) \geq p\left(f\mathrm{k}'\, g\right). \tag{IV.40}$$

Thus, we have $\mathrm{p}(\mathrm{k}') \geq \mathrm{p}(\mathrm{k})$.

Notice that if k is strongly diagonally HS, then $v_n = 1$ for n large enough, so that inequalities (IV.39) and (IV.40) are in fact equalities and thus $\mathrm{p}(\mathrm{k}') = \mathrm{p}(\mathrm{k})$. □

Proof of Proposition IV.5.6. We only prove (ii) as the proof of (i) is similar and easier for the last part. We keep notations from the proof of Lemma IV.5.12. Assume that $\mathrm{p}(\mathrm{k}') = 1$. We deduce from (IV.38) and from Section IV.5.4 that $\mathfrak{R}_e[w_{n,m}\, \mathrm{k}'\, w_{n,m}]$ is concave. We deduce from Corollary IV.3.2, letting m goes to infinity, that $\mathfrak{R}_e[\sqrt{fg}\, v_n\, \mathrm{k}'\, v_n \sqrt{fg}]$ is concave. Use Lemma IV.4.1 (ii) with $h = \sqrt{f/g}\, v_n + (1 - v_n)$ to obtain that $\mathfrak{R}_e[f v_n\, \mathrm{k}'\, v_n g]$ is concave. Then, letting n goes to infinity and using again Corollary IV.3.2, we deduce that $\mathfrak{R}_e[f\mathrm{k}' g] = \mathfrak{R}_e[\mathrm{k}]$ is concave.

Use also Lemma IV.5.12 to get $\mathrm{p}(\mathrm{k}) \leq \mathrm{p}(\mathrm{k}')$. Now if $\mathrm{p}(\mathrm{k}) = 0$, then we have that $\mathfrak{R}_0[\mathrm{k}] = 0$ which is equivalent to $\mathfrak{R}_0[\mathbb{1}_{\{\mathrm{k}>0\}}] = 0$. Since $\{\mathrm{k} > 0\} = \{\mathrm{k}' > 0\}$, this is also equivalent to $\mathfrak{R}_0[\mathrm{k}'] = 0$. As this is ruled out because $\mathrm{p}(\mathrm{k}') = 1$, we deduce that $\mathrm{p}(\mathrm{k}) = 1$. □

Proof of Theorem IV.5.5. The result is an immediate consequence of Proposition IV.5.6 and the second part of Lemma IV.5.12. □

IV.6 Three properties of the Pareto and anti-Pareto frontiers

We introduce in Section IV.6.1 the bi-objective minimization problem, where one tries to minimize simultaneously the cost of the vaccination and the effective reproduction number, and recall results from Chapter III on the Pareto and anti-Pareto optimal strategies and frontiers. Then, we derive in Section IV.6.2 there existence of Pareto optimal rays as soon as there exists a Pareto optimal strategy uniformly strictly bounded from above by 1. We prove in Section IV.6.3 that creating a cordon sanitaire is not the worst idea in the sense that it is not anti-Pareto optimal (and it can be Pareto optimal or not). Eventually, in Section IV.6.4 we give a characterization of $c_\star = C_\star(0)$ using the notion of independent set from graph theory.

IV.6.1 Pareto and anti-Pareto frontiers

We quantify the cost of the vaccination strategy $\eta \in \Delta$ by a function $C : \Delta \to \mathbb{R}^+$, and we assume that $C(\mathbb{1}) = 0$ (doing nothing costs nothing), C is non-increasing (doing more costs more) and continuous for the weak topology on Δ defined in Section IV.3.5. Recall that $1 - \eta$ represents the proportion of the population which has been vaccinated when using the strategy η. One natural choice is the uniform cost function $C = C_{\mathrm{uni}}$ defined for $\eta \in \Delta$ by:

$$C_{\mathrm{uni}}(\eta) = \int_\Omega (1 - \eta)\, \mathrm{d}\mu. \tag{IV.41}$$

In Chapter III, we formalized and study the problem of optimal allocation strategies for a perfect vaccine. This question may be viewed as a bi-objective minimization problem, where one tries to minimize simultaneously the cost of the vaccination and the effective reproduction number:

$$\min_\Delta (C, \mathfrak{R}_e). \tag{IV.42}$$

We briefly summarize the results from Chapter III. We shall assume that the kernel k has a finite double norm, the loss function is given by the effective reproduction function $\mathfrak{R}_e[\mathrm{k}]$, and the cost function C is furthermore *decreasing* (this is the case of the uniform cost), that is, for any $\eta_1, \eta_2 \in \Delta$:

$$\eta_1 \leq \eta_2 \quad \text{and} \quad \int_\Omega \eta_1\, \mathrm{d}\mu < \int_\Omega \eta_2\, \mathrm{d}\mu \implies C(\eta_1) > C(\eta_2).$$

To be precise, the next results can be found in Propositions III.5.4 and III.5.5 (notice in particular, that Assumptions 4 and 5 holds thanks to Lemma III 5.13). By definition, we have $\mathfrak{R}_0 = \max_\Delta \mathfrak{R}_e$ and we set $c_{\max} = \max_\Delta C$ which is positive as C is decreasing. Related to the minimization problem (IV.42), we shall consider $\mathfrak{R}_{e\star}$ the *optimal loss* function and $C_\star$ the *optimal cost* function defined by:

$$\begin{aligned} \mathfrak{R}_{e\star}(c) &= \min \{\mathfrak{R}_e(\eta) \,:\, \eta \in \Delta,\, C(\eta) \leq c\} \quad \text{for } c \in [0, c_{\max}], \\ C_\star(\ell) &= \min \{C(\eta) \,:\, \eta \in \Delta,\, \mathfrak{R}_e(\eta) \leq \ell\} \quad \text{for } \ell \in [0, \mathfrak{R}_0]. \end{aligned}$$

We have $C_\star(\mathfrak{R}_0) = 0$ and $\mathfrak{R}_{e\star}(0) = \mathfrak{R}_0$ since C is decreasing. For convenience, we write $c_\star$ for the minimal cost required to completely stop the transmission of the disease:

$$c_\star = C_\star(0) = \inf\{c \in [0, c_{\max}] \,:\, \mathfrak{R}_{e\star}(c) = 0\}. \tag{IV.43}$$

The function $\mathfrak{R}_{e\star}$ is continuous, decreasing on $[0, c_\star]$ and zero on $[c_\star, 1]$; the function $C_\star$ is continuous and decreasing on $[0, \mathfrak{R}_0]$; and the functions $\mathfrak{R}_{e\star}$ and $C_\star$ are the inverse of each other, that is, $\mathfrak{R}_{e\star} \circ C_\star(\ell) = \ell$ for $\ell \in [0, \mathfrak{R}_0]$ and $C_\star \circ \mathfrak{R}_{e\star}(c) = c$ for $c \in [0, c_\star]$.

We define the Pareto optimal strategies $\mathcal{P}$ as the "best" solutions of the minimization problem (IV.42) (we refer the reader to Chapter III for a precise justification of this terminology):

$$\mathcal{P} = \{\eta \in \Delta \,:\, C(\eta) = C_\star(\mathfrak{R}_e(\eta)) \quad \text{and} \quad \mathfrak{R}_e(\eta) = \mathfrak{R}_{e\star}(C(\eta))\},$$

and the Pareto frontier as their outcomes:

$$\mathcal{F} = \{(C(\eta), \mathfrak{R}_e(\eta)) \,:\, \eta \in \mathcal{P}\}.$$

The set $\mathcal{P}$ is a non empty compact (for the weak topology) in Δ and furthermore the Pareto frontier can be easily represented using the graph of the optimal loss function or cost function:

$$\mathcal{F} = \{(C_\star(\ell), \ell) \,:\, \ell \in [0, \mathfrak{R}_0]\} = \{(c, \mathfrak{R}_{e\star}(c)) \,:\, c \in [0, c_\star]\}.$$

It is also of interest to consider the “worst” strategies which can be viewed as solutions to the bi-objective maximization problem:c3-

$$\max_{\Delta}(C, \mathfrak{R}_e). \tag{IV.44}$$

To be precise, the next results can be found in Propositions III.5.8 and III.5.9 (notice in particular that Assumption 6 holds in general but that Assumption 7 holds under the stronger condition that the kernel k is monatomic, see Section 5.4 in Chapter III). Related to the maximization problem (IV.44), we shall consider $\mathfrak{R}_e^\star$ the *optimal loss* function and $C^\star$ the *optimal cost* function defined by:

$$\begin{aligned} \mathfrak{R}_e^\star(c) &= \max\,\{\,\mathfrak{R}_e(\eta)\;:\;\eta \in \Delta,\, C(\eta) \geq c\,\} \quad \text{for } c \in [0, c_{\max}], \\ C^\star(\ell) &= \max\,\{\,C(\eta)\;:\;\eta \in \Delta,\, \mathfrak{R}_e(\eta) \geq \ell\,\} \quad \text{for } \ell \in [0, \mathfrak{R}_0]. \end{aligned}$$

We have $C^\star(0) = c_{\max}$ and $\mathfrak{R}_e^\star(c_{\max}) = 0$ since C is decreasing and $C(\mathbb{0}) = c_{\max}$. For convenience, we write $c^\star$ for the maximal cost of totally inefficient strategies:

$$c^\star = C^\star(\mathfrak{R}_0) = \max\{c \in [0, c_{\max}]\;:\;\mathfrak{R}_e^\star(c) = \mathfrak{R}_0\}. \tag{IV.45}$$

The function $C^\star$ is decreasing on $[0, \mathfrak{R}_0]$; the function $\mathfrak{R}_e^\star$ is constant equal to $\mathfrak{R}_0$ on $[0, c^\star]$; we have $\mathfrak{R}_e^\star \circ C^\star(\ell) = \ell$ for $\ell \in [0, \mathfrak{R}_0]$. This latter property implies that the function $\mathfrak{R}_e^\star$ is continuous.

We define the anti-Pareto optimal strategies $\mathcal{P}^{\text{Anti}}$ as the “worst” strategies, that is solutions of the maximization problem (IV.44):

$$\mathcal{P}^{\text{Anti}} = \{\eta \in \Delta\;:\;C(\eta) = C^\star(\mathfrak{R}_e(\eta)) \quad \text{and} \quad \mathfrak{R}_e(\eta) = \mathfrak{R}_e^\star(C(\eta))\},$$

and the anti-Pareto frontier as their outcomes:

$$\mathcal{F}^{\text{Anti}} = \left\{(C(\eta), \mathfrak{R}_e(\eta))\;:\;\eta \in \mathcal{P}^{\text{Anti}}\right\}.$$

The set $\mathcal{P}$ is non empty and furthermore the Pareto frontier can be easily represented using the graph of the optimal cost function:

$$\mathcal{F}^{\text{Anti}} = \{(C^\star(\ell), \ell)\;:\;\ell \in [0, \mathfrak{R}_0]\}. \tag{IV.46}$$

We also have that the feasible region or set of possible outcomes for $(C, \mathfrak{R}_e)$:

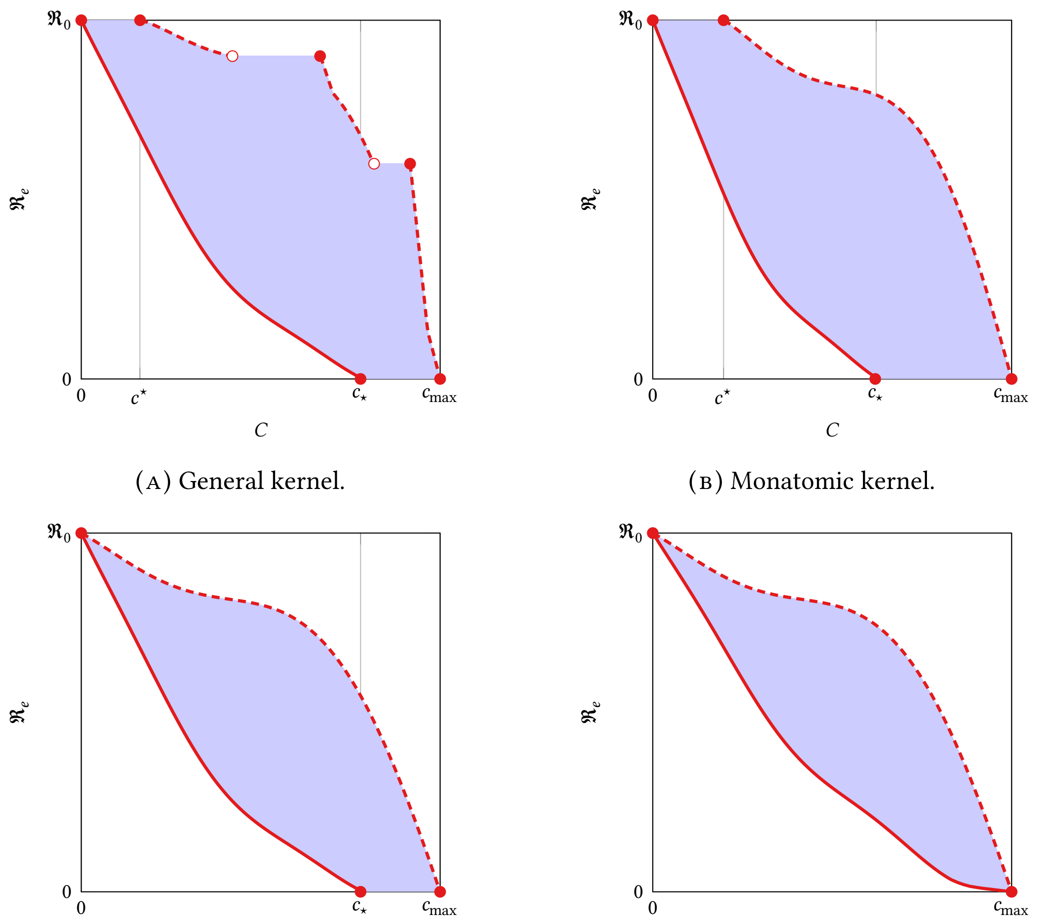
$$\mathbf{F} = \{(C(\eta), \mathfrak{R}_e(\eta))\;:\;\eta \in \Delta\}$$

is compact, path connected, and its complement is connected in $\mathbb{R}^2$. It is the whole region between the graphs of the one-dimensional value functions:

$$\begin{aligned} \mathbf{F} &= \{(c, \ell) \in [0, c_{\max}] \times [0, \mathfrak{R}_0]\;:\;\mathfrak{R}_{e\star}(c) \leq \ell \leq \mathfrak{R}_e^\star(c)\} \\ &= \{(c, \ell) \in [0, c_{\max}] \times [0, \mathfrak{R}_0]\;:\;C_\star(\ell) \leq c \leq C^\star(\ell)\}. \end{aligned}$$

If furthermore k is monatomic with atom Ω_a, then thanks to Lemma III.5.13, we have $c^\star = C(\mathbb{1}_{\Omega_a})$ (which is 0 if k is irreducible); the function $\mathfrak{R}_e^\star$ is continuous, decreasing on $[c^\star, c_{\max}]$; the function $C^\star$ is continuous and decreasing on $[0, \mathfrak{R}_0]$; the functions $\mathfrak{R}_e^\star$ and $C^\star$ are the inverse of each other, that is, $\mathfrak{R}_e^\star \circ C^\star(\ell) = \ell$ for $\ell \in [0, \mathfrak{R}_0]$ and $C^\star \circ \mathfrak{R}_e^\star(c) = c$ for $c \in [c^\star, c_{\max}]$; and the set $\mathcal{P}^{\text{Anti}}$ is compact and $\mathcal{F}^{\text{Anti}} = \{(c, \mathfrak{R}_e^\star(c))\;:\;c \in [c^\star, c_{\max}]\}$.

We plotted in Figure IV.6 the typical Pareto and anti-Pareto frontiers for a general kernel (notice the anti-Pareto frontier is not connected *a priori*), a monatomic kernel (notice the anti-Pareto frontier is connected), and a positive kernel. In the latter case, the properties of the frontiers are stated in the next lemma.



(A) General kernel. (B) Monatomic kernel.

(C) Irreducible kernel. (D) Kernel strictly positive almost surely.

Figure IV.6: Generic appearance of the feasible region (light blue), the Pareto frontier (thick red line) and the anti-Pareto frontiers (dashed red line) for the cost function $\mathfrak{R}_e[\mathrm{k}]$, with kernel k, and a continuous decreasing cost function C.

Lemma IV.6.1. *Suppose that the cost function C is continuous decreasing with $C(\mathbb{1}) = 0$ and consider the loss function $\mathfrak{R}_e[\mathrm{k}]$, with k a finite double norm kernel such that a.s. $\mathrm{k} > 0$. Then, we have $\mathfrak{R}_0[\mathrm{k}] > 0$, $c^\star = 0$, $c_\star = c_{\max}$ and the strategy $\mathbb{1}$ (resp. $\mathbb{0}$) is the only Pareto optimal as well as the only anti-Pareto optimal strategy with cost $c = 0$ (resp. $c = 1$).*

Proof. Since $\mathrm{k} > 0$, we get that k is irreducible (and thus monatomic) and $\mathfrak{R}_0 > 0$, thanks to Lemma IV.3.4. We get that $c^\star = 0$. This implies that the strategy $\mathbb{1}$ is anti-Pareto optimal. As C is decreasing, we also get that the strategy $\mathbb{1}$ is Pareto optimal.

Let $\eta \in \Delta$ be different from $\mathbb{0}$. We get that the kernel $\mathrm{k}\eta$ restricted to the set of positive μ-measure $\{\eta > 0\}$ is positive, thus $\mathrm{k}\eta$ is monatomic (take $\Omega_{\mathrm{a}} = \{\eta > 0\}$ and $\Omega_{\mathrm{i}} = \Omega_{\mathrm{a}}^{\mathrm{c}}$). Thanks Lemma IV.3.4, we get that $\mathfrak{R}_e(\eta) > 0$. This readily implies that $c_\star = c_{\max}$ and that the strategy $\mathbb{0}$ is Pareto optimal. As C is decreasing, we also get that the strategy $\mathbb{0}$ is anti-Pareto optimal. □

IV.6.2 Optimal ray

As the loss function $\mathfrak{R}_e$ is convex and homogeneous, and if the cost function is affine, then the set $\mathcal{P}$ of Pareto optimal strategies may contains a non-trivial optimal ray $\{\lambda\eta \,:\, \lambda \in [0, 1]\}$. This optimal ray has already been observed in finite dimension, see [130].

Proposition IV.6.2 (Optimal ray)**.** *Suppose that the cost function C is continuous decreasing and affine with $C(\mathbb{1}) = 0$, and that the loss function $\mathfrak{R}_e[\mathrm{k}]$, with k a finite double norm kernel, is convex.*

If there exists a Pareto optimal strategy $\eta_\star \in \mathcal{P}$ such that $\sup_\Omega \eta_\star \in (0,1)$*, then the strategies $\lambda\eta_\star$ are Pareto optimal for all $\lambda \in [0, 1/\sup_\Omega \eta_\star]$.*

Remark IV.6.3. Suppose assumptions of Proposition IV.6.2 hold so that there is an optimal ray $\{\lambda\eta_\star : \lambda \in [0,1]\} \subset \mathcal{P}$, where $\sup \eta_\star = 1$. Then, by homogeneity of the loss function, the Pareto frontier has a linear part (from $(C(\eta_\star), \mathfrak{R}_e(\eta_\star))$ to $(c_{\max}, 0)$).

Proof of Proposition IV.6.2. Assume there exists a Pareto optimal strategy $\eta_\star$ such that $\sup_\Omega \eta_\star$ belongs to $(0,1)$. Let $\lambda \in (0, 1/\sup_\Omega \eta_\star]$, so that $\lambda\eta_\star \in \Delta$, and let $\eta \in \Delta$ such that $\mathfrak{R}_e(\eta) \leq \mathfrak{R}_e(\lambda\eta_\star)$, and thus $\mathfrak{R}_e(\eta) \leq \lambda\mathfrak{R}_e(\eta_\star)$. Since $\sup_\Omega \eta_\star < 1$, there exists $s \in (0,1]$ such that $(1-s)\eta_\star + s\eta/\lambda \in \Delta$. Using the homogeneity and the convexity of $\mathfrak{R}_e$, we get:

$$\begin{aligned}\mathfrak{R}_e((1-s)\eta_\star + s\eta/\lambda) &= \frac{1}{\lambda}\mathfrak{R}_e((1-s)\lambda\eta_\star + s\eta)\\ &\leq (1-s)\mathfrak{R}_e(\lambda\eta_\star)/\lambda + s\mathfrak{R}_e(\eta)/\lambda\\ &\leq \mathfrak{R}_e(\eta_\star).\end{aligned}$$

Since $\eta_\star$ is Pareto optimal, we deduce that $C((1-s)\eta_\star + s\eta/\lambda) \geq C(\eta_\star)$. Since C is affine, we get that $C(\eta) \geq C(\lambda\eta_\star)$. Hence, $\lambda\eta_\star$ is solution of the problem $\min C(\eta)$ for $\eta \in \Delta$ such that $\mathfrak{R}_e(\eta) \leq \ell$ with $\ell = \mathfrak{R}_e(\lambda\eta_\star)$. We conclude that $\lambda\eta_\star$ is Pareto optimal using Proposition III.5.5 (ii). Use that the Pareto optimal set is closed, see Corollary III.5.7 to get that $\lambda\eta_\star$ is Pareto optimal for $\lambda = 0$. □

IV.6.3 Creating a cordon sanitaire is not the worst idea

We say a strategy $\eta \in \Delta$ is a *cordon sanitaire* or *disconnecting* (for the kernel k) if $\eta \neq \mathbb{0}$ and the kernel k restricted to the set $\{\eta > 0\}$ is not connected. We make some elementary comments on disconnecting strategies.

Remark IV.6.4. Let k be a kernel.

(i) The strategy $\eta = \mathbb{1}$ is disconnecting if and only if k is not connected.

(ii) A strategy η is disconnecting if and only if the strategy $\mathbb{1}_{\{\eta>0\}}$ is disconnecting.

(iii) If $\mathrm{k} > 0$, then there does not exist a disconnecting strategy.

The next proposition states that if the strategy η is anti-Pareto optimal for a kernel k and non zero, then the kernel k restricted to $\{\eta > 0\}$ is irreducible and thus the kernel $\mathbb{1}_{\{\eta>0\}}\mathrm{k}\mathbb{1}_{\{\eta>0\}}$ is quasi-irreducible. Let us remark that in general none of those implications are equivalences.

Proposition IV.6.5 (A cordon sanitaire is never the worst idea)**.** *Suppose that the cost function C is continuous decreasing and consider the loss function $\mathfrak{R}_e[\mathrm{k}]$, with k a finite double norm kernel such that $\mathfrak{R}_0[\mathrm{k}] > 0$. Then, a disconnecting strategy is not anti-Pareto optimal.*

In the non-oriented cycle graph from Example IV.1.2, this property is illustrated in Figure IV.1 as the disconnecting strategy “one in 4”, see Figure IV.2, is not anti-Pareto.

The proof of the proposition relies on the next lemma which is a direct application of [138, Lemma 11] to our setting. For $A \in \mathscr{F}$, let $\mathrm{m}(\lambda, \mathrm{k}, A)$ be the multiplicity (possibly equal to 0) of the eigenvalue $\lambda \in \mathbb{C}^*$ for the integral operator $T_{\mathrm{k}\mathbb{1}_A}$ associated to the kernel $\mathrm{k}\mathbb{1}_A$.

Lemma IV.6.6. *Let* k *be kernel with finite double norm. Let $A, B \in \mathscr{F}$ be such that $A \cap B = \emptyset$ a.s. and $\mathrm{k}(B, A) = 0$. For all $\lambda \in \mathbb{C}^*$, we have:*

$$\mathrm{m}(\lambda, \mathrm{k}, A \cup B) = \mathrm{m}(\lambda, \mathrm{k}, A) + \mathrm{m}(\lambda, \mathrm{k}, B),$$

and thus

$$\mathfrak{R}_e[\mathrm{k}](\mathbb{1}_A + \mathbb{1}_B) = \max\left(\mathfrak{R}_e[\mathrm{k}](\mathbb{1}_A), \mathfrak{R}_e[\mathrm{k}](\mathbb{1}_B)\right). \tag{IV.47}$$

We are now in a position to prove Proposition IV.6.5.

Proof of Proposition IV.6.5. Let η be a disconnecting strategy, and thus $\eta \neq \mathbb{0}$. Since η is disconnecting, that is, k restricted to $\{\eta > 0\}$ is not irreducible, we deduce there exists $A, B \in \mathscr{F}$ such that $\mu(A) > 0$, $\mu(B) > 0$, $(\mathrm{k}\eta)(B, A) = 0$ and a.s. $A \cup B = \{\eta > 0\}$ and $A \cap B = \emptyset$. In particular (IV.47) holds with k replaced by kη. First assume that $\mathfrak{R}_e[\mathrm{k}\eta](\mathbb{1}_A) \geq \mathfrak{R}_e[\mathrm{k}\eta](\mathbb{1}_B)$, so that (IV.47) yields:

$$\mathfrak{R}_e[\mathrm{k}](\eta) = \mathfrak{R}_e[\mathrm{k}\eta](\mathbb{1}_A + \mathbb{1}_B) = \mathfrak{R}_e[\mathrm{k}\eta](\mathbb{1}_A).$$

For $\theta \in [0, 1]$, define the strategy $\eta_\theta = \eta\mathbb{1}_A + \theta\eta\mathbb{1}_B$. We deduce that:

$$\begin{aligned}\mathfrak{R}_e[\mathrm{k}](\eta_\theta) = \mathfrak{R}_e[\mathrm{k}\eta_\theta](\mathbb{1}_A + \mathbb{1}_B) &= \max(\mathfrak{R}_e[\mathrm{k}\eta_\theta](\mathbb{1}_A), \mathfrak{R}_e[\mathrm{k}\eta_\theta](\mathbb{1}_B))\\ &= \max(\mathfrak{R}_e[\mathrm{k}\eta](\mathbb{1}_A), \theta\mathfrak{R}_e[\mathrm{k}\eta](\mathbb{1}_B))\\ &= \mathfrak{R}_e[\mathrm{k}\eta](\mathbb{1}_A)\\ &= \mathfrak{R}_e[\mathrm{k}](\eta),\end{aligned}$$

where we used (IV.47) with k replaced by kη_θ for the second equality as $(\mathrm{k}\eta_\theta)(B, A) = 0$, and the homogeneity of the spectral radius in the third. Thus, the map $\theta \mapsto \mathfrak{R}_e[\mathrm{k}](\eta_\theta)$ is constant on $[0, 1]$. Since $\mu(B) > 0$ and C is decreasing, we get that $\theta \mapsto C(\eta_\theta)$ is decreasing. This implies that η_θ is worse than η for any $\theta \in [0, 1)$, and thus η is not anti-Pareto optimal.

The case $\mathfrak{R}_e[\mathrm{k}\eta](\mathbb{1}_B) \geq \mathfrak{R}_e[\mathrm{k}\eta](\mathbb{1}_A)$ is handled similarly. □

IV.6.4 A characterization of $c_\star = C_\star(0)$ when the support of k is symmetric

We characterize the Pareto optimal strategies which minimize $\mathfrak{R}_e$ when the kernel k has a symmetric support; and we get a very simple representation of $C_\star(0)$ when the cost is uniform $C = C_{\mathrm{uni}}$.

Let us first recall a notion from graph theory. If $G = (V, E)$ is an non-oriented graph with vertices set V and edge set E, an *independent* set of G is a subset $A \subset V$ of vertices which are pairwise not adjacent, that is, $i, j \in A$ implies $ij \notin E$. The *independence number* of a graph G, denoted by $\alpha(G)$, is the maximum of $\sharp A/\sharp G$, over all the independent sets A of G. Following [84], we generalize this definition to kernels.

Definition IV.6.7 (Independent sets for kernels)**.** *Let* k *be a kernel on* Ω*. A measurable set* $A \in \mathscr{F}$ *is an independent set of* k *if* $\mathrm{k} = 0$ $\mu^{\otimes 2}$*-a.s. on* $A \times A$*. The independence number* $\alpha(\mathrm{k})$ *of the kernel* k *is:*

$$\alpha(\mathrm{k}) = \sup\{\mu(A) : \ A \textit{ is an independent set of } \mathrm{k}\}.$$

A compactness argument will show that the supremum defining α is reached.

Proposition IV.6.8 (Existence of a maximal independent set)**.** *For any kernel* k *on* Ω*, there exists an independent set* A *of* k *that is* maximal*, in the sense that* $\mu(A) = \alpha(\mathrm{k})$*.*

Proof. First, notice that the independent sets and maximal independent sets of a kernel k depends only on the support $\{\mathrm{k} > 0\}$ of k. Therefore, the maximal independent sets of the kernel k and of the kernel $\mathbb{1}_{\{\mathrm{k}>0\}}$ are the same. In particular, we can assume without loss of generality that the kernel k is bounded.

Let $(A_n,\ n \in \mathbb{N})$ be a sequence of independent sets for k such that:

$$\lim_{n\to\infty} \mu(A_n) = \alpha(\mathrm{k}).$$

Since Δ is sequentially compact for the weak topology according to Lemma IV.3.5, up to taking a sub-sequence, we may assume that the sequence $(\mathbb{1}_{A_n},\ n \in \mathbb{N})$ converges weakly to some function $g \in \Delta$. Since k is bounded, the integral operator T_k is well defined. We deduce that $T_\mathrm{k}(\mathbb{1}_{A_n})$ belongs to Δ and converges a.s. towards $T_\mathrm{k}(g)$. This implies that $\mathbb{1}_{A_n} T_\mathrm{k}(\mathbb{1}_{A_n})$ converges weakly towards $gT_\mathrm{k}(g)$. We deduce that:

$$\int_\Omega gT_\mathrm{k}(g)\,\mathrm{d}\mu = \lim_{n\to\infty} \int_\Omega \mathbb{1}_{A_n} T_\mathrm{k}(\mathbb{1}_{A_n})\,\mathrm{d}\mu = \lim_{n\to\infty} \mathrm{k}(A_n, A_n) = 0.$$

As $g \in \Delta$, this implies that $\{ g > 0 \}$ is an independent set of k and thus $\mu(g > 0) \leq \alpha(\mathrm{k})$. Besides, since $(\mathbb{1}_{A_n}, n \in \mathbb{N})$ converges weakly to g, we get:

$$\int_\Omega g \, \mathrm{d}\mu = \lim_{n\to\infty} \mu(A_n) = \alpha(\mathrm{k}).$$

This implies that $\mu(g > 0) \geq \int_\Omega g \, \mathrm{d}\mu = \alpha(\mathrm{k})$. We deduce that $\mu(g > 0) = \alpha(\mathrm{k})$, and since $\{ g > 0 \}$ is an independent set, it is also maximal. □

In the following result, we prove that maximum independent sets provides an optimal Pareto strategy for the loss function $\mathfrak{R}_e$ and the cost function C_{uni} given by (IV.41) corresponding to the cost $c_\star = C_\star(0)$, see also Remark IV.6.10 for a general cost function. This property is illustrated in Figure IV.1 where the Pareto frontier of the non-oriented cycle graph from Example IV.1.2 is plotted. It is possible to prevent infections without vaccinating the whole population as $c_\star = 1/2 < 1 = c_{\mathrm{max}}$.

Proposition IV.6.9. *Let* k *be a kernel with finite double norm such that its support,* $\{\mathrm{k} > 0\}$*, is a.s. a symmetric subset of* Ω^2*. We consider the cost* $C = C_{\mathrm{uni}}$ *given by* (IV.41)*. For any maximal independent set* $A_\star$ *of* k*, the strategy* $\mathbb{1}_{A_\star}$ *is Pareto optimal for the loss* $\mathfrak{R}_e[\mathrm{k}]$ *and we have:*

$$c_\star = C_\star(0) = C(\mathbb{1}_{A_\star}) = 1 - \alpha(\mathrm{k}). \tag{IV.48}$$

Remark IV.6.10. Definition IV.6.7 on maximal independent set is in fact associated to the uniform cost $C = C_{\mathrm{uni}}$. More generally, we could define the independence number $\alpha_C(\mathrm{k})$ of the kernel k with respect to a decreasing continuous cost function C (recall the convention $C(\mathbb{1}) = 0$ and $c_{\mathrm{max}} = C(\mathbb{0})$) as:

$$\alpha_C(\mathrm{k}) = \sup\{c_{\mathrm{max}} - C(\mathbb{1}_A) \; : \; A \text{ is an independent set of k}\}.$$

The notations are consistent as $\alpha_C = \alpha$ for $C = C_{\mathrm{uni}}$. Adapting the proof of Proposition IV.6.8, we get that for any kernel k on Ω, there exists an independent set A of k that is C-maximal, in the sense that $\alpha_C(\mathrm{k}) = c_{\mathrm{max}} - C(\mathbb{1}_A)$. Following the proof of Proposition IV.6.9, we then get that if the kernel k has a finite double norm whose support, $\{\mathrm{k} > 0\}$, is a.s. a symmetric subset of Ω^2, then for any C-maximal independent set $A_\star$ of k, the strategy $\mathbb{1}_{A_\star}$ is Pareto optimal for the loss $\mathfrak{R}_e[\mathrm{k}]$ and the cost C. Furthermore, we have:

$$c_\star = C_\star(0) = C(\mathbb{1}_{A_\star}) = \min\{C(\mathbb{1}_A) \; : \; A \text{ is an independent set of k}\}.$$

Proof of Proposition IV.6.9. The existence of a maximum independent set A is given by Proposition IV.6.8. The effective reproduction number obviously vanishes for the strategy $\mathbb{1}_A$ with cost $1-\alpha(\mathrm{k})$ as $(T_{\mathrm{k}\mathbb{1}_A})^2 = T_{\mathrm{k}}\, T_{\mathbb{1}_A \mathrm{k} \mathbb{1}_A} = 0$. Now, let $\eta \in \Delta$ be such that $\mathfrak{R}_e[\mathrm{k}](\eta) = 0$. To complete the proof of the proposition, it is enough to prove that $C_{\mathrm{uni}}(\eta) \geq 1 - \alpha(\mathrm{k})$.

Since $\mathfrak{R}_e[\mathrm{k}](\eta) = 0$, the spectral radius of $T_{\mathrm{k}\eta}$ is equal to 0. Let $\varepsilon > 0$ and consider the kernel k_ε defined on Ω by:

$$\mathrm{k}_\varepsilon(x, y) = \mathbb{1}_{\{ \mathrm{k}(x,y)>\varepsilon \}}.$$

Since $T_{\mathrm{k}\eta} - \varepsilon T_{\mathrm{k}_\varepsilon \eta}$ is a positive operator, we deduce from (IV.9) that $\varepsilon \rho(T_{\mathrm{k}_\varepsilon \eta}) = \rho(\varepsilon T_{\mathrm{k}_\varepsilon \eta}) \leq \rho(T_{\mathrm{k}\eta}) = 0$ and thus $\rho(T_{\mathrm{k}_\varepsilon \eta}) = 0$. Set $\mathrm{k}' = \mathbb{1}_{\{ \mathrm{k}>0 \}}$. Since $\lim_{\varepsilon\to 0+} \| \mathrm{k}_\varepsilon - \mathrm{k}' \|_{p,q} = 0$, we deduce from Proposition IV.3.7 on the stability of $\mathfrak{R}_e$ that $\rho(T_{\mathrm{k}'\eta}) = \mathfrak{R}_e[\mathrm{k}'](\eta) = \lim_{\varepsilon\to 0+} \mathfrak{R}_e[\mathrm{k}_\varepsilon](\eta) = \lim_{\varepsilon\to 0+} \rho(T_{\mathrm{k}_\varepsilon \eta}) = 0$. As the support of k is symmetric, we deduce that the kernel k' is symmetric. According to (IV.12), we have:

$$\rho(T_{\mathrm{k}''}) = \rho(T_{\mathrm{k}'\eta}) = 0,$$

with $\mathrm{k}'' = \sqrt{\eta}\, \mathrm{k}' \sqrt{\eta} = \sqrt{\eta}\, \mathbb{1}_{\{ \mathrm{k}>0 \}} \sqrt{\eta}$. Since the kernel k'' is symmetric, non-negative and bounded by 1, this implies that $\mathrm{k}'' = 0 \; \mathrm{d}\mu^{\otimes 2}$-a.s., and thus $\{ \eta > 0 \}$ is an independent set for k. This gives $\mu(\eta > 0) \leq \alpha(\mathrm{k})$. Therefore, we have the following lower bound for the cost $C_{\mathrm{uni}}(\eta)$:

$$C_{\mathrm{uni}}(\eta) = 1 - \int_\Omega \eta \, \mathrm{d}\mu \geq 1 - \mu(\eta > 0) \geq 1 - \alpha(\mathrm{k}).$$

This ends the proof of the proposition. □

IV.7 Pareto and anti-Pareto frontiers for reducible kernels

When the kernel k is "truly reducible" (corresponding to the set of indices I below to be such that $\sharp I \geq 2$), it is natural to ask whether the frontiers of the subsystems entirely characterize the frontiers for k, and in what sense the optimization problems can be "reduced" to the separate study of each irreducible component.

We can achieve an elementary description of the anti-Pareto frontier when the kernel is not reducible using a Frobenius decomposition, see [89, 155] and [138] or the "super diagonal" form, see [45, Part II.2]. For convenience, we follow [138], see also [17, Lemma 5.17] in the case k symmetric.

Let k be a kernel on Ω with finite double norm. Let $\mathscr{A}$ be the set of k-invariant sets, and notice that $\mathscr{A}$ is stable by countable unions and countable intersections. Let $\sigma(\mathscr{A})$ be the σ-field generated by $\mathscr{A}$, and we denote by $(\Omega_i, i \in I)$ the at most countable (but possibly empty) collection of atoms with respect to the measure μ. Notice that the atoms are define up to an a.s. equivalence and can be chosen to be pair-wise disjoint. For $i \in I$, we set:

$$\mathrm{k}_i = \mathbb{1}_{\Omega_i} \mathrm{k} \mathbb{1}_{\Omega_i}, \tag{IV.49}$$

which is a kernel on Ω with finite double norm. Set $\Omega_0 = (\cup_{i\in I} \Omega_i)^{\mathsf{c}}$ (and assume the set of indices I has been chosen so that it does not contain 0). Thanks to [138, Lemma 12] or [155, Section II], there exists a total order, say $\preccurlyeq$, on I (not unique in general) such that for all $i, j \in I$:

(i) $j \prec i$ implies $\mathrm{k}(\Omega_i, \Omega_j) = 0$. In the epidemiology setting, $j \prec i$ means that the sub-population Ω_j can not infect the sub-population Ω_i.

(ii) $\mu(\Omega_i) > 0$ and k restricted to Ω_i is irreducible and has positive spectral radius, that is k_i is quasi-irreducible, and $\mathfrak{R}_e[\mathrm{k}](\mathbb{1}_{\Omega_i}) = \mathfrak{R}_0[\mathrm{k}_i] > 0$.

(iii) k reduced to Ω_0 is quasi-nilpotent, that is $\mathfrak{R}_e[\mathrm{k}](\mathbb{1}_{\Omega_0}) = 0$.

(iv) For all $\lambda \in \mathbb{C}^*$:
$$\mathrm{m}(\lambda, \mathrm{k}) = \sum_{i\in I} \mathrm{m}(\lambda, \mathrm{k}_i). \tag{IV.50}$$

The next remark gives some elementary results related to the Frobenius decomposition.

Remark IV.7.1. Recall $\mathfrak{R}_0[\mathrm{k}]$ denote the spectral radius of the integral operator with kernel k. Recall that $\{\mathrm{k} \equiv 0\} = \{x \in \Omega : \mathrm{k}(x, \Omega) + \mathrm{k}(\Omega, x) = 0\}$. We have:

(i) If the spectral radius of the kernel k is positive, then I is non-empty.

(ii) If the kernel k is quasi-irreducible, then $\Omega_0 = \{\mathrm{k} \equiv 0\}$ and I is a singleton.

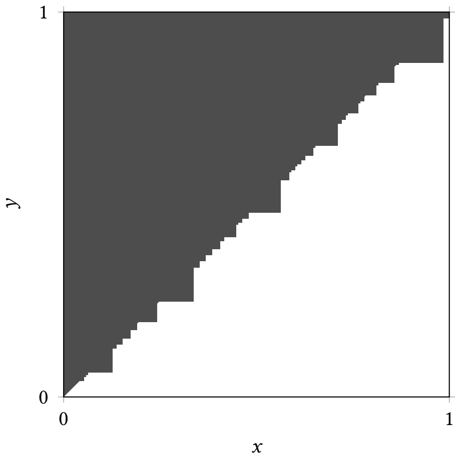
(iii) The kernel k is monatomic if and only if I is a singleton, say $I = \{a\}$. Then the set Ω_a is the atom of k.

(iv) If A invariant implies A^{c} invariant, then we have $\Omega_0 = \{\mathrm{k} \equiv 0\}$ and $\mathrm{k} = \sum_{i\in I} \mathrm{k}_i$ (k reduced to Ω_0 is zero and intuitively k is block diagonal).

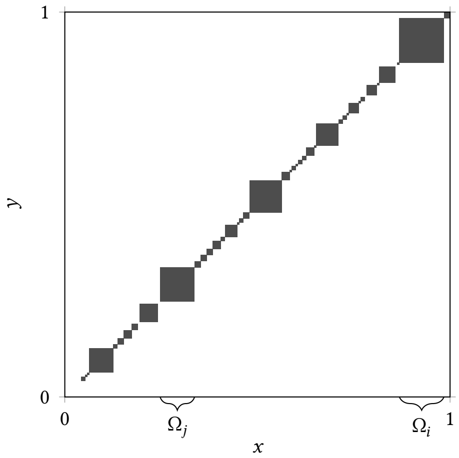
(v) The cardinal of set of indices $i \in I$ such that $\mathfrak{R}_0[\mathrm{k}_i] = \mathfrak{R}_0[\mathrm{k}]$ is exactly equal to the multiplicity of $\mathfrak{R}_0[\mathrm{k}]$ for T_k, that is $\mathrm{m}(\mathfrak{R}_0[\mathrm{k}], \mathrm{k})$.

(vi) An eigenvalue λ of T_k is *distinguished* if its distinguished multiplicity $\sharp\{i \in I : \mathfrak{R}_0[\mathrm{k}_i] = \lambda\}$ is positive. Notice that $\mathfrak{R}_0[\mathrm{k}]$ is distinguished with its distinguished multiplicity equal to its multiplicity. Indeed if $\mathfrak{R}_0[\mathrm{k}]$ is an eigenvalue of k_i, then it is its spectral radius and thus has multiplicity one as k_i is quasi-irreducible. We also deduce that $\mathrm{m}(\mathfrak{R}_0[\mathrm{k}], \mathrm{k}_i) \in \{0, 1\}$ for all $i \in I$.

For $i \in I$ and $\eta \in \Delta$, we set $\eta_i = \eta \mathbb{1}_{\Omega_i}$ and recall that $\mathrm{k}_i = \mathbb{1}_{\Omega_i} \mathrm{k} \mathbb{1}_{\Omega_i}$. We now give the decomposition of $\mathfrak{R}_e[\mathrm{k}]$ according to the quasi-irreducible components $(\mathrm{k}_i, i \in I)$ of k.



(A) A representation of the kernel k.



(B) A representation of the kernel $\tilde{\mathrm{k}} = \sum_{i\in I} \mathrm{k}_i$. We have $\mathrm{Spec}[\mathrm{k}] = \mathrm{Spec}[\tilde{\mathrm{k}}]$ and thus $\mathfrak{R}_e[\mathrm{k}] = \mathfrak{R}_e[\tilde{\mathrm{k}}]$.

Figure IV.7: Example of a kernel k with the white zone included in $\{\mathrm{k} = 0\}$ and the kernel $\tilde{\mathrm{k}} = \sum_{i\in I} \mathrm{k}_i$, with $\mathrm{k}_i = \mathbb{1}_{\Omega_i}\mathrm{k}\mathbb{1}_{\Omega_i}$ and $\mathrm{k}(\Omega_i, \Omega_j) = 0$ for $j \prec i$.

Lemma IV.7.2. *Let* k *be a kernel with a finite double norm and positive spectral radius. We have for* $\eta \in \Delta$*:*

$$\mathfrak{R}_e[\mathrm{k}](\eta) = \max_{i\in I} \mathfrak{R}_e[\mathrm{k}_i](\eta_i) = \max_{i\in I} \mathfrak{R}_e[\mathrm{k}](\eta\mathbb{1}_{\Omega_i}). \tag{IV.51}$$

Proof. For $A \in \mathscr{F}$, recall $\mathrm{m}(\lambda, \mathrm{k}, A)$ denotes the multiplicity (possibly equal to 0) of the eigenvalue $\lambda \in \mathbb{C}^*$ for the integral operator $T_{\mathrm{k}\mathbb{1}_A}$ associated to the kernel $\mathrm{k}\mathbb{1}_A$. Let $A, B \in \mathscr{F}$ be such that $A \cap B = \emptyset$ a.s. and $\mathrm{k}(B, A) = 0$. Let $\eta \in \Delta$. Clearly we have $(\mathrm{k}\eta)(B, A) = 0$, and thus Lemma IV.6.6 gives that for all $\eta \in \Delta$:

$$\mathrm{m}(\lambda, \mathrm{k}\eta, A \cup B) = \mathrm{m}(\lambda, \mathrm{k}\eta, A) + \mathrm{m}(\lambda, \mathrm{k}\eta, B).$$

Then, an immediate adaptation of the proof of [138, Theorem 7] gives that for all $\lambda \in \mathbb{C}^*$:

$$\mathrm{m}(\lambda, \mathrm{k}\eta, \Omega) = \sum_{i\in I} \mathrm{m}(\lambda, \mathrm{k}\eta, \Omega_i). \tag{IV.52}$$

By definition of $\mathrm{m}(\lambda, \cdot, \cdot)$, we get $\mathfrak{R}_e[\mathrm{k}](\eta) = \max\{|\lambda| : \mathrm{m}(\lambda, \mathrm{k}\eta, \Omega) > 0\}$ and $\mathfrak{R}_e[\mathrm{k}\mathbb{1}_{\Omega_i}](\eta) = \max\{|\lambda| : \mathrm{m}(\lambda, \mathrm{k}\eta, \Omega_i) > 0\}$. This gives that:

$$\mathfrak{R}_e[\mathrm{k}](\eta) = \max_{i\in I} \mathfrak{R}_e[\mathrm{k}\mathbb{1}_{\Omega_i}](\eta).$$

To conclude, notice that $\mathfrak{R}_e[\mathrm{k}](\eta\mathbb{1}_{\Omega_i}) = \mathfrak{R}_e[\mathrm{k}\mathbb{1}_{\Omega_i}](\eta) = \mathfrak{R}_e[\mathbb{1}_{\Omega_i}\mathrm{k}\mathbb{1}_{\Omega_i}](\eta) = \mathfrak{R}_e[\mathrm{k}_i](\eta_i)$, where we used Lemma IV.4.1 (i) for the second equality. ☐

Set $\tilde{\mathrm{k}} = \sum_{i\in I} \mathrm{k}_i$. As a consequence of (IV.52), we have that:

$$\mathrm{Spec}[\mathrm{k}] = \mathrm{Spec}[\tilde{\mathrm{k}}] \quad \text{and} \quad \mathfrak{R}_e[\mathrm{k}] = \mathfrak{R}_e[\tilde{\mathrm{k}}].$$

In view of Section IV.4.1, we get an other transformation of the kernel k which leaves the function Spec[k] unchanged. We represent in Figure IV.7(A) an example of a kernel k with its atomic decomposition using $\preccurlyeq$ as a partial order on Ω and in Figure IV.7(B) the corresponding kernel $\tilde{\mathrm{k}}$.

We set $\mathfrak{R}_0 = \mathfrak{R}_0[\mathrm{k}]$. For $i \in I$, we consider the loss $\mathfrak{R}_e[\mathrm{k}_i]$ and the corresponding optimal loss function $R_i^\star$ defined on $[0, c_{\max}]$ and optimal cost function $C_i^\star$. For convenience the function $C_i^\star$ which is defined on $[0, \mathfrak{R}_0[\mathrm{k}_i]]$ is extended to $[0, \mathfrak{R}_0]$ by setting $C_i^\star = 0$ on $(\mathfrak{R}_0[\mathrm{k}_i], \mathfrak{R}_0]$. We eventually give the description of the anti-Pareto frontier. Notice also that $\{\mathrm{k}_i \equiv 0\} = \Omega_i^{\mathsf{c}}$. Recall that $c_{\max} = C(\mathbb{1})$.

Corollary IV.7.3. *Suppose that the cost function C is continuous decreasing with $C(\mathbb{1}) = 0$ and consider the loss function $\mathfrak{R}_e[\mathrm{k}]$, with k a finite double norm kernel. Assume that $\mathfrak{R}_0 = \mathfrak{R}_0[\mathrm{k}] > 0$. We have:*

$$\mathfrak{R}_e^\star = \max_{i\in I} R_i^\star \quad \text{(on } [0, c_{\max}]\text{), and} \quad C^\star = \max_{i\in I} C_i^\star \quad \text{(on } [0, \mathfrak{R}_0]\text{);}$$

the maximal cost of totally inefficient strategies is given by:

$$c^\star = C^\star(\mathfrak{R}_0) = \max_{i\in I}\{C(\mathbb{1}_{\Omega_i}) \;:\; \mathfrak{R}_0[\mathrm{k}_i] = \mathfrak{R}_0[\mathrm{k}]\};$$

and the anti-Pareto frontier is given by:

$$\mathcal{F}^{\mathrm{Anti}} = \left\{\left(\max_{i\in I} C_i^\star(\ell), \ell\right) \;:\; \ell \in [0, \mathfrak{R}_0]\right\}. \tag{IV.53}$$

Furthermore, we have for $\ell \in [0, \mathfrak{R}_0]$:

$$C_\star(\ell) = C(\eta_\star) \quad \text{with} \quad \eta_\star = \mathbb{1}_{\Omega_0} + \sum_{i\in I} \eta_{i,\star},$$

where $\eta_\star$ is Pareto optimal with $\mathfrak{R}_e[\mathrm{k}](\eta_\star) = \ell$, and, for $i \in I$, the strategy $\eta_{i,\star} = \eta_\star \mathbb{1}_{\Omega_i}$ restricted to Ω_i is Pareto optimal for the kernel k_i restricted to Ω_i, with $\mathfrak{R}_e[\mathrm{k}_i](\eta_{i,\star}) = \min(\ell, \mathfrak{R}_0[\mathrm{k}_i])$. We also have an upper bound for the minimal cost which ensures that no infection occurs at all:

$$c_\star = C_\star(0) \leq C(\mathbb{1}_{\Omega_0}).$$

Proof. Equation (IV.51) and the definition of $\mathfrak{R}_e^\star$ readily implies that $\mathfrak{R}_e^\star = \max_{i\in I} R_i^\star$.

We set $\mathfrak{R}_0 = \mathfrak{R}_0[\mathrm{k}]$ and recall that $\mathfrak{R}_e[\mathrm{k}_i](\mathbb{1}) = \mathfrak{R}_0[\mathrm{k}_i]$. Let $\ell \in (0, \mathfrak{R}_0]$. Notice that (IV.50) implies that there is a finite number of indices $i \in I$ such that $\mathfrak{R}_0[\mathrm{k}_i] \geq \ell$. This and (IV.51) readily implies that $C^\star(\ell) = \max_{i\in I} C_i^\star(\ell)$ for $\ell > 0$. Use that $C^\star(0) = C_i^\star(0) = c_{\max}$ to deduce that the equality $C^\star = \max_{i\in I} C_i^\star$ holds on $[0, \mathfrak{R}_0]$. The formula for $c^\star = C^\star(\mathfrak{R}_0)$ is a consequence of (IV.51), Lemma III.5.13 and Remark IV.7.1 (v). The formula (IV.53) for $\mathcal{F}^{\mathrm{Anti}}$ is then a consequence of (IV.46).

Eventually, if $\eta_\star$ is Pareto optimal with $\mathfrak{R}_e[\mathrm{k}](\eta_\star) = \ell$, we deduce from (IV.51) that $\mathfrak{R}_e[\mathrm{k}](\eta_\star \mathbb{1}_{\Omega_0^c})$ is also equal to ℓ, and since C is decreasing, this implies that $\eta_\star \geq \mathbb{1}_{\Omega_0}$ and thus $\eta_\star = \mathbb{1}_{\Omega_0} + \sum_{i\in I} \eta_{i,\star}$ with $\eta_{i,\star} = \eta_\star \mathbb{1}_{\Omega_i}$. Now if $\eta_{i,\star}$ were not Pareto optimal for the kernel k_i restricted to Ω_i or if $\mathfrak{R}_e[\mathrm{k}_i](\eta_{i,\star}) < \min(\ell, \mathfrak{R}_0[\mathrm{k}_i])$, we could increase $\eta_\star$ on Ω_i without changing the value of $\mathfrak{R}_e[\mathrm{k}]$, and thus $\eta_\star$ would not be Pareto optimal. Thus, we get that $\eta_{i,\star}$ is Pareto optimal for the kernel k_i restricted to Ω_i, that is, $\eta_{i,\star} + \mathbb{1}_{\Omega_0}$ is Pareto optimal for the kernel k_i, and that $\mathfrak{R}_e[\mathrm{k}_i](\eta_{i,\star}) = \min(\ell, \mathfrak{R}_0[\mathrm{k}_i])$. From the inequality $\eta_\star \geq \mathbb{1}_{\Omega_0}$, we deduce that $c_\star = C_\star(0) \leq C(\mathbb{1}_{\Omega_0})$. □

Remark IV.7.4. If k is not monatomic, then Assumption 7 in Chapter III (that is any local maximum of the loss function is also a global maximum) may or may not be satisfied for the loss function $\mathfrak{R}_e = \mathfrak{R}_e[\mathrm{k}]$, see the case of the two population model in [37]. In the former case the function $C^\star$ is continuous and the anti-Pareto frontier is connected, whereas in the latter case the function $C^\star$ may have jumps and then the anti-Pareto frontier has more than one connected component.

Chapter V

Optimal vaccination: various (counter) intuitive examples

Chapter Abstract

In previous the previous chapters, we formalized the problem of optimal allocation strategies for a (perfect) vaccine in an infinite-dimensional metapopulation SIS model. The aim of the current chapter is to illustrate this theoretical framework with multiple examples where one can derive the analytic expression of the optimal strategies. In particular, we investigate three questions: whether or not it is possible to vaccinate optimally when the vaccine doses are given one at a time; the effect of assortativity (the tendency to have more contacts with similar individuals) on the shape of optimal vaccination strategies; the particular case where everybody has the same number of neighbours.

The material for this chapter has been released in [38].

Chapter Content:

V.1 Introduction

V.1.1 Motivation

The basic reproduction number, denoted by $\mathfrak{R}_0$, plays a fundamental role in epidemiology as it determines the long-term behavior of an epidemic. For an homogeneous model, it is defined as the number of secondary cases generated by an infected individual in an otherwise susceptible population. When this number is below 1, an infected individual causes less than one infection before its recovery in average; the disease therefore declines over time until it eventually dies out. On the contrary, when the reproduction number is greater than 1, the disease will invade the population. From this property, we deduce that a proportion equal to $1 - 1/\mathfrak{R}_0$ of the population should be immunized in order to stop the outbreak. We refer the reader to the monograph of Keeling and Rohani [93] for a reminder of these basic properties on the reproduction number.

In heterogeneous generalizations of classical compartmental models, see [14, 102] and Chapter II, the population is stratified into homogeneous groups sharing the same characteristics (time to recover from the disease, interaction with the other groups, ...). For these models, it is still possible to define a meaningful reproduction number $\mathfrak{R}_0$, as the number of secondary cases generated by a typical infectious individual when all other individuals are uninfected; see [42]. With this definition, it is still true that the outbreak dies out if $\mathfrak{R}_0$ is smaller than 1 and invade the population otherwise; see [79, 150, 151, 153] and Chapter II for instance.

Suppose now that we have at our disposal a vaccine with *perfect efficacy*, that is, vaccinated individuals are completely immunized to the disease. After a vaccination campaign, let η denote the (non necessarily homogeneous) proportion of **non-vaccinated** individuals in the population: we will call η a *vaccination strategy*. For any strategy η, let us denote by $\mathfrak{R}_e(\eta)$ the corresponding reproduction number of the non-vaccinated population. The choice of η naturally raises a question that may be expressed as the following informal constrained optimization problem:

$$\begin{cases} \textbf{Minimize:} & \text{the quantity of vaccine to administrate} \\ \textbf{Such that:} & \text{herd immunity is reached, that is, } \mathfrak{R}_e \leq 1. \end{cases} \tag{V.1}$$

For practical reasons, we will instead look at the problem the other way around. If the vaccine is only available in limited quantities, the decision makers could try to allocate the doses so as to maximize efficiency; a natural indicator of this efficiency is the effective reproduction number. This reasoning leads to the following constrained problem:

$$\begin{cases} \textbf{Minimize:} & \text{the effective reproduction number } \mathfrak{R}_e \\ \textbf{Such that:} & \text{a given quantity of available vaccine.} \end{cases} \tag{V.2}$$

In accordance with Chapter III, we will denote by $\mathfrak{R}_{e\star}$ the value of this problem: it is a function of the quantity of available vaccine. The graph of this function is called the *Pareto frontier*. In order to measure how bad a vaccination strategy can be, we will also be interested in maximizing the effective reproduction number given a certain quantity of vaccine:

$$\begin{cases} \textbf{Maximise:} & \text{the effective reproduction number } \mathfrak{R}_e \\ \textbf{subject to:} & \text{a given quantity of available vaccine.} \end{cases} \tag{V.3}$$

The value function corresponding to this problem is denoted by $\mathfrak{R}_e^\star$ and its graph is called the *anti-Pareto frontier*. We will quantify the "quantity of available vaccine" for the vaccination strategy η by a cost $C(\eta)$. Roughly speaking the "best" (resp. "worst") vaccination strategies are solutions to Problem (V.2) (resp. Problem (V.3)) and will be called *Pareto optimal* (resp. *anti-Pareto optimal*) strategies.

The problem of optimal vaccine allocation has been studied mainly in the metapopulation setting where the population is divided into a finite number of subgroups with the same characteristics. Longini, Ackerman and Elverback were the first interested in the question of optimal vaccine distribution given a limited quantity of vaccine supply [110]. Using the concept of next-generation matrix introduced by Diekmann, Heesterbeek and Metz [42], Hill and Longini reformulated this problem thanks to the reproduction number [81]. Several theoretical and numerical studies followed focusing on Problem (V.1) and/or Problem (V.2) in the metapopulation setting [47, 67, 75, 130]. We also refer the reader to the introduction of Chapter III for a detailed review of the bibliography.

In the two previous Chapters, we provided a framework where Problems (V.2) and (V.3) are well-posed, justified that the optimizers are indeed Pareto optimal and studied in detail the Pareto and anti-Pareto frontiers. Since there is no closed form for the effective reproduction number, Problems (V.2) and (V.3) are hard to solve in full generality: our goal here is to exhibit examples where one can derive analytic expressions for the optimal vaccination strategies. The simple models we study give a gallery of examples and counter-examples to natural questions or conjectures, and may help understanding common rules of thumb for choosing vaccination policies. We will in particular be interested in the following two notions.

(i) (**Greedy parametrization of the frontiers**). For the decision maker it is important to know if global optimization and sequential optimization are the same as one cannot unvaccinate people and redistribute the vaccine once more doses become available. More precisely, there is a natural order on the vaccination strategies: let us write $\eta' \leq \eta$ if all the people that are vaccinated when following the strategy η are also vaccinated when following the strategy η'. Let η be an optimal solution of (V.2) for cost $c = C(\eta)$, that is, $\mathfrak{R}_e(\eta) = \mathfrak{R}_{e\star}(c)$. If, for $c' > c$, we can find a strategy $\eta' \leq \eta$ such that $\mathfrak{R}_e(\eta') = \mathfrak{R}_{e\star}(c')$, then the optimization may be, at least in principle, found in a greedy way: giving sequentially each new dose of vaccine so as to minimize $\mathfrak{R}_e$ gives, in the end, an optimal strategy for any quantity of vaccine. By analogy with the corresponding notion for algorithms we will say in this case that there exists a *greedy parametrization* of the Pareto frontier. The existence of such a greedy parametrization was already discussed by Cairns in [29] and is examined for each model throughout this chapter.

(ii) (**Assortative/Disassortative network**). The second notion is a property of the network called *assortativity*: a network is called assortative when the nodes tend to attach to others that are similar in some way and *disassortative* otherwise. The assortativity or disassortativity of a network is an important property that helps to understand its topology. It has been oberved that social networks are usually disassortative while biological and technological networks are disassortative, see for example [120]. The optimal vaccination strategies can differ dramatically in the case of assortative versus disassortative mixing, see Galeotti and Rogers [65] for a study in a population composed of two groups.

V.1.2 Main results

Section V.2 is dedicated to metapopulation models. We present two simple models that, despite being seemingly very similar, display totally different behaviors: the asymmetric and symmetric circle graphs. For the first one, we derive a greedy parametrization of the Pareto frontier. On the second one, we observe numerically that the Pareto frontier is much more complicated, and in particular cannot be parametrized greedily.

After Section V.3, where we recall the kernel setting used in Chapter III, we focus in Section V.4 on the effect of assortativity on optimal vaccination strategies. We define a simple kernel model that may be assortative or disassortative depending on the value of two parameters. We describe completely the optimal vaccination strategies, and show that the best strategies for the assortative case are the worst ones if the mixing pattern is disassortative, and vice-versa. We also prove that all the Pareto and anti-Pareto frontiers admit greedy parametrizations, and that Pareto optimal strategies prioritize individuals that in some sense have the highest degree, that is, are the most connected.

Targeting individuals that are the most connected is a common approach used to prevent an epidemic in a complex network [126]. In [40], we show that these strategies are optimal for the so-called monotonic kernel models in which we can naturally order individuals according to their connectivity. The question then arises of how to treat a model in which all individuals share the same degree. This motivates Section V.5 which is devoted to constant degree kernels. Those are the analogue of regular graphs in the infinite-dimensional setting. We prove a general result that gives a sufficient conditions on the kernel to make the uniform strategies either the "best" or the "worst" possible strategies. In Section V.6, we study in detail one particular case of a constant degree kernel. We also provide an example of kernel for which the set of optimal strategies has an infinite number of connected components. In this particular case, there is no greedy parametrization of the Pareto frontier.

V.2 First examples in the discrete setting

In this section, we use the framework developped by Hill and Longini in [81] for metapopulation models and provide optimal vaccination strategies for two very simple examples. Despite their simplicity, these examples showcase a number of interesting behaviors, that will occur a in much more general setting, as we will see in the rest of the chapter.

V.2.1 The reproduction number in metapopulation models

In metapopulation models, the population is divided into $N \geq 2$ different subpopulations and we suppose that individuals within a same subpopulation share the same characteristics. The different groups are labeled $0, 1, \ldots, N-1$. We denote by $\mu_0, \mu_1, \ldots, \mu_{N-1}$ their respective size (in proportion with respect to the total size) and we suppose that those do not change over time. By the linearization of the dynamic of the epidemic at the disease-free equilibrium, we obtain the so-called *next-generation matrix* K, see [153], which is a $N \times N$ matrix with non-negative coefficients. For a detailled discussion on the biological interpretation of the coefficients of the next-generation matrix, we refer the reader to Section IV.2. We also refer to [37] for an extensive treatment of the case $N = 2$.

The basic reproduction number is equal to the spectral radius of the next-generation matrix:

$$\mathfrak{R}_0 = \rho(K), \tag{V.4}$$

where ρ denotes the spectral radius. Since the matrix K has non-negatives entries, Perron-Frobenius theory implies that $\mathfrak{R}_0$ is also an eigenvalue of K. If $\mathfrak{R}_0 > 1$, the epidemic process grows away from zero infectives while if $\mathfrak{R}_0 < 1$, the disease cannot invade the population; see [153, Theorem 2].

Now, let us introduce the effect of vaccination. Suppose that we have at our disposal a vaccine with perfect efficacy, *i.e.*, vaccinated individuals are completely immunized to the infection. We denote by $\eta = (\eta_0, \ldots, \eta_{N-1})$ the vector of the proportions of **non-vaccinated** individuals in the different groups. We shall call η a vaccination strategy and denote by $\Delta = [0, 1]^N$ the set of all possible vaccination strategies. According to Chapters III and IV, the next-generation matrix corresponding to the dynamic with vaccination is equal to the matrix K multiplied by the matrix $\mathrm{Diag}(\eta)$ on the right, where $\mathrm{Diag}(\eta)$ is the $N \times N$ diagonal matrix with coefficients $\eta \in \Delta$. We call the spectral radius of this matrix the *effective reproduction number*:

$$\mathfrak{R}_e(\eta) = \rho\left(K \cdot \mathrm{Diag}(\eta)\right). \tag{V.5}$$

The effective reproduction number accounts for the vaccinated (and immunized) people in the population, as opposed to the basic reproduction number, which corresponds to a fully susceptible population. When nobody is vaccinated, that is $\eta = \mathbb{1} = (1, \ldots, 1)$, $\mathrm{Diag}(\eta)$ is equal to the identity matrix, the next-generation matrix is unchanged and $\mathfrak{R}_e(\eta) = R(\mathbb{1}) = \mathfrak{R}_0$.

We suppose that the *cost* of a vaccination strategy is, up to an irrelevant multiplicative constant, equal to the total proportion of vaccinated people and is therefore given by:

$$C(\eta) = \sum_{i=0}^{N-1} (1 - \eta_i)\mu_i = 1 - \sum_{i=0}^{N-1} \eta_i \mu_i, \tag{V.6}$$

where $\eta = (\eta_0, \dots, \eta_{N-1}) \in \Delta$. We refer to Section III.5.1 and Remark III.5.2. for considerations on more general cost functions.

Example V.2.1 (Uniform vaccination). The uniform strategy of cost c consists in vaccinating the same proportion of people in each group: $\eta = (1-c)\mathbb{1}$. By homogeneity of the spectral radius, the reproduction number $\mathfrak{R}_e(\eta)$ is then equal to $(1-c)\mathfrak{R}_0$.

V.2.2 Optimal allocation of vaccine doses

As mentioned in the introduction and recalled in Section V.2.1, reducing the reproduction number is fundamental in order to control and possibly eradicate the epidemic. However, the vaccine may only be available in a limited quantity, and/or the decision maker may wish to limit the cost of the vaccination policy. This motivates our interest in the following related problem:

$$\begin{cases} \min & \mathfrak{R}_e(\eta), \\ \text{such that} & C(\eta) = c. \end{cases} \tag{V.7}$$

According to Chapter III, one can replace the constraint $\{C(\eta) = c\}$ by $\{C(\eta) \leq c\}$ without modifying the solutions. The opposite problem consists in finding out the *worst possible way* of allocating vaccine. While this does not seem at first sight to be as important, a good understanding of bad vaccination strategies may also provide rules of thumb in terms of anti-patterns. In order to estimate how bad a vaccination strategy can be, we therefore also consider the following problem:

$$\begin{cases} \max & \mathfrak{R}_e(\eta), \\ \text{such that} & C(\eta) = c. \end{cases} \tag{V.8}$$

According to Chapter III, one can replace the constraint $\{C(\eta) = c\}$ by $\{C(\eta) \geq c\}$ without modifying the solutions.

Since the coefficients of the matrix $K\mathrm{Diag}(\eta)$ depend continuously on η, it is classical that its eigenvalues also depend continuously on η (see for example [87, Appendix D]) and in particular the function $\mathfrak{R}_e$ is continuous on $\Delta = [0,1]^N$. Since the function C is also continuous on Δ, the compactness of Δ ensures the existence of solutions for Problems (V.7) and (V.8). For $c \in [0,1]$, $\mathfrak{R}_{e\star}(c)$ (resp. $\mathfrak{R}_e^\star(c)$) stands for the minimal (resp. maximal) value taken by $\mathfrak{R}_e$ on the set of all vaccination strategies η such that $C(\eta) = c$:

$$\mathfrak{R}_{e\star}(c) = \min\{\mathfrak{R}_e(\eta) \; : \; \eta \in \Delta \text{ and } C(\eta) = c\}, \tag{V.9}$$

$$\mathfrak{R}_e^\star(c) = \max\{\mathfrak{R}_e(\eta) \; : \; \eta \in \Delta \text{ and } C(\eta) = c\}. \tag{V.10}$$

It is easy to check that the functions $\mathfrak{R}_{e\star}$ and $\mathfrak{R}_e^\star$ are non increasing. Indeed, if η^1 and η^2 are two vaccination strategies such that $\eta^1 \leq \eta^2$ (where $\leq$ stands for the pointwise order), then $\mathfrak{R}_e(\eta^1) \leq \mathfrak{R}_e(\eta^2)$ according to the Perron-Frobenius theory. This easily implies that $\mathfrak{R}_{e\star}$ and $\mathfrak{R}_e^\star$ are non-increasing. We refer to Chapters III and IV for more properties on those functions; in particular they are also continuous. For the vaccination strategy $\eta = \mathbb{0} = (0, \dots, 0)$ (everybody is vaccinated) with cost $C(\mathbb{0}) = 1$, the transmission of the disease in the population is completely stopped, *i.e.*, the reproduction number is equal to 0. In the examples below, we will see that for some next-generation matrices K, this may be achieved with a strategy η with cost $C(\eta) < 1$. Hence, let us denote by $c_\star$ the minimal cost required to completely stop the transmission of the disease:

$$c_\star = \inf\{c \in [0,1] \; : \; \mathfrak{R}_{e\star}(c) = 0\}. \tag{V.11}$$

In a similar fashion, we define by symmetry the maximal cost of totally inefficient vaccination strategies:

$$c^\star = \sup\{c \in [0,1] \; : \; \mathfrak{R}_e^\star(c) = \mathfrak{R}_0\}. \tag{V.12}$$

According to Lemma III.5.13, we have $c^\star = 0$ if the matrix K is irreducible, *i.e.*, not similar via a permutation to a block upper triangular matrix. The two matrices considered below in this section are irreducible.

Following Chapter III, the *Pareto frontier* associated to the "best" vaccination strategies, solution to Problem V.7, is defined by:

$$\mathcal{F} = \{(c, \mathfrak{R}_{e\star}(c)) \; : \; c \in [0, c_\star]\}. \tag{V.13}$$

The set of "best" vaccination strategies, called *Pareto optimal* strategies, is defined by:

$$\mathcal{P} = \{\eta \in \Delta \; : \; , (C(\eta), \mathfrak{R}_e(\eta)) \in \mathcal{F}\}. \tag{V.14}$$

When $c^\star = 0$ (which will be the case for all the examples considered in this chapter), the *anti-Pareto frontier* associated to the "worst" vaccination strategies, solution to Problem V.8, is defined by:

$$\mathcal{F}^{\text{Anti}} = \{(c, \mathfrak{R}_e^\star(c)) \; : \; c \in [0, 1]\}. \tag{V.15}$$

The set of "worst" vaccination strategies, called *anti-Pareto optimal* strategies, is defined by:

$$\mathcal{P}^{\text{Anti}} = \{\eta \in \Delta \; : \; (C(\eta), \mathfrak{R}_e(\eta)) \in \mathcal{F}^{\text{Anti}}\}. \tag{V.16}$$

The set of uniform strategies will play a role in the sequel:

$$\mathcal{S}^{\text{uni}} = \{t\mathbb{1} \; : \; t \in [0, 1]\}. \tag{V.17}$$

We denote by $\mathbf{F} = \{(C(\eta), \mathfrak{R}_e(\eta)) \; : \; \eta \in \Delta\}$ the set of all possible outcomes. According to Section 6.1 in Chapter III, the set $\mathbf{F}$ is a subset of $[0, 1] \times [0, \mathfrak{R}_0]$ delimited below by the graph of $\mathfrak{R}_{e\star}$ and above by the graph of $\mathfrak{R}_e^\star$; it is compact, path connected and its complement is connected in $\mathbb{R}^2$.

A *path* of vaccination strategies is a measurable function $\gamma : [a, b] \to \Delta$ where $a < b$. It is *monotone* if for all $a \leq s \leq t \leq b$ we have $\gamma(s) \geq \gamma(t)$, where $\leq$ denotes the pointwise order. A *greedy parametrization* of the Pareto (resp. anti-Pareto) frontier is a monotone continuous path γ such that the image of $(C \circ \gamma, \mathfrak{R}_e \circ \gamma)$ is equal to $\mathcal{F}$ (resp. $\mathcal{F}^{\text{Anti}}$). If such a path exists, then its image can be browsed by a greedy algorithm which performs infinitesimal locally optimal steps.

Remark V.2.2. Let K be the next-generation matrix and let $\lambda \in \mathbb{R}_+ \backslash \{0\}$. By homogeneity of the spectral radius, we have $\rho(\lambda K \cdot \text{Diag}(\eta)) = \lambda \rho(K \cdot \text{Diag}(\eta))$. Thus, the solutions of Problems (V.7) and (V.8) and the value of $c_\star$ are invariant by scaling of the matrix K. As for the functions $\mathfrak{R}_{e\star}$ and $\mathfrak{R}_e^\star$, they are scaled by the same quantity. Hence, in our study, the value of $\mathfrak{R}_0$ will not matter. Our main concern will be to find the best and the worst vaccination strategies for a given cost and compare them to the uniform strategy.

V.2.3 The fully asymmetric circle model

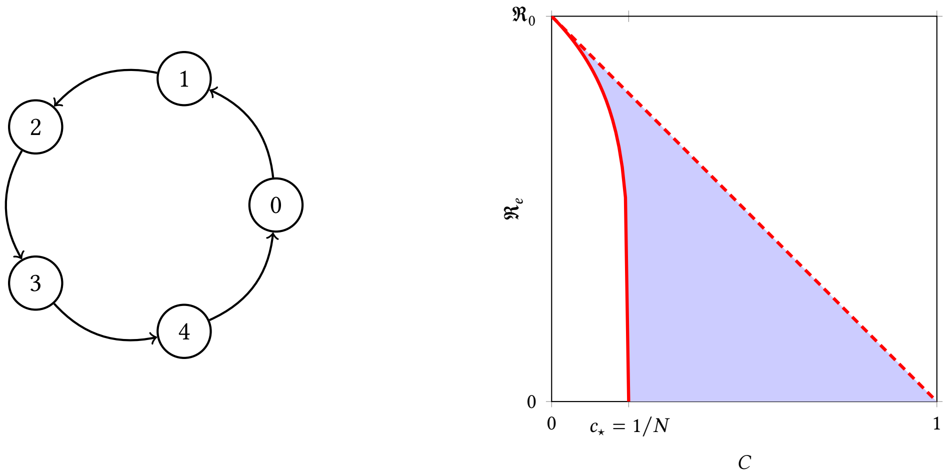
We consider a model of $N \geq 2$ equal subpopulations (*i.e.* $\mu_0 = \cdots = \mu_{N-1} = 1/N$) where each subpopulation only contaminates the next one. The next-generation matrix, which is equal to the cyclic permutation matrix, and the effective next generation matrix are given by:

$$K = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ 0 & & & 0 & 1 \\ 1 & 0 & & & 0 \end{pmatrix} \quad \text{and} \quad K \cdot \text{Diag}(\eta) = \begin{pmatrix} 0 & \eta_0 & & & \\ & 0 & \eta_1 & & \\ & & \ddots & \ddots & \\ 0 & & & 0 & \eta_{N-2} \\ \eta_{N-1} & 0 & & & 0 \end{pmatrix}, \tag{V.18}$$

where $\eta = (\eta_0, \ldots, \eta_{N-1}) \in [0, 1]^N$. The next-generation matrix can be interpreted as the adjacency matrix of a directed graph. In this case, this is the fully asymmetric cyclic graph; see Figure V.1(A).

By an elementary computation, the characteristic polynomial of the matrix $K \cdot \text{Diag}(\eta)$ is equal to $X^N - \prod_{0 \leq i \leq N-1} \eta_i$. Hence, the effective reproduction number can be computed via an explicit formula, it corresponds to the geometric mean:

$$\mathfrak{R}_e(\eta) = \left(\prod_{i=0}^{N-1} \eta_i \right)^{1/N}. \tag{V.19}$$



(A) Graphical representation of the transmission of the disease.

(B) Red line: the Pareto frontier $\mathcal{F}$; dashed line: the anti-Pareto frontier $\mathcal{F}^{\text{Anti}}$ (which corresponds to the uniform strategies); blue region: all possible outcomes $\mathbf{F}$.

Figure V.1: Example of optimization for the fully asymmetric circle model with $N = 5$ subpopulations.

The Pareto and anti-Pareto frontier are totally explicit for this elementary example, and given by the following proposition. For additional comments on this example see also Example V.5.9 below.

Proposition V.2.3 (Asymmetric circle). *For the fully asymmetric circle model, we have:*

(i) The least quantity of vaccine necessary to completely stop the propagation of the disease is $c_\star = 1/N$. Pareto optimal strategies have a cost smaller than $c_\star$, and correspond to giving all the available vaccine to one subpopulation:

$$\mathcal{P} = \left\{\eta = (\eta_0, \dots, \eta_{N-1}) \in [0,1]^N \ : \ \eta_i = 1 \textit{ for all } i \textit{ but at most one}\right\}.$$

The Pareto frontier is given by the graph of the function $\mathfrak{R}_{e\star}$ on $[0, c_\star]$, where $\mathfrak{R}_{e\star}$ is given by:

$$\mathfrak{R}_{e\star}(c) = (1 - Nc)_+^{1/N} \quad \textit{for} \quad c \in [0,1].$$

(ii) The maximal cost of totally inefficient vaccination strategies is $c^\star = 0$. The anti-Pareto optimal strategies consist in vaccinating uniformly the population, i.e.*:*

$$\mathcal{P}^{\text{Anti}} = \mathcal{S}^{\text{uni}}.$$

The anti-Pareto frontier is given by the graph of the function $\mathfrak{R}_e^\star \ : \ c \mapsto 1 - c$.

Remark V.2.4 (Greedy parametrization). From Proposition V.2.3, we see that there exists a greedy parametrization of the Pareto frontier, which consists in giving all the available vaccine to one subpopulation until its complete immunization. Similarly, the anti-Pareto frontier is greedily parametrized by the uniform strategies.

Proof. We first prove (i). Suppose that $c \geq 1/N$. There is enough vaccine to protect entirely one of the groups and obtain $\mathfrak{R}_e(\eta) = 0$ thanks to Equation (V.19). This gives $c_\star \leq 1/N$ and $\mathfrak{R}_{e\star}(c) = 0$ for $c \geq 1/N$.

Let $0 \leq c < 1/N$. According to [23, Section 3.1.5], the map $\eta \mapsto \mathfrak{R}_e(\eta)$ is concave. According to Bauer's maximum principle [121, Corollary A.3.3], $\mathfrak{R}_e$ attains its minimum on $\{\eta \in [0,1]^N \ : \ C(\eta) = c\}$ at some extreme point of this set. These extreme points are strategies $\eta \in [0,1]^N$ such that $\eta_i = 1 - Nc$ for some i and $\eta_j = 1$ for all $j \neq i$. Since $\mathfrak{R}_e$ is a symmetric function of its N variables, it takes the same value $(1-Nc)^{1/N}$ on all these strategies, so they are all minimizing, which proves Point (i).

We give another elementary proof of (i) when $c < 1/N$. Let η be a solution of Problem (V.7). Assume without loss of generality that $\eta_0 \leq \cdots \leq \eta_{N-1}$. Suppose for a moment that $\eta_1 < 1$, and let $\varepsilon > 0$ be small enough to ensure $\eta_0 > \varepsilon$ and $\eta_1 < 1-\varepsilon$. Then the vaccination strategy $\tilde{\eta} = (\eta_0 - \varepsilon, \eta_1 + \varepsilon, \eta_2, \dots, \eta_{N-1})$ is admissible, and:

$$\mathfrak{R}_e(\tilde{\eta})^N = \mathfrak{R}_e(\eta)^N - (\varepsilon(\eta_1 - \eta_0) + \varepsilon^2) \prod_{i=2}^{N-1} \eta_i < \mathfrak{R}_e(\eta)^N,$$

contradicting the optimality of η. Therefore the Pareto-optimal strategies have only one term different from 1, and must be equal to $((1-Nc), 1, \dots, 1)$, up to a permutation of the indices.

Now, let us prove (ii). Let η such that $C(\eta) = c$. According to the inequality of arithmetic and geometric means:

$$\mathfrak{R}_e(\eta) \leq \frac{\eta_0 + \cdots + \eta_{N-1}}{N} = 1 - c.$$

By Example V.2.1, the right hand side is equal to the effective reproduction number of the uniform vaccination at cost c. This ends the proof of the proposition. □

In Figure V.1(b), we have plotted the Pareto and the anti-Pareto frontiers corresponding to asymmetric circle model with $N = 5$ subpopulations.

V.2.4 Fully symmetric circle model

We now consider the case where each of the N subpopulation may infect both of their neighbours. The next-generation matrix and the effective next-generation matrix are given by:

$$K = \begin{pmatrix} 0 & 1 & & 0 & 1 \\ 1 & 0 & 1 & & 0 \\ & 1 & \ddots & \ddots & \\ 0 & & \ddots & 0 & 1 \\ 1 & 0 & & 1 & 0 \end{pmatrix} \quad \text{and} \quad K \cdot \mathrm{Diag}(\eta) = \begin{pmatrix} 0 & \eta_0 & & 0 & \eta_0 \\ \eta_1 & 0 & \eta_1 & & 0 \\ & \eta_2 & \ddots & \ddots & \\ 0 & & \ddots & 0 & \eta_{N-2} \\ \eta_{N-1} & 0 & & \eta_{N-1} & 0 \end{pmatrix}. \tag{V.20}$$

Again, we can represent this model as a graph; see Figure V.2(a). There is no closed-form formula to express $\mathfrak{R}_e$ for $N \geq 5$ and the optimization is way harder than the asymmetric case. Since K is irreducible, we have $c^\star = 0$. Our only analytical result for this model is the computation of $c_\star$.

Proposition V.2.5 (Optimal strategy for stopping the transmission). *For the fully symmetric circle model, the strategy $\eta^* = \mathbb{1}_{i\ odd}$ is Pareto optimal for the fully symmetric circle and $\mathfrak{R}_e(\eta^*) = 0$. In particular, $c_\star$ is equal to $C(\eta^*) = \lceil N/2 \rceil / N$.*

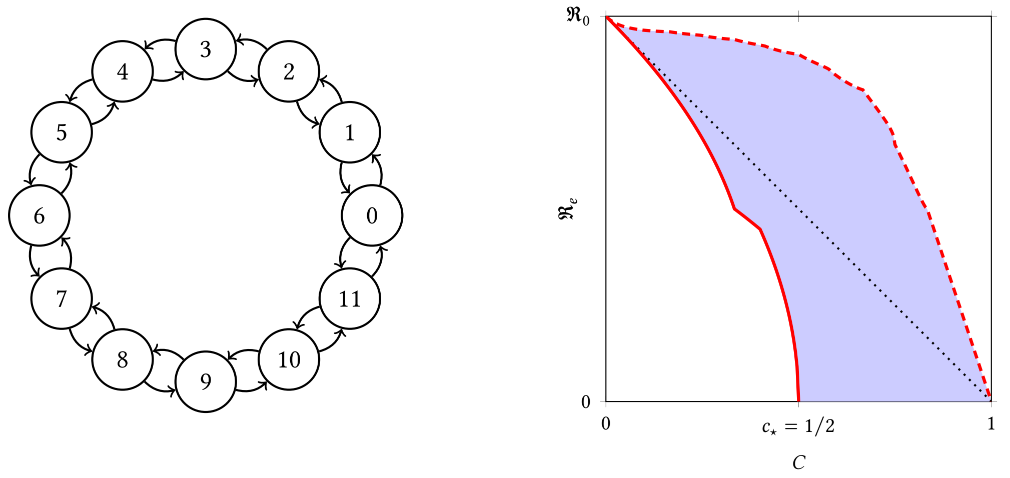
Proof. The term X^{N-2} of the characteristic polynomial of $K\mathrm{Diag}(\eta)$ has a coefficient equal to the sum of all principal minors of size 2:

$$\eta_0\eta_1 + \eta_1\eta_2 + \dots + \eta_{N-2}\eta_{N-1} + \eta_{N-1}\eta_0. \tag{V.21}$$

If η is such that $NC(\eta) < \lceil N/2 \rceil$, then at least one of the term above is not equal to 0, proving that the sum is positive. Hence, there is at least one eigenvalue of $K\mathrm{Diag}(\eta)$ different from 0, and $\mathfrak{R}_e(\eta) > 0$. We deduce that $c_\star \geq \lceil N/2 \rceil / N$.

Now, let η^* such that $\eta_i^* = 0$ for all even i and $\eta_i^* = 1$ for all odd i, so that $C(\eta^*) = \lceil N/2 \rceil / N$. The matrix $K\mathrm{Diag}(\eta^*)$ is nilpotent as its square is 0. Since the spectral radius of a nilpotent matrix is equal 0, we get $\mathfrak{R}_e(\eta^*) = 0$. This ends the proof of the proposition.

We can give another proof of the proposition: it is enough to notice that the nodes labelled with an odd number form a maximal independent set of the cyclic graph. Taking η^* equal to the indicator function of this set, we deduce from Section IV.6.4 that η^* is Pareto optimal, $\mathfrak{R}_e(\eta^*) = 0$ and $c_\star = C(\eta^*)$. □



(A) Graphical representation of the transmission of the disease.

(B) Red line: the Pareto frontier $\mathcal{F}$; dashed line: the anti-Pareto frontier $\mathcal{F}^{\text{Anti}}$; dotted line: path of the uniform strategies; blue region: all possible outcomes **F**.

Figure V.2: Example of optimization for the fully symmetric circle model with $N = 12$ subpopulations.

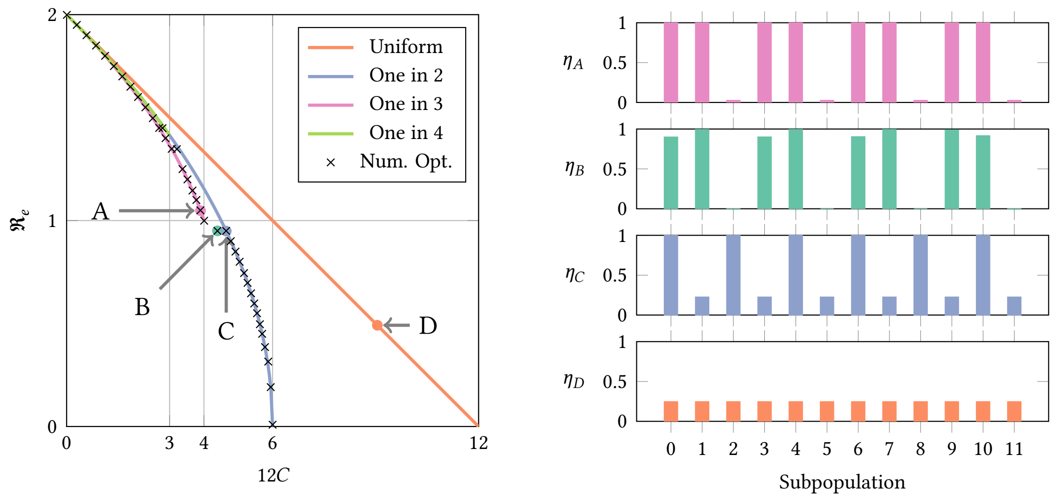
We pursue the analysis of this model with numerical computations. We choose $N = 12$ subpopulations, and compute an approximate Pareto frontier, using the Borg multiobjective evolutionary algorithm[1]. The results are plotted in Figure V.3. We represent additionnally the curves $(c, R(\eta(c)))$ where the vaccination strategy $\eta(c)$ for a given cost c are given by deterministic path of "meta-strategies":

- **Uniform strategy:** distribute the vaccine uniformly to all N subpopulations;
- **"One in j" strategy:** vaccinate one in j subpopulation, for $j = 2, 3, 4$.

Let us follow the scatter plot of $\mathfrak{R}_{e\star}$ in Figure V.3(A), starting from the upper left.

1. In the beginning nobody is vaccinated, and $\mathfrak{R}_0$ is equal to 2.
2. For small costs all strategies have similar efficiency. Zooming shows that the (numerically) optimal strateges split the available vaccine equally between four subpopulations that are separed from each other by two subpopulations. This corresponds to the "one in 3" meta-strategies path. As represented in Figure V.3(B), η_A with outcome point $A = (C(\eta_A), \mathfrak{R}_e(\eta_A))$ belongs to this path. In particular, note that disconnecting the graph is not Pareto optimal for $12c = 3$ as the disconnecting "one in 4" strategy gives values $\mathfrak{R}_e = \sqrt{2} \simeq 1.41$ opposed to the value $\mathfrak{R}_e \simeq 1.37$ for the "one in 3" strategy with same cost. However, note that, in agreement with Proposition IV.6.5, this disconnecting "one in 4" strategy is also not anti-Pareto optimal, since it performs better than the uniform strategy with the same cost.
3. When $12c = 4$ the circle has been split in four "islands" of two interacting subpopulations. There is a small interval of values of c for which it is (numerically) optimal to split the additional vaccine uniformly between the four "islands", and give it entirely to one subpopulation in each island: see point B and the associated strategy η_B.
4. Afterwards (see point C), it is in fact better to try and vaccinate all the (say) even numbered subpopulations. Therefore, the optimal vaccinations *do not vary monotonously* with respect to the amount of available vaccine; in other words, distributing vaccine in a greedy way is

[1]The algorithm is described in [71]; we use the version coded in the `BlackBoxOptim` package for the Julia programming language.



(A) Effective reproduction number $\mathfrak{R}_e$ against vaccination cost c for various meta-strategies.

(B) Vaccination strategies corresponding to the four labelled points.

Figure V.3: Pareto frontier and computation of the outcomes for the paths of the four meta-strategies. Some meta-strategies $\{\eta_A, \eta_B, \eta_C, \eta_D\}$ are represented on the right with their corresponding outcome points $\{A, B, C, D\}$ on the left.

not optimal. This also suggests that, even though the frontier is continuous (in the objective space (c, r)), the set of optimal *strategies* may not be connected: the "one in two" vaccination strategy of point C cannot be linked to "no vaccination" strategy by a continuous path of optimal strategies. In particular, the Pareto frontier cannot be greedily parametrized. The disconnectedness of the set of optimal strategies will be established rigorously in Section V.6 for another model.

5. For $12c = 6$, that is $c = c_\star$ as stated in Proposition V.2.5, it is possible to vaccinate completely all the (say) even numbered subpopulations, thereby disconnecting the graph completely. The infection cannot spread at all.

6. Even though the problem is symmetric and all subpopulations play the same role, the proportional allocation of vaccine is far from optimal; on the contrary, the optimal allocations focus on some subpopulations.

Using the same numerical algorithm, we have also computed the anti-Pareto frontier for this model; see the dashed line in Figure V.2(B). Although we do not give a formal proof, the anti-Pareto frontier seems to be perfectly given by the following greedy parametrization:

1. Distribute all the available vaccine supply to one group until it is completely immunized.

2. Once this group is fully vaccinated, distribute the vaccine doses to one of its neighbour.

3. Continue this procedure by vaccinating the neighbour of the last group that has been immunized.

4. When there are only two groups left, the vaccine shall be split equitably between these two.

V.3 The kernel model

In order to get a finer description of the heterogeneity, we could divide the population into a growing number of subgroups $N \to \infty$. The recent advances in graph limits theory [12, 111] justify describing the transmission of the disease by a kernel defined on a probability space.

Let $(\Omega, \mathscr{F}, \mu)$ be a probability space that represents the population: the individuals have features labeled by Ω and the infinitesimal size of the population with feature x is given by

$\mu(\mathrm{d}x)$. Let $L^2(\mu)$ (L^2 for short) be the space of real-valued measurable functions f defined on Ω such that $\| f\|_2 = (\int_\Omega f^2 \, \mathrm{d}\mu)^{1/2}$ is finite, where functions which agree μ-a.s. are identified. Let $L^2_+ = \{f \in L^2 \, : \, f \geq 0\}$ be the subset of non-negative functions of L^2. We define a *kernel* on Ω as a $\mathbb{R}_+$-valued measurable function defined on $(\Omega^2, \mathscr{F}^{\otimes 2})$. We will only consider kernels with finite double-norm on L^2:

$$\| \mathrm{k} \|_{2,2} = \int_{\Omega\times\Omega} \mathrm{k}(x,y)^2 \, \mu(\mathrm{d}x)\mu(\mathrm{d}y) < +\infty. \tag{V.22}$$

To a kernel k with finite doubel norm on L^2, we associate the integral operator T_k on L^2 defined by:

$$T_\mathrm{k}(g)(x) = \int_\Omega \mathrm{k}(x,y)g(y) \, \mu(\mathrm{d}y) \quad \text{for } g \in L^2 \text{ and } x \in \Omega. \tag{V.23}$$

The operator T_k is bounded with operator norm $\| T_\mathrm{k} \|_{L^2}$ such that:

$$\| T_\mathrm{k} \|_{L^2} \leq \| \mathrm{k} \|_{2,2} \, . \tag{V.24}$$

According to [68, p. 293], T_k is actually compact (and even Hilbert-Schmidt). A kernel is said to be symmetric if $\mathrm{k}(x,y) = \mathrm{k}(y,x)$, $\mu(\mathrm{d}x)\mu(\mathrm{d}y)$ almost surely. It is said to be *irreducible* if for all $A \in \mathscr{F}$, we have:

$$\int_{A\times A^\mathrm{c}} \mathrm{k}(x,y) \, \mu(\mathrm{d}x)\mu(\mathrm{d}y) = 0 \implies \mu(A) \in \{0,1\}. \tag{V.25}$$

If k is not irreducible, it is called *reducible*.

By analogy with the discrete setting and also based on Chapters II and IV, we define the basic reproduction number in this context thanks to the following formula:

$$\mathfrak{R}_0 = \rho(T_\mathrm{k}), \tag{V.26}$$

where ρ stands for the spectral radius of an operator. According to the Krein-Rutman theorem, $\mathfrak{R}_0$ is an eigenvalue of T_k. Besides, there exists left and right eigenvectors associated to this eigenvalue in L^2_+; such functions are called Perron eigenfunctions.

For f, g two non-negative bounded measurable functions defined on Ω and k a kernel on Ω with finite double norm on L^2, we denote by $f\mathrm{k}g$ the kernel on Ω defined by:

$$(f\mathrm{k}g)(x,y) = f(x)\,\mathrm{k}(x,y)g(y). \tag{V.27}$$

Since f and g are bounded, the kernel $f\mathrm{k}g$ has also a finite double norm on L^2.

Denote by Δ the set of measurable functions defined on Ω taking values in $[0,1]$. A function η in Δ represents a vaccination strategy: $\eta(x)$ represents the proportion of **non-vaccinated** individuals with feature x. In particular $\eta = \mathbb{1}$ (the constant function equal to 1) corresponds to no vaccination and $\eta = \mathbb{0}$ (the constant function equal to 0) corresponds to the whole population vaccinated. The uniform strategies are given by:

$$\eta^{\mathrm{uni}} = t\mathbb{1}$$

for some $t \in [0,1]$, and we denote by $\mathcal{S}^{\mathrm{uni}} = \{t\mathbb{1} \, : \, t \in [0,1]\}$ the set of uniform strategies.

The (uniform) cost of the vaccination strategy $\eta \in \Delta$ is given by the total proportion of vaccinated people, that is:

$$C(\eta) = \int_\Omega (1-\eta) \, \mathrm{d}\mu = 1 - \int_\Omega \eta \, \mathrm{d}\mu. \tag{V.28}$$

The measure $\eta \, \mathrm{d}\mu$ corresponds to the *effective population*, that is the individuals who effectively play a role in the dynamic of the epidemic. The effective reproduction number is defined by:

$$\mathfrak{R}_e(\eta) = \rho(T_{\mathrm{k}\eta}), \tag{V.29}$$

We consider the weak topology on Δ, so that with a slight abuse of notation we identify Δ with $\{\eta \in L^2 \, : \, 0 \leq \eta \leq 1\}$. According to Theorem III.4.2, the function $\mathfrak{R}_e : \eta \mapsto \mathfrak{R}_e(\eta)$ is continuous on Δ equipped with the weak topology. Since Δ is compact with respect to this topology, we have deduced the existence of solutions for Problems (V.7) and (V.8). We will conserve the same notation and definitions as in the discrete setting for: the value functions $\mathfrak{R}_{e\star}$ and $\mathfrak{R}_e^\star$, the minimal/maximal costs $c_\star$ and $c^\star$, the various sets of strategies $\mathcal{S}^{\text{uni}}$, $\mathcal{P}$ and $\mathcal{P}^{\text{Anti}}$, and the various frontiers $\mathcal{F}$ and $\mathcal{F}^{\text{Anti}}$, see Equations (V.9)-(V.17) in Section V.2.2.

We shall also use the following result from Corollary IV.6.1 (recall that a vaccination strategy is defined up the a.s. equality).

Lemma V.3.1. *Let* k *be a kernel on* Ω *with finite double norm on* L^2 *such that a.s.* $\mathrm{k} > 0$. *Then, we have* $c^\star = 0$, $c_\star = 1$ *and the strategy* $\mathbb{1}$ *(resp.* $\mathbb{0}$*) is the only Pareto optimal as well as the only anti-Pareto optimal strategy with cost* $c = 0$ *(resp.* $c = 1$*).*

Example V.3.2 (Discrete and continuous representations of a metapopulation model). We recall the natural correspondence between metapopulation models (discrete models) and kernel models (continuous models) from Section 7.4 in Chapter III. Consider a metapopulation model with N groups given by a finite set $\Omega_{\mathrm{d}} = \{0, 1, \dots, N-1\}$ equipped with a probability measure μ_{d} giving the relative size of each group and a next generation matrix $K = (K_{ij},\ i, j \in \Omega_{\mathrm{d}})$. the discrete σ-algebra. The corresponding discrete kernel k_{d} on Ω_{d}) is defined by:

$$K_{ij} = \mathrm{k}_{\mathrm{d}}(i, j)\mu_j \quad \text{where} \quad \mu_i = \mu_{\mathrm{d}}(\{i\}). \tag{V.30}$$

Then, the matrix $K \cdot \mathrm{Diag}(\eta)$ is the matrix representation of the endomorphism $T_{\mathrm{k_d}\eta}$ in the canonical basis of $\mathbb{R}^N$.

Following Chapter III, we can also consider a continuous representation on the state space $\Omega_{\mathrm{c}} = [0, 1)$ equipped with the Lebesgue measure μ_{c}. Let $I_0 = [0, \mu_0)$, $I_1 = [\mu_0, \mu_0 + \mu_1)$, ..., $I_{N-1} = [1 - \mu_{N-1}, 1)$, so that the intervals $(I_n,\ 0 \leq n < N)$ form a partition of Ω. Now define the kernel:

$$\mathrm{k}_{\mathrm{c}} = \sum_{1 \leq i,j < N} \mathrm{k}_{\mathrm{d}}(i, j) \mathbb{1}_{I_i \times I_j}. \tag{V.31}$$

Denote by $\mathfrak{R}_e^{\mathrm{d}}$ and $\mathfrak{R}_e^{\mathrm{c}}$ the effective reproduction number in the discrete and continuous representation models. In the same manner, the uniform cost in each model is denoted by C^{d} and C^{c}. According to Chapter III, these functions are linked through the following relation:

$$\mathfrak{R}_e^{\mathrm{d}}(\eta^{\mathrm{d}}) = \mathfrak{R}_e^{\mathrm{c}}(\eta^{\mathrm{c}}), \quad \text{and} \quad C^{\mathrm{d}}(\eta^{\mathrm{d}}) = C^{\mathrm{c}}(\eta^{\mathrm{c}}),$$

for all $\eta^{\mathrm{d}} : \Omega_{\mathrm{d}} \to [0, 1]$ and $\eta^{\mathrm{c}} : \Omega_{\mathrm{c}} \to [0, 1]$ such that:

$$\eta^{\mathrm{d}}(i) = \frac{1}{\mu_i} \int_{I_i} \eta^{\mathrm{c}} \, \mathrm{d}\mu_{\mathrm{c}} \quad \text{for all} \quad i \in \Omega_{\mathrm{d}}.$$

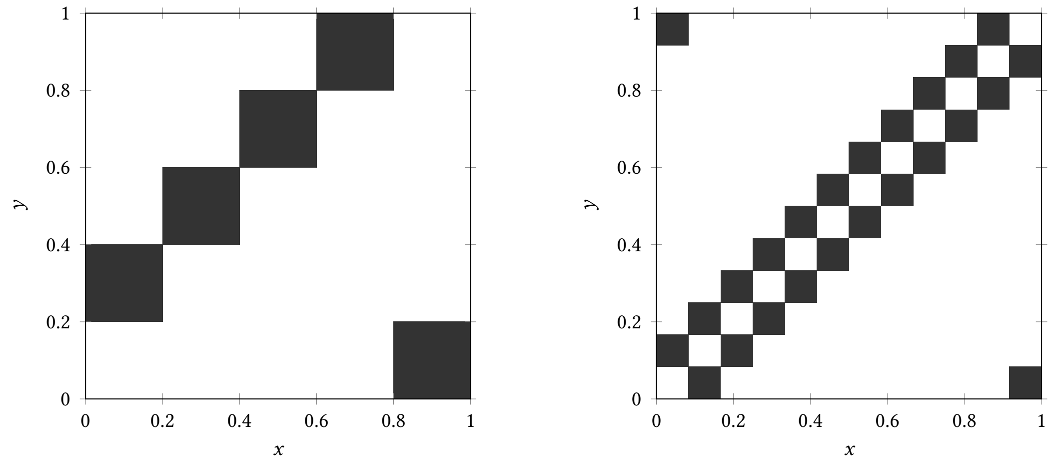
Let us recall that the Pareto and anti-Pareto frontiers for the two models are the same.

In Figure V.4, we have plotted the kernels of the continuous models associated to the asymmetric and symmetric circles models from Sections V.2.3 and V.2.4.

V.4 Assortative versus disassortative mixing

V.4.1 Motivation

We consider a population divided into an at most countable number of groups. Individuals within the same group interact with intensity a and individuals in different groups interact with intensity b. Hence, the model is entirely determined by the coefficients a and b and the size of the different groups. This simple model allows to study the effect of assortativity, that is, the tendancy for individuals to connect with individuals belonging to their own subgroup. The mixing pattern is called *assortative* (higher interaction in the same subgroup) if $a > b$, and *disassortative* (lower interaction in the same subgroup) when $b < a$. Our result illustrate how assortativity affects optimal vaccination strategies. This has been already studied by Galeotti and Rogers [65] in a



(A) The asymmetric circle.

(B) The symmetric circle.

Figure V.4: Kernels $\mathrm{k_c}$ (equal to 0 in the white zone and to 1 in the black zone) on $\Omega_\mathrm{c} = [0, 1)$ and μ_c the Lebesgue measure of the continuous model associated to discrete metapopulation models.

population composed of two groups. They observed that the optimal immunization strategies can differ dramatically in the case of assortative versus disassortative mixing.

When the number of subgroups is finite and a is equal to 0, the next-generation matrix of this model corresponds, up to a multiplicative constant, to the adjacency matrix of a complete multipartite graph. Recall that an m-partite graph is a graph that can be colored with m different colors, so that no two endpoints of an edge have the same color. When $m = 2$ these are the so-called bipartite graphs. A complete multipartite graph is a m-partite graph (for some $m \in \mathbb{N}^*$) in which there is an edge between every pair of vertices from different colors.

The complete multipartite graphs have interesting spectral properties. Indeed, Smith [145] showed that a graph with at least one edge has its spectral radius as its only positive eigenvalue if and only if its non-isolated vertices induce a complete multipartite graph. In [54], Esser and Harary proved that two complete m-partite graphs with the same number of nodes are isomorphic if and only if they have the same spectral radius. More precisely, they obtained a comparison of the spectral radii of two complete m-partite graphs by comparing the sizes of the sets in their partitions through majorization; see [54, Lemma 3].

The goal of this section is to generalize and complete these results and give a full picture of the Pareto and anti-Pareto frontiers for the assortative and the disassortative models.

V.4.2 Spectrum and convexity

We will use an integer intervals notation to represent the considered kernels. For $i, j \in \mathbb{N} \cup \{+\infty\}$, we set $[\![i, j]\!]$ (resp. $[\![i, j[\![$) for $[i, j] \cap (\mathbb{N} \cup \{+\infty\})$ (resp. $[i, j) \cap \mathbb{N}$). Let $N \in [\![2, +\infty]\!]$ and $\Omega = [\![0, N[\![$. The set Ω is endowed with the discrete σ-algebra $\mathscr{F} = \mathcal{P}(\Omega)$ and a probability measure μ. To simplify the notations, we write μ_i for $\mu(\{i\})$ and $f_i = f(i)$ for a function f defined on Ω. Without loss of generality, we can suppose that $\mu_i \geq \mu_j > 0$ for all $0 \leq i \leq j < N$. We consider the kernel k defined for $i, j \in \Omega$ by:

$$\mathrm{k}(i, j) = \begin{cases} a & \text{if} \quad i = j, \\ b & \text{otherwise}, \end{cases} \tag{V.32}$$

where a and b are two non-negative real numbers.

If $b = 0$, then the kernel is reducible, see Section IV.7, and the effective reproduction number is given by the following formula: $\mathfrak{R}_e(\eta) = a \max_{i \in \Omega} \eta_i \, \mu_i$, for all $\eta = (\eta_i, i \in \Omega) \in \Delta$. This is sufficient to treat this case and we have $c^\star = 1 - \mu_0$.

From now on, we assume that $b > 0$. For $n \in \mathbb{N}$, we consider the symmetric matrix M_n of size $(n+1) \times (n+1)$ given by:

$$M_n(i,j) = \begin{cases} a & \text{if} \quad i = j, \\ b & \text{otherwise.} \end{cases}$$

The matrix M_n is the sum of b times the all-ones matrix and $a - b$ times the identity matrix. Hence, M_n has two distinct eigenvalues: $nb + a$ with multiplicity 1 and $a - b$ with multiplicity n.

The next two results describe the spectrum of T_k in both the assortative and disassortative case. Notice the spectrum of T_k is real as k is symmetric.

Proposition V.4.1 (Concavity of $\mathfrak{R}_e$ in the disassortative case)**.** *Let* k *be given by* (V.32)*, with* $b \geq a \geq 0$ *and* $b > 0$*. Then,* $\mathfrak{R}_0$ *is the only positive eigenvalue of* T_k*, and it has multiplicity one. Furthermore, the function* $\mathfrak{R}_e$ *is concave.*

Proof. We give a direct proof when N is finite, and uses an approximation procedure for $N = \infty$. We first assume that N is finite. For $n < N$, let $v_n = \mathbb{1}_{[\![0,n]\!]}$ and set $T_n = T_{v_n \mathrm{k} v_n}$. The non-null eigenvalues of T_n (with their multiplicity) are the eigenvalues of the matrix $M_n \cdot \mathrm{Diag}_n(\mu)$, where $\mathrm{Diag}_n(\mu)$ is the diagonal $(n+1) \times (n+1)$-matrix with $(\mu_0, \ldots, \mu_n)$ on the diagonal. Thanks to [87, Theorem 1.3.22], these are also the eigenvalues of the matrix $Q_n = \mathrm{Diag}_n(\mu)^{1/2} \cdot M_n \cdot \mathrm{Diag}_n(\mu)^{1/2}$. By Sylvester's law of inertia [87, Theorem 4.5.8], the matrix Q_n has the same signature as the symmetric matrix M_n. In particular, since we have supposed $a - b \leq 0$, M_n has only one positive eigenvalue. Thus, Q_n has only one positive eigenvalue, and thanks to Perron-Frobenius theory, it is its spectral radius. This concludes the proof when N is finite by choosing $n = N - 1$.

If $N = \infty$, we consider the limit $n \to N$. Since

$$\lim_{n \to \infty} \| \mathrm{k} - v_n \mathrm{k} v_n \|_{2,2} = 0,$$

the spectrum of T_n converges to the spectrum of T_k, with respect to the Hausdorff distance, and the multiplicity on the non-zero eigenvalues also converge, see Corollary IV.3.2. This shows that $\rho(T_\mathrm{k})$ is the only positive eigenvalue of T_k, and it has multiplicity one. Since k is symmetric, we deduce the concavity of the function $\mathfrak{R}_e$ from Theorem IV.5.5. □

Now, let us examine the assortative case. Adapting the proof of Proposition V.4.1, we could directly conclude that the spectrum of T_k is a subset of $\mathbb{R}_+$ when $a \geq b$. Below, we give a more direct proof of this result relying on Chapter IV.

Proposition V.4.2 (Convexity of $\mathfrak{R}_e$ in the assortative case)**.** *Let* k *be given by* (V.32)*, with* $a \geq b > 0$*. Then the operator* T_k *is positive semi-definite and the function* $\mathfrak{R}_e$ *is convex.*

Proof. Since for any $g \in L^2$, we have:

$$\int_{\Omega \times \Omega} g(x) \mathrm{k}(x,y) g(y)\, \mu(\mathrm{d}x)\mu(\mathrm{d}y) = a \sum_{i \in \Omega} g_i^2 \mu_i^2 + b \sum_{i \neq j} g_i g_j\, \mu_i \mu_j \geq b \| g \|_2^2 .$$

This implies that T_k is positive semi-definite. Thus, as k is symmetric, the fonction $\mathfrak{R}_e$ is convex, thanks to Theorem IV.5.5. □

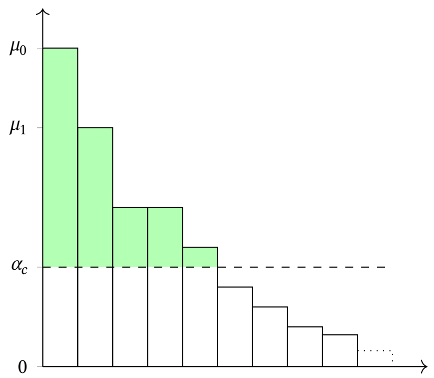
V.4.3 Explicit description of the Pareto and anti-Pareto frontiers

For $c \in [0,1]$, we define a "*horizontal vaccination*" $\eta^\mathrm{h}(c) \in \Delta$ with cost c in the following manner. It will be convenient to define first the quantity, rather than the proportion, of vaccinated people of same type. For all $\alpha \in [0, \mu_0]$, let $\xi^\mathrm{h}(\alpha)$ be defined by

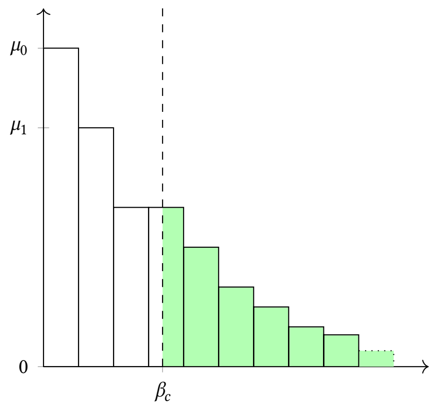
$$\xi_i^\mathrm{h}(\alpha) = \min(\alpha, \mu_i), \quad i \in \Omega. \tag{V.33}$$

For all $i \in \Omega$, $\xi_i^\mathrm{h}(\alpha)$ is a non-decreasing and continuous function of α. The map $\alpha \mapsto \sum_i \xi_i^\mathrm{h}(\alpha)$ is continuous and increasing from $[0, \mu_0]$ to $[0,1]$, so for any $c \in [0,1]$, there exists a unique α_c such that $\sum_i \xi_i^\mathrm{h}(\alpha_c) = 1 - c$. We then define the horizontal vaccination profile $\eta^\mathrm{h}(c)$ by:

$$\eta_i^\mathrm{h}(c) = \xi_i^\mathrm{h}(\alpha_c)/\mu_i, \quad i \in \Omega. \tag{V.34}$$



(A) Representation of the greedy path $\mathscr{P}_{\mathrm{h}}$.



(B) Representation of the greedy path $\mathscr{P}_{\mathrm{v}}$.

Figure V.5: Greedy parametrization of the (anti-)Pareto front. The bar plot represents the measure μ. The proportion of green in each bar correspond to the proportion of vaccinated individuals in each subpopulation.

In words, it consists in vaccinating in such a way that the quantity of the non-vaccinated individuals $\xi_i^{\mathrm{h}} = \eta_i \mu_i$ in each subpopulation is always less than the "horizontal" threshold α: see Figure V.5(A). Note that $\eta^{\mathrm{h}}(0) = \mathbb{1}$ (no vaccination), whereas $\eta^{\mathrm{h}}(1) = \mathbb{0}$ (full vaccination), and that the path $c \mapsto \eta^{\mathrm{h}}(c)$is greedy. We denote its range by $\mathscr{P}_{\mathrm{h}}$.

For $c \in [0, 1]$, we define similarly a "*vertical vaccination*" $\eta^{\mathrm{v}}(c) \in \Delta$ with cost c. First let us define for $\beta \in [0, N]$:

$$\xi_i^{\mathrm{v}}(\beta) = \mu_i \cdot \min(1, (\beta - i)_+), \quad i \in \Omega. \tag{V.35}$$

The map $\beta \mapsto \sum_i \xi_i^{\mathrm{v}}(\beta)$ is increasing and continuous from $[0, N]$ to $[0, 1]$, so for any $c \in [0, 1]$ there exists a unique β_c such that $\sum_i \xi_i^{\mathrm{v}}(\beta_c) = 1 - c$. We then define the vertical vaccine profile $\eta^{\mathrm{v}}(c)$ by:

$$\eta_i^{\mathrm{v}}(c) = \xi_i^{\mathrm{v}}(\beta_c)/\mu_i, \quad i \in \Omega. \tag{V.36}$$

In words, if $\lfloor \beta \rfloor = \ell$, this consists in vaccinating all subpopulations j for $j > \ell$, and a fraction $(\ell + 1 - \beta)$ of the subpopulation ℓ, see Figure V.5(B) for a graphical representation.

For all $i \in \Omega$, $\eta_i^{\mathrm{v}}(c)$ is a non-increasing and continuous function of c. In this case also, $\eta^{\mathrm{v}}(0) = \mathbb{1}$ (no vaccination), while $\eta^{\mathrm{v}}(1) = \mathbb{0}$ (full vaccination). The path $c \mapsto \eta^{\mathrm{v}}(\beta(c))$ is also greedy. We denote its range by $\mathscr{P}_{\mathrm{v}}$.

These two paths give a greedy parametrization of the Pareto and anti-Pareto frontiers for the assortative and disassortative models: more explicitly, we have the following result, whose proof can be found in Section V.4.4.

Theorem V.4.3 (Assortative vs disassortative)**.** *Let* k *be given by* (V.32), *with* $b > 0$ *and* $a \geq 0$.

(i) *(**Assortative model.**) If $a \geq b > 0$, then $\mathscr{P}_{\mathrm{v}}$ and $\mathscr{P}_{\mathrm{h}}$ are greedy parametrizations of the anti-Pareto and Pareto frontiers respectively.*

(ii) *(**Disassortative model.**) If $b \geq a > 0$, then $\mathscr{P}_{\mathrm{v}}$ and $\mathscr{P}_{\mathrm{h}}$ are greedy parametrizations of the Pareto and anti-Pareto frontiers respectively.*

(iii) *(**Complete multipartite model.**) If $a = 0$ and $b > 0$, then $\mathscr{P}_{\mathrm{h}}$ is a greedy parametrization of the anti-Pareto frontier and the subset of strategies $\eta \in \mathscr{P}_{\mathrm{v}}$ such that $C(\eta) \leq 1 - \mu_0$ is a greedy parametrization of the Pareto frontier. In particular, we have $c_\star = 1 - \mu_0$ and $c^\star = 0$.*

Notice that $c^\star = 0$ and $c_\star = 1$ in cases (i) and (ii) as k is positive.

Remark V.4.4 (Highest Degree vaccination). The effective degree function of a symmetric kernel k at $\eta \in \Delta$ is the function on Ω is defined by:

$$\deg_\eta(x) = \int_\Omega \mathrm{k}(x, y)\eta(y)\, \mu(\mathrm{d}y). \tag{V.37}$$

When $\eta = \mathbb{1}$, it is simply called the degree of k and is denoted by deg. In our model, the effective degree of the subgroup i is given by

$$\deg_\eta(i) = a\eta_i\mu_i + b\sum_{\ell\neq i}\eta_\ell\mu_\ell, \tag{V.38}$$

and thus the degree of the subgroup i is given by $\deg(i) = (a-b)\mu_i + b$. As $\mu_i \geq \mu_j$ for $0 \leq i < j < N$, we deduce that the degree function in monotone: non-increasing in the assortative model and non-decreasing in the disassortative model. So the group with the highest degree corresponds to the largest group in the assortative model and the smallest group (if it exists) in the disassortative model.

Consider the assortative model where all the groups have different size, *i.e.*, $\mu_0 > \mu_1 > \ldots$ Following the parametrization $c \mapsto \eta^{\mathrm{h}}(c)$, starting from $c = 0$, will first decrease the effective size of the group 0 (the group with the highest degree) until it reaches the effective degree of group 1 (with the second highest degree). Once these two groups share the same effective degree which corresponds to reaching $\mu_0\eta_0^{\mathrm{h}} = \mu_1$, they are vaccinated uniformly (that is, ensuring that they keep the same effective degree: using (V.38) this corresponds to $\mu_0\eta_0^{\mathrm{h}} = \mu_1\eta_1^{\mathrm{h}}$) until their effective degree is equal to the third highest degree, and so on and so forth.

In the disassortative model, the function $\deg_\eta$ remains (strictly) increasing when the vaccination strategies in $\mathscr{P}_{\mathrm{v}}$ are applied. In particular, if $\mu_0 > \mu_1 > \ldots$, then the optimal strategies prioritize the groups with the higher effective degree until they are completely immunized. If multiple groups share the same degree, it is optimal to give all available doses to one group.

In conclusion, in both models, the optimal vaccination consists in vaccinating the groups with the highest effective degree in priority if this group is unique. But if multiple groups share the same degree (*i.e.*, have the same size), the optimal strategies *differ* between the assortative and the disassortative case. In the assortative case, groups with the same size must be vaccinated uniformly while in the disassortive case, all the vaccine doses shall be given to one group until it is completely vaccinated.

Example V.4.5 (Group sizes following a dyadic distribution). Let $N = \infty$, $\Omega = \mathbb{N}$ and $\mu_i = 2^{-(i+1)}$ for all $i \in \Omega$. Following Section 7.4.1 in Chapter III, we will couple this discrete model with a continuum model for a better visualization on the figures. Let $\Omega_c = [0, 1)$ be equipped with the Borel σ-field $\mathscr{F}_c$ and the Lebesgue measure μ_c. The set Ω_c is partitionned into a countable number of intervals $I_i = [1 - 2^{-i}, 1 - 2^{-i-1})$, for $i \geq 0$, so that $\mu_c(I_i) = \mu_i$. The kernel of the continuous model corresponding to k in (V.32) is given by:

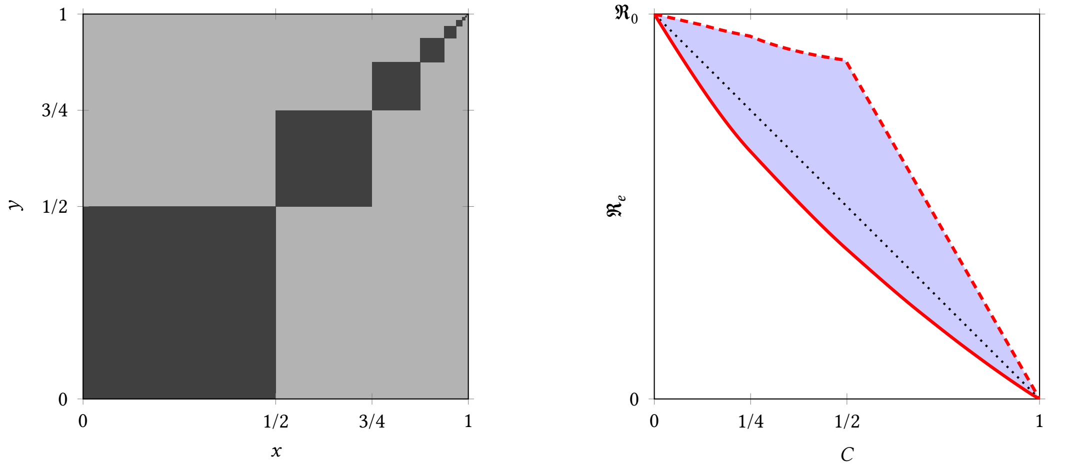
$$\mathrm{k}_c = (a-b)\sum_{i\in\mathbb{N}}\mathbb{1}_{I_i\times I_i} + b\mathbb{1}. \tag{V.39}$$

The kernel k_c is plotted in Figures V.6(A), V.7(A) and V.8(A) for different values of a and b corresponding respectively to the assortative, the disassortative and the complete multipartite case corresponding to points (i), (ii) and (iii) of Theorem V.4.3 respectively. Their respective Pareto and anti-Pareto frontiers are plotted in Figures V.6(B), V.7(B) and V.8(B), using a finite-dimensional approximation of the kernel k and the power iteration method. In Figure V.8(B), the value of $c_\star$ is equal to $1 - \mu_0 = 1/2$. With this continuous representation of the population, the set $\mathscr{P}_{\mathrm{v}}$ corresponds to the strategies of the form $\mathbb{1}_{[0,t)}$ for $t \in [0, 1]$.

Notice that the Pareto frontier in the assortative case is convex. This is consistent with Proposition III.6.6 since the cost function is affine and $\mathfrak{R}_e$ is convex when $a \geq b$; see Proposition V.4.2. In the same manner, the anti-Pareto frontier in the disassortative and the multipartite cases is concave. Once again, this is consistent with Proposition III.6.6 since the cost function is affine and $\mathfrak{R}_e$ is concave when $a \leq b$; see Proposition V.4.1.

V.4.4 Proof of Theorem V.4.3

After recalling known facts of majorization theory, we first consider the finite dimension models, and then the general case by an approximation argument.



(A) Grayplot of the kernel k_c from (V.39) on $\Omega_c = [0, 1)$ with $\mathrm{k}_c = a = 5$ on the dark gray zone, and $\mathrm{k}_c = b = 2$ on the light gray zone.

(B) Red line: the Pareto frontier $\mathcal{F}$; dashed line: the anti-Pareto frontier $\mathcal{F}^{\mathrm{Anti}}$; dotted line: path of the uniform strategies; blue region: all possible outcomes $\mathbf{F}$.

Figure V.6: An example of assortative model.

Majorization

In this section, we recall briefly some definitions and results from majorization theory, and refer to [9, 115] for an extensive treatment of this topic.

Let $n \geq 2$ and $\xi, \chi \in \mathbb{R}_+^n$. We denote by $\xi^\downarrow$ and $\chi^\downarrow$ their respective order statistics, that is the vectors in $\mathbb{R}_+^n$ with the same components, but sorted in descending order. We say that ξ is *majorized* by χ, and write $\xi \prec \chi$, if:

$$\sum_{j=1}^{i} \xi_j^\downarrow \leq \sum_{j=1}^{i} \chi_j^\downarrow \quad \text{for all } i \in \{1, \ldots, n\}, \text{ and } \quad \sum_{j=1}^{n} \xi_j = \sum_{j=1}^{n} \chi_j. \tag{V.40}$$

Among the various characterizations of majorization, we will use the following by Hardy, Littlewood and Pòlya; see [115, Proposition I.4.B.3]:

$$\xi \prec \chi \iff \sum_i (\xi_i - t)_+ \leq \sum_i (\chi_i - t)_+ \quad \text{for all} \quad t \in \mathbb{R}_+, \tag{V.41}$$

where $u_+ = \max(u, 0)$, for all $u \in \mathbb{R}$. A real-valued function Θ defined on $\mathbb{R}_+^n$ is called *Schur-convex* if it is non-decreasing with respect to $\prec$, that is, $\xi \prec \chi$ implies $\Theta(\xi) \leq \Theta(\chi)$. A function Θ is called *Schur-concave* if $(-\Theta)$ is Schur-convex.

Shur convexity and concavity of the spectral radius in finite dimension

We define the function Θ_n on $\mathbb{R}_+^{n+1}$ by:

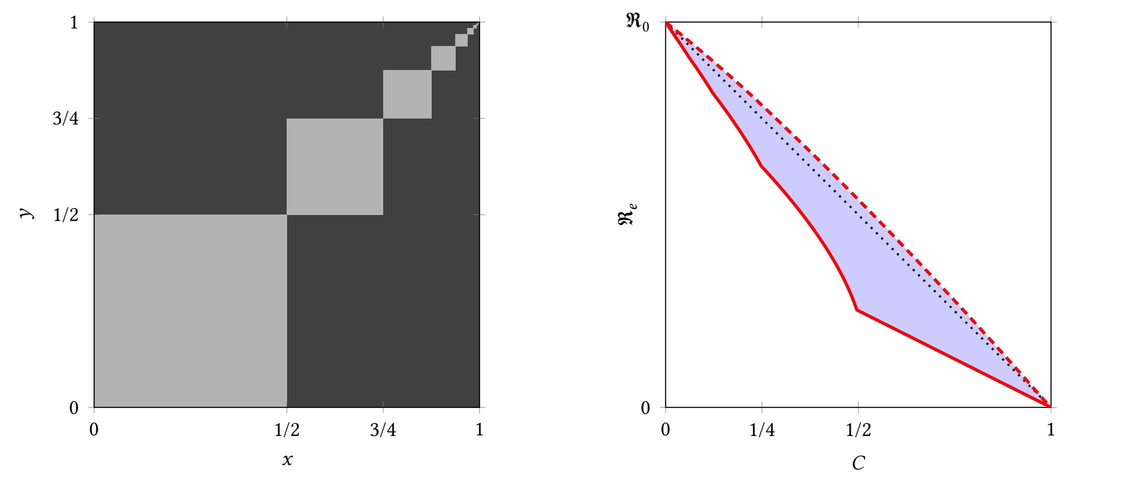
$$\Theta_n(\xi) = \rho(M_n \cdot \mathrm{Diag}(\xi)),$$

where $\mathrm{Diag}(\xi)$ is the diagonal $(n+1) \times (n+1)$-matrix with ξ on the diagonal. By construction, for $\eta = (\eta_0, \ldots, \eta_n, 0, \ldots)$, we have:

$$\mathfrak{R}_e(\eta) = \Theta_n(\eta_0 \mu_0, \ldots, \eta_n \mu_n). \tag{V.42}$$

The key property below will allow us to identify the optimizers.

Lemma V.4.6 (Schur-concavity and Schur-convexity)**.** *Let $b > 0$ and $a \geq 0$. The function Θ_n is Schur-convex if $a \geq b$, and Schur-concave if $a \leq b$.*



(A) Grayplot of the kernel k_c from (V.39) on $\Omega_c = [0, 1)$ with $k_c = a = 2$ on the light gray zone, and $k_c = b = 5$ on the dark gray zone.

(B) Red line: the Pareto frontier $\mathcal{F}$; dashed line: the anti-Pareto frontier $\mathcal{F}^{\text{Anti}}$; dotted line: path of the uniform strategies; blue region: all possible outcomes $\mathbf{F}$.

Figure V.7: An exmple of disassortative model.

Proof. Let us consider the disassortative case where $a \leq b$. By a classical result of majorization theory [115, Proposition I.3.C.2.], it is enough to show that Θ_n is symmetric and concave.

To prove that Θ_n is symmetric, consider σ be a permutation of $\{0, 1, \dots, n\}$ and P_σ the associated permutation matrix of size $(n+1) \times (n+1)$. Since $P_\sigma M_n P_\sigma^{-1} = M_n$, we deduce that $\Theta_n(\xi_\sigma) = \Theta_n(\xi)$, where ξ_σ is the σ-permutation of $\xi \in \mathbb{R}_+^{n+1}$. Thus Θ_n is symmetric.

We now prove that Θ_n is concave on $\mathbb{R}_+^{n+1}$. Since $\mathfrak{R}_e$ is concave thanks to Proposition V.4.1, we deduce from (V.42), that the function Θ_n is concave on $[0, \mu_0] \times \ldots \times [0, \mu_n]$. Since Θ_n is homogeneous, it is actually concave on the whole domain $\mathbb{R}_+^n$. This concludes the proof when $a \leq b$.

The proof is the same for the assortative case $a \geq b$, replacing the reference to Proposition V.4.1 by Proposition V.4.2. □

Extreme vaccinations for fixed cost

Let us show that the horizontal and vertical vaccinations give extreme points for the preorder $\prec$ on finite sets, when the quantity of vaccine is fixed. Recall that ξ^{h} and ξ^{v} are defined in (V.33) and (V.35) respectively.

Proposition V.4.7 (Extreme vaccinations). *Let $n \in [\![0, N[\![$, $\beta \in [0, n)$ and $\alpha \in [0, \mu_0]$. Let $\xi^{\mathrm{v},n} = (\xi_0^{\mathrm{v}}(\beta), \dots, \xi_n^{\mathrm{v}}(\beta))$, and $\xi^{\mathrm{h},n} = (\xi_0^{\mathrm{h}}(\alpha), \dots, \xi_n^{\mathrm{h}}(\alpha))$. For any $\xi = (\xi_0, \dots, \xi_n) \in [0, \mu_0] \times \cdots \times [0, \mu_n]$, we have:*

$$\left(\sum_{i=0}^{n} \xi_i = \sum_{i=0}^{n} \xi_i^{\mathrm{v},n}\right) \implies \xi \prec \xi^{\mathrm{v},n}, \quad \textit{and} \quad \left(\sum_{i=0}^{n} \xi_i = \sum_{i=0}^{n} \xi_i^{\mathrm{h},n}\right) \implies \xi^{\mathrm{h},n} \prec \xi.$$

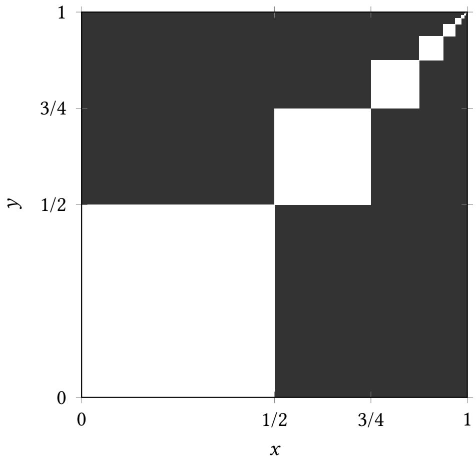
Proof. Let $\xi \in [0, \mu_0] \times \cdots \times [0, \mu_n]$ be such that $\sum_{i=0}^n \xi_i = \sum_{i=0}^n \xi_i^{\mathrm{v},n}$. The reordered vector $\xi^\downarrow$ clearly satisfies the same conditions, so without loss of generality we may assume that ξ is sorted in descending order. Using Equation (V.35), we get:

$$\sum_{i=0}^{\ell} \xi_i \leq \sum_{i=0}^{\ell} \mu_i = \sum_{i=0}^{\ell} \xi_i^{\mathrm{v},n}, \qquad 0 \leq \ell < \lfloor \beta \rfloor.$$

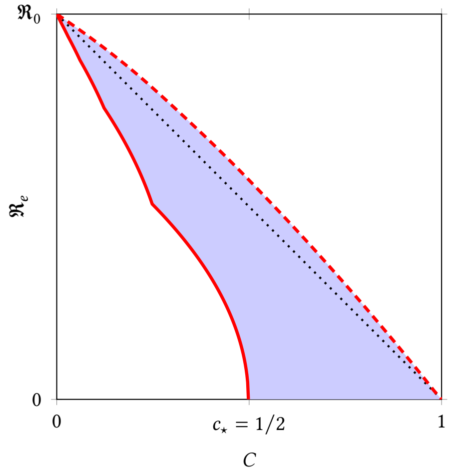
We also have:

$$\sum_{i=0}^{\ell} \xi_i \leq \sum_{i=0}^{n} \xi_i = \sum_{i=0}^{n} \xi_i^{\mathrm{v},n} = \sum_{i=0}^{\ell} \xi_i^{\mathrm{v},n}, \qquad \ell \geq \lfloor \beta \rfloor.$$

Therefore, $\xi \prec \xi^{\mathrm{v},n}$, by the definition of $\prec$.



(A) Grayplot of the kernel k_c from (V.39) on $\Omega_c = [0, 1)$ with $\mathrm{k}_c = a = 0$ (on the white zone) and $\mathrm{k}_c = b = 6$ (on the black zone).



(B) Red line: the Pareto frontier $\mathcal{F}$; dashed line: the anti-Pareto frontier $\mathcal{F}^{\mathrm{Anti}}$; dotted line: path of the uniform strategies; blue region: all possible outcomes $\mathbf{F}$.

Figure V.8: An example of complete multipartite model.

Similarly, let $\xi \in [0, \mu_0] \times \cdots \times [0, \mu_n]$ be such that $\sum_{i=0}^n \xi_i = \sum_{i=0}^n \xi_i^{\mathrm{h},n}$. If $t \geq \alpha$ then

$$\sum_i (\xi_i^{\mathrm{h},n} - t)_+ = 0 \leq \sum_i (\xi_i - t)_+,$$

while if $t \in [0, \alpha)$, using the fact that $\sum_{i=0}^n \xi_i = \sum_{i=0}^n \xi_i^{\mathrm{h},n}$, the expression $\xi_i^{\mathrm{h}} = \min(\alpha, \mu_i)$, and the inequalities $\xi_i \leq \mu_i$, we get:

$$\begin{aligned}
\sum_{i=0}^n (\xi_i^{\mathrm{h},n} - t)_+ &= \sum_{i=0}^n (\xi_i^{\mathrm{h},n} - t) + \sum_{i=0}^n (t - \xi_i^{\mathrm{h},n})_+ \\
&= \sum_{i=0}^n (\xi_i - t) + \sum_{i=0}^n (t - \mu_i)_+ \\
&\leq \sum_{i=0}^n (\xi_i - t) + \sum_{i=0}^n (t - \xi_i)_+ \\
&= \sum_{i=0}^n (\xi_i - t)_+.
\end{aligned}$$

This gives $\xi^{\mathrm{h},n} \prec \xi$, by the characterization (V.41). □

"Vertical" Pareto optima in the disassortative case

We consider here the disassortative model $b \geq a \geq 0$ and $b > 0$. Let $c \in (0, 1)$ and set $D(c) = \{\eta \in \Delta \,:\, C(\eta) = c\}$. We will solve the constrained optimization Problem (V.7) that corresponds to:

$$\begin{cases} \min & \mathfrak{R}_e(\eta), \\ \text{such that} & \eta \in D(c). \end{cases} \tag{V.43}$$

Recall the definitions of β_c and $\eta^{\mathrm{v}}(c)$ given page 153. Let $\eta \in \Delta$ be any strategy with cost c. Let n be large enough so that $\sum_{j>n} \mu_j < 1 - c$ so that $\sum_{j \leq n} \eta_j \mu_j > 0$, and assume that $n > \beta$. Let $\eta^{(n)} \in \Delta$ be defined by:

$$\eta_i^{(n)} = \frac{\sum_{j \leq n} \eta_j^{\mathrm{v}}(c) \mu_j}{\sum_{j \leq n} \eta_j \mu_j} \mathbb{1}_{\{i \leq n\}} \eta_i.$$

Note that since $C(\eta^{\mathrm{v}}(c)) = c = C(\eta)$, we have $\lim_{n\to N} \eta^{(n)} = \eta$ (pointwise and in L^2). Let $\xi^n = (\eta_0^{(n)}\mu_0, \dots, \eta_n^{(n)}\mu_n)$ and $\xi^{\mathrm{v},n}$ be defined as in Proposition V.4.7 with $\beta = \beta_c$. By construction, we have $\sum_{i=0}^n \xi_i^n = \sum_{i=0}^n \xi_i^{\mathrm{v},n}$, so by Proposition V.4.7, we get $\xi^n \prec \xi^{\mathrm{v},n}$. This implies that:

$$\mathfrak{R}_e(\eta^{(n)}) = \Theta_n(\xi^n) \geq \Theta_n(\xi^{\mathrm{v},n}) = \mathfrak{R}_e(\eta^{\mathrm{v}}(c)),$$

where the inequality follows from the Schur concavity of Θ_n in the disassortative case (see Lemma V.4.6) and where the last equality is holds as $n \geq \lceil \beta_c \rceil$. Since $\mathfrak{R}_e$ is continuous and $\eta^{(n)}$ converges to η, we get $\mathfrak{R}_e(\eta) \geq \mathfrak{R}_e(\eta^{\mathrm{v}})$. This implies that η^{v} is a solution of Problem (V.43).

If $a > 0$, then k is positive everywhere, and we deduce from Lemma V.3.1 that $c_\star = 1$. If $a = 0$, it is easy to prove that $\{0\}$ is a maximal independant set of k; this gives that $c_\star = 1 - \mu_0$, thanks to Section 6.4 in Chapter IV. Since for all $c \in [c_\star, 1]$ there exists $\eta \in \mathscr{P}_{\mathrm{v}}$ such that $C(\eta) = c$, we also get that $\mathscr{P}_{\mathrm{v}} \cap \{\eta \in \Delta \,:\, C(\eta) \leq c_\star\}$ is a parametrization of the Pareto frontier. This gives the parametrization of the Pareto frontier using $\mathscr{P}_{\mathrm{v}}$ from Theorem V.4.3 (ii) and (iii).

"Horizontal" anti-Pareto optima in the disassortative case

We still consider $b \geq a \geq 0$ and $b > 0$. Let $c \in (0, 1)$. We now turn to the anti-Pareto frontier by studying the constrained maximization Problem (V.8) that corresponds to:

$$\begin{cases} \max & \mathfrak{R}_e(\eta), \\ \text{such that} & \eta \in D(c). \end{cases} \tag{V.44}$$

Recall the definitions of α_c and $\eta^{\mathrm{h}}(c)$ given page 152. Let η be such that $C(\eta) = c$. Let n be large enough so that $\sum_{j>n} \mu_j < 1 - c$ so that $\sum_{j\leq n} \eta_j\mu_j > 0$. Define $\eta^{(n)} \in \Delta$ by:

$$\eta_i^{(n)} = \frac{\sum_{j\leq n} \eta_j^{\mathrm{h}}(c)\mu_j}{\sum_{j\leq n} \eta_j\mu_j} \mathbb{1}_{\{i\leq n\}}\, \eta_i.$$

Let $\xi^n = (\eta_0^{(n)}\mu_0, \eta_1^{(n)}\mu_1, \dots, \eta_n^{(n)}\mu_n)$ and let $\xi^{\mathrm{h},n}$ be defined as in Proposition V.4.7 with $\alpha = \alpha_c$. By construction, we have $\sum_{i=0}^n \xi_i^n = \sum_{i=0}^n \xi_i^{\mathrm{h},n}$, so by Proposition V.4.7, we obtain $\xi^{\mathrm{h},n} \prec \xi^n$. This implies that:

$$\mathfrak{R}_e(\eta^{(n)}) = \Theta_n(\xi^n) \leq \Theta_n(\xi^{\mathrm{h},n}) = \mathfrak{R}_e(\eta^{\mathrm{h}}(c)\, \mathbb{1}_{[\![0,n]\!]}),$$

where the inequality follows from the Schur concavity of Θ_n.

Now, as n goes to infinity $\eta^{(n)}$ converges pointwise and in L^2 to η, and $\eta^{\mathrm{h}}(c)\,\mathbb{1}_{[\![0,n]\!]}$ converges pointwise and in L^2 to $\eta^{\mathrm{h}}(c)$, so by continuity of $\mathfrak{R}_e$ we get $\mathfrak{R}_e(\eta) \leq \mathfrak{R}_e(\eta^{\mathrm{h}}(c))$, and $\eta^{\mathrm{h}}(c)$ is solution of the Problem (V.44) and is thus anti-Pareto optimal for $c \in (0, 1)$ as $c^\star = 0$. Since $c^\star = 0$, we also deduce from Propsotion III.5.8 (iii) that $\mathbb{0}$ and $\mathbb{1}$ are anti-Pareto optimal. Since for all $c \in [0, 1]$ there exists $\eta \in \mathscr{P}_{\mathrm{h}}$ such that $C(\eta) = c$, we deduce that $\mathscr{P}_{\mathrm{h}}$ is a parametrization of the anti-Pareto frontier.

The assortative case

The case $a \geq b > 0$, corresponding to point (i) in Proposition V.4.3, is handled similarly, replacing concavity by convexity, minima by maxima and vice versa.

V.5 Constant degree kernels and unifom vaccinations

V.5.1 Motivation

We have seen in the previous section an example of model where vaccinating individuals with the highest degree is the best strategy. A similar phenomenon is studied in [40], where under monotonicity arguments on the kernel, vaccinating individuals with the highest (resp. lowest) degree is Pareto (resp. anti-Pareto) optimal. However, in case multiple individuals share the same maximal degree, the optimal strategies completely differs between the assortative and the disassortative models: the Pareto optimal strategies for one model correspond to the anti-Pareto optimal strategies for the other and vice versa. Motivated by this curious symmetry, we investigate further into constant degree kernels in this section.

In graph theory, a regular graph is a graph where all vertices have the same number of in-neighbors, and the same number of out-neighbors. In other words all vertices have the same in-degree and the same out-degree. Limits of undirected regular graphs have been studied in details by Backhausz and Szegedy [12] and Kunszenti-Kovács, Lovász and Szegedy [100]. When the graphs are dense, their limit can be represented as a regular graphon, that is a symmetric kernel with a constant degree function. In this section, we do not require the kernel to be symmetric and the condition for a general kernel to be constant degree is given by Definition V.5.1.

In this section, we examine the situation where all the individuals have the same number of connections. In Section V.5.2, we first introduce the constant degree kernels and give the main result (see Proposition V.5.4) on the optimality of the uniform strategies with respect to the convex or concave property of $\mathfrak{R}_e$. Section V.5.3 is devoted to the proof of this main result. We study in more detail the optimal strategies in an example of constant degree symmetric kernel of rank two in Section V.6.

V.5.2 On the uniform strategies for constant degree kernels

We now give the definition of constant degree kernels and then study in detail some examples. For a kernel k on Ω, we set, for all $z \in \Omega$ and $A \in \mathscr{F}$:

$$\mathrm{k}(z, A) = \int_A \mathrm{k}(z, y)\, \mu(\mathrm{d}y) \quad \text{and} \quad \mathrm{k}(A, z) = \int_A \mathrm{k}(x, z)\, \mu(\mathrm{d}x).$$

For $z \in \Omega$, its in-degree is $\mathrm{k}(z, \Omega)$ and its out-degree is $\mathrm{k}(\Omega, z)$.

Definition V.5.1 (Constant degree kernel)**.** *A kernel* k *with a finite L^2 double-norm and a positive spectral radius $\mathfrak{R}_0 > 0$ is called* constant degree *if all the in-degrees and all the out-degrees have the same value, that is, the maps $x \mapsto \mathrm{k}(x, \Omega)$ and $y \mapsto \mathrm{k}(\Omega, y)$ defined on Ω are constant, and thus equal.*

Remark V.5.2. Let k be a constant degree kernel with spectral radius $\mathfrak{R}_0 > 0$. Notice the condition "all the in-degrees and out-degrees have the same value" is also equivalent to $\mathbb{1}$ being a left and right eigenfunction of T_k. We now check that the corresponding eigenvalue is $\mathfrak{R}_0$.

Let $h \in L^2_+(\Omega)\backslash\{\mathbb{0}\}$ be a left Perron-eigenfunction. Denote by λ the eigenvalue associated to $\mathbb{1}$. Then, we have:

$$\lambda \int_\Omega h(x)\, \mu(\mathrm{d}x) = \int_\Omega h(x)\mathrm{k}(x, y)\mu(\mathrm{d}x)\mu(\mathrm{d}y) = \mathfrak{R}_0 \int_\Omega h(y)\, \mu(\mathrm{d}y),$$

where the first equality follows from the regularity of k and from the fact that h is a left Perron-eigenfunction of T_k. Since h is non-negative and not equal to $\mathbb{0}$ almost everywhere, we get that $\lambda = \mathfrak{R}_0$ and $\mathbb{1}$ is a right Perron-eigenvector of T_k. With a similar proof, we show that $\mathbb{1}$ is a left Perron-eigenvector of T_k. In particular, if k is constant degree, then the reproduction number is expressed:

$$\mathfrak{R}_0 = \int_{\Omega\times\Omega} \mathrm{k}(x, y)\, \mu(\mathrm{d}x)\mu(\mathrm{d}y). \tag{V.45}$$

Example V.5.3. We now give examples of constant degree kernels.

(i) Let $G = (E, V)$ be a finite non-oriented simple graph, and μ the uniform probability measure on the vertices V. The degree of a vertex $x \in V$ is given by $\deg(x) = \sharp\{y \in V : (x, y) \in E\}$. The graph G is constant degree if all its vertices have the same degree, say $d \geq 1$. Then the kernel defined on the finite space $\Omega = V$ by the adjacency matrix is constant degree with $\mathfrak{R}_0 = d$. Notice it is also symmetric.

(ii) Let $G = (E, V)$ be a finite directed graph, and μ be the uniform probability measure on the vertices V. The in-degree of a vertex $x \in V$ is given by $\deg_{\text{in}}(x) = \sharp\{y \in V : (y, x) \in E\}$ and the out-degree is given by $\deg_{\text{out}}(x) = \sharp\{y \in V : (x, y) \in E\}$. The graph G is regular if all its vertices have the same in-degree and out-degree, say $d \geq 1$. Then the kernel defined on the finite space $\Omega = V$ by the adjacency matrix is regular with $\mathfrak{R}_0 = d$. Notice it might not be symmetric.

(iii) Let $\Omega = (\mathbb{R}/(2\pi\mathbb{Z}))^n$ be the n-dimensional torus endowed with its Borel σ-field $\mathscr{F}$ and the normalized Lebesgue measure μ. Let f be a measurable square-integrable non-negative function defined on Ω. We consider the geometric kernel defined by, for $x, y \in \Omega$:

$$\mathrm{k}_f(x, y) = f(x - y).$$

The operator T_{k_f} corresponds to the convolution by f, and its spectral radius is given by $\mathfrak{R}_0 = \int_\Omega f \mathrm{d}\mu$. Then the kernel k_f is constant degree as soon as f is not equal to 0 almost surely.

(iv) More generally, let $(\Omega, \cdot)$ be a compact topological group and let μ be its left Haar probability measure. Let f be non-negative square-integrable function on Ω. Then the kernel $\mathrm{k}_f(x, y) = f(y^{-1} \cdot x)$ is constant degree.

We summarize our main result in the next proposition, whose proof is given in Section V.5.3. We recall that a strategy is called uniform if it is constant over Ω.

Proposition V.5.4. *Let* k *be a constant degree kernel on* Ω.

*(i) If the map $\mathfrak{R}_e$ defined on Δ is convex, then all uniform strategies are Pareto optimal (*i.e. $\mathcal{S}^{\text{uni}} \subset \mathcal{P}$*). Consequently, $c_\star = 1$, the Pareto frontier is the segment joining* $(0, \mathfrak{R}_0)$ *to* $(1, 0)$, *and for all* $c \in [0, 1]$:

$$\mathfrak{R}_{e\star}(c) = (1 - c)\mathfrak{R}_0.$$

(ii) If k *is irreducible and the map $\mathfrak{R}_e$ defined on Δ is concave, then all uniform strategies are anti-Pareto optimal (*i.e. $\mathcal{S}^{\text{uni}} \subset \mathcal{P}^{\text{Anti}}$*). Consequently, $c^\star = 0$, the anti-Pareto frontier is the segment joining* $(0, \mathfrak{R}_0)$ *to* $(1, 0)$, *and for all* $c \in [0, 1]$:

$$\mathfrak{R}_e^\star(c) = (1 - c)\mathfrak{R}_0.$$

In Section 5.2 in Chapter IV, we give sufficient condition on the spectrum of T_k to be either concave or convex. Combining this result with Proposition V.5.4, we get the following corollary.

Corollary V.5.5. *Let* k *be a constant degree symmetric kernel.*

*(i) If the eigenvalues of T_k are non-negative, then the uniform vaccination strategies are Pareto optimal and $c_\star = 1$ (*i.e. $\mathcal{S}^{\text{uni}} \subset \mathcal{P}$*).*

(ii) If k *is irreducible and the eigenvalues of T_k different from $\mathfrak{R}_0$ are non-positive, then the uniform vaccination strategies are anti-Pareto optimal and $c^\star = 0$ (*i.e. $\mathcal{S}^{\text{uni}} \subset \mathcal{P}^{\text{Anti}}$*).*

Remark V.5.6 (Equivalent conditions). Let k be a constant degree symmetric kernel. The eigenvalues of the operator T_k are non-negative if and only if T_k is semi-definite positive, that is:

$$\int_{\Omega\times\Omega} \mathrm{k}(x, y) g(x) g(y) \mu(\mathrm{d}x)\mu(\mathrm{d}y) \geq 0 \quad \text{for all} \quad g \in L^2. \tag{V.46}$$

Similarly, the condition given in Corollary V.5.5 (ii) that implies the concavity of $\mathfrak{R}_e$ is equivalent to the semi-definite negativity of T_k on the orthogonal of $\mathbb{1}$:

$$\int_{\Omega\times\Omega} \mathrm{k}(x, y) g(x) g(y) \mu(\mathrm{d}x)\mu(\mathrm{d}y) \leq 0 \quad \text{for all} \quad g \in L^2 \quad \text{such that} \quad \int_\Omega g \,\mathrm{d}\mu = 0. \tag{V.47}$$

Remark V.5.7 (Comparison with a result from [130]). Poghotanyan, Feng, Glasser and Hill [130, Theorem 4.7] obtained a similar result in finite dimension using a result from Friedland [64]: if the next-generation non-negative matrix K of size $N \times N$ satisfies the following conditions

(i) $\sum_{j=0}^{N-1} K_{ij}$ does not depend on $i \in [\![0, N-1]\!]$ (which corresponds the parameters a_i in [130, Equation (2.4)] being all equal),

(ii) $\mu_i K_{ij} = \mu_j K_{ji}$ for all $i, j \in [\![0, N-1]\!]$ where μ_i denote the relative size of population i (which corresponds to [130, Equation (2.4)]),

(iii) K is not singular and its inverse is an M-matrix (*i.e.*, its non-diagonal coefficients are non-positive),

then the uniform strategies are Pareto optimal (*i.e.*, they minimize the reproduction number among all strategies with same cost). Actually, this can be seen as a direct consequence of Corollary V.5.5 (i). Indeed, the corresponding kernel $\mathrm{k_d}$ defined by (V.30) in the discrete probability space $\Omega = [\![0, N-1]\!]$ endowed with the discrete probability measure μ_d also defined by (V.30) has constant degree thanks to Point (i) and is symmetric thanks to Point (ii). Since K^{-1} is an M-matrix, its real eigenvalues are positive according to [15, Chapter 6 Theorem 2.3]. The eigenvalues of $T_{\mathrm{k_d}}$ and K are actually the same as K is the representation matrix of $T_{\mathrm{k_d}}$ in the canonic basis of $\mathbb{R}^N$. We conclude that the operator $T_{\mathrm{k_d}}$ is positive definite. Hence Corollary V.5.5 (i) can be applied to recover that the uniform strategies are Pareto optimal.

However, the converse is not true. As a counter-example, consider a population divided in $N = 3$ groups of same size (*i.e*, $\mu_0 = \mu_1 = \mu_2 = 1/3$) and the following next-generation matrix:

$$K = \begin{pmatrix} 3 & 2 & 0 \\ 2 & 2 & 1 \\ 0 & 1 & 4 \end{pmatrix} \qquad \text{with inverse} \quad K^{-1} = \begin{pmatrix} 1.4 & -1.6 & 0.4 \\ -1.6 & 2.4 & -0.6 \\ 0.4 & -0.6 & 0.4 \end{pmatrix}.$$

Clearly Points (i) and (ii) hold and Point (iii) fails as K^{-1} is not an M-matrix. Nevertheless, the matrix K is definite positive as its eigenvalues $\sigma(K) = \{5, 2+\sqrt{3}, 2-\sqrt{3}\}$ are positive. And thus, thanks to Corollary V.5.5 (i), we get that the uniform strategies are Pareto optimal. Hence, Corollary V.5.5 (i) is a strict generalization of [130, Theorem 4.7] even for finite metapopulation models.

Remark V.5.8. We also refer the reader to the paper of Friedland and Karlin [63]. From the Inequality (7.10) therein, we can obtain Corollary V.5.5 (i) when Ω is a compact set of $\mathbb{R}^n$, μ is a finite measure, k is a continuous symmetrizable kernel such that $\mathrm{k}(x, x) > 0$ for all $x \in \Omega$.

Below, we give examples of metapopulation models from the previous sections where Proposition V.5.4 applies. For continuous models, we refer the reader to Sections V.6.

Example V.5.9 (Fully asymmetric cycle model). We consider the fully asymmetric circle model with $N \geq 2$ vertices developed in Section V.2.3. Since the in and out degree of each vertex is exactly one, the adjacency matrix is regular according to Example V.5.3 (ii). In this case the effective spectral radius $\mathfrak{R}_e$ is given by formula (V.19), which corresponds to the geometric mean. According to [23, Section 3.1.5], the map $\eta \mapsto \mathfrak{R}_e(\eta)$ is concave. So Proposition V.5.4 (ii) applies.

The spectrum of the adjency matrix is given by the Nth roots of unity. So, for $N \geq 3$ in this example, we get that the function $\mathfrak{R}_e$ is concave whereas its spectrum does not lie in $\mathbb{R}_- \cup \{\mathfrak{R}_0\}$. This proves that the condition given in Section 5.2 in Chapter IV to get $\mathfrak{R}_e$ concave (mainly the kernel or adjacency matrix is diagonally symmetrizable and has only one simple positive eigenvalue) is only sufficient.

Example V.5.10 (Finite assortative and disassortative model). Let $\Omega = \{0, 1, \ldots, N-1\}$ and μ be the uniform probability on Ω. Let $a, b \in \mathbb{R}_+$. We consider the kernel from the models developed in Section V.4:

$$\mathrm{k}(i, j) = a\mathbb{1}_{i=j} + b\mathbb{1}_{i\neq j}.$$

Since μ is uniform, the kernel k is constant degree; provided its spectral radius is positive, *i.e.*, a or b is positive.

In the assortative model $0 < b \leq a$, according to Proposition V.4.2, the eigenvalues of the symmetric operator T_k are non-negative. Hence, Corollary V.5.5 (i) applies: the Pareto optimal strategies are the uniform ones. This is consistent with Theorem V.4.3 (i).

In the dissortative model, we have $0 \leq a \leq b$ and $b > 0$. According to Proposition V.4.1, the eigenvalues of T_k different from its spectral radius are non-positive. Thanks to Section 5.2 in Chapter IV, the function $\mathfrak{R}_e$ is concave. Hence, Corollary V.5.5 (ii) applies: the uniform strategies are anti-Pareto. This is consistent with Theorem V.4.3 (ii) and (iii).

V.5.3 Proof of Proposition V.5.4

We start with a series of lemmas.

Lemma V.5.11 (Reduction)**.** *If* k *is a constant degree kernel on* Ω*, then there exists a finite partition* $(\Omega_i, i \in I)$ *of* Ω *in measurable subsets such that a.e.* $\mathrm{k} = \sum_{i\in I} \mathrm{k}_i$*, where* $\mathrm{k}_i = \mathbb{1}_{\Omega_i}\mathrm{k}\mathbb{1}_{\Omega_i}$, $\mathfrak{R}_0[\mathrm{k}_i] = \mathfrak{R}_0[\mathrm{k}]$, *and* k_i *restricted to* Ω_i *is irreducible and constant degree.*

Proof. We recall that a set $A \in \mathscr{F}$ is invariant if $\mathrm{k}(A^\mathrm{c}, A) = 0$, where for $A, B \in \mathscr{F}$:

$$\mathrm{k}(B, A) = \int_{B\times A} \mathrm{k}(x, y)\, \mu(\mathrm{d}x)\mu(\mathrm{d}y).$$

Since k is constant degree, we get $\mathrm{k}(\Omega, \Omega) = \mathfrak{R}_0$, $\mathrm{k}(A^\mathrm{c}, \Omega) = \mathfrak{R}_0\, \mu(A^\mathrm{c})$ and $\mathrm{k}(\Omega, A) = \mathfrak{R}_0\, \mu(A)$. Assume that A is an invariant set, then:

$$\mathrm{k}(A, A^\mathrm{c}) = \mathrm{k}(\Omega, \Omega) - \mathrm{k}(\Omega, A) - \mathrm{k}(A^\mathrm{c}, \Omega) + \mathrm{k}(A^\mathrm{c}, A) = 0.$$

This gives that A^c is also invariant. According to Section 7 in Chapter IV there exists an at most countable partition $(\Omega_i, i \in I)$ of Ω with positive measure such that $\mathrm{k} = \sum_{i\in I} \mathrm{k}_i$, with $\mathrm{k}_i = \mathbb{1}_{\Omega_i}\mathrm{k}\mathbb{1}_{\Omega_i}$ and k_i restricted to Ω_i is irreducible or quasi-nilpotent. Since $\mathbb{1}$ is an eigenvector of T_k associated to the eigenvalue $\mathfrak{R}_0$ and the sets $(\Omega_i, i \in I)$ are pairwise disjoint, we deduce that $\mathbb{1}_{\Omega_i}$ is an eigenvector of T_{k_i} with eigenvalue $\mathfrak{R}_0 > 0$, for all $i \in I$. Hence, all the kernels k_i restricted to Ω_i are irreducible and regular and the cardinal of I is equal to the multiplicity of $\mathfrak{R}_0$ (for T_k). Since k has finite L^2 double-norm, the operator T_k is compact, and the multiplicity of $\mathfrak{R}_0 > 0$, and thus the cardinal of I, is finite. □

Lemma V.5.12. *Let* k *be a constant degree irreducible kernel on* Ω*. Then the uniform strategy is a critical point for* $\mathfrak{R}_e$ *among all the strategies with the same cost in* $(0, 1)$*, and more precisely: for all* η *with the same cost in* $(0, 1)$ *as* η^uni *and* $\varepsilon > 0$ *small enough,*

$$\mathfrak{R}_e((1 - \varepsilon)\eta^\mathrm{uni} + \varepsilon\eta) = \mathfrak{R}_e(\eta^\mathrm{uni}) + O(\varepsilon^2).$$

Proof. Let η^uni be the uniform strategy with cost $c \in (0, 1)$. Since k is irreducible, we get that $(1 - c)\mathfrak{R}_0$ is a simple isolated eigenvalue of $\mathrm{k}\eta^\mathrm{uni}$, whose corresponding left and right eigenvector are $\mathbb{1}$ as $\mathrm{k}\eta^\mathrm{uni}$ is also constant degree. For $\eta \in \Delta$, we get that $T_{\mathrm{k}((1-\varepsilon)\eta^\mathrm{uni}+\varepsilon\eta)}$ converges to $T_{\mathrm{k}\eta^\mathrm{uni}}$ (in operator norm, thanks to (V.24)) as ε goes down to 0. Notice that:

$$\left\| T_{\mathrm{k}(\eta^\mathrm{uni}+\varepsilon(\eta-\eta^\mathrm{uni}))} - T_{\mathrm{k}\eta^\mathrm{uni}} \right\|_{L^2} = O(\varepsilon^2).$$

According to [98, Theorem 2.6], we get that for any $\eta \in \Delta$ and $\varepsilon > 0$ small enough:

$$\begin{aligned}\mathfrak{R}_e((1 - \varepsilon)\eta^\mathrm{uni} + \varepsilon\eta) - \mathfrak{R}_e(\eta^\mathrm{uni}) &= \varepsilon \int_\Omega \mathrm{k}(x, y)(\eta(y) - \eta^\mathrm{uni}(y))\, \mu(\mathrm{d}x)\mu(\mathrm{d}y) + O(\varepsilon^2)\\ &= \varepsilon\mathfrak{R}_0 \int_\Omega (\eta(y) - \eta^\mathrm{uni}(y))\, \mu(\mathrm{d}y) + O(\varepsilon^2),\end{aligned}$$

where for the last equality we used that k is constant degree. In particular, if η and η^uni have the same cost $c \in (0, 1)$, then $\mathfrak{R}_e((1 - \varepsilon)\eta^\mathrm{uni} + \varepsilon\eta) - \mathfrak{R}_e(\eta^\mathrm{uni}) = O(\varepsilon^2)$, which means that the uniform strategy is a critical point for $\mathfrak{R}_e$ among all the strategies with cost $c \in (0, 1)$. □

Proof of Proposition V.5.4. We prove (i), and thus consider $\mathfrak{R}_e$ convex and k constant degree. We first consider the case where k irreducible. For any η, Lemma V.5.12 and the convexity of $\mathfrak{R}_e$ imply that $\mathfrak{R}_e(\eta^{\text{uni}}) + O(\varepsilon^2) = \mathfrak{R}_e((1-\varepsilon)\eta^{\text{uni}} + \varepsilon\eta) \leq (1-\varepsilon)\mathfrak{R}_e(\eta^{\text{uni}}) + \varepsilon\mathfrak{R}_e(\eta)$, where η^{uni} the uniform strategy with the same cost as η. Sending ε to 0, we get $\mathfrak{R}_e(\eta) \geq \mathfrak{R}_e(\eta^{\text{uni}})$, so $\mathfrak{R}_e$ is minimal at η^{uni}.

Since $C(\eta^{\text{uni}}) = c$ and $\mathfrak{R}_e(\eta^{\text{uni}}) = (1-c)\mathfrak{R}_0$, we deduce that $\mathfrak{R}_{e\star}(c) = \mathfrak{R}_0(1-c)$ and thus, the Pareto frontier is a segment given by $\mathcal{F} = \{(c, (1-c)\mathfrak{R}_0) \,:\, c \in [0,1]\}$.

If k is not irreducible, then use the representation from Lemma V.5.11, to get that $\mathfrak{R}_e[\mathrm{k}] = \max_{i\in I} \mathfrak{R}_e[\mathrm{k}_i]$. Since the cost is affine, we deduce that a strategy η with $\mathfrak{R}_e[\mathrm{k}](\eta) = \ell \in [0, \mathfrak{R}_0]$ is optimal if and only if, for all $i \in I$, the strategies $\eta_i = \eta\mathbb{1}_{\Omega_i}$ are optimal for the kernel k restricted to Ω_i and $\mathfrak{R}_e[\mathrm{k}_i](\eta_i) = \ell$. Then the first step of the proof yields that $\eta_i = \ell\mathbb{1}_{\Omega_i}$ and thus the uniform strategy $\eta^{\text{uni}} = \ell\mathbb{1}_\Omega$ is optimal. This ends the proof of (i).

The proof of (ii) is similar to the case $\mathfrak{R}_e$ convex when k is irreducible. □

V.6 Constant degree symmetric kernels of rank two

V.6.1 Pareto and anti-Pareto frontiers

Any constant degree symmetric kernel may be decomposed spectrally as

$$\mathrm{k}(x,y) = \mathfrak{R}_0 + \sum_{n\in\mathbb{N}^*} \varepsilon_n \alpha_n(x)\alpha_n(y),$$

with $\varepsilon_n \in \{-,+\}$, $(\alpha_n, n \in \mathbb{N}^*)$ an orthogonal family of L^2 also orthogonal to $\mathbb{1}$. As an application of the results from the previous section, we will treat the case of symmetric constant degree kernel operators of rank 2, where one can explicitely minimize and maximize $\mathfrak{R}_e$ among all strategies of a given cost.

We suppose that $\Omega = [0,1)$ is equipped with the Borel σ-field $\mathscr{F}$ and a probability measure μ such that its cumulative distribution function φ, defined by $\varphi(x) = \mu([0,x])$ for $x \in \Omega$, is continuous and increasing. We consider the following two kernels on Ω:

$$\mathrm{k}^\varepsilon(x,y) = \mathfrak{R}_0 + \varepsilon\alpha(x)\alpha(y), \quad \text{with} \quad \varepsilon \in \{-,+\}, \tag{V.48}$$

$\mathfrak{R}_0 > 0$ and $\alpha \in L^2$ is increasing and satisfies:

$$\sup_\Omega \alpha^2 \leq \mathfrak{R}_0 \quad \text{and} \quad \int_\Omega \alpha \,\mathrm{d}\mu = 0. \tag{V.49}$$

Remark V.6.1 (Generality). We note that this particular choice of Ω may be made without loss of generality, and that the strict monotonicity assumption is almost general: we refer the interested reader to Section V.6.2 for further discussion of this point.

For $\varepsilon \in \{-,+\}$, the kernel k^ε is symmetric and constant degree. Furthermore, $\mathfrak{R}_0$ and $\varepsilon\int_\Omega \alpha^2\,\mathrm{d}\mu$ are the only eigenvalues (and their multiplicity is one) of $T_{\mathrm{k}^\varepsilon}$ with corresponding eigen-vector $\mathbb{1}$ and α. Since $\alpha^2 \leq \mathfrak{R}_0$, we also get that $\mathfrak{R}_0$ is indeed the spectral radius of $T_{\mathrm{k}^\varepsilon}$.

The Pareto (resp. anti-Pareto) frontier is already greedily parametrized by the uniform strategies for the kernel k^+ (resp. k^-), see Corollary V.5.5. The following result restricts the choice of anti-Pareto (resp. Pareto) optimal strategies to two extreme strategies. Hence, in order to find them, it is enough to compute and compare the two values of $\mathfrak{R}_e$ for each cost.

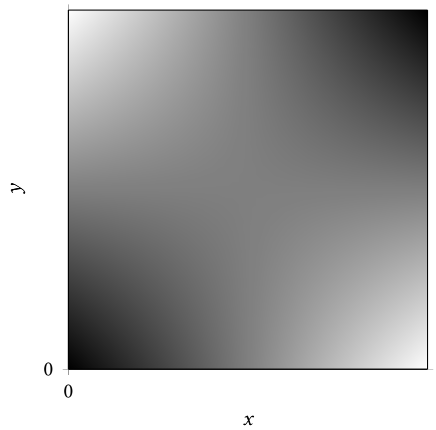
We recall the set of uniform strategies $\mathcal{S}^{\text{uni}} = \{t\mathbb{1} \,:\, t \in [0,1]\}$ and consider the following set of extremal strategies:

$$\mathcal{S}_0 = \left\{\mathbb{1}_{[0,t)} \,:\, t \in [0,1]\right\} \quad \text{and} \quad \mathcal{S}_1 = \left\{\mathbb{1}_{[t,1)} \,:\, t \in [0,1]\right\}$$

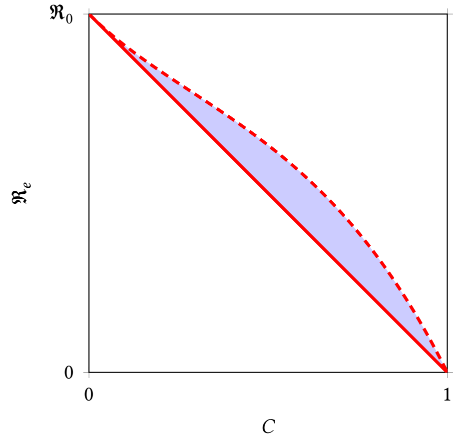
as well as the following set of strategies which contains $\mathcal{S}^{\text{uni}}$ thanks to (V.49):

$$\mathcal{S}^{\perp\alpha} = \left\{\eta \in \Delta \,:\, \int_\Omega \alpha\eta\,\mathrm{d}\mu = 0\right\}.$$

Recall that strategies are defined up to the a.s. equality. The proof of the next proposition is given is Section V.6.3



(A) Grayplot of the kernel $k^+(x, y) = 1 + (2x - 1)(2y - 1)$ on $[0, 1)$.



(B) Red line: the Pareto frontier $\mathcal{F}$ (which corresponds to the uniform strategies); dashed line: the anti-Pareto frontier $\mathcal{F}^{\text{Anti}}$; blue region: all possible outcomes $\mathbf{F}$.

Figure V.9: An example of a constant degree kernel operator of rank 2.

Proposition V.6.2 (Optima are uniform or on the sides)**.** *Let the kernel* k^ε *on* $[0, 1)$ *endowed with a probability measure whose cumulative distribution function is increasing and continuous, be given by* (V.48) *with* $\mathfrak{R}_0 > 0$ *and* α *a strictly increasing function on* $[0, 1)$ *such that* (V.49) *holds.*

The kernel k^+ *A strategy is Pareto optimal if and only if it belongs to* $\mathcal{S}^{\perp\alpha}$*. In particular, for any* $c \in [0, 1]$*, the strategy* $(1 - c)\mathbb{1}$ *costs* c *and is Pareto optimal. The only possible anti-Pareto strategies of cost* c *are* $\mathbb{1}_{[0,1-c)}$ *and* $\mathbb{1}_{[c,1)}$*. In other words,*

$$\mathcal{P} = \mathcal{S}^{\perp\alpha} \quad \textit{and} \quad \mathcal{P}^{\text{Anti}} \subset \mathcal{S}_0 \cup \mathcal{S}_1.$$

The kernel k^- *A strategy is anti-Pareto optimal if and only if it belongs to* $\mathcal{S}^{\perp\alpha}$*. In particular, for any* $c \in [0, 1]$*, the strategy* $(1 - c)\mathbb{1}$ *costs* c *and is anti-Pareto optimal. The only possible Pareto strategies of cost* c *are* $\mathbb{1}_{[0,1-c)}$ *and* $\mathbb{1}_{[c,1)}$*. In other words,*

$$\mathcal{P} \subset \mathcal{S}_0 \cup \mathcal{S}_1 \quad \textit{and} \quad \mathcal{P}^{\text{Anti}} = \mathcal{S}^{\perp\alpha}.$$

In both cases, we have $c^\star = 0$ *and* $c_\star = 1$*.*

Remark V.6.3. Intuitively, the populations $\{\alpha < 0\}$ and $\{\alpha > 0\}$ behave in an assortative way for k^+ and in a disassortative way for k^-. As in Section V.4, the uniform strategies are Pareto optimal in the “assortative” k^+ case and anti-Pareto optimal in the “disassortative” k^- case.

Remark V.6.4. Under the assumptions of Proposition V.6.2, if furthermore α is anti-symmetric with respect to $1/2$, that is $\alpha(x) = -\alpha(1 - x)$ for $x \in (0, 1)$, and μ is symmetric with respect to $1/2$, that is $\mu([0, x]) = \mu([1 - x, 1))$, then it is easy to check from the proof of Proposition V.6.2 that the strategies from $\mathcal{S}_0$ and $\mathcal{S}_1$ are both optimal: $\mathcal{P}^{\text{Anti}} = \mathcal{S}_0 \cup \mathcal{S}_1$ for k^+ and $\mathcal{P} = \mathcal{S}_0 \cup \mathcal{S}_1$ for k^-. We plotted such an instance of k^+ and the corresponding Pareto and anti-Pareto frontiers in Figure V.9. We refer to Section V.6.4 for an instance where α is not symmetric and $\mathcal{P} \neq \mathcal{S}_0 \cup \mathcal{S}_1$ for k^-.

V.6.2 On the choice of Ω and on the monotonicity assumption

Using a reduction model technique from Section 7 in Chapter III, let us first see that there is no loss of generality by considering the kernel $k^\varepsilon = \mathfrak{R}_0 + \varepsilon\alpha \otimes \alpha$ on $\Omega = [0, 1)$ endowed with the Lebesgue measure and with α non-decreasing.

Suppose that the function α in (V.48) is replaced by an $\mathbb{R}$-valued measurable function α_0 defined on a general probability space $(\Omega_0, \mathscr{F}_0, \mu_0)$ such that (V.49) holds. Thus, with obvious notations, for $\varepsilon \in \{-,+\}$, the kernel $\mathfrak{R}_0 + \varepsilon\alpha_0 \otimes \alpha_0$ is a kernel on Ω_0. Denote by F the repartition function of α_0 (that is, $F(r) = \mu_0(\alpha_0 \leq r)$ for $r \in \mathbb{R}$) and take α as the quantile function of α_0, that is, the right continuous inverse of F. Notice the function α is defined on the probability space $(\Omega, \mathscr{F}, \mu)$ is non-decreasing and satisfies (V.49). Consider the probability kernel $\kappa : \Omega_0 \times \mathscr{F} \to [0,1]$ defined by $\kappa(x,\cdot) = \delta_{F(\alpha_0(x))}(\cdot)$, with δ the Dirac mass, if α is continuous at $\alpha_0(x)$ (that is, $F(\alpha_0(x)-) = F(\alpha_0(x))$) and the uniform probability measure on $[F(\alpha_0(x)-), F(\alpha_0(x))]$ otherwise. On the measurable space $(\Omega_0 \times \Omega, \mathscr{F}_0 \otimes \mathscr{F})$, we consider the probability measure $\nu(\mathrm{d}x_1, \mathrm{d}x_2) = \mu_0(\mathrm{d}x_1)\kappa(x_1, \mathrm{d}x_2)$, whose marginals are exactly μ_0 and μ. Then, for $\varepsilon \in \{-,+\}$, we have that $\nu(\mathrm{d}x_1, \mathrm{d}x_2) \otimes \nu(\mathrm{d}y_1, \mathrm{d}y_2)$-a.s.:

$$\mathfrak{R}_0 + \varepsilon\alpha_0(x_1)\alpha_0(y_1) = \mathfrak{R}_0 + \varepsilon\alpha(x_2)\alpha(y_2).$$

According to Section 7.3 of Chapter III, see in particular Proposition III.7.3, the kernels $\mathfrak{R}_0 + \varepsilon\alpha_0 \otimes \alpha_0$ and $\mathfrak{R}_0 + \varepsilon\alpha \otimes \alpha$ are coupled and there is a correspondence between the corresponding (anti-)Pareto optimal strategies and their (anti-)Pareto frontiers are the same.

Hence, there is no loss in generality in assuming that the function α in (V.48) is indeed defined on $[0,1)$ and is non-decreasing.

On the contrary, one cannot assume in full generality that α is *strictly* increasing, and when it is only non-decreasing, the situation is more complicated. Indeed, let us take the parameters $\mathfrak{R}_0 = 1$ and $\alpha = \mathbb{1}_{[0,0.5)} - \mathbb{1}_{[0.5,1)}$. Then, the kernel k^- is complete bi-partite: $\mathrm{k}^- = \mathbb{1}_{[0,0.5)\times[0.5,1)} + \mathbb{1}_{[0.5,1)\times[0,0.5)}$. Hence, according to Theorem V.4.3 (iii), we have $c_\star = 0.5$ for the kernel k^-. In a similar fashion, one can see that $\mathrm{k}^+ = \mathbb{1}_{[0,0.5)\times[0,0.5)} + \mathbb{1}_{[0.5,1)\times[0.5,1)}$ is assortative and reducible; it is then easy to check that $c^\star = 0.5$ for the kernel k^+. However, it is still true that, for all costs c:

- $\mathbb{1}_{[0,1-c)}$ or $\mathbb{1}_{[c,1)}$ is solution of Problem (V.8) when the kernel k^+ is considered,
- $\mathbb{1}_{[0,1-c)}$ or $\mathbb{1}_{[c,1)}$ is solution of Problem (V.7) when the kernel k^- is considered,

From the proof of Proposition V.6.2, we can not expect to have strict inequalities in (V.59) if α is only non-decreasing, and thus one can not expect $\mathcal{S}_0 \cup \mathcal{S}_1$ to contain $\mathcal{P}^{\mathrm{Anti}}$ for the kernel k^+ or $\mathcal{P}$ for the kernel k^-.

V.6.3 Proof of Proposition V.6.2

We assume that $\mathfrak{R}_0 > 0$ and α is a strictly increasing function defined on Ω such that (V.49) holds. Without loss of generality, we shall assume that $\mathfrak{R}_0 = 1$ unless otherwise specified. We write $\mathfrak{R}_e^\varepsilon$ for the effective reproduction function associated to the kernel k^ε. We shall also write εa for a if $\varepsilon = +$ and $-a$ if $\varepsilon = -1$. We first rewrite $\mathfrak{R}_e$ in two different ways in Section V.6.3. Then, we consider the kernel k^- in Section V.6.3 and the kernel k^+ in Section V.6.3.

Two expressions of the effective reproduction function

We provide an explicit formula for the function $\mathfrak{R}_e$, and an alternative variational formulation, both of which will be needed below.

Lemma V.6.5. *Assume $\mathfrak{R}_0 = 1$ and α is a strictly increasing function defined on Ω such that* (V.49) *holds. We have for $\varepsilon \in \{+,-\}$ and $\eta \in \Delta$:*

$$2\mathfrak{R}_e^\varepsilon(\eta) = \int \eta \,\mathrm{d}\mu + \varepsilon \int \alpha^2 \eta \,\mathrm{d}\mu + \sqrt{\left(\int \eta \,\mathrm{d}\mu - \varepsilon \int \alpha^2 \eta \,\mathrm{d}\mu\right)^2 + 4\varepsilon \left(\int \alpha\eta \,\mathrm{d}\mu\right)^2}. \tag{V.50}$$

Alternatively, $\mathfrak{R}_e$ is the solution of the variational problem:

$$\mathfrak{R}_e^\varepsilon(\eta) = \sup_{h \in B_+^\eta} \left(\int_0^1 h\,\eta \,\mathrm{d}\mu\right)^2 + \varepsilon\left(\int_0^1 h\,\alpha\eta \,\mathrm{d}\mu\right)^2, \tag{V.51}$$

where

$$B_+^\eta = \left\{ h \in L_+^2 \ : \ \int_0^1 h^2\, \eta \, \mathrm{d}\mu = 1 \right\}.$$

The supremum in (V.51) *is reached for the right Perron eigenfunction of* $T_{\mathrm{k}^+\eta}$ *chosen in* B_+^η*.*

Proof. We first prove (V.50). For all $\eta \in \Delta$, the rank of the kernel operator $T_{\mathrm{k}^\varepsilon\eta}$ is smaller or equal to 2 and $\mathrm{Im}(T_{\mathrm{k}^\varepsilon\eta}) \subset \mathrm{Vect}(\mathbb{1}, \alpha)$. The matrix of $T_{\mathrm{k}^\varepsilon\eta}$ in the basis $(\mathbb{1}, \alpha)$ of the range of $T_{\mathrm{k}^\varepsilon\eta}$ is given by:

$$\begin{pmatrix} \int \eta \, \mathrm{d}\mu & \int \alpha\eta \, \mathrm{d}\mu \\ \varepsilon \int \alpha\eta \, \mathrm{d}\mu & \varepsilon \int \alpha^2\, \eta \, \mathrm{d}\mu \end{pmatrix}. \tag{V.52}$$

An explicit computation of the spectrum of this matrix yields Equation (V.50) for its largest eigenvalue.

The variational formula (V.51) is a direct consequence of general Lemma V.6.6 below. □

Lemma V.6.6 (Variational formula for $\mathfrak{R}_e$ when k is symmetric)**.** *Suppose that* k *is a symmetric kernel on* Ω *with a finite double norm in* L^2*. Then, we have that for all* $\eta \in \Delta$*:*

$$\mathfrak{R}_e(\eta) = \sup_{h \in B_+^\eta} \int_{\Omega\times\Omega} h(x)\eta(x)\, \mathrm{k}(x,y)\, h(y)\eta(y)\, \mu(\mathrm{d}x)\mu(\mathrm{d}y), \tag{V.53}$$

where

$$B_+^\eta = \left\{ h \in L_+^2 \ : \ \int_\Omega h^2\, \eta \, \mathrm{d}\mu = 1 \right\}.$$

The supremum in (V.53) *is reached for the right Perron eigenfunction of* $T_{\mathrm{k}^+\eta}$ *chosen in* B_+^η*.*

Proof. For a finite measure ν on $(\Omega, \mathscr{F})$, as usual, we denote by $L^2(\nu)$ the set of measurable real-valued functions f such that $\int_\Omega f^2 \mathrm{d}\nu < +\infty$ endowed with the usual scalar product, so that $L^2(\nu)$ is an Hilbert space. Let $\eta \in \Delta$. We denote by $\mathcal{T}_{\mathrm{k}\eta}$ the integral operator associated to the kernel $\mathrm{k}\eta$ seen as an operator on the Hilbert space $L^2(\eta \mathrm{d}\mu)$: for $g \in L^2(\eta\mathrm{d}\mu)$ and $x \in \Omega$ we have $\mathcal{T}_{\mathrm{k}\eta}(g)(x) = \int_\Omega \mathrm{k}(x,y)g(y)\, \mu(\mathrm{d}y)$. The operator $\mathcal{T}_{\mathrm{k}\eta}$ is self-adjoint and compact since the double-norm of k in $L^2(\eta\mathrm{d}\mu)$ is finite. It follows from the Krein-Rutman theorem and the Courant–Fischer–Weyl min-max principle that its spectral radius is given by the variational formula:

$$\rho(\mathcal{T}_{\mathrm{k}\eta}) = \sup_{h \in B_+^\eta} \int_{\Omega\times\Omega} h(x)\, \mathrm{k}(x,y)\, h(y)\, \eta(x)\mu(\mathrm{d}x)\, \eta(y)\mu(\mathrm{d}y).$$

Besides, the set $L^2(\mu)$ is densely and continuously embedded in $L^2(\eta\mathrm{d}\mu)$ and the restriction of $\mathcal{T}_{\mathrm{k}\eta}$ to $L^2(\mu)$ is equal to $T_{\mathrm{k}\eta}$. Thanks to Lemma III.3.2 (iii), we deduce that $\rho(T_{\mathrm{k}\eta})$ is equal to $\rho(\mathcal{T}_{\mathrm{k}\eta})$, which gives (V.53).

Let h_0 be the right Perron eigenfunction of $T_{\mathrm{k}^+\eta}$ chosen such that $h_0 \in B_+^\eta$. We get:

$$\int_{\Omega\times\Omega} \eta(x)h_0(x)\, \mathrm{k}(x,y)\, \eta(y)h_0(y)\, \mu(\mathrm{d}x)\mu(\mathrm{d}y) = \mathfrak{R}_e(\eta) \int_\Omega \eta(x)h_0(x)^2\, \mu(\mathrm{d}x) = \mathfrak{R}_e(\eta).$$

Thus, the supremum in (V.53) is reached for $h = h_0$. □

The kernel k^-

Since α is increasing, we have $\mu(\alpha^2 = \mathfrak{R}_0) = 0$ and thus the symmetric kernel k^+ is $\mu^{\otimes 2}$-a.s. positive. It follows from Remark V.3.1 that $c^\star = 0$ and $c_\star = 1$, and the strategy $\mathbb{1}$ (resp. $\mathbb{0}$) is the only Pareto optimal as well as the only anti-Pareto optimal strategy with cost $c = 0$ (resp. $c = 1$). Since the kernel k^- is constant degree and symmetric, and the non-zero eigenvalues of T_{k^-} are given by $\mathfrak{R}_0 = 1$ and $-\int \alpha^2\, \mathrm{d}\mu$, the latter being negative, we deduce from Corollary V.5.5 (ii) that $\mathcal{S}^{\mathrm{uni}} \subset \mathcal{P}^{\mathrm{Anti}}$. On the one hand, if η is anti-Pareto optimal, one can use that $\mathfrak{R}_e^-(\eta) = \int \eta \, \mathrm{d}\mu$ (as

$\mathfrak{R}_e^-(\eta^{\mathrm{uni}}) = \int \eta^{\mathrm{uni}}\, \mathrm{d}\mu$) and (V.50) to deduce that $\eta \in \mathcal{S}^{\perp\alpha}$. On the other hand, if η belongs to $\mathcal{S}^{\perp\alpha}$, we deduce from (V.50) that $\mathfrak{R}_e(\eta) = \int \eta\, \mathrm{d}\mu$, and thus η is anti-Pareto optimal. In conclusion, we get $\mathcal{P}^{\mathrm{Anti}} = \mathcal{S}^{\perp\alpha}$.

We now study the Pareto optimal strategies. We first introduce a notation inspired by the stochastic order of real valued random variables: we say that $\eta_1, \eta_2 \in \Delta$ with the same cost are in *stochastic order*, and we write $\eta_1 \leq_{\mathrm{st}} \eta_2$ if:

$$\int_0^t \eta_1\, \mathrm{d}\mu \geq \int_0^t \eta_2\, \mathrm{d}\mu \quad \text{for all} \quad t \in [0,1]. \tag{V.54}$$

We also write $\eta_1 <_{\mathrm{st}} \eta_2$ if the inequality in (V.54) is strict for at least one $t \in (0,1)$. If $\eta_1 <_{\mathrm{st}} \eta_2$ and h is an increasing bounded function defined on $[0,1)$, then we have:

$$\int_\Omega h\,\eta_1\, \mathrm{d}\mu < \int_\Omega h\,\eta_2\, \mathrm{d}\mu. \tag{V.55}$$

Let $c \in (0,1)$ be fixed. Define the vaccination strategies with cost c:

$$\eta_0 = \mathbb{1}_{[0,1-c)} \quad \text{and} \quad \eta_1 = \mathbb{1}_{[c,1)}. \tag{V.56}$$

In particular we have $\eta_0 <_{\mathrm{st}} \eta_1$ as μ has no atom and Ω as full support. Let η be a vaccination strategy with cost c not equal to η_0 or η_1. We get:

$$\eta_0 <_{\mathrm{st}} \eta <_{\mathrm{st}} \eta_1.$$

We now rewrite the function $\mathfrak{R}_e^-$ in order to use the stochastic order on the vaccination strategies. We deduce from (V.50) that:

$$4\mathfrak{R}_e^-(\eta) = 4\int \eta\, \mathrm{d}\mu - H(\eta)^2 \quad \text{with} \quad H(\eta) = \sqrt{\int (1+\alpha)^2 \eta\, \mathrm{d}\mu} - \sqrt{\int (1-\alpha)^2 \eta\, \mathrm{d}\mu}. \tag{V.57}$$

Then, using that α is increasing and $[-1,1]$-valued, we deduce from (V.55) (with $h = (1+\alpha)^2$ and $h = -(1-\alpha)^2$) and the definition of H in (V.57) that:

$$H(\eta_0) < H(\eta) < H(\eta_1).$$

This readily implies that $\mathfrak{R}_e^-(\eta) > \min(\mathfrak{R}_e^-(\eta_0), \mathfrak{R}_e^-(\eta_1))$: among strategies of cost c, the only possible Pareto optimal ones are η_0 and η_1. We deduce that $\mathcal{P} \subset \mathcal{S}_0 \cup \mathcal{S}_1$.

The kernel k^+

Arguing as for k^-, we get that $c^\star = 0$ and $c_\star = 1$, and the strategy $\mathbb{1}$ (resp. $\mathbb{0}$) is the only Pareto optimal as well as the only anti-Pareto optimal strategy with cost $c = 0$ (resp. $c = 1$). Since the kernel k^+ is constant degree and symmetric, and the non-zero eigenvalues of T_{k^+} given by $\mathfrak{R}_0$ and $\int_\Omega \alpha^2\, \mathrm{d}\mu$ are positive, we deduce from Corollary V.5.5 (i) that $\mathcal{S}^{\mathrm{uni}} \subset \mathcal{P}$.

Arguing as in Section V.6.3 for the identification of the anti-Pareto optima based on (V.50) (with $\varepsilon = +$ instead of $\varepsilon = -$) and using that $\mathcal{S}^{\mathrm{uni}} \subset \mathcal{P}$ (instead of $\mathcal{S}^{\mathrm{uni}} \subset \mathcal{P}^{\mathrm{Anti}}$), we deduce that $\mathcal{P} = \mathcal{P}^\top$.

We now consider the anti-Pareto optima. Let $c \in (0,1)$. We first start with some comparison of integrals with respect to the vaccination strategies, with cost c, η_0 and η_1 defined by (V.56). Let η be a strategy of cost c not equal to η_0 or η_1 (recall that a strategy is defined up to the a.s. equality). Consider the monotone continuous non-negative functions defined on $[0,1]$:

$$\phi_0 \;:\; x \mapsto \varphi^{-1}\left(\int_{[0,x)} \eta\, \mathrm{d}\mu\right), \quad \text{and} \quad \phi_1 \;:\; x \mapsto \varphi^{-1}\left(1 - \int_{[x,1)} \eta\, \mathrm{d}\mu\right).$$

Let $i \in \{0,1\}$. Let ϕ_i^{-1} denote the generalized left-continuous inverse of ϕ_i. Notice that $\eta(x)\,\mu(\mathrm{d}x)$-a.s. $\phi_i^{-1} \circ \phi_i(x) = x$. The measure $\eta_i \mathrm{d}\mu$ is the push-forward of $\eta\,\mathrm{d}\mu$ through ϕ_i, so that for h bounded measurable:

$$\int h\,\eta\,\mathrm{d}\mu = \int h_i\,\eta_i\,\mathrm{d}\mu \quad \text{with} \quad h_i = h \circ \phi_i^{-1}. \tag{V.58}$$

Since η is not equal to η_0 a.s., there exists $x_0 < 1-c$ such that, $\phi_0(x) = x$ for $x \in [0, x_0]$ and $\phi_0(x) < x$ for $x \in (x_0, 1]$. Thus, we deduce that $\phi_0^{-1}(y) = y$ for all $y \in [0, x_0]$ and $\phi_0^{-1}(y) > y$ for all $y \in (x_0, 1-c]$. Similarly, since η is not equal to η_1 almost surely, there exists $x_1 > c$ such that $\phi_1^{-1}(y) = y$ for all $y \in (x_1, 1]$ and $\phi_1^{-1}(y) < y$ for all $y \in [c, x_1)$. Since α is non-decreasing as μ has no atom and Ω as full support, we deduce from from (V.58), applied to $h\alpha$, that if h is a.s. positive bounded measurable, then:

$$\int h_0\,\alpha\eta_0\,\mathrm{d}\mu < \int h\,\alpha\eta\,\mathrm{d}\mu < \int h_1\,\alpha\eta_1\,\mathrm{d}\mu. \tag{V.59}$$

Let h be the right Perron eigenfunction of $T_{\mathrm{k}^+\eta}$ chosen such that $h \in B_+^{\eta}$. Since k^+ is a.s. positive and thus irreducible, we have that h is a.s. positive. Thanks to Lemma V.6.5, we have:

$$\mathfrak{R}_e^+(\eta) = \left(\int h\,\eta\,\mathrm{d}\mu\right)^2 + \left(\int h\,\alpha\eta\,\mathrm{d}\mu\right)^2 \quad \text{and} \quad \int h^2\,\eta\,\mathrm{d}\mu = 1. \tag{V.60}$$

We deduce from (V.58) that for $i \in \{0,1\}$:

$$\int h\,\eta\,\mathrm{d}\mu = \int h_i\,\eta_i\,\mathrm{d}\mu \quad \text{and} \quad 1 = \int h^2\,\eta\,\mathrm{d}\mu = \int h_i^2\,\eta_i\,\mathrm{d}\mu.$$

In particular h_i belongs to $B_+^{\eta_i}$. Using that a.s. $h > 0$, we then deduce from (V.60) and (V.59) that:

$$\mathfrak{R}_e^+(\eta) < \max_{i\in\{0,1\}} \left(\int h_i\,\eta_i\,\mathrm{d}\mu\right)^2 + \left(\int h_i\,\alpha\eta_i\,\mathrm{d}\mu\right)^2 \le \max_{i\in\{0,1\}} \mathfrak{R}_e(\mathbb{1}_{A_i}).$$

We conclude that only η_0 or η_1 can maximize $\mathfrak{R}_e^+$ among the strategies of cost $c \in (0,1)$. We deduce that $\mathcal{P}^{\mathrm{Anti}} \subset \mathcal{S}_0 \cup \mathcal{S}_1$.

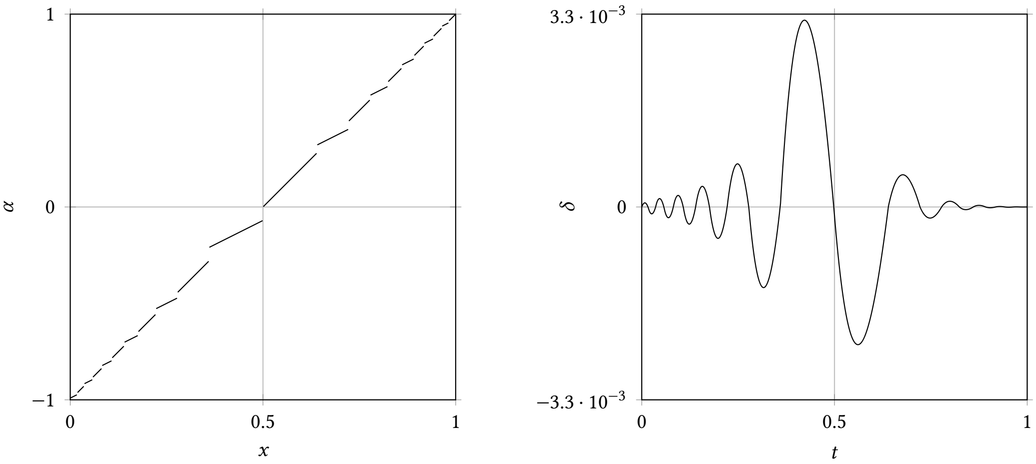
V.6.4 An example where all parametrizations of the Pareto frontier have an infinite-number of discontinuities

The purpose of this section is to give a particular example of kernel on a continuous model where we rigorously prove that the Pareto frontier cannot be greedily parametrized, that is, parametrized by a continuous path in Δ (as in the fully symmetric circle), and that all the parametrizations have an arbitrary large number of discontinuities (possibly countably infinite).

We keep the setting from Section V.6. Without loss of generality, we assume that $\mathfrak{R}_0 = 1$, and we consider the kernel $\mathrm{k}^- = 1 - \alpha \otimes \alpha$ on $\Omega = [0,1)$ endowed with its Lebesgue measure. We know from the previous section that, for any cost, either η_0 or η_1 are Pareto optimal, and that all other strategies are non-optimal. The idea is then to build an α in such a way that for some costs, one must vaccinate “on the left” and for other costs “on the right”.

Let $N \in [\![2, +\infty]\!]$. Consider an increasing sequence $(x_n,\, n \in [\![0, N]\!])$ such that $x_0 = 1/2$, $x_N = 1$ and $\lim_{n\to\infty} x_n = 1$ if $N = \infty$. For $0 \le n < N$, let $p_n = x_{n+1} - x_n$ and assume that $p_{n+1} < p_n$ for $n \in [\![0, N[\![$. For $n \ge 1$, let x_{-n} be the symmetric of x_n with respect to $1/2$, *i.e.*, $x_{-n} = 1 - x_n$. The function α is piecewise linear defined on $(0,1)$ by:

$$\alpha(x) = \begin{cases} 2x - 1, & \text{for } x \in [x_{2m}, x_{2m+1}), \\ \\ x - 1 + \frac{x_{2m-1}+x_{2m}}{2} & \text{for } x \in [x_{2m-1}, x_{2m}). \end{cases} \tag{V.61}$$



(A) Graph of the function α defined by Equation (V.61) and Example V.6.9.

(B) Graph of the corresponding function δ defined in Equation (V.64).

Figure V.10: Plots of the functions of interest in Section V.6.4.

See Figure V.10(A) for an instance of the graph of α given in Example V.6.9. Note that for all $n \in [\![0, N[\![$, we have:

$$\int_{x_n}^{x_{n+1}} \alpha \,\mathrm{d}\mu = -\int_{x_{-n-1}}^{x_{-n}} \alpha \,\mathrm{d}\mu. \tag{V.62}$$

This proves that the integral of α over $[0, 1)$ is equal to 0. Of course, $\sup_{[0,1)} \alpha^2 = 1 = \mathfrak{R}_0$. Hence, α satisfies Condition (V.49).

We recall that a function $\gamma : [0, c_\star] \mapsto \Delta$ is a prametrization of the Pareto frontier if for all $c \in [0, c_\star]$ the strategy $\gamma(c)$ is Pareto optimal with cost $C(\gamma(c)) = c$. Now we can prove there exists no greedy parametrization of the Pareto frontier of the kernel k^- and even impose an arbitrary large lower bound for the number of discontinuities.

Proposition V.6.7. *Let $N \in [\![2, +\infty]\!]$. Consider the kernel $\mathrm{k}^- = 1 - \alpha \otimes \alpha$ from* (V.48) *on $\Omega = [0, 1)$ endowed with its Lebesgue measure, with α given by* (V.61). *Then, any parametrization of the Pareto frontier has at least $2N - 2$ and at most $20N - 2$ discontinuities.*

The proof is given at the end of this section, and relies on the following technical lemma based on the comparison of the following monotone paths γ_0 and γ_1 from $[0, 1]$ to Δ:

$$\gamma_0(t) = \mathbb{1}_{[0,t)}, \quad \text{and} \quad \gamma_1(t) = \mathbb{1}_{[1-t,1)}, \quad t \in [0, 1] \tag{V.63}$$

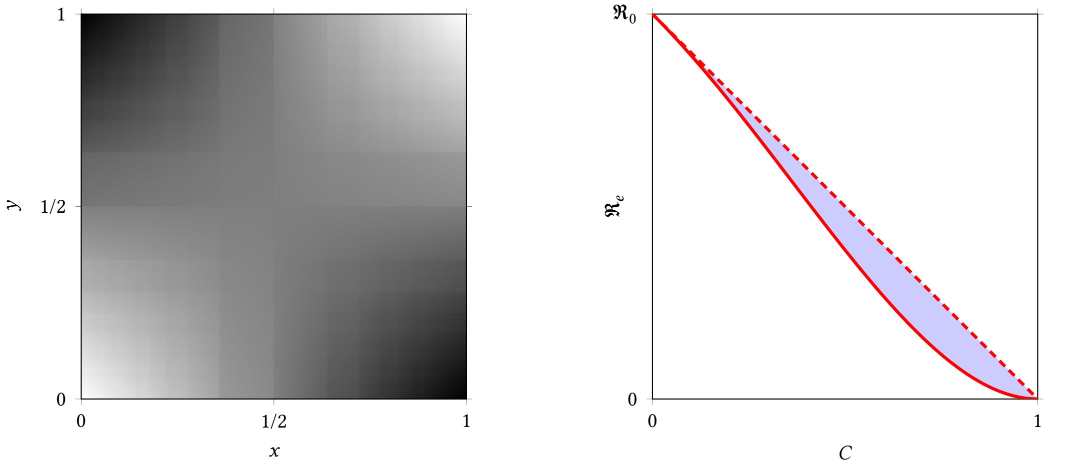
which parameterizes $\mathcal{S}_0$ and $\mathcal{S}_1$ as $\gamma_0([0, 1]) = \mathcal{S}_0$ and $\gamma_1([0, 1]) = \mathcal{S}_1$. Notice that strategies $\gamma_0(t)$ and $\gamma_1(t)$ have the same cost $1 - t$.

Consider the function $\delta : [0, 1] \to \mathbb{R}$ which, according to Proposition V.6.2, measures the difference between the effective reproduction numbers at the extreme strategies:

$$\delta(t) = \mathfrak{R}_e(\gamma_0(t)) - \mathfrak{R}_e(\gamma_1(t)). \tag{V.64}$$

The function δ is continuous and $\delta(0) = \delta(1) = 0$. We say that $t \in (0, 1)$ is a zero crossing of δ if $\delta(t) = 0$ and there exists $\varepsilon > 0$ such that $\delta(t + r)\delta(t - r) < 0$ for all $r \in (0, \varepsilon)$. The following result gives some information on the zeros of the function δ.

Lemma V.6.8. *Under the assumptions of Proposition V.6.7, the function δ defined in* (V.64) *has at least $2N - 2$ zero-crossing in $(0, 1)$ and at most $20N$ zeros in $[0, 1]$. Besides, if $N = \infty$, 0 and 1 are the only accumulation points of the set of zeros of δ.*



(A) Grayplot of the kernel $\mathrm{k}^-(x,y) = 1 - \alpha(x)\alpha(y)$ where α is plotted in Figure V.10(A).

(B) Red line: the Pareto frontier $\mathcal{F}$; dashed line: the anti-Pareto frontier $\mathcal{F}^{\text{Anti}}$ (which corresponds to the uniform strategies); blue region: all possible outcomes $\mathbf{F}$.

Figure V.11: An example of a constant degree kernel operator of rank 2.

Proof. Using the explicit representation of $\mathfrak{R}_e^-$ from Lemma V.6.5, see (V.50) with $\varepsilon = -$, we get the function δ can be expressed as:

$$2\delta(t) = V_1(t) - V_0(t) + \sqrt{V_0(t)^2 - M_0(t)^2} - \sqrt{V_1(t)^2 - M_1(t)^2}, \tag{V.65}$$

where, as $\int \alpha \, \mathrm{d}\mu = 0$:

$$M_0(t) = 2\int_0^t \alpha \, \mathrm{d}\mu, \quad V_0(t) = t + \int_0^t \alpha^2 \, \mathrm{d}\mu, \quad M_1(t) = M_0(1-t) \ \text{ and } \ V_1(t) = t + \int_{1-t}^1 \alpha^2 \, \mathrm{d}\mu.$$

Elementary computations give that for all $n \in [\![0, N[\![$:

$$\int_{x_n}^{x_{n+1}} \alpha(x)^2 \, \mathrm{d}x - \int_{x_{-n-1}}^{x_{-n}} \alpha(x)^2 \, \mathrm{d}x = \frac{(-1)^n p_n^3}{4}, \tag{V.66}$$

where we recall that $p_n = x_{n+1} - x_n$. Hence, we obtain that for all $n \in [\![-N, N]\!]$:

$$V_1(x_n) - V_0(x_n) = \frac{1}{4} \sum_{i=|n|}^{\infty} (-1)^i p_i^3. \tag{V.67}$$

Since the sequence $(p_n, n \in [\![0, N[\![)$ is decreasing, we deduce that the sign of $V_1(x_n) - V_0(x_n)$ alternates depending on the parity of $n \in]\!] - N, N[\![$: it is positive for odd n and negative for even n. The same result holds for the numbers $\delta(x_n)$ since $M_0(x_n) = M_1(1 - x_n)$ for all $n \in [\![-N, N]\!]$ according to (V.62) (use that, with $b > 0$, the function $x \mapsto x - \sqrt{x^2 - b^2}$ is decreasing for $x \geq \sqrt{b}$ as its derivative is negative). This implies that δ has at least $2N - 2$ zero-crossing in $(0, 1)$.

We now prove that δ has at most $20N$ zeros in $[0, 1]$ and that 0 and 1 are the only possible accumulation points of the set of zeros of δ. It is enough to prove that δ has at most 10 zeros on $[x_n, x_{n+1}]$ for all finite $n \in [\![-N, N[\![$. On such an interval $[x_n, x_{n+1}]$, the function α is a polynomial of degree one. Consider first n odd and non-negative, so that for $t \in [x_n, x_{n+1}]$, we get that with $a = 1 - (x_n + x_{n+1})/2$:

$$\begin{aligned} M_0(t) &= 2t^2 - 2t + b_1, \quad V_0(t) = \frac{4}{3}t^3 - 2t^2 + 2t + b_2, \\ M_1(t) &= t^2 - 2at + b_3, \quad V_1(t) = -\frac{1}{3}t^3 + at^2 + (1 - a^2)t + b_4, \end{aligned}$$

where b_i are constants. If t is a zero of δ, then it is also a zero of the polynomial P given by:

$$P = 4(V_1 - V_0)\left(V_0 M_1^2 - V_1 M_0^2\right) - \left(M_0^2 - M_1^2\right)^2 .$$

Since the degree of P is exactly 10, it has at most 10 zeros. Thus δ has at most 10 zeros on $[x_n, x_{n+1}]$. This ends the proof. ☐

Proof of Proposition V.6.7. According to Proposition V.6.2, the Pareto strategy of cost $c = 1 - t \in [0, 1]$ is $\gamma_0(t)$ or $\gamma_1(t)$. Then, a zero crossing of the function δ on $(0, 1)$ corresponds to a discontinuity of any parametrization of the Pareto frontier. We deduce from Lemma V.6.8 that in $(0, 1)$ there are at least $2N - 2$ and at most $20N - 2$ zeros crossing and thus discontinuities of any parametrization of the Pareto frontier. ☐

Example V.6.9. In Figure V.10(A), we have represented the function α defined by (V.61) where:

$$x_n = \frac{1}{2}\log_{12}(12(n+1)), \quad 0 \le n \le N = 11.$$

Hence, the mesh $(x_n, -N \le n \le N)$ is composed by $2N + 1 = 23$ points. The graph of the corresponding function δ defined in (V.64) is drawn in Figure V.10(B). The grayplot of the kernel $\mathrm{k}^- = 1 - \alpha \otimes \alpha$ is given in Figure V.11(A) and the associated Pareto and anti-Pareto frontiers are plotted in Figure V.11(B).

Bibliography

[1] E. Abbé. “Community detection and stochastic block models: recent developments”. In: *The Journal of Machine Learning Research* 18 (2018).

[2] D. Aldous and R. Lyons. “Processes on unimodular random networks”. In: *Electronic Journal of Probability* 12 (2007), pp. 1454–1508.

[3] L. Almeida, P.-A. Bliman, G. Nadin, B. Perthame, and N. Vauchelet. “Final size and convergence rate for an epidemic in heterogeneous populations”. In: *Mathematical Models and Methods in Applied Sciences* 31.5 (2021), pp. 1021–1051.

[4] H. Andersson and T. Britton. *Stochastic Epidemic Models and Their Statistical Analysis*. Vol. 151. Lecture Notes in Statistics. Springer-Verlag, 2000.

[5] P. M. Anselone and T. W. Palmer. “Collectively compact sets of linear operators”. In: *Pacific Journal of Mathematics* 25.3 (1968), pp. 417–422.

[6] P. M. Anselone. *Collectively compact operator approximation theory and applications to integral equations*. Prentice-Hall, 1971.

[7] P. M. Anselone and J. W. Lee. “Spectral properties of integral operators with nonnegative kernels”. In: *Linear Algebra and its Applications* 9 (1974), pp. 67–87.

[8] J. Arino and P. van den Driessche. “Disease spread in metapopulations”. In: *Nonlinear dynamics and evolution equations*. Vol. 48. Fields Institute Communications. Amererican Mathematical Society, 2006, pp. 1–12.

[9] B. C. Arnold. *Majorization and the Lorenz Order: A Brief Introduction*. Vol. 43. Lecture Notes in Statistics. Springer-Verlag, 1987.

[10] K. B. Athreya and P. E. Ney. *Branching Processes*. Vol. 196. Grundlehren der mathematischen Wissenschaften. Springer Verlag, 1972.

[11] C. Atkinson and G. E. H. Reuter. “Deterministic epidemic waves”. In: *Mathematical Proceedings of the Cambridge Philosophical Society* 80.2 (1976), pp. 315–330.

[12] Á. Backhausz and B. Szegedy. “Action convergence of operators and graphs”. In: *Canadian Journal of Mathematics* (2020), pp. 1–50.

[13] M. Benaïm and M. W. Hirsch. “Differential and stochastic epidemic models”. In: *Differential Equations with Applications to Biology*. Vol. 21. Fields Institute Communications. American Mathematical Society and Fields Institute, 1999, pp. 31–44.

[14] E. Beretta and V. Capasso. “Global Stability Results for a Multigroup SIR Epidemic Model”. In: *Mathematical Ecology*. Proceedings of the Autumn Course Research Seminars (Trieste, Italy, Nov. 24–Dec. 12, 1986). World Scientific, 1986, pp. 317–342.

[15] A. Berman and R. J. Plemmons. *Nonnegative Matrices in the Mathematical Sciences*. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics, 1994.

[16] D. Bernoulli. “Essai d’une nouvelle analyse de la mortalité causée par la petite vérole, et des avantages de l’inoculation pour la prevenir”. In: *Mémoire de mathématiques et de physique de l’Académie des Sciences* (1766), pp. 1–45.

[17] B. Bollobás, S. Janson, and O. Riordan. “The phase transition in inhomogeneous random graphs”. In: *Random Structures Algorithms* 31.1 (2007), pp. 3–122.

[18] C. Borgs, J. Chayes, L. Lovász, V. Sós, and K. Vesztergombi. "Convergent sequences of dense graphs I: Subgraph frequencies, metric properties and testing". In: *Advances in Mathematics* 219.6 (2008), pp. 1801–1851.

[19] C. Borgs, J. Chayes, H. Cohn, and Y. Zhao. "An L^p theory of sparse graph convergence I: Limits, sparse random graph models, and power law distributions". In: *Transactions of the American Mathematical Society* 372.5 (2019), pp. 3019–3062.

[20] C. Borgs, J. Chayes, L. Lovász, V. Sós, and K. Vesztergombi. "Convergent sequences of dense graphs II. Multiway cuts and statistical physics". In: *Annals of Mathematics* 176.1 (2012), pp. 151–219.

[21] C. Borgs, J. T. Chayes, H. Cohn, and Y. Zhao. "An L^p theory of sparse graph convergence II: LD convergence, quotients and right convergence". In: *The Annals of Probability* 46.1 (2018), pp. 337–396.

[22] A. Boussaïri and B. Chergui. "A transformation that preserves principal minors of skew-symmetric matrices". In: *Electronic Journal of Linear Algebra* 32 (2017), pp. 131–137.

[23] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge University Press, 2004.

[24] F. Brauer and C. Castillo-Chavez. *Mathematical Models in Population Biology and Epidemiology*. 2nd ed. Vol. 40. Texts in Applied Mathematics. Springer-Verlag, 2012.

[25] T. Britton, F. Ball, and P. Trapman. "A mathematical model reveals the influence of population heterogeneity on herd immunity to SARS-CoV-2". In: *Science* 369.6505 (2020), pp. 846–849.

[26] D. Burago, Y. Burago, and S. Ivanov. *A Course in Metric Geometry*. Vol. 33. Graduate Studies in Mathematics. American Mathematical Society, 2001.

[27] L. Burlando. "Continuity of spectrum and spectral radius in Banach algebras". In: *Banach Center Publications* 30 (1994), pp. 53–100.

[28] S. N. Busenberg, M. Iannelli, and H. R. Thieme. "Global behavior of an age-structured epidemic model". In: *SIAM Journal on Mathematical Analysis* 22.4 (1991), pp. 1065–1080.

[29] A. J. G. Cairns. "Epidemics in Heterogeneous Populations: Aspects of Optimal Vaccination Policies". In: *Mathematical Medicine and Biology* 6.3 (1989), pp. 137–159.

[30] J. B. Conway. *A course in functional analysis*. 2nd ed. Vol. 96. Graduate Texts in Mathematics. Springer-Verlag, 1990.

[31] U. L. Daleckij and M. G. Krejn. *Stability of Solutions of Differential Equations in Banach Space*. Vol. 43. Translation of Mathematical Monographs. American Mathematical Society, 2005.

[32] Y. De Castro, C. Lacour, and T. M. P. Ngoc. "Adaptive Estimation of Nonparametric Geometric Graphs". In: *Mathematical Statistics and Learning* 2.3/4 (2019), pp. 217–274.

[33] K. Deimling. *Nonlinear Functional Analysis*. Springer-Verlag, 1985.

[34] K. Deimling. *Ordinary Differential Equations in Banach Spaces*. Vol. 596. Lecture Notes in Mathematics. Springer-Verlag, 1977.

[35] J.-F. Delmas, D. Dronnier, and P.-A. Zitt. *An Infinite-Dimensional SIS Model*. Version 1. 2020. arXiv: `2006.08241 [math.DS]`.

[36] J.-F. Delmas, D. Dronnier, and P.-A. Zitt. *Effective reproduction number: Convexity, invariance and cordons sanitaires*. Version 1. 2021. arXiv: `2110.12693 [math.OC]`.

[37] J.-F. Delmas, D. Dronnier, and P.-A. Zitt. "Optimal vaccination for a 2 sub-populations SIS model". Forthcoming.

[38] J.-F. Delmas, D. Dronnier, and P.-A. Zitt. *Optimal vaccination: Various (counter) intuitive examples*. Version 1. 2021. arXiv: `2112.08756 [math.OC]`.

[39] J.-F. Delmas, D. Dronnier, and P.-A. Zitt. *Targeted Vaccination Strategies for an Infinite-dimensional SIS model*. Version 2. 2021. arXiv: `2103.10330v2 [math.PR]`.

[40] J.-F. Delmas, D. Dronnier, and P.-A. Zitt. “Vaccinating higly connected people is (sometimes) optimal”. Forthcoming.

[41] L. Di Domenico, G. Pullano, C. E. Sabbatini, P.-Y. Boëlle, and V. Colizza. “Impact of lockdown in Île-de-France and possible exit strategies”. In: *BMC Medicine* 18 (2020).

[42] O. Diekmann, J. A. P. Heesterbeek, and J. A. J. Metz. “On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations”. In: *Journal of Mathematical Biology* 28.4 (1990), pp. 365–382.

[43] K. Dietz. “Transmission and control of arbovirus diseases”. In: *Epidemiology* 104 (1975), pp. 104–121.

[44] K. Dietz and J. Heesterbeek. “Daniel Bernoulli’s epidemiological model revisited”. In: *Mathematical Biosciences* 180 (1-2 2002), pp. 1–21.

[45] H. R. Dowson. *Spectral theory of linear operators*. Vol. 12. London Mathematical Society Monographs. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], London-New York, 1978.

[46] E. Duijzer, W. van Jaarsveld, J. Wallinga, and R. Dekker. “The most efficient critical vaccination coverage and its equivalence with maximizing the herd effect”. In: *Mathematical Biosciences* 282 (2016), pp. 68–81.

[47] L. E. Duijzer, W. L. van Jaarsveld, J. Wallinga, and R. Dekker. “Dose-optimal vaccine allocation over multiple populations”. In: *Production and Operations Management* 27.1 (2018), pp. 143–159.

[48] N. Dunford and J. T. Schwartz. *Linear Operators. Part I*. Wiley Classics Library. John Wiley & Sons, 1988.

[49] D. Edmunds and D. Evans. *Spectral Theory and Differential Operators*. 2nd ed. Vol. 1. Oxford University Press, 2018.

[50] G. Elek. “Note on limits of finite graphs”. In: *Combinatorica* 27.4 (2007), pp. 503–507.

[51] L. Elsner and K. P. Hadeler. “Maximizing the spectral radius of a matrix product”. In: *Linear Algebra and its Applications* 469 (2015), pp. 153–168.

[52] S. Enayati and O. Y. Özaltın. “Optimal influenza vaccine distribution with equity”. In: *European Journal of Operational Research* 283.2 (2020), pp. 714–725.

[53] G. M. Engel and H. Schneider. “Matrices diagonally similar to a symmetric matrix”. In: *Linear Algebra and its Applications* 29 (1980), pp. 131–138.

[54] F. Esser and F. Harary. “On the spectrum of a complete multipartite graph”. In: *European Journal of Combinatorics* 1.3 (1980), pp. 211–218.

[55] A. Fall, A. Iggidr, G. Sallet, and J. J. Tewa. “Epidemiological models and Lyapunov functions”. In: *Mathematical Modelling of Natural Phenomena* 2.1 (2007), pp. 62–83.

[56] Z. Feng, A. N. Hill, A. T. Curns, and J. W. Glasser. “Evaluating targeted interventions via meta-population models with multi-level mixing”. In: *Mathematical Biosciences* 287 (2017), pp. 93–104.

[57] Z. Feng, A. N. Hill, P. J. Smith, and J. W. Glasser. “An elaboration of theory about preventing outbreaks in homogeneous populations to include heterogeneity or preferential mixing”. In: *Journal of Theoretical Biology* 386 (2015), pp. 177–187.

[58] Z. Feng, W. Huang, and C. Castillo-Chavez. “Global behavior of a multi-group SIS epidemic model with age structure”. In: *Journal of Differential Equations* 218.2 (2005), pp. 292–324.

[59] F. Fenner, D. A. Henderson, I. Arita, Z. Jezek, I. D. Ladnyi, et al. *Smallpox and its eradication*. Vol. 6. History of International Public Health. World Health Organization, 1988.

[60] P. Fine, K. Eames, and D. L. Heymann. “Herd Immunity: A Rough Guide”. In: *Clinical Infectious Diseases* 52.7 (2011), pp. 911–916.

[61] P. E. M. Fine. “Herd Immunity: History, Theory, Practice”. In: *Epidemiologic Reviews* 15.2 (1993), pp. 265–302.

[62] K.-H. Förster and B. Nagy. “On the Collatz-Wielandt numbers and the local spectral radius of a nonnegative operator”. In: *Linear Algebra and its Applications* 120 (1989), pp. 193–205.

[63] S. Friedland and S. Karlin. “Some inequalities for the spectral radius of non-negative matrices and applications”. In: *Duke Mathematical Journal* 42.3 (1975), pp. 459–490.

[64] S. Friedland. “Convex spectral functions”. In: *Linear and Multilinear Algebra* 9.4 (1981), pp. 299–316.

[65] A. Galeotti and B. W. Rogers. “Strategic Immunization and Group Structure”. In: *American Economic Journal: Microeconomics* 5.2 (2013), pp. 1–32.

[66] A. Ganesh, L. Massoulie, and D. Towsley. “The effect of network topology on the spread of epidemics”. In: *Proceedings IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies.* (Miami, FL, USA). IEEE, 2005.

[67] E. Goldstein, A. Apolloni, B. Lewis, J. C. Miller, M. Macauley, S. Eubank, M. Lipsitch, and J. Wallinga. “Distribution of vaccine/antivirals and the ‘least spread line’ in a stratified population”. In: *Journal of The Royal Society Interface* 7.46 (2010), pp. 755–764.

[68] J. J. Grobler. “Compactness conditions for integral operators in Banach function spaces”. In: *Indagationes Mathmaticae (Proceedings)* 32 (1970), pp. 287–294.

[69] J. J. Grobler. “Spectral theory in Banach lattices”. In: *Operator Theory in Function Spaces and Banach Lattices.* Vol. 75. Oper. Theory Adv. Appl. Birkhäuser, 1995, pp. 133–172.

[70] M. Haber, I. M. Longini, and M. E. Halloran. “Measures of the effects of vaccination in a randomly mixing population”. In: *International Journal of Epidemiology* 20.1 (1991), pp. 300–310.

[71] D. Hadka and P. Reed. “Borg: An Auto-Adaptive Many-Objective Evolutionary Computing Framework”. In: *Evolutionary Computation* 21.2 (2013), pp. 231–259.

[72] C. J. A. Halberg Jr. and A. E. Taylor. “On the spectra of linked operators”. In: *Pacific Journal of Mathematics* 6 (1956), pp. 283–290.

[73] E. Halley. “VI. An estimate of the degrees of the mortality of mankind; drawn from curious tables of the births and funerals at the city of Breslaw; with an attempt to ascertain the price of annuities upon lives”. In: *Philosophical Transactions of the Royal Society of London* 17.196 (1693), pp. 596–610.

[74] M. E. Halloran, I. M. Longini, and C. J. Struchiner. “Design and interpretation of vaccine field studies”. In: *Epidemiologic Reviews* 21.1 (1999), pp. 73–88.

[75] L. Hao, J. W. Glasser, Q. Su, C. Ma, Z. Feng, Z. Yin, J. L. Goodson, N. Wen, C. Fan, H. Yang, L. E. Rodewald, Z. Feng, and H. Wang. “Evaluating vaccination policies to accelerate measles elimination in China: a meta-population modelling study”. In: *International Journal of Epidemiology* 48.4 (2019), pp. 1240–1251.

[76] T. E. Harris. “Contact interactions on a lattice”. In: *The Annals of Probability* 2.6 (1974), pp. 969–988.

[77] D. J. Hartfiel and R. Loewy. “On matrices having equal corresponding principal minors”. In: *Linear Algebra and its Applications* 58 (1984), pp. 147–167.

[78] J. Heesterbeek. “The law of mass-action in epidemiology: a historical perspective”. In: *Ecological Paradigms Lost.* Elsevier, 2005. Chap. 5, pp. 81–105.

[79] H. W. Hethcote and H. R. Thieme. “Stability of the endemic equilibrium in epidemic models with subpopulations”. In: *Mathematical Biosciences* 75.2 (1985), pp. 205–227.

[80] H. W. Hethcote and J. A. Yorke. *Gonorrhea Transmission Dynamics and Control.* Ed. by S. Levin. Vol. 56. Lecture Notes in Biomathematics. Springer-Verlag, 1984.

[81] A. N. Hill and I. M. Longini Jr. "The critical vaccination fraction for heterogeneous epidemic models". In: *Mathematical Biosciences* 181.1 (2003), pp. 85–106.

[82] M. W. Hirsch. "The dynamical systems approach to differential equations". In: *Bulletin of the American Mathematical Society* 11.1 (1984), pp. 1–64.

[83] M. W. Hirsch and H. L. Smith. "Monotone dynamical systems". In: vol. 2. Handbook of Differential Equations: Ordinary Differential Equations. Elsevier, 2006, pp. 239–357.

[84] J. Hladký and I. Rocha. "Independent sets, cliques, and colorings in graphons". In: *European Journal of Combinatorics.* Selected papers of EuroComb17 88 (2020), p. 103108.

[85] P. W. Holland, K. B. Laskey, and S. Leinhardt. "Stochastic blockmodels: First steps". In: *Social Networks* 5.2 (1983), pp. 109–137.

[86] F. Hoppensteadt. "An age dependent epidemic model". In: *Journal of the Franklin Institute* 297.5 (1974), pp. 325–333.

[87] R. A. Horn and C. R. Johnson. *Matrix analysis.* 2nd ed. Cambridge University Press, 2013.

[88] S. Huggett and D. Jordan. *A topological aperitif.* 2nd ed. Springer-Verlag, 2009.

[89] R.-J. Jang-Lewis and H. D. Victory Jr. "On the ideal structure of positive, eventually compact linear operators on Banach lattices". In: *Pacific Journal of Mathematics* 157.1 (1993), pp. 57–85.

[90] S. Janson. *Graphons, cut norm and distance, couplings and rearrangements.* NYJM Monographs. New York Journal of Mathematics, 2013.

[91] O. Kallenberg. *Foundations of modern probability.* 3rd ed. Vol. 99. Probability Theory and Stochastic Modelling. Springer-Verlag, 2021, p. 946.

[92] T. Kato. *Perturbation theory for linear operators.* Vol. 132. Springer-Verlag, 2013.

[93] M. J. Keeling and P. Rohani. *Modeling infectious diseases in humans and animals.* Princeton University Press, 2008.

[94] D. G. Kendall. "Discussion on Professor Bartlett's paper". In: *Journal of the Royal Statistical Society. Series A (General)* 120.1 (1957), pp. 64–67.

[95] W. Kermack and A. McKendrick. "A contribution to the mathematical theory of epidemics". In: *Proceedings of the Royal Society of London.* Series A, Containing Papers of a Mathematical and Physical Character 115.772 (1927), pp. 700–721.

[96] W. Kermack and A. McKendrick. "Contributions to the mathematical theory of epidemics II. The problem of endemicity". In: *Proceedings of the Royal Society of London.* Series A, Containing Papers of a Mathematical and Physical Character 138.834 (1932), pp. 55–83.

[97] W. Kermack and A. McKendrick. "Contributions to the mathematical theory of epidemics III. Further studies of the problem of endemicity". In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 141.843 (1933), pp. 94–122.

[98] B. R. Kloeckner. "Effective perturbation theory for simple isolated eigenvalues of linear operators". In: *Journal of Operator Theory* 81.1 (2019), pp. 175–194.

[99] H. König. *Eigenvalue distribution of compact operators.* Vol. 16. Operator Theory: Advances and Applications. Birkhäuser Verlag, 1986.

[100] D. Kunszenti-Kovács, L. Lovász, and B. Szegedy. "Measures on the square as sparse graph limits". In: *Journal of Combinatorial Theory, Series B* 138 (2019), pp. 1–40.

[101] T. G. Kurtz. "Solutions of ordinary differential equations as limits of pure jump markov processes". In: *Journal of Applied Probability* 7.1 (1970), pp. 49–58.

[102] A. Lajmanovich and J. A. Yorke. “A deterministic model for gonorrhea in a nonhomogeneous population”. In: *Mathematical Biosciences* 28.3 (1976), pp. 221–236.

[103] K. Latrach. “Essential spectra on spaces with the Dunford-Pettis property”. In: *Journal of Mathematical Analysis and Applications* 233.2 (1999), pp. 607–622.

[104] M. Y. Li and J. S. Muldowney. “Global stability for the SEIR model in epidemiology”. In: *Mathematical Biosciences* 125.2 (1995), pp. 155–164.

[105] T. M. Liggett. *Interacting Particle Systems.* Springer-Verlag, 1985.

[106] T. M. Liggett. *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes.* Springer-Verlag, 1999.

[107] X. Lin and J. W.-H. So. “Global stability of the endemic equilibrium and uniform persistence in epidemic models with subpopulations”. In: *The ANZIAM Journal* 34.3 (1993), pp. 282–295.

[108] J. Lindquist, J. Ma, P. van den Driessche, and F. H. Willeboordse. “Effective degree network disease models”. In: *Journal of Mathematical Biology* 62.2 (2011), pp. 143–164.

[109] R. Loewy. “Principal minors and diagonal similarity of matrices”. In: *Linear Algebra and its Applications* 78 (1986), pp. 23–64.

[110] I. M. Longini, E. Ackerman, and L. R. Elveback. “An optimization model for influenza A epidemics”. In: *Mathematical Biosciences* 38.1-2 (1978), pp. 141–157.

[111] L. Lovász. *Large networks and graph limits.* American Mathematical Society colloquium publications 60. American Mathematical Society, 2012.

[112] L. Lovàsz and B. Szegedy. “Limits of dense graph sequences”. In: *Journal of Combinatorial Theory, Series B* 96.6 (2006), pp. 933–957.

[113] G. MacDonald. “The analysis of equilibrium in malaria”. In: *Tropical Diseases Bulletin* 49.9 (1952), pp. 813–829.

[114] I. Marek. “Frobenius theory of positive operators: comparison theorems and applications”. In: *SIAM Journal on Applied Mathematics* 19.3 (1970), pp. 607–628.

[115] A. W. Marshall, I. Olkin, and B. C. Arnold. *Inequalities: Theory of Majorization and Its Applications.* 2nd ed. Springer Series in Statistics. Springer-Verlag, 2011.

[116] L. Matrajt and I. M. Longini. “Critical immune and vaccination thresholds for determining multiple influenza epidemic waves”. In: *Epidemics* 4 (1 2012), pp. 22–32.

[117] K. Miettinen. *Nonlinear multiobjective optimization.* Springer-Verlag, 1998.

[118] D. Mollison. “Possible velocities for a simple epidemic”. In: *Advances in Applied Probability* 4.2 (1972), pp. 233–257.

[119] J. D. Newburgh. “The variation of spectra”. In: *Duke Mathematical Journal* 18.1 (1951), pp. 165–176.

[120] M. E. J. Newman. “Assortative Mixing in Networks”. In: *Physical Review Letters* 89.20 (2002).

[121] C. P. Niculescu and L.-E. Persson. *Convex Functions and Their Applications.* CMS Books in Mathematics. Springer-Verlag, 2006.

[122] R. D. Nussbaum. “Convexity and log convexity for the spectral radius”. In: *Linear Algebra and its Applications* 73 (1986), pp. 59–122.

[123] R. D. Nussbaum. “Eigenvectors of nonlinear positive operators and the linear Krein-Rutman theorem”. In: *Fixed Point Theory* (Sherbrooke, Canada, June 2–21, 1980). Vol. 886. Lecture Notes in Mathematics. Springer-Verlag, 1981, pp. 309–330.

[124] R. D. Nussbaum. “The radius of the essential spectrum”. In: *Duke Mathematical Journal* 37.3 (1970), pp. 473–478.

[125] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani. “Epidemic processes in complex networks”. In: *Reviews of Modern Physics* 87.3 (2015), pp. 925–979.

[126] R. Pastor-Satorras and A. Vespignani. “Immunization of complex networks”. In: *Physical Review E* 65.3 (2002).

[127] G. K. Pedersen. *C^*-algebras and their automorphism groups.* 2nd ed. Pure and Applied Mathematics. London San Diego Cambridge Oxford: Academic Press, 2018.

[128] D. E. Pelinovsky. *Localization in periodic potentials. From Schrödinger operators to the Gross-Pitaevskii equation.* Vol. 390. London Mathematical Society Lecture Note Series. Cambridge University Press, 2011.

[129] M. Penrose. *Random geometric graphs.* Vol. 5. Oxford Studies in Probability. Oxford University Press, 2003.

[130] G. Poghotanyan, Z. Feng, J. W. Glasser, and A. N. Hill. “Constrained minimization problems for the reproduction number in meta-population models”. In: *Journal of Mathematical Biology* 77.6 (2018), pp. 1795–1831.

[131] M. Preziosi. “Effects of pertussis vaccination on transmission: vaccine efficacy for infectiousness”. In: *Vaccine* 21.17-18 (2003), pp. 1853–1861.

[132] J. G. Restrepo, E. Ott, and B. R. Hunt. “Characterizing the Dynamical Importance of Network Nodes and Links”. In: *Physical Review Letters* 97.9 (2006).

[133] S. Ruan. “Spatial-Temporal Dynamics in Nonlocal Epidemiological Models”. In: Biological and Medical Physics, Biomedical Engineering. Springer Berlin Heidelberg, 2007. Chap. 5, pp. 97–122.

[134] S. Ruan and D. Xiao. “Stability of steady states and existence of travelling waves in a vector-disease model”. In: *Proceedings of the Royal Society of Edinburgh: Section A Mathematics* 134 (5 2004), pp. 991–1011.

[135] S. Saha, A. Adiga, B. A. Prakash, and A. K. S. Vullikanti. “Approximation Algorithms for Reducing the Spectral Radius to Control Epidemic Spread”. In: *Proceedings of the 2015 SIAM International Conference on Data Mining*. Society for Industrial and Applied Mathematics, 2015.

[136] H. H. Schaefer. *Banach lattices and positive operators.* Vol. 215. Grundlehren der mathematischen Wissenschaften. Springer-Verlag, 1974.

[137] H. Schaefer. “On the singularities of an analytic function with values in a Banach space”. In: *Archiv der Mathematik* 11.1 (1960), pp. 40–43.

[138] J. Schwartz. “Compact positive mappings in Lebesgue spaces”. In: *Communications on Pure and Applied Mathematics* 14 (1961), pp. 693–705.

[139] J. F. Seward, J. X. Zhang, T. J. Maupin, L. Mascola, and A. O. Jumaan. “Contagiousness of varicella in vaccinated cases: a household contactc study”. In: *Journal of the American Medical Association* 292.6 (2004), pp. 704–708.

[140] E. Shim and A. P. Galvani. “Distinguishing vaccine efficacy and effectiveness”. In: *Vaccine* 30.47 (2012), pp. 6700–6705.

[141] C. E. G. Smith. “Prospects for the control of infectious disease”. In: *Proceedings of the Royal Society of Medicine* 63.11 Pt 2 (1970), pp. 1181–1190.

[142] H. L. Smith. “Cooperative systems of differential equations with concave nonlinearities”. In: *Nonlinear Analysis: Theory, Methods & Applications* 10.10 (1986), pp. 1037–1052.

[143] H. L. Smith. *Monotone Dynamical Systems.* Vol. 41. Mathematical Surveys and Monographs. American Mathematical Society, 1995.

[144] H. L. Smith and H. R. Thieme. “Strongly order preserving semiflows generated by functional-differential equations”. In: *Journal of Differential Equations* 93.2 (1991), pp. 332–363.

[145] J. H. Smith. “Some properties of the spectrum of a graph”. In: *Combinatorial Structures and their Applications.* Proceedings of the Calgary International Conference on Combinatorial Strucutres and their Applications (Calgary, Canada, June 1969). Gordon and Breach, 1970, pp. 403–406.

[146] P. G. Smith, L. C. Rodrigues, and P. E. M. Fine. “Assessment of the protective efficacy of vaccines against common diseases using case-control and cohort studies”. In: *International Journal of Epidemiology* 13.1 (1984), pp. 87–93.

[147] P. G. Smith. “Concepts of herd protection and immunity”. In: *Procedia in Vaccinology* 2.2 (2010), pp. 134–139.

[148] M. Somerville, K. Kumaran, and R. Anderson. *Public Health and Epidemiology at a Glance.* 2nd ed. at a Glance. John Wiley & Sons, 2016.

[149] D. Stevanović, I. Gutman, and M. U. Rehman. “On spectral radius and energy of complete multipartite graphs”. In: *Ars Mathematica Contemporanea* 9.1 (2015), pp. 109–113.

[150] H. R. Thieme. “Global stability of the endemic equilibrium in infinite dimension: Lyapunov functions and positive operators”. In: *Journal of Differential Equations* 250.9 (2011), pp. 3772–3801.

[151] H. R. Thieme. “Local Stability in Epidemic Models for Heterogeneous Populations”. In: *Mathematics in Biology and Medicine.* Proceedings of an International Conference (Bari, Italy, July 18–22, 1983). Vol. 57. Lecture Notes in Biomathematics. Springer-Verlag, 1985, pp. 185–189.

[152] H. R. Thieme. “Spectral Bound and Reproduction Number for Infinite-Dimensional Population Structure and Time Heterogeneity”. In: *SIAM Journal on Applied Mathematics* 70.1 (2009), pp. 188–211.

[153] P. Van Den Driessche and J. Watmough. “Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission”. In: *Mathematical Biosciences* 180.1 (2002), pp. 29–48.

[154] P. Van Mieghem, D. Stevanović, F. Kuipers, C. Li, R. van de Bovenkamp, D. Liu, and H. Wang. “Decreasing the spectral radius of a graph by link removals”. In: *Physical Review E* 84.1 (2011).

[155] H. D. Victory Jr. “On linear integral operators with nonnegative kernels”. In: *Journal of Mathematical Analysis and Applications* 89.2 (1982), pp. 420–441.

[156] R. Vizuete, P. Frasca, and F. Garin. “Graphon-Based Sensitivity Analysis of SIS Epidemics”. In: *IEEE Control Systems Letters* 4.3 (2020), pp. 542–547.

[157] P. Volkmann. “Gewöhnliche Differentialungleichungen mit quasimonoton wachsenden Funktionen in topologischen Vektorräumen”. In: *Mathematische Zeitschrift* 127.2 (1972), pp. 157–164.

[158] K. Yosida and E. Hewitt. “Finitely Additive Measures”. In: *Transactions of the American Mathematical Society* 72.1 (1952), pp. 46–66.

[159] A. C. Zaanen. *Introduction to Operator Theory in Riesz Spaces.* Springer-Verlag, 1997.

[160] A. C. Zaanen. *Linear analysis: measure and integral, Banach and Hilbert space, linear integral equations.* Vol. 2. Bibliotheca Matematica. Interscience Publishers, 1953.

[161] H. Zhao and Z. Feng. “Identifying optimal vaccination strategies via economic and epidemiological modeling”. In: *Journal of Biological Systems* 27.4 (2019), pp. 423–446.