TARGETED VACCINATION STRATEGIES FOR AN INFINITE DIMENSIONAL SIS MODEL

JEAN-FRANÇOIS DELMAS, DYLAN DRONNIER, AND PIERRE-ANDRÉ ZITT

Abstract. We formalize and study the problem of optimal allocation strategies for a (perfect) vaccine in the infinite-dimensional 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 proportion of the infected population in the endemic state. We prove the existence of Pareto optimal strategies, describe the corresponding Pareto frontier in both cases, and study its convexity and stability properties. 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.

1. Introduction

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 [18]. 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 article, we establish a theoretical framework to study targeted vaccination strategies for the deterministic infinite-dimensional SIS model introduced in [7], that encompasses as particular cases the SIS model on graphs or stochastic block models. In the companion papers [8, 9], we provide a series of general and specific examples that complete and illustrate the present work; see Section 1.5 for more detail.

1.2. **Herd immunity and targeted vaccination strategies.** Let us start by recallling a few classical results in mathematical epidemiology; we refer to Keeling and Rohani's book [32] 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 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.

Date: March 17, 2021.

2010 *Mathematics Subject Classification.* 92D30, 47B34, 47A25, 58E17, 34D20.

Key words and phrases. SIS Model, infinite dimensional ODE, kernel operator, vaccination strategy, effective reproduction number, multi-objective optimization, Pareto frontier, maximal independent set.

This work is partially supported by Labex Bézout reference ANR-10-LABX-58.

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 $R_0 < 1$ and invade the population if $R_0 > 1$. For many classical mathematical models of epidemiology, such as SIS or S(E)IR, this intuition can be made rigorous: the quantity R_0 may be computed from the parameters of the model, and the threshold phenomenon occurs.

Assuming $R_0 > 1$ in an homogeneous population, suppose now that only a proportion η^{uni} of the population can catch the disease, the remainder being immunized. An infected individual will now only generate $\eta^{\text{uni}} R_0$ 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 $R_e(\eta^{\text{uni}}) = \eta^{\text{uni}} R_0$. This fact led to the recognition by Smith in 1970 [43] and Dietz in 1975 [11] of a simple threshold theorem: the incidence of an infection declines if the proportion of non-immune individuals is reduced below $\eta^{\text{uni}}_{\text{crit}} = 1/R_0$. 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 [44, 45].

It is of course unrealistic to depict human populations as homogeneous, and many generalizations of the homogeneous model have been studied; see [32, Chapter 3] for examples and further references. For most of these generalizations, it is still possible to define a meaningful reproduction number R_0, as the number of secondary cases generated by a *typical* infectious individual when all other individuals are uninfected; see [10]. 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 $R_e(\eta)$ denote 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 optimization problem:

$$\begin{cases} \textbf{Minimize:} & \text{the quantity of vaccine to administrate} \\ \textbf{subject to:} & \text{herd immunity is reached, that is, } R_e \leq 1. \end{cases}$$

If there is not enough vaccine available, one may also be interested in the closely related problem:

$$\begin{cases} \textbf{Minimize:} & \text{the effective reproduction number } R_e \\ \textbf{subject to:} & \text{a given quantity of available vaccine.} \end{cases}$$

Interestingly enough, the strategy $\eta^{\text{uni}}_{\text{crit}}$ which consists in delivering the vaccine *uniformly* to the population, without taking inhomogeneity into account, leaving a proportion $\eta^{\text{uni}}_{\text{crit}} = 1/R_0$ of the population unprotected is admissible (in fact critical) for the optimization problem (1.2) in the sense that $R_e(\eta^{\text{uni}}_{\text{crit}}) = 1$.

However, herd immunity may be achieved even if the proportion of unprotected people is *greater* than $1/R_0$, by targeting certain group(s) within the population; see Figure 3.3 in [32]. For example, the discussion of vaccination control of gonorrhea in [26, Section 4.5] suggests that it may be better to prioritize the vaccination of people that have already caught the disease: this strategy, guided by the equilibrium state, will be denoted by η^{equi} and will be defined formally below. Let us mention here an observation in the same vein made by Britton, Ball and Trapman in [4]. 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 $1 - 1/R_0$. This can be reformulated in term of targeted vaccination strategies that prioritize the individuals that are more likely to get or stay infected in a S(E)IR epidemic may perfrom better than the uniform allocation of vaccine.

The main goal of this paper is two-fold: formalize the optimization problems (1.2) and (1.2) 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 epidemy 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}.

1.3. **Literature on targeted vaccination strategies.** Targeted vaccination problems have mainly been studied using two different mathematical frameworks.

1.3.1. *On meta-populations models.* Problems (1.2) and (1.2) 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, 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 discussion in [27, Section 10]. 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 R_e.

Problem (1.2) has been studied in this setting by Hill and Longini [27]. These authors study the geometric properties of the so-called threshold hypersurface, that is the vaccination allocations for which $R_e = 1$. They also compute the vaccination belonging to this surface with minimal cost for an Influenza A model. Making structural assumptions on the mixing parameters, Poghotayan, Feng, Glasser and Hill in [40] derive an analytical formula for the solutions of Problem (1.2) in two-groups population and obtained interesting properties in higher dimension such as the convexity of the function R_e. Many papers also contain numerical studies of the optimization problems (1.2) and (1.2) on real-world data using gradient techniques or similar methods; see for example [21, 17, 12, 16, 49].

Finally, the effective reproduction number is not the only reasonable way of quantifying a population's vulnerability to an infection. For a 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 [12, 13] for further discussion on this topic.

1.3.2. *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 [20, 41]. Since

Van Mieghem *et al.* proved in [47] 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 [42] and references therein.

1.4. **Main results.** The differential equations governing the epidemic dynamics in metapopulation SIS models were developed by Lajmanovich and Yorke in their pioneer paper [33]. In [7], we introduced a natural generalization of their equation, which can also be viewed as the limit equation of the stochastic SIS dynamic on network, in an infinite-dimensional 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.

1.4.1. *Regularity of the effective reproduction function R_e.* We define the function R_e in a general operator framework, which we call the kernel model. Let $\mathsf{k} : \Omega \times \Omega \to \mathbb{R}_+$ be some measurable non-negative kernel defined on a probability space $(\Omega, \mathscr{F}, \mu)$ and T_k the corresponding integral operator:

$$T_\mathsf{k}(h)(x) = \int_\Omega \mathsf{k}(x,y)h(y)\,\mu(\mathrm{d}y).$$

In the setting of [7] (see in particular Equation (11) therein), T_k is the so-called *next generation operator*, where the kernel k is defined in terms of a transmission rate kernel $k(x,y)$ and a recovery rate function γ by the product $\mathsf{k}(x,y) = k(x,y)/\gamma(y)$. This setting and the necessary technical assumptions on k and γ are formalized in Assumption 2 on page 12.

Under a technical integrability assumption on the kernel k (formalized on page 11 as Assumption 1), the operator T_k is compact from $L^p(\Omega,\mu)$ to itself for some $p \in (1, +\infty)$. The reproduction number R_0 is then the spectral radius of T_k.

Following [7, 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

$$R_e(\eta) = \rho(T_{\mathsf{k}\eta}),$$

where ρ stands for the spectral radius and $\mathsf{k}\eta$ stands for the kernel $(\mathsf{k}\eta)(x,y) = \mathsf{k}(x,y)\eta(y)$. We will see below how to define similarly the proportion of infected population at the endemic equilibrium; this quantity also depends on the vaccination and will be denoted by $\mathfrak{I}(\eta)$. A vaccination strategy η is called *critical* if it achieves the herd immunity threshold, that is $R_e(\eta) \leq 1$, or equivalently $\mathfrak{I}(\eta) = 0$.

In particular, the "strategy" that consists in vaccinating no one corresponds to $\eta \equiv 1$, and of course $R_e(1) = R_0$. As the spectral radius is positively homogeneous, we also get, when $R_0 \geq 1$, that the uniform strategy that corresponds to the constant function:

$$\eta^{\text{uni}}_{\text{crit}} \equiv \frac{1}{R_0}$$

is critical, as $R_e(\eta^{\text{uni}}_{\text{crit}}) = 1$, and thus achieves the herd immunity threshold. This result is consistent with the homogeneous population results given in Section 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 (1.2) or (1.2) is to prove that the function 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 R_e as the composition of the spectral radius ρ and the map $\eta \mapsto T_{\mathsf{k}\eta}$. The spectral radius is indeed continuous at compact operators if we endow the set of bounded operators with the operator norm topology; see [38, 5]. However, this would require choosing the uniform topology on Δ, which would then not be compact.

We therefore endow Δ with the weak topology, see Section 3.1.1, for which compactness holds; see Lemma 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 [31, p. 431]. However, the family of operators $(T_{\mathsf{k}\eta}, \eta \in \Delta)$ satisfies a nice property called *collective compactness* which enables us to recover continuity, using results by Anselone [1]. This leads to the following result, proved in Theorem 3.5 below. We recall that Assumption 1, formulated on page 11, provides an integrability condition on the kernel k.

Theorem 1.1 (Continuity of the spectral radius)**.** *Under Assumption 1 on the kernel* k*, the function* $R_e : \Delta \to \mathbb{R}_+$ *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 $R_e[\mathsf{k}]$ to stress the dependence of the function R_e in the kernel k. In Proposition 3.6, we prove the stability of R_e, by giving natural sufficient conditions on a sequence of kernels $(\mathsf{k}_n, n \in \mathbb{N})$ converging to k which imply that $R_e[\mathsf{k}_n]$ converges uniformly towards $R_e[\mathsf{k}]$. This result has both theoretical and practical interest: the next-generation operator is unknown in practice, and has to be estimated from data. Thanks to this result, the value of R_e computed from the estimated operator should converge to the true value.

We also prove the convexity of R_e when k is a symmetric positive semi-definite kernel (or even symmetrisable into a positive semi-definite kernel); see Proposition 3.8. This answers partially a conjecture formulated by Hill and Longini in finite dimension [27, Conjecture 8.1]; see Section 1.5 for further results in this direction. Finally, we give in Section 5.5 conditions on two kernels k and k' to define the same function R_e on Δ (in the terminology of linear algebra, this correspond to identify all the preservers of the map $\mathsf{k} \mapsto R_e[\mathsf{k}]$). In finite dimension, these conditions essentially state that the kernels (encoded as matrices in this case) are diagonally similar up to transposition; see Hartfiel and Loewy [24].

1.4.2. *On the maximal endemic equilibrium in the SIS model.* The parameters of the SIS model considered in [7] are the probability state 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 shall $[(\Omega, \mathscr{F}, \mu), k, \gamma]$ for the parameters or $[k, \gamma]$ in a more compact form when there is no ambiguity on the probability state space. We suppose in the following that the technical Assumption 2, formulated on page 12, holds, so that the SIS dynamical evolution may be 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:

$$\partial_t u_t = F(u_t) \quad \text{for } t \in \mathbb{R}_+, \quad \text{where} \quad F(g) = (1-g)\mathcal{T}_k(g) - \gamma g \quad \text{for } g \in \Delta, \tag{1}$$

with an initial condition $u_0 \in \Delta$ and with $\mathcal{T}_k$ the integral operator corresponding to the kernel k acting on the set of bounded measurable functions. It is proved in [7] that such a solution u exists and is unique under Assumption 2 on the model parameters $[(\Omega, \mathscr{F}, \mu), k, \gamma]$. An *equilibrium* of (1) is a function $g \in \Delta$ such that $F(g) = 0$. According to [7], 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 $R_0 \leq 1$ 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; furthermore $\mathfrak{g}$ is the long-time proportion of infected population: $\lim_{t\to+\infty} u_t = \mathfrak{g}$ as soon as the initial condition is non-zero; see [7, Theorem 4.14].

As hinted in [26, 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; see Proposition 7.2. Recall that Assumption 2, formulated page 12, provides technical conditions on the parameters k and γ of the SIS model.

Theorem 1.2 (The maximal equilibrium yields a critical vaccination)**.** *Suppose Assumption 2 holds. If $R_0 \geq 1$, then $R_e(\eta^{\text{equi}}) = 1$.*

Finally, let us describe informally another consequence of Proposition 7.2. We were able to prove in [7, Theorem 4.14] that, in the connected case, if $R_0 > 1$, the disease-free equilibrium $u = 0$ is unstable. Proposition 7.2 gives spectral information on the formal linearization of the dynamics (1) near any equilibrium h; in particular if $h \neq \mathfrak{g}$ then h is linearly unstable.

1.4.3. *Regularity of the total proportion of infected population function $\mathfrak{I}$.* According to [7, Section 5.3.], the SIS equation with vaccination strategy η is given by (1), 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 $\eta \equiv 1$ gives $\mathfrak{g} = \mathfrak{g}_1$), 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:

$$\mathfrak{I}(\eta) = \int_\Omega \mathfrak{g}_\eta\,\eta\,\mathrm{d}\mu = \int_\Omega \mathfrak{g}_\eta(x)\,\eta(x)\,\mu(\mathrm{d}x).$$

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 [12, 13]. 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 R_e. Recall that Δ is endowed with the weak topology of $L^p(\Omega,\mu)$ with $p \in (1,+\infty)$. Assumption 2 on page 12, made about the model parameters $[(\Omega,\mathscr{F},\mu),k,\gamma]$, ensures that the infinite-dimensional SIS model, given by equation (1), is well defined. The next theorem corresponds to Theorem 3.12.

Theorem 1.3 (Continuity of the equilibrium infection size)**.** *Under Assumption 2, the function $\mathfrak{I} : \Delta \to \mathbb{R}_+$ is continuous with respect to the weak topology on Δ.*

In Proposition 3.13, we prove the stability of $\mathfrak{I}$, by giving natural sufficient condition 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 $\mathrm{L} =$

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 3.4 and 3.11.

1.4.4. *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 4.2). We shall also assume that C is continuous with respect to the weak topology on Δ, and that doing nothing costs nothing, that is, $C(1) = 0$. A simple and natural choice is the 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 Section 4.1 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 R_e or $\mathfrak{I}$. Following classical terminology for multi-objective optimisation problems [37], 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:

$$0 = C(1) \leq C(\eta) \leq C(0) = \text{cost of vaccinating the whole population},$$
$$0 = \mathrm{L}(0) \leq \mathrm{L}(\eta) \leq \mathrm{L}(1) = \text{loss incurred in the absence of vaccination}.$$

Let $\mathrm{L}^\star$ be the *optimal loss* function and $C^\star_\mathrm{L}$ the *optimal* cost function defined by:

$$\mathrm{L}^\star(c) = \inf\,\{\,\mathrm{L}(\eta) : \eta \in \Delta,\, C(\eta) \leq c\,\} \quad \text{for } c \in [0, C(0)],$$
$$C^\star_\mathrm{L}(\ell) = \inf\,\{\,C(\eta) : \eta \in \Delta,\, \mathrm{L}(\eta) \leq \ell\,\} \quad \text{for } \ell \in [0, \mathrm{L}(1)].$$

Proposition 4.7 and Theorem 4.8 states that the Pareto frontier is non empty and has a continuous parametrization, and that the functions $\mathrm{L}^\star$ and $C^\star_\mathrm{L}$ are minima and not infima; see Figure 1(b) below for a visualization of the frontier. Assumption 3 on page 21, gives general regularity condition on the cost and loss functions, which are satisfied for the cost C_{uni} and the loss R_e and $\mathfrak{I}$.

Theorem 1.4 (Properties of the Pareto frontier)**.** *Under Assumption 3, the function $C^\star_\mathrm{L}$ is continuous and decreasing on $[0, \mathrm{L}(1)]$, the function $\mathrm{L}^\star$ is continuous on $[0, C(0)]$ decreasing on $[0, C^\star_\mathrm{L}(0)]$ and zero on $[C^\star_\mathrm{L}(0), C(0)]$; furthermore the Pareto frontier is continuous and:*

$$\mathcal{F}_\mathrm{L} = \{(c, \mathrm{L}^\star(c)) : c \in [0, C^\star_\mathrm{L}(0)]\} = \{(C^\star_\mathrm{L}(\ell), \ell) : \ell \in [0, \mathrm{L}(1)]\}.$$

We also establish that $\mathcal{P}_\mathrm{L}$ is compact in Δ for the weak topology in Corollary 4.9; and that the Pareto frontier is convex if C and L are convex in Proposition 4.14. We study in Proposition 4.10 the stability of the Pareto frontier and the set of Pareto optima when the parameters vary.

Section 6 is devoted to the characterization of the quantity $C^\star_{R_e}(0)$, when the cost is the uniform cost, which is the minimal cost which ensures that $R_e = 0$, that is, no infection occurs at all. In the case of a finite state space Ω, and when the support of k is symmetric, this is related to maximal independent sets of the graph with vertices Ω and edges given by the support of k; see Proposition 6.3.

Remark 1.5 (Eradication strategies do not depend on the loss). In [7], we proved that, for all $\eta \in \Delta$, the equilibrium infection size $\mathfrak{I}(\eta)$ is non zero if and only if $R_e(\eta) > 1$. First, this implies that $\mathcal{P}_\mathfrak{I}$ is a subset of $\{\eta \in \Delta : \ R_e(\eta) \geq 1\}$. Secondly, a vaccination strategy $\eta \in \Delta$ is Pareto optimal for the objectives (R_e, C) and satisfies $R_e(\eta) = 1$ if and only if η is Pareto optimal for the objectives $(\mathfrak{I}, C)$ and satisfies $\mathfrak{I}(\eta) = 0$:

$$\eta \in \mathcal{P}_{R_e} \text{ and } R_e(\eta) = 1 \quad \Longleftrightarrow \quad \eta \in \mathcal{P}_\mathfrak{I} \text{ and } \mathfrak{I}(\eta) = 0. \tag{2}$$

Remark 1.6 (Minimal cost of eradication). Assume $R_0 > 1$. The equivalence (2) implies directly that:

$$C^\star_{R_e}(1) = C^\star_\mathfrak{I}(0).$$

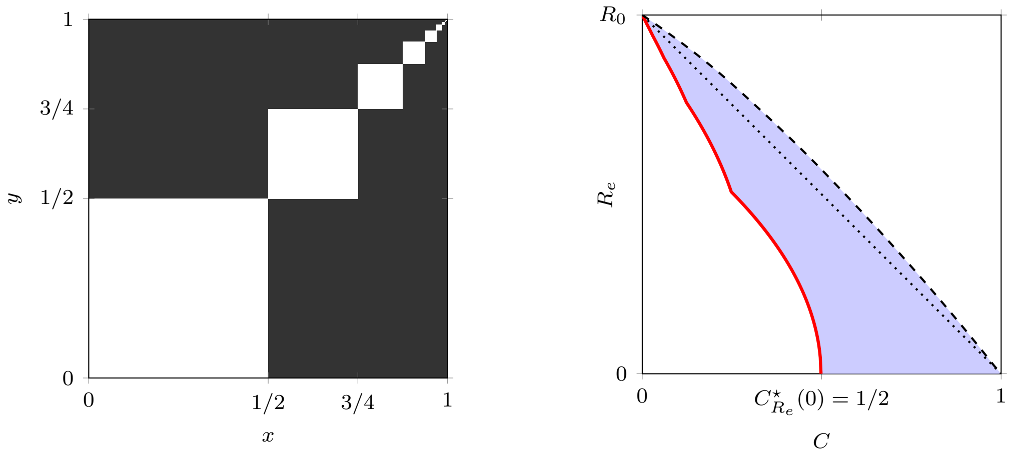
Thus, this latter quantity can be seen as the minimal cost (or minimum percentage of people that shall be vaccinated) required to eradicate the infection. Recall the critical vaccination strategies $\eta^{\text{uni}}_{\text{crit}} \equiv 1/R_0$ and $\eta^{\text{equi}} = 1 - \mathfrak{g}$ (as $R_e(\eta^{\text{uni}}_{\text{crit}}) = R_e(\eta^{\text{equi}}) = 1$). Consider the simple affine cost $C = C_{\text{uni}}$ given by (31). Since $C(\eta^{\text{uni}}_{\text{crit}}) = 1 - 1/R_0$ and $C(\eta^{\text{equi}}) = \int_\Omega \mathfrak{g} \, \mathrm{d}\mu = \mathfrak{I}(1)$, we obtain the following upper bounds of the minimal cost required to eradicate the infection:

$$C^\star_{R_e}(1) = C^\star_\mathfrak{I}(0) \leq \min\left(1 - \frac{1}{R_0}\,, \int_\Omega \mathfrak{g} \, \mathrm{d}\mu\right).$$

Let us illustrate some of our results on an example, which will be discussed in details in a forthcoming companion paper.

Example 1.7 (Multipartite graphon). Graphs that can be colored with ℓ colors, so that no two endpoints of an edge have the same color are known as ℓ-partite graphs. In a biological setting, this corresponds to a population of ℓ groups, such that individuals in a group can not contaminate individuals of the same group. 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 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). We can represent this model by using a 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 k is then given by $\mathsf{k} = \kappa(1 - \sum_{n\in\mathbb{N}^*} \mathbb{1}_{I_n\times I_n})$; it is represented in Figure 1(a).

Consider the loss $\mathrm{L} = R_e$ and the cost $C = C_{\text{uni}}$ giving the overall proportion of vaccinated individuals. Based on the results of [15, 46], we prove in [8] 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 $\deg(x) = \int_\Omega \mathsf{k}(x, y) \, \mu(\mathrm{d}y)$, 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 1(b), the corresponding Pareto frontier is drawn as the solid red line; the colored zone corresponds to all the possible values of $(C(\eta), R_e(\eta))$, where η ranges over Δ; the dotted line corresponds to the outcome of the uniform vaccination strategy $\eta \equiv c$, that is $(C(\eta), R_e(\eta)) = (c, (1-c)R_0)$ where c ranges over $[0, 1]$; and the dashed curve corresponds to the outcome of the worst



(A) Grayplot of the kernel k, with $\Omega = [0, 1]$ and μ the Lebesgue measure (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 dashed line.

FIGURE 1. Example of optimization with $\mathrm{L} = R_e$.

vaccination strategies (for this model, those strategies correspond to the uniform vaccination of the nodes with the updated lower degree; see [8]). We conclude the discussion of this example with a series of remarks:

- The value $C^\star_{R_e}(0) = 1/2$ is the Lebesgue measure of the maximal independent set I_1; see Proposition 6.3.
- The Pareto frontier is not convex, thus the function R_e is not convex on Δ.
- 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 the Pareto frontier. The latter has been computed numerically using the power iteration method. In particular, we obtained the following value: $R_0 \simeq 0.697\kappa$.

1.5. **On the companion papers.** In two companion papers, we illustrate the theoretical framework exposed here and answer through examples some natural questions on the Pareto optima and Pareto frontier.

The paper [9] is dedicated to the complete treatment of the two groups model, $\Omega = \{1, 2\}$, when $\mathrm{L} = R_e$ and $C = C_{\text{uni}}$ is the overall proportion of vaccinated individuals. We also give some partial results when the loss is $\mathrm{L} = \mathfrak{I}$. Despite its apparent simplicity, the derivation of formulae for the Pareto optimal strategies is non trivial. In addition, this model is rich enough to give examples of many different 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 people with highest contacts.* 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.
- *Connectedness of the set of Pareto optima.* The set $\mathcal{P}_{R_e}$ of optimal strategies for the loss R_e may be path-connected and ordered, so that there is a continuous parametrization the Pareto frontier. However, for other parameters, this set is not connected, and there is no continuous (for the topology of the simple convergence and thus the $L^1(\mu)$ topology) parametrization the Pareto frontier.
- *Stability of the set of Pareto optima for the loss function R_e.* The extended Pareto frontier $\{(c, R_e[\mathsf{k}]^\star(c)) : c \in [0, C(0)]\}$ is a continuous function of the next generation matrix k; however the set of Pareto optima presents some discontinuities.
- *Dependence of the Pareto optimum on the choice of the loss function.* For examples where $R_0 > 1$, the optimal strategies for minimizing $\mathfrak{I}$ and R_e may coincide, so that $\mathcal{P}_{\mathfrak{I}} = \mathcal{P}_{R_e} \cap \{\eta \in \Delta : R_e(\eta) \geq 1\}$, or may be entirely different, $\mathcal{P}_{\mathfrak{I}} \cap \mathcal{P}_{R_e} \cap \{\eta \in \Delta : 1 < R_e(\eta) < R_0\} = \emptyset$, depending on the parameters.

In [8], with the uniform cost $C = C_{\text{uni}}$, we study kernels in infinite dimension for which it is possible to derive the Pareto optimal strategies. In some natural examples such as the configuration model (that is $T_{k/\gamma}$ has rank one), which corresponds to proportionate mixing in finite dimensional models, we prove that $\mathcal{P}_{\mathfrak{I}} = \mathcal{P}_{R_e} \cap \{\eta \in \Delta : R_e(\eta) \geq 1\}$; and $C(\eta^{\text{equi}}) \leq C(\eta^{\text{uni}}_{\text{crit}})$, that is: the critical vaccination given by the maximal equilibrium is more efficient than the critical uniform vaccination. This inequality is strict in some cases, and is in the spirit of the observation made in [4] that we discussed at the end of Section 1.3.1. In the configuration model, we prove that the function R_e is convex. We also provide non-trivial infinite dimensional examples for which the critical strategies $\eta^{\text{uni}}_{\text{crit}}$ and η^{equi} are equal and are Pareto optimal for the loss R_e: $C(\eta^{\text{uni}}_{\text{crit}}) = C(\eta^{\text{equi}}) = C^\star_{R_e}(1)$.

1.6. **Structure of the paper.** Section 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 R_e and $\mathfrak{I}$. Their regularity is established in Section 3. We present the multi-objective optimization problem in Section 4 under general condition on the loss function L and cost function C and prove the results on the Pareto frontier. In Section 5, we discuss the equivalent representation of models with different parameters. In particular, we study the transformations on the next-generation operator k that keep the function R_e unchanged, with some more detailed results in the finite dimensional case. Section 6 is devoted to the characterization of the Pareto-optimal vaccination strategies that prevent any contamination and the computation of $C^\star_{R_e}(0)$. Proofs of a few technical results are gathered Section 7.

2. Setting and notation

2.1. **Spaces and operators.** 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 usual its 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 $\int_\Omega fg\,\mathrm{d}\mu$ 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:

$$\rho(T) = \mathrm{rad}(\mathrm{Spec}(T)) = \lim_{n\to\infty} \|T^n\|_E^{1/n}. \tag{3}$$

If E is also a functional space, for $g \in E$, we denote by M_g the multiplication (possibly unbounded) operator defined by $M_g(h) = gh$ for all $h \in E$.

2.2. **Kernels and 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 $f\mathsf{k}g$ the kernel defined by:

$$f\mathsf{k}g : (x, y) \mapsto f(x)\,\mathsf{k}(x, y)g(y). \tag{4}$$

When γ is a positive measurable function defined on Ω, we write k/γ for $\mathsf{k}\gamma^{-1}$, and remark that it may differ from $\gamma^{-1}\mathsf{k}$.

For $p \in (1, +\infty)$, we define the double norm of a signed kernel k by:

$$\|\mathsf{k}\|_{p,q} = \left(\int_\Omega \left(\int_\Omega |\mathsf{k}(x, y)|^q\,\mu(\mathrm{d}y)\right)^{p/q} \mu(\mathrm{d}x)\right)^{1/p} \quad \text{with } q \text{ given by } \quad \frac{1}{p} + \frac{1}{q} = 1. \tag{5}$$

Assumption 1 (On the kernel model $[(\Omega, \mathscr{F}, \mu), \mathsf{k}]$)**.** *Let $(\Omega, \mathscr{F}, \mu)$ be a probability space. The kernel k on Ω has a* finite double-norm*, that is, $\|\mathsf{k}\|_{p,q} < +\infty$ for some $p \in (1, +\infty)$.*

To a kernel k such that $\|\mathsf{k}\|_{p,q} < +\infty$, we associate the positive integral operator T_k on L^p defined by:

$$T_\mathsf{k}(g)(x) = \int_\Omega \mathsf{k}(x, y)g(y)\,\mu(\mathrm{d}y) \quad \text{for } g \in L^p \text{ and } x \in \Omega \tag{6}$$

According to [22, p. 293], T_k is compact. It is well known and easy to check that:

$$\|T_\mathsf{k}\|_{L^p} \leq \|\mathsf{k}\|_{p,q}. \tag{7}$$

For $\eta \in \Delta$, the kernel $\mathsf{k}\eta$ has also a finite double norm on L^p and the operator M_η is bounded, so that the operator $T_{\mathsf{k}\eta} = T_\mathsf{k} M_\eta$ is compact. We can define the effective spectrum function $\mathrm{Spec}[\mathsf{k}]$ from Δ to $\mathscr{K}$ by:

$$\mathrm{Spec}[\mathsf{k}](\eta) = \mathrm{Spec}(T_{\mathsf{k}\eta}), \tag{8}$$

the *effective reproduction number* function $R_e[\mathsf{k}] = \mathrm{rad} \circ \mathrm{Spec}[\mathsf{k}]$ from Δ to $\mathbb{R}_+$ by:

$$R_e[\mathsf{k}](\eta) = \mathrm{rad}(\mathrm{Spec}(T_{\mathsf{k}\eta})) = \rho(T_{\mathsf{k}\eta}), \tag{9}$$

and the corresponding *reproduction number*:

$$R_0[\mathsf{k}] = R_e[\mathsf{k}](1) = \rho(T_\mathsf{k}). \tag{10}$$

Following the framework of [7], for $q \in (1, +\infty)$, we also consider the following norm for the kernel k:

$$\|\mathsf{k}\|_{\infty,q} = \sup_{x\in\Omega} \left(\int_\Omega \mathsf{k}(x,y)^q \, \mu(\mathrm{d}y) \right)^{1/q}.$$

Clearly, we have that $\|\mathsf{k}\|_{\infty,q}$ finite implies that $\|\mathsf{k}\|_{p,q}$ is also finite, with p such that $1/p + 1/q = 1$. When $\|\mathsf{k}\|_{\infty,q} < +\infty$, the corresponding positive bounded linear integral operator $\mathcal{T}_\mathsf{k}$ on $\mathscr{L}^\infty$ is similarly defined by:

$$\mathcal{T}_\mathsf{k}(g)(x) = \int_\Omega \mathsf{k}(x,y) g(y) \, \mu(\mathrm{d}y) \quad \text{for } g \in \mathscr{L}^\infty \text{ and } x \in \Omega. \tag{11}$$

Notice that the integral operators $\mathcal{T}_\mathsf{k}$ and T_k corresponds to the operators T_k and $\hat{T}_\mathsf{k}$ in [7].

According to [7, Lemma 3.7], the operator $\mathcal{T}_\mathsf{k}^2$ on $\mathscr{L}^\infty$ is compact and $\mathcal{T}_\mathsf{k}$ has the same spectral radius as T_k:

$$\rho(\mathcal{T}_\mathsf{k}) = \rho(T_\mathsf{k}). \tag{12}$$

Following [7], we consider the following assumption. Recall that $k/\gamma = k\gamma^{-1}$.

Assumption 2 (On the SIS model $[(\Omega, \mathscr{F}, \mu), k, \gamma]$)**.** *Let $(\Omega, \mathscr{F}, \mu)$ be a probability space. The recovery rate function γ is a function which belongs to $\mathscr{L}^\infty_+$ and the transmission rate kernel k on Ω^2 is such that $\|k/\gamma\|_{\infty,q} < +\infty$ for some $q \in (1, +\infty)$.*

Assumption 2 implies Assumption 1 with $\mathsf{k} = k/\gamma$. Under Assumption 2, we also consider the bounded operators $\mathcal{T}_{k/\gamma}$ on $\mathscr{L}^\infty$, as well as $T_{k/\gamma}$ on L^p, which are the so called *next-generation operator.*

2.3. **Dynamics for the SIS model and equilibria.** The SIS dynamics considered in [7] (under Assumption 2) follows the vector field F defined on $\mathscr{L}^\infty$ by:

$$F(g) = (1 - g)\mathcal{T}_k(g) - \gamma g. \tag{13}$$

More precisely, we consider $u = (u_t, t \in \mathbb{R})$, where $u_t \in \Delta$ for all $t \in \mathbb{R}_+$ such that:

$$\partial_t u_t = F(u_t) \quad \text{for } t \in \mathbb{R}_+, \tag{14}$$

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 [7] that such a solution u exists and is unique.

An *equilibrium* of (14) is a function $g \in \Delta$ such that $F(g) = 0$. According to [7], 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* R_0 associated to the SIS model given by (14) is the spectral radius of the next-generation operator, so that using the definition of the effective reproduction number (9), (12) and (10), this amounts to:

$$R_0 = \rho(\mathcal{T}_{k/\gamma}) = R_0[k/\gamma] = R_e[k/\gamma](1). \tag{15}$$

If $R_0 \leq 1$ (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 $R_0 > 1$ (super-critical case), then 0 is still an equilibrium but different from the maximal equilibrium $\mathfrak{g}$, as $\int_\Omega \mathfrak{g} \, \mathrm{d}\mu > 0$.

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 $f\mathrm{k}g$ from (4). For $\eta \in \Delta$, the kernels $k\eta/\gamma$ and $k\eta$ have finite norm $\|\cdot\|_{\infty,q}$ under Assumption 2, so we can consider the bounded positive operators $\mathcal{T}_{k\eta/\gamma}$ and $\mathcal{T}_{k\eta}$ on $\mathscr{L}^\infty$. According to [7, Section 5.3.], the SIS equation with vaccination strategy η is given by (14), where F is replaced by F_η defined by:

$$F_\eta(g) = (1-g)\mathcal{T}_{k\eta}(g) - \gamma g. \tag{16}$$

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 unvaccinated individual of feature x to be infected at time t. Since the effective reproduction number is the spectral radius of $\mathcal{T}_{k\eta/\gamma}$, we recover (9) as $\rho(\mathcal{T}_{k\eta/\gamma}) = \rho(T_{k\eta/\gamma}) = R_e[k/\gamma](\eta)$. We denote by $\mathfrak{g}_\eta$ the corresponding maximal equilibrium (so that $\mathfrak{g} = \mathfrak{g}_1$). In particular, we have:

$$F_\eta(\mathfrak{g}_\eta) = 0. \tag{17}$$

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:

$$\mathfrak{I}(\eta) = \int_\Omega \mathfrak{g}_\eta \, \eta \, \mathrm{d}\mu = \int_\Omega \mathfrak{g}_\eta(x)\,\eta(x)\,\mu(\mathrm{d}x). \tag{18}$$

We deduce from (16) and (17) that $\mathfrak{g}_\eta \eta = 0, \mu$ a.s. is equivalent to $\mathfrak{g}_\eta = 0$. Applying the results of [7] to the kernel $k\eta$, we deduce that:

$$\mathfrak{I}(\eta) > 0 \iff R_e(\eta) > 1. \tag{19}$$

We conclude this section with a result on the maximal equilibrium $\mathfrak{g}$ that completes what is known from [7]. Notice that, if $R_0 > 1$, then Property (ii) implies that the strategy $1-\mathfrak{g}$ is critical.

Proposition 2.1 (On the maximal equilibrium)**.** *Suppose Assumption 2 holds.*

(i) For any $h \in \Delta$, $h = \mathfrak{g}$ if and only if $F(h) = 0$ and $R_e(1-h) \leq 1$.
(ii) If $\mathfrak{g} \neq 0$, then $R_e(1-\mathfrak{g}) = 1$.

Proposition 2.1 is a consequence of Proposition 7.2 proved in Section 7.

3. Properties of the loss functions

3.1. Preliminaries.

3.1.1. *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:

$$\lim_{n\to\infty} \int_\Omega h g_n \, \mathrm{d}\mu = \int_\Omega h g \, \mathrm{d}\mu. \tag{20}$$

Notice that (20) 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 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 [6, Theorem V.4.2], its unit ball is weakly compact. The set Δ is closed and convex, therefore it is weakly closed; see [6, 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 [6, 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 a sequentially continuous function is continuous. □

3.1.2. *On the continuity of the spectral radius.* We recall some well known facts on operators. Let $(E, \|\cdot\|)$ be a Banach space. Let $A, B \in \mathcal{L}(E)$. According to [39, Appendix A1], we have:

$$\mathrm{Spec}(AB) \cup \{0\} = \mathrm{Spec}(BA) \cup \{0\} \quad \text{and thus} \quad \rho(AB) = \rho(BA). \tag{21}$$

Furthermore, if A is compact then 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. In conclusion, we get:

$$A, B \in \mathcal{L}(E) \text{ and } A \text{ compact} \quad \Longrightarrow \quad \mathrm{Spec}(AB) = \mathrm{Spec}(BA). \tag{22}$$

According to [36, Theorem 4.2], if A, B and $A - B$ are positive operators, then:

$$\rho(A) \geq \rho(B). \tag{23}$$

Let $(E', \|\cdot\|')$ be a Banach space such that E' is continuously and densely embedded in E. Let $T \in \mathcal{L}(E)$ such that $T(E') \subset E'$, and denote by T' its restriction on E' seen as an operator on E'. According to [23, Corollary 1 and Section 6], we have:

$$T \text{ and } T' \text{ compact} \quad \Longrightarrow \quad \mathrm{Spec}(T) = \mathrm{Spec}(T'). \tag{24}$$

We recall some definitions. A sequence $(T_n, n \in \mathbb{N})$ of elements of $\mathcal{L}(E)$ converges strongly to $T \in \mathcal{L}(E)$ if $\lim_{n\to\infty} \|T_n x - Tx\| = 0$ for all $x \in E$. Following [1], 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.

Lemma 3.2 (Continuity of the spectrum and the spectral radius). *Let $(T_n, n \in \mathbb{N})$ and T be elements of $\mathcal{L}(E)$. If one of the two following conditions holds:*

(i) the operator T is compact and $\lim_{n\to\infty} \|T_n - T\|_E = 0$;

(ii) the sequence $(T_n, n \in \mathbb{N})$ is collectively compact and converges strongly to T;

then $\lim_{n\to\infty} \mathrm{Spec}(T_n) = \mathrm{Spec}(T)$ in $(\mathscr{K}, d_{\mathrm{H}})$ and $\lim_{n\to} \rho(T_n) = \rho(T)$.

Notice that if the sequence $(T_n, n \in \mathbb{N})$ is collectively compact and converges strongly to T, then T is compact.

Proof. Since the spectrum of a compact operator is totally disconnected, we deduce from [38, Theorem IV.3] that $\lim_{n\to\infty} \mathrm{Spec}(T_n) = \mathrm{Spec}(T)$ holds under Condition (i). See also the survey [5].

Notice that T compact is a direct consequence of Condition (ii). We deduce from (d) and (e) in [2, Section 3] (see also [1, Theorems 4.8 and 4.16]) that $\lim_{n\to\infty} \mathrm{Spec}(T_n) = \mathrm{Spec}(T)$ holds under Condition (ii). Then use that the function rad is continuous to deduce the convergence of the spectral radius from the convergence of the spectra. □

If k is a kernel on Ω with finite double norm on L^p, recall T_k denotes the corresponding integral compact operator on L^p. The next corollary is a direct consequence of Lemma 3.2 with Condition (i), as the convergence for the finite double norm on L^p implies the convergence in operator norm of the corresponding integral operator; see (7).

Corollary 3.3. *Let $p \in (1, +\infty)$. Let $(\mathsf{k}_n, n \in \mathbb{N})$ and k be kernels on Ω with finite double norms on L^p. If $\lim_{n\to\infty} \|\mathsf{k}_n - \mathsf{k}\|_{p,q} = 0$, then $\lim_{n\to\infty} \mathrm{Spec}(T_{\mathsf{k}_n}) = \mathrm{Spec}(T_\mathsf{k})$ in $(\mathscr{K}, d_\mathrm{H})$ and $\lim_{n\to} \rho(T_{\mathsf{k}_n}) = \rho(T_\mathsf{k})$.*

3.2. **First properties and continuity of the effective reproduction number R_e.** Let k be a kernel on Ω with finite double norm. Recall the effective reproduction number function $R_e[\mathsf{k}]$ defined on Δ by (9): $R_e[\mathsf{k}](\eta) = \rho(T_\mathsf{k} M_\eta)$ and the reproduction number $R_0[\mathsf{k}] = \rho(T_\mathsf{k})$. When there is no risk of confusion on the kernel k, we simply write R_e and R_0 for $R_e[\mathsf{k}]$ and $R_0[\mathsf{k}]$.

Proposition 3.4. *Suppose Assumption 1 holds. The function $R_e = R_e[\mathsf{k}]$ satisfies the following properties:*

- *(i) $R_e(\eta_1) = R_e(\eta_2)$ if $\eta_1 = \eta_2$, μ a.s., and $\eta_1, \eta_2 \in \Delta$,*
- *(ii) $R_e(0) = 0$ and $R_e(1) = R_0$,*
- *(iii) $R_e(\eta_1) \leq R_e(\eta_2)$ for all $\eta_1, \eta_2 \in \Delta$ such that $\eta_1 \leq \eta_2$,*
- *(iv) $R_e(\lambda\eta) = \lambda R_e(\eta)$, for all $\eta \in \Delta$ and $\lambda \in [0, 1]$.*

Proof. If $\eta_1 = \eta_2$ μ-a.s., then we have that $T_{\mathsf{k}\eta_1} = T_{\mathsf{k}\eta_2}$, and thus $R_e(\eta_1) = R_e(\eta_2)$. This gives point (i). Point (ii) is a direct consequence of the definition of 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 $\eta_1 \leq \eta_2 \in \Delta$, then the operator $T_{\mathsf{k}\eta_2} - T_{\mathsf{k}\eta_1}$ is positive; according to (23), we get: $\rho(T_{\mathsf{k}\eta_1}) \leq \rho(T_{k\eta_2})$. This concludes the proof of point (iii). □

We generalize a continuity property on the spectral radius originally stated in [7] by weakening the topology.

Theorem 3.5 (Continuity of $R_e[\mathsf{k}]$ and $\mathrm{Spec}[\mathsf{k}]$)**.** *Suppose Assumption 1 holds. Then, the functions* $\mathrm{Spec}[\mathsf{k}]$ *and* $R_e[\mathsf{k}]$ *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 k takes negative values.

Proof. Let B denote the unit ball in L^p, with $p \in (1, +\infty)$ from Assumption 1. Since the operator T_k is compact, the set $T_\mathsf{k}(B)$ is relatively compact. For all $\eta \in \Delta$, set $\eta B =$

$\{\eta g : g \in B\}$. As $\eta B \subset B$, we deduce that $T_{\mathsf{k}\eta}(B) = T_{\mathsf{k}}(\eta B) \subset T_{\mathsf{k}}(B)$. This implies that the family $(T_{\mathsf{k}\eta}, \eta \in \Delta)$ 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 $(T_{\mathsf{k}\eta_n}(g), n \in \mathbb{N})$ converges μ-a.s. to $T_{\mathsf{k}\eta}(g)$. Consider the function $K(x) = \left(\int_\Omega \mathsf{k}(x,y)^q \, \mu(\mathrm{d}y)\right)^{1/q}$ which belongs to L^p, thanks to (5). Since for all x,

$$|T_{\mathsf{k}\eta_n}(g)(x)| = T_{\mathsf{k}}(|\eta_n g|)(x) \leq K(x) \, \|\eta_n g\|_p \leq K(x) \, \|g\|_p,$$

we deduce, by dominated convergence, that the convergence holds also in L^p:

$$\lim_{n\to\infty} \|T_{\mathsf{k}\eta_n}(g) - T_{\mathsf{k}\eta}(g)\|_p = 0, \tag{25}$$

so that $T_{\mathsf{k}\eta_n}$ converges strongly to $T_{\mathsf{k}\eta}$. Using Lemma 3.2 (ii) (with $T_n = T_{\mathsf{k}\eta_n}$ and $T = T_{\mathsf{k}\eta}$) on the continuity of the spectrum, we get that $\lim_{n\to\infty} \mathrm{Spec}[\mathsf{k}](\eta_n) = \mathrm{Spec}[\mathsf{k}](\eta)$. The function $\mathrm{Spec}[\mathsf{k}]$ is thus sequentially continuous, and, thanks to Lemma 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 $R_e[\mathsf{k}]$ then follows from its definition (3) as the composition of the continuous functions rad and $\mathrm{Spec}[\mathsf{k}]$. □

We complete Corollary 3.3 on the stability property of the spectrum and spectral radius with respect to the kernel k.

Proposition 3.6 (Stability of $R_e[\mathsf{k}]$ and $\mathrm{Spec}[\mathsf{k}]$)**.** *Let $p \in (1, +\infty)$. Let $(\mathsf{k}_n, n \in \mathbb{N})$ and k be kernels on Ω with finite double norms on L^p. If $\lim_{n\to\infty} \|\mathsf{k}_n - \mathsf{k}\|_{p,q} = 0$, then we have:*

$$\lim_{n\to\infty} \sup_{\eta\in\Delta} \left| R_e[\mathsf{k}_n](\eta) - R_e[\mathsf{k}](\eta) \right| = 0 \quad \textit{and} \quad \lim_{n\to\infty} \sup_{\eta\in\Delta} d_{\mathrm{H}}\Big(\mathrm{Spec}[\mathsf{k}_n](\eta), \mathrm{Spec}[\mathsf{k}](\eta) \Big) = 0. \tag{26}$$

Proof. We first prove that $\lim_{n\to\infty} \mathrm{Spec}[\mathsf{k}_n](\eta_n) = \mathrm{Spec}[\mathsf{k}](\eta)$, where the sequence $(\eta_n, n \in \mathbb{N})$ is any sequence in Δ which converges weakly to $\eta \in \Delta$.

The operators $\mathscr{A} = \{T_{\mathsf{k}}\} \cup \{T_{\mathsf{k}_n} : n \in \mathbb{N}\}$ are compact, and we deduce from (7) that:

$$\lim_{n\to\infty} \|T_{\mathsf{k}_n} - T_{\mathsf{k}}\|_{L^p} = 0.$$

The family $\mathscr{A}$ is then easily seen to be collectively compact. This implies, see [1, 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. We deduce the sequence $(T_n = T_{\mathsf{k}_n \eta_n} = T_{\mathsf{k}_n} M_{\eta_n}, n \in \mathbb{N})$ of elements of $\mathscr{A}'$ is collectively compact and that $T = T_{\mathsf{k}\eta} = T_{\mathsf{k}} M_\eta$ is compact.

Let $g \in L^p$. We have:

$$\|T_n(g) - T(g)\|_p \leq \|T_{\mathsf{k}_n} - T_{\mathsf{k}}\|_{L^p} \, \|g\|_p + \|T_{\mathsf{k}\eta_n}(g) - T_{\mathsf{k}\eta}(g)\|_p.$$

Using $\lim_{n\to\infty} \|T_{\mathsf{k}_n} - T_{\mathsf{k}}\|_{L^p} = 0$ and (25), we deduce that $\lim_{n\to\infty} \|T_n(g) - T(g)\|_p$, that is $(T_n, n \in \mathbb{N})$ converges strongly to T. With Lemma 3.2 (ii), we get that $\lim_{n\to\infty} \mathrm{Spec}(T_n) = \mathrm{Spec}(T)$, that is $\lim_{n\to\infty} \mathrm{Spec}[\mathsf{k}_n](\eta_n) = \mathrm{Spec}[\mathsf{k}](\eta)$.

Then, as the function $\eta \mapsto d_{\mathrm{H}}\Big(\mathrm{Spec}[\mathsf{k}_n](\eta), \mathrm{Spec}[\mathsf{k}](\eta) \Big)$ is continuous on the compact set Δ, thanks to Theorem 3.5, 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

$$\begin{aligned}\sup_{\eta\in\Delta} d_{\mathrm{H}}\Big(\,\mathrm{Spec}[\mathsf{k}_n](\eta),\mathrm{Spec}[\mathsf{k}](\eta)\Big)&\\ &= d_{\mathrm{H}}\Big(\,\mathrm{Spec}[\mathsf{k}_n](\eta_n),\mathrm{Spec}[\mathsf{k}](\eta_n)\Big)\\ &\leq d_{\mathrm{H}}\Big(\,\mathrm{Spec}[\mathsf{k}_n](\eta_n),\mathrm{Spec}[\mathsf{k}](\eta)\Big)+d_{\mathrm{H}}\Big(\,\mathrm{Spec}[\mathsf{k}](\eta_n),\mathrm{Spec}[\mathsf{k}](\eta)\Big),\end{aligned}$$

using the continuity of $\mathrm{Spec}[\mathsf{k}]$, we deduce that along this sub-sequence the right hand side converges to 0. Since this results holds for any converging sub-sequence, we get the second part of (26). The first part then follows from the definition (3) of R_e as a composition, and the Lipschitz continuity of the function rad. □

3.3. **A sufficient condition for the convexity of R_e.** 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 $M = DM'D^{-1}$. The matrix M is *diagonally symmetrizable* (resp. *diagonally positive semi-definite*) if it is diagonally similar to a symmetric (resp. symmetric positive semi-definite) matrix, or, equivalently, if M admits a decomposition $M = DS$ (or $M = SD$), where D is a diagonal matrix with positive diagonal entries and S is a symmetric (resp. symmetric positive semi-definite) matrix.

We say a kernel k' is an Hilbert-Schmidt symmetric positive semi-definite kernel if $\|\mathsf{k}'\|_{2,2} < +\infty$ and the corresponding integral operator $T_{\mathsf{k}'}$ on L^2 is symmetric positive definite. Following [48, Example A, p252] we give the analogue of diagonally positive semi-definite for operators.

Definition 3.7 (Diagonally HS positive semi-definite kernel)**.** *A kernel k on Ω is said to be diagonally HS positive semi-definite if there exist an Hilbert-Schmidt symmetric positive semi-definite kernel k' on Ω and two measurable $[0, +\infty)$-valued functions f, g defined on Ω such that $\mathsf{k} = f\mathsf{k}'g$, that is:*

$$\mathsf{k}(x, y) = f(x)\,\mathsf{k}'(x, y)\,g(y) \quad \textit{for all } x, y \in \Omega. \tag{27}$$

Let us mention that in finite dimension (Ω finite), it is possible to take $g = 1$ (or $f = 1$) and thus recover the decomposition $M = DS$ (or $M = SD$). However, this is no more possible in infinite dimension in general because of the integrability condition (5) with $p = q = 2$ on the symmetric kernel.

In [27], Hill and Longini state a conjecture (Conjecture 8.1) that may be generalized in our infinite dimensional setting as follows: if all the eigenvalues of T_{k} are non-negative, is $R_e[\mathsf{k}]$ necessarily convex? While we do not entirely settle this conjecture, we prove that diagonally HS positive semi definite kernels give rise to a convex R_e (Proposition 3.8) and have necessarily non-negative eigenvalues (Lemma 3.9).

Proposition 3.8 (Convexity of R_e)**.** *Suppose that Assumption 1 holds and that k is a diagonally HS positive semi-definite kernel. Then, the function $R_e[\mathsf{k}]$ defined on Δ is convex.*

Proof. The proof relies on an idea which appears in [19] just before Theorem 4.3.

By definition, there exists an Hilbert-Schmidt symmetric positive semi-definite kernel k' on Ω and two measurable $[0, +\infty)$-valued functions f, g defined on Ω such that $\mathsf{k} = f\mathsf{k}'g$. Let T denote the integral operator on L^2 with kernel k'. Recall that for a real-valued function u defined on Ω, M_u denotes the multiplication by u operator.

We first assume that f, g are bounded, and thus belongs to $\mathscr{L}_+^\infty$. By hypothesis, the kernel k has finite double norm on L^p. It has also a finite double norm on L^2; and thus the corresponding integral operator $\tilde{T}_\mathsf{k}$ on L^2 is also compact. We have for $\eta \in \Delta$:

$$R_e[\mathsf{k}](\eta) = \rho(T_\mathsf{k} M_\eta) = \rho(\tilde{T}_\mathsf{k} M_\eta) = \rho(M_f\, T\, M_{g\eta}),$$

where we used (24) (with $E' = L^{p\vee 2}$ and $E = L^{p\wedge 2}$) for the second equality. Since T is a symmetric positive semi-definite operator on L^2, there exists a symmetric positive semi-definite operator Q on L^2 such that $Q^2 = T$. Using (21) twice, we get:

$$R_e[\mathsf{k}](\eta) = \rho\left(M_f\, Q^2\, M_{g\eta}\right) = \rho\left(Q^2\, M_{fg\eta}\right) = \rho\left(Q\, M_{fg\eta}\, Q\right).$$

Since the symmetric operator $Q\, M_{fg\eta}\, Q$ (on L^2) is also positive semi-definite, we deduce from the Courant-Fischer-Weyl min-max principle that:

$$R_e[\mathsf{k}](\eta) = \rho\left(Q\, M_{fg\eta}\, Q\right) = \sup_{u \in L^2 \setminus \{0\}} \frac{\langle u, Q\, M_{fg\eta}\, Qu\rangle}{\langle u, u\rangle}.$$

Since the map $\eta \mapsto \langle u, Q\, M_{fg\eta}\, Qu\rangle$ defined on Δ is linear, we deduce that $\eta \mapsto R_e[\mathsf{k}](\eta)$ is convex as a supremum of linear functions.

We then remove the bound condition on f, g using an approximation scheme. Let k', f, g and $\mathsf{k} = f\mathsf{k}'g$ be as in Definition 3.7. For $n \in \mathbb{N}^*$, set $\mathsf{k}_n = f_n \mathsf{k}' g_n$, where $h_n = h \wedge n$ for $h \in \{f, g\}$. As $\mathsf{k} \geq \mathsf{k}_n$ and $\lim_{n\to\infty} \mathsf{k}_n = \mathsf{k}$ pointwise, we get by dominated convergence that $\lim_{n\to\infty} \|\mathsf{k} - \mathsf{k}_n\|_{p,q} = 0$. We deduce from the first part of the proof that the functions $R_e[\mathsf{k}_n]$, are convex. Then using the first part of (26) in Proposition 3.6, we deduce that $R_e[\mathsf{k}]$ is convex as limit of $R_e[\mathsf{k}_n]$. □

Lemma 3.9. *Suppose that Assumption 1 holds and that* k *is a diagonally HS positive semi-definite kernel. Then, we have* $\mathrm{Spec}(T_\mathsf{k}) \subset \mathbb{R}_+$.

Proof. In a first step, assume that the representation (27) from Definition 3.7 is such that f and g are bounded and bounded away from 0. Then we have with $h = \sqrt{f/g}$ and $\mathsf{k}'' = \sqrt{fg}\, \mathsf{k}' \sqrt{fg}$ that $\mathsf{k} = h\mathsf{k}''h^{-1}$. Since k' is an Hilbert-Schmidt symmetric positive semi-definite kernel and $\sqrt{fg}$ is bounded, we deduce that k'' is also an Hilbert-Schmidt symmetric positive semi-definite kernel. Let T denote the corresponding operator on L^2 with kernel k''. Since T is positive semi-definite, we deduce that $\mathrm{Spec}(T) \subset \mathbb{R}_+$. Notice that k has finite double norm in L^2, and denote by $\tilde{T}_\mathsf{k}$ the corresponding integral operator in L^2. Notice that M_h and $M_{1/h}$ are bounded operators, and that T, and thus $M_h T$, are compact operators. We deduce from (22) (with $A = M_h T$ and $B = M_{1/h}$) that $\mathrm{Spec}(\tilde{T}_\mathsf{k}) = \mathrm{Spec}(T) \subset \mathbb{R}_+$. Then use (24) (with $E' = L^{p\vee 2}$ and $E = L^{p\wedge 2}$) to deduce that $\mathrm{Spec}(T_\mathsf{k}) = \mathrm{Spec}(\tilde{T}_\mathsf{k}) \subset \mathbb{R}_+$.

We then remove the bound condition on f, g using a similar approximation scheme as in the proof of Proposition 3.8. Let k', f, g and $\mathsf{k} = f\mathsf{k}'g$ be as in Definition 3.7. For $n \in \mathbb{N}^*$, set $\mathsf{k}_n = f_n\, (v_n \mathsf{k}' v_n)\, g_n$, where $v_n = \mathbb{1}_{\{f \geq 1/n \text{ and } g \geq 1/n\}}$ and $h_n = n^{-1} \vee (h \wedge n)$ for $h \in \{f, g\}$. As $\mathsf{k} \geq \mathsf{k}_n$ and $\lim_{n\to\infty} \mathsf{k}_n = \mathsf{k}$ pointwise, we get by dominated convergence that $\lim_{n\to\infty} \|\mathsf{k} - \mathsf{k}_n\|_{p,q} = 0$. Finally, using Proposition 3.6, we obtain that $\mathrm{Spec}(T_\mathsf{k}) = \lim_{n\to\infty} \mathrm{Spec}(T_{\mathsf{k}_n})$ is thus a subset of $\mathbb{R}_+$. □

3.4. **Properties of the asymptotic proportion of the infected population $\mathfrak{I}$.** Suppose that Assumption 2 holds. Recall from (18) 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 $F_\eta(h) = 0$. The following preliminary result states informally that, starting from a state higher than this maximal equilibrium, the epidemics must decrease everywhere.

Lemma 3.10. *Let $\eta, g \in \Delta$. If $F_\eta(g) \geq 0$, then we have $g \leq \mathfrak{g}_\eta$.*

Proof. According to [7, Proposition 2.10], the solution u_t of the SIS model with vaccination $\partial_t u_t = F_\eta(u_t)$ and initial condition $u_0 = g$ is non-decreasing since $F_\eta(g) \geq 0$. According to [7, Proposition 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 $g \leq \mathfrak{g}_\eta$. □

We may now state the main properties of the function $\mathfrak{I}$.

Proposition 3.11 (Basic properties of $\mathfrak{I}$)**.** *Suppose that Assumption 2 holds. The function $\mathfrak{I}$ has the following properties:*

(i) $\mathfrak{I}(\eta_1) = \mathfrak{I}(\eta_2)$ if $\eta_1 = \eta_2$ μ-a.s. and $\eta_1, \eta_2 \in \Delta$.
(ii) $\mathfrak{I}(\eta) = 0$ if and only if $R_e(\eta) \leq 1$.
(iii) $\mathfrak{I}(\eta_1) \leq \mathfrak{I}(\eta_2)$ for all $\eta_1, \eta_2 \in \Delta$ such that $\eta_1 \leq \eta_2$.
(iv) $\mathfrak{I}(\lambda\eta) \leq \lambda\mathfrak{I}(\eta)$ for all $\eta \in \Delta$ and $\lambda \in [0, 1]$.

Proof. If $\eta_1 = \eta_2$ μ-a.s., 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 (19).

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 (17), we get $F_{\eta_2}(\mathfrak{g}_{\eta_1}) \geq 0$. By Lemma 3.10 this implies $\mathfrak{g}_{\eta_1} \leq \mathfrak{g}_{\eta_2}$. To sum up,

$$\eta_1 \leq \eta_2 \quad \Longrightarrow \quad \mathfrak{g}_{\eta_1} \leq \mathfrak{g}_{\eta_2}. \tag{28}$$

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 now consider point (iv). Since $\lambda \in [0, 1]$, we deduce from (28) 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 7.

Theorem 3.12 (Continuity of $\mathfrak{I}$)**.** *Suppose that Assumption 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 3.13 (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 2. Assume furthermore there exists $p' \in (1, +\infty)$ such that $\mathsf{k} = \gamma^{-1}k$ and $(\mathsf{k}_n = \gamma_n^{-1}k_n, n \in \mathbb{N})$ have finite double norm in $L^{p'}$ and that $\lim_{n\to\infty} \|\mathsf{k}_n - \mathsf{k}\|_{p',q'} = 0$. Then we have:*

$$\lim_{n\to\infty} \sup_{\eta\in\Delta} \left| \mathfrak{I}[k_n, \gamma_n](\eta) - \mathfrak{I}[k, \gamma](\eta) \right| = 0. \tag{29}$$

4. The optimization problem

4.1. 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 $C(1) = 0$. A simple cost model is the affine cost given by:

$$C_{\text{aff}}(\eta) = \int_\Omega (1 - \eta(x))\, c_{\text{aff}}(x)\, \mu(\mathrm{d}x), \tag{30}$$

  where $c_{\text{aff}}(x)$ is the cost of vaccination of population of feature x, with $c_{\text{aff}} \in L^1$ positive. Without loss of generality we can assume that $\int c_{\text{aff}}\, \mathrm{d}\mu = 1$, so that $C_{\text{aff}}(0) = 1$. The real cost of the vaccination may be a more complicated function $\psi(C_{\text{aff}}(\eta))$, 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. Furthermore, if we assume that c_{aff} is bounded and bounded away from 0 (that is c_{aff} and $1/c_{\text{aff}}$ belongs to $\mathscr{L}^\infty_+$), then replacing the probability measure μ by $\mu_0(\mathrm{d}x) = c_{\text{aff}}(x)\, \mu(\mathrm{d}x)$ and the transmission rate kernel k by $k_0 = k/c_{\text{aff}}$ will not affect the optimal strategies (notice that if Assumption 2 holds for the parameters $[(\Omega, \mathscr{F}, \mu), k, \gamma]$, then it also holds for $[(\Omega, \mathscr{F}, \mu_0), k_0, \gamma]$). Therefore, we will consider in the applications without any real loss of generality the simple affine cost $C = C_{\text{uni}}$ defined by:

$$C_{\text{uni}}(\eta) = \int_\Omega (1 - \eta)\, \mathrm{d}\mu. \tag{31}$$

- ***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 assumptions for L; see below. These assumptions are in particular satisfied if the loss is the effective reproduction number R_e, or if we choose $\mathrm{L} = \mathfrak{I}$ the asymptotic proportion of infected individuals.

We consider the following bi-objective minimization problem on vaccination strategies:

$$\begin{cases} \textbf{Minimize:} & (C(\eta), \mathrm{L}(\eta)) \\ \textbf{subject to:} & \eta \in \Delta. \end{cases}$$

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 that we defined in the introduction.

Definition 4.1 (Pareto optimal strategy). *A strategy $\eta^\star \in \Delta$ is said to be* Pareto optimal *for the bi-objective optimization problem* (4.1) *if no other strategy η is strictly better than $\eta^\star$; in other words, for all $\eta \in \Delta$ the following two conditions hold:*

$$C(\eta) < C(\eta^\star) \implies \mathrm{L}(\eta) > \mathrm{L}(\eta^\star), \tag{32a}$$

$$\mathrm{L}(\eta) < \mathrm{L}(\eta^\star) \implies C(\eta) > C(\eta^\star). \tag{32b}$$

The set of Pareto optimal strategies for the loss L *is denoted by $\mathcal{P}_{\mathrm{L}}$. The* Pareto frontier *is defined as the set of Pareto optimal outcomes:*

$$\mathcal{F}_{\mathrm{L}} = \{(C(\eta^\star), \mathrm{L}(\eta^\star)) :\ \eta^\star \in \mathcal{P}_{\mathrm{L}}\}. \tag{33}$$

From the definition, we see that a Pareto optimal strategy $\eta^\star \in \mathcal{P}_{\mathrm{L}} \subset \Delta$ necessarily solves the following optimization problems, with $\ell = \mathrm{L}(\eta^\star)$ and $c = C(\eta^\star)$:

(34a) **Minimize:** $\mathrm{L}(\eta)$

(34b) **subject to:** $\eta \in \Delta$, $C(\eta) \leq c$,

as well as

(35a) **Minimize:** $C(\eta)$

(35b) **subject to:** $\eta \in \Delta$, $\mathrm{L}(\eta) \leq \ell$.

A converse statement will be given in Corollary 4.8.

4.2. **Assumption and results.** In order to state our assumptions on the cost and the loss functions, let us first define a few terms.

Definition 4.2. *We say that a real-valued function H defined on Δ endowed with the weak topology is:*

- **Weak-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 $\lambda \in [0, 1]$.*

Assumption 3 (On the cost function and loss function)**.** *The cost function C and the loss function L are weak-continuous functions defined on Δ. The cost function is decreasing and normalized such that $C(1) = 0$. The loss function is non-decreasing and sub-homogeneous (in particular $\mathrm{L}(0) = 0$).*

Remark 4.3 (Good cost and loss functions). The cost functions C_{aff} defined by (30) (recall c_{aff} is positive), and thus C_{uni} as well, satisfy Assumption 3. We deduce from Propositions 3.4 and 3.11 and Theorems 3.5 and 3.12 that the functions R_e and $\mathfrak{I}$ are loss functions satisfying Assumption 3.

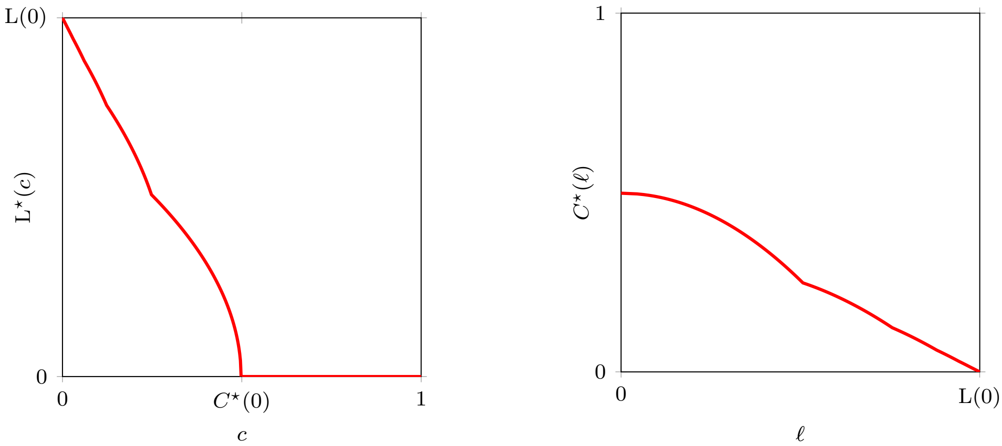
Remark 4.4 (Equivalent vaccinations). Since Δ in endowed with the weak topology, we will consider the set of Pareto optimal vaccination modulo the μ-a.s. equality.

Remark 4.5 (Not vaccinating is Pareto-optimal). As C is decreasing, we deduce from Remark 4.4 that, under Assumption 3, the constant strategy $\eta = 1$ (which consists in no vaccination) is Pareto optimal since it is the unique strategy that minimizes the cost function C. In particular, the set $\mathcal{P}_{\mathrm{L}}$ of Pareto optimal strategies is non-empty.

We have the following fundamental result.

Proposition 4.6 (Optimal solutions for fixed cost or fixed loss)**.** *Suppose that Assumption 3 holds. For any cost $c \in [0, C(0)]$, there exists a minimizer of the loss under the cost constraint $C(\cdot) \leq c$, that is, a solution to Problem* (34)*. Similarly, for any loss $\ell \in [0, \mathrm{L}(1)]$, there exists a minimizer of the cost under the loss constraint $\mathrm{L}(\cdot) \leq \ell$, that is a solution to Problem* (35)*.*

Proof. Let $c \in [0, C(0)]$. The set $\{\eta \in \Delta : C(\eta) \leq c\}$ is non-empty as it contains 1 since $C(1) = 0$. 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



(A) Graph of $\mathrm{L}^\star$. (B) The graph of $C^\star$.

FIGURE 2. Graph of the functions $L^\star$ and $C^\star$ from Example 1.7, with the uniform cost ($C = C_{\text{uni}}$) and the loss $\mathrm{L} = R_e$.

get that L restricted to this compact set reaches its minimum. Thus, Problem (34) has a solution. The proof is similar for the existence of a solution to Problem (35), using that $\mathrm{L}(0) = 0$, which is a consequence of Assumption 3. □

According to Proposition 4.6, we can define the following optimal loss function $\mathrm{L}^\star$ and optimal cost function $C^\star_{\mathrm{L}}$ which are non-increasing functions corresponding to the value functions of Problem (34) and (35) respectively:

$$\mathrm{L}^\star(c) = \min \{ \mathrm{L}(\eta) : \eta \in \Delta, C(\eta) \leq c \} \quad \text{for } c \in [0, C(0)], \tag{36}$$

$$C^\star_{\mathrm{L}}(\ell) = \min \{ C(\eta) : \eta \in \Delta, \mathrm{L}(\eta) \leq \ell \} \quad \text{for } \ell \in [0, \mathrm{L}(1)]. \tag{37}$$

We shall write $C^\star$ for $C^\star_{\mathrm{L}}$ when there is no ambiguity on the loss function L. In Figure 2, we represented the functions $\mathrm{L}^\star$ and $C^\star$ corresponding to the loss $\mathrm{L} = R_e$ and the uniform cost $C = C_{\text{uni}}$ (which satisfy Assumption 3) from Example 1.7.

We now study some regularity properties of $\mathrm{L}^\star$ and $C^\star$ and prove that the constraints in Problems (34) and (35) are binding. Recall that decreasing means strictly decreasing.

Proposition 4.7. *Suppose that Assumption 3 holds. The functions $C^\star$ and $\mathrm{L}^\star$ satisfy the following properties:*

(i) We have $C^\star(0) \in [0, C(0)]$ and $\mathrm{L}^\star(C^\star(0)) = 0$.
(ii) The optimal cost $C^\star$ is continuous decreasing on $[0, \mathrm{L}(1)]$.
(iii) The optimal loss $\mathrm{L}^\star$ is continuous decreasing on $[0, C^\star(0)]$ and 0 on $[C^\star(0), C(0)]$.
(iv) We have $\mathrm{L}^\star \circ C^\star(\ell) = \ell$ for $\ell \in [0, \mathrm{L}(1)]$ and $C^\star \circ \mathrm{L}^\star(c) = c$ for $c \in [0, C^\star(0)]$.

Furthermore, the constraints in Problem (34) and in Problem (35) are binding:

(v) If η solves Problem (34) for some $c \in [0, C^\star(0)]$, then $C(\eta) = c$.
(vi) If η solves Problem (35) for some $\ell \in [0, \mathrm{L}(1)]$, then $\mathrm{L}(\eta) = \ell$.

Proof. The proof of (i) is immediate.

We now prove point (vi) and the fact that $C^\star$ is decreasing on $[0, \mathrm{L}(1)]$, which is part of (ii). Let $0 \leq \ell < \mathrm{L}(1)$ and let $\eta \in \Delta$ be a solution of Problem (35), that is: $\mathrm{L}(\eta) \leq \ell$ and $C(\eta) = C^\star(\ell)$. Let us first note that, since $\mathrm{L}(\eta) \leq \ell < \mathrm{L}(1)$, η cannot be equal to 1 μ-a.s. as L is defined on Δ endowed with the weak topology; we also have $C^\star(\ell) = C(\eta) > C(1) = 0$ as C is decreasing.

For any $\theta \in [0, 1)$, consider the strategy $\eta_\theta = \theta\eta + (1 - \theta)$. Since η is not equal to 1 μ-a.s., we get $\eta_\theta \geq \eta$ and $\int_\Omega \eta_\theta \,\mathrm{d}\mu > \int_\Omega \eta \,\mathrm{d}\mu$. Since C is decreasing, we obtain:

$$C(\eta_\theta) < C(\eta) = C^\star(\ell).$$

As L is weakly-continuous, the map $\phi : \theta \mapsto \mathrm{L}(\eta_\theta)$ is continuous with $\phi(0) = \mathrm{L}(\eta_0) = \mathrm{L}(1)$ and $\phi(1) = \mathrm{L}(\eta) \leq \ell$. By the intermediate value theorem, for any $\ell' \in (\mathrm{L}(\eta), \mathrm{L}(1))$, there exists $\theta \in (0, 1)$ such that $\phi(\theta) = \mathrm{L}(\eta_\theta) = \ell'$. Since η_θ is admissible for Problem (35) with loss constraint ℓ', we deduce that $C^\star(\ell') \leq C(\eta_\theta) < C^\star(\ell)$. This proves that $\mathrm{L}(\eta) = \ell$, and thus (vi) holds for $\ell \in [0, \mathrm{L}(1))$, and that $C^\star$ is decreasing first on $[0, \mathrm{L}(1))$ and then on $[0, \mathrm{L}(1)]$. Since $\eta = 1$ is Pareto optimal (see Remark 4.5), point (vi) also holds for $\ell = \mathrm{L}(1)$.

We prove (v) and the fact that $\mathrm{L}^\star$ is 0 on $[C^\star(0), C(0)]$ and decreasing on $[0, C^\star(0)]$, which is part of (iii). Since $\mathrm{L}^\star$ is non-increasing by definition, (i) implies that $\mathrm{L}^\star$ is 0 on $[C^\star(0), C(0)]$. Let $0 \leq c < C^\star(0)$ and let $\eta \in \Delta$ be a solution of Problem (34), that is: $C(\eta) \leq c$ and $\mathrm{L}(\eta) = \mathrm{L}^\star(c)$. Since $c < C^\star(0)$, we deduce from the definition of $\mathrm{L}^\star$ and $C^\star$ that $\mathrm{L}^\star(c) > 0$. For any $\theta \in [0, 1)$, since the loss is sub-homogeneous, we get:

$$\mathrm{L}(\theta\eta) \leq \theta \mathrm{L}(\eta) = \theta \mathrm{L}^\star(c).$$

As the function C is weakly-continuous, the map $\varphi : \theta \mapsto C(\theta\eta)$ is continuous with $\varphi(0) = C(0)$, and $\varphi(1) = C(\eta) \leq c$. By the intermediate value theorem, for any $c' \in (C(\eta), C(0))$, there exists $\theta \in (0, 1)$ such that $\varphi(\theta) = C(\theta\eta) = c'$. Since $\theta\eta$ is admissible for Problem (34) with cost constraint c', we deduce that $\mathrm{L}^\star(c') \leq \mathrm{L}(\theta\eta) \leq \theta\mathrm{L}^\star(c)$. This gives that $\mathrm{L}^\star(c') < \mathrm{L}^\star(c)$ as $\mathrm{L}^\star(c) > 0$. This proves that $C(\eta) = c$, and thus (v) holds for $c \in [0, C^\star(0))$, and that $\mathrm{L}^\star$ is decreasing first on $[0, C^\star(0))$ and then on $[0, C^\star(0)]$. It is also immediate to check that (v) also holds for $c = C^\star(0)$, thanks to (i).

We now prove the first part of (iv), that is: $\mathrm{L}^\star \circ C^\star(\ell) = \ell$ for $\ell \in [0, \mathrm{L}(1)]$. Let $\ell \in [0, \mathrm{L}(1)]$. Let η be a solution to Problem (35). We have that $C(\eta) = C^\star(\ell)$ and, according to (vi), that $\mathrm{L}(\eta) = \ell$. By definition of $\mathrm{L}^\star$, we deduce that $\mathrm{L}^\star \circ C^\star(\ell) \leq \ell$. Let η' be a solution to Problem (34) with $c = C^\star(\ell)$. Then, we have $\mathrm{L}(\eta') = \mathrm{L}^\star \circ C^\star(\ell) \leq \ell$ and, thanks to (v), we get $C(\eta') = C^\star(\ell)$. Thus η' also solves Problem (35), and thanks to (vi), we have $\mathrm{L}(\eta') = \ell$. This gives $\mathrm{L}^\star \circ C^\star(\ell) = \ell$. The proof of the second part of (iv) is similar and left to the reader. This ends the proof of (iv).

To conclude the proof, it remains to check that $C^\star$ and $\mathrm{L}^\star$ are continuous. The range of the function $C^\star$ is a subset of $[0, C^\star(0)]$. We deduce from (iv) and the definition of L that $[0, L(1)] = \mathrm{L}^\star([0, C^\star(0)])$. Since $\mathrm{L}^\star$ is decreasing on $[0, C^\star(0)]$, we deduce that $\mathrm{L}^\star$ is continuous on $[0, C^\star(0)]$ and thus on $[0, C(0)]$ thanks to (i). The proof of the continuity of $C^\star$ is similar and left to the reader. □

The main result of this section states that all the solutions of the optimization Problems (34) or (35) are Pareto optimal, and gives a description of the Pareto frontier $\mathcal{F}_\mathrm{L}$ as a graph. In Figure 2, we have plotted the Pareto frontier from Example 1.7 which is given in Figure 2(A) by the part of the graph of $L^\star$ restricted to the interval $[0, C^\star(0)]$, or in

Figure 2(B) by the graph of $C^\star$ up to the symmetry with respect to the diagonal, as $L^\star$ is a left-inverse of $C^\star$.

Theorem 4.8 (Pareto optimum and Pareto frontier)**.** *Suppose that Assumption 3 holds. We have:*

$$\begin{aligned}\eta \in \mathcal{P}_{\mathrm{L}} &\iff \eta \text{ solves (34) for some } c \in [0, C^\star(0)]\\ &\iff \eta \text{ solves (35) for some } \ell \in [0, \mathrm{L}(1)].\end{aligned}$$

The Pareto frontier is given by:

$$\mathcal{F}_{\mathrm{L}} = \{(C^\star(\ell), \ell) : \ell \in [0, \mathrm{L}(1)]\} = \{(c, \mathrm{L}^\star(c)) : c \in [0, C^\star(0)]\}. \tag{38}$$

Proof. Let $\ell \in [0, \mathrm{L}(1)]$ and let $\eta \in \Delta$ be solution of Problem (35), that is: $\mathrm{L}(\eta) \leq \ell$ and $C(\eta) = C^\star(\ell)$. Thanks to Proposition 4.7 (vi), we have $\mathrm{L}(\eta) = \ell$. Since $C^\star$ is decreasing, thanks to Proposition 4.7 (ii), if $\eta' \in \Delta$ satisfies $\mathrm{L}(\eta') < \ell$, then we have:

$$C(\eta') \geq C^\star(\mathrm{L}(\eta')) > C^\star(\ell) = C(\eta).$$

Hence, Condition (32b) is satisfied. Condition (32a) is satisfied because η is solution of Problem (35). This proves that η is Pareto optimal. The converse assertion is straightforward and has already been mentioned after Definition 4.1. The proof is similar when $\eta \in \Delta$ is solution of Problem (34) with $c \in [0, C^\star(0)]$. □

We give an immediate consequence of the continuity of the functions $\mathrm{L}^\star$ and $C^\star$.

Corollary 4.9. *Suppose that Assumption 3 holds. The set of Pareto optimal strategies $\mathcal{P}_{\mathrm{L}}$ is compact (for the weak topology).*

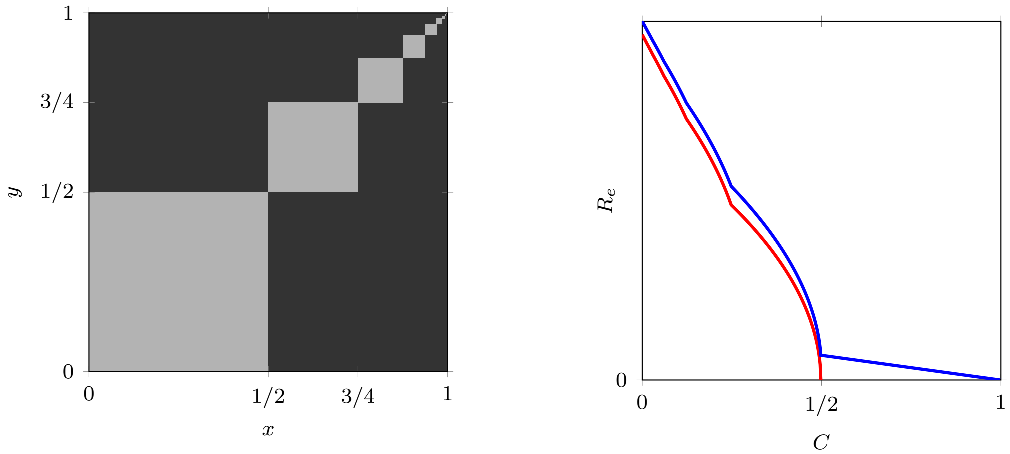
Proof. Since $C^\star$ is continuous, we deduce that $\mathcal{F}_{\mathrm{L}}$, which is given by (38), is compact and thus closed. Since $\mathcal{P}_{\mathrm{L}} = f^{-1}(\mathcal{F}_{\mathrm{L}})$, where the function $f = (C, \mathrm{L})$ defined on Δ is continuous, we deduce that $\mathcal{P}_{\mathrm{L}}$ is closed and thus compact as Δ is compact. □

We can consider the stability of the Pareto frontier and the set of Pareto optima. Recall that, thanks to (38), the graph $\{(c, \mathrm{L}^\star(c)) : c \in [0, C(0)]\}$ of $\mathrm{L}^\star$ is the union of the Pareto frontier and the straight line joining $(0, C^\star(0))$ to $(0, C(0))$ 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.

Proposition 4.10. *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 Assumption 3 holds for the cost C and the loss functions L_n, $n \in \mathbb{N}$, and L. Then $L_n^\star$ converges uniformly to $L^\star$. 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 $\mathrm{L}(\eta) > 0$, then we have $\eta \in \mathcal{P}_{\mathrm{L}}$.*

It might happen that some elements of $\mathcal{P}_{\mathrm{L}}$ are not weak limit of sequence of elements of $\mathcal{P}_{\mathrm{L}_n}$; see [9] for such discontinuity.

Remark 4.11 (On the continuity of the Pareto Frontier). 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 $C^\star_{\mathrm{L}}(0)$. This situation is represented in Figure 3. In Figure 3(A), we have plotted a perturbation $\mathsf{k}_\varepsilon = \mathsf{k} + \varepsilon \sum_{n \in \mathbb{N}^\star} \mathbb{1}_{I_n \times I_n}$ of the multipartite kernel k defined in Example 1.7 for $\varepsilon > 0$ small. According to Proposition 3.6, $R_e[\mathsf{k}_\varepsilon]$ converges uniformly to $R_e[\mathsf{k}]$ when ε vanishes.



(A) Grayplot of the kernel k_ε, with $\Omega = [0, 1]$ and μ the Lebesgue measure (k_ε is equal to the constant $\kappa > 0$ on the black zone and to $\varepsilon\kappa$ on the gray zone, with $\varepsilon > 0$ small).

(B) In red, the Pareto frontier of the kernel k represented in Figure 1(A) compared to the Pareto frontier of the kernel k_ε in blue line.

FIGURE 3. On the stability of the Pareto frontier

However, the Pareto optimal strategies for k_ε that cost more than $1/2$ do not converge to some Pareto optimal strategies for k. This can be seen in Figure 3(B), where the Pareto frontier of k_ε (in blue) corresponding to costs larger than $1/2$ does not have a counterpart in the Pareto frontier of k (in red).

If the cost function is affine, then there is a nice geometric property of the Pareto frontier.

Lemma 4.12. *Suppose that Assumption 3 holds and that the cost function is affine (*i.e. $C = C_{\text{aff}}$ *given by* (30)*). Then, we have* $\mathrm{L}^\star(\theta c + (1-\theta)C(0)) \leq \theta \mathrm{L}^\star(c)$ *for all* $c \in [0, C(0)]$ *and* $\theta \in [0, 1]$.

Remark 4.13. Geometrically, Lemma 4.12 means that the graph of the loss $\mathrm{L}^\star : [0, C(0)] \to [0, \mathrm{L}(1)]$ is below its chords with end point $(1, \mathrm{L}^\star(1)) = (1, 0)$. See Figures 1(B) or for a typical representation of the Pareto frontier (red solid line).

Proof. Let $c \in [0, C(0)]$ and $\theta \in [0, 1]$. Let $\eta \in \mathcal{P}_\mathrm{L}$ with cost $C(\eta) = c$ and thus $\mathrm{L}(\eta) = \mathrm{L}^\star(c)$. We have:

$$C(\theta\eta) = \theta C(\eta) + (1-\theta)C(0) \leq \theta c + (1-\theta)C(0).$$

Therefore, $\theta\eta$ is admissible for Problem (34) with cost constraint $C(\cdot) \leq \theta c + (1-\theta)C(0)$. This implies that $\mathrm{L}^\star(\theta c + (1-\theta)C(0)) \leq \mathrm{L}(\theta\eta) \leq \theta\mathrm{L}^\star(c)$, thanks to the sub-homogeneity of the loss function L. □

In some case, we shall prove that the considered loss function is convex. In this case, choosing a convex cost function implies that the Pareto frontier is convex. We provide a short proof of this result.

Proposition 4.14. *Suppose that Assumption 3 holds. If the cost function* C *and the loss function* L *are convex, then the Pareto frontier is convex, i.e., the function* $C^\star$ *and* $\mathrm{L}^\star$ *are convex.*

Proof. Since, according to Proposition 4.7, the functions $C^\star$ and $\mathrm{L}^\star$ are continuous, decreasing on the segment where they are positive, and $\mathrm{L}^\star$ is the inverse of $C^\star$ on $[0, \mathrm{L}(1)]$, it is enough to prove that $C^\star$ is convex.

Let $\ell_0, \ell_1 \in [0, \mathrm{L}(1)]$. By Corollary 4.8, there exist η_0, η_1 such that $\mathrm{L}(\eta_i) = \ell_i$ 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 $C^\star((1-\theta)\ell_0 + \theta\ell_1) \leq C(\eta) \leq (1-\theta)C^\star(\ell_0) + \theta C^\star(\ell_1)$, and $C^\star$ is convex. □

Remark 4.15. Using Propositions 3.8 and 4.14, we get that, when the next generation operator k is a diagonally HS positive semi-definite kernel (see Definition 3.7) satisfying Assumption 1, then the Pareto frontier for the loss R_e and the cost $C = C_{\text{uni}}$ given by (31) (or the slightly more general cost C_{aff} given by (30)) is convex.

See Example 1.7 where the Pareto frontier $\mathcal{F}_{R_e}$ (and thus R_e) is not convex.

5. Two notions of equivalence between models

5.1. **Motivation.**

The kernel model under Assumption 1, where only the functions R_e is considered, and the SIS model under Assumption 2, where the function R_e and $\mathfrak{I}$ are considered, are completely characterized by their parameters respectively given by $\text{Param} = [(\Omega, \mathscr{F}, \mu), \mathsf{k}]$ and $\text{Param} = [(\Omega, \mathscr{F}, \mu), k, \gamma]$, where $(\Omega, \mathscr{F}, \mu)$ is a probability space, k and k are non-negative kernels on Ω and γ is a non-negative function on Ω. In full generality, the cost function, or its density c_{aff} in the affine case given in (30), should also be considered as a parameter. For simplicity and following the comment at the beginning of Section 4.1, we will only consider the uniform cost C_{uni} given in (31). The loss function $\mathrm{L} = \mathfrak{I}$ is defined only in the SIS model, whereas the loss function $\mathrm{L} = R_e$, which can be considered in the SIS model or the more general kernel model, will implicitly be considered in the more general latter model.

In order to emphasize the dependence of a quantity H on the parameters Param of the model, we shall write $H[\text{Param}]$ for H. For example we write: $\Delta[\text{Param}]$ for the set of functions $\{\eta \in \mathscr{L}^\infty(\Omega, \mathscr{F}) : 1 \geq \eta \geq 0\}$ which clearly depends on the parameters $\text{Param} = [(\Omega, \mathscr{F}, \mu), \mathsf{k}]$ (or $\text{Param} = [(\Omega, \mathscr{F}, \mu), k, \gamma]$); the effective reproduction function $R_e[\text{Param}]$, the set of Pareto optimal strategies $\mathcal{P}_{\mathrm{L}[\text{Param}]}$ for the loss $\mathrm{L} \in \{R_e, \mathfrak{I}\}$, and the corresponding Pareto frontier $\mathcal{F}_{\mathrm{L}[\text{Param}]}$. For example, under Assumption 2, we have the equality of the following functions: $R_e[(\Omega, \mathscr{F}, \mu), k, \gamma] = R_e[(\Omega, \mathscr{F}, \mu), k/\gamma, 1] = R_e[(\Omega, \mathscr{F}, \mu), k/\gamma]$, where for the last equality the left hand-side refers to the SIS model and the right hand-side refers to the kernel model. Using (21), if $\inf \gamma > 0$, then we also have $R_e[(\Omega, \mathscr{F}, \mu), k/\gamma] = R_e[(\Omega, \mathscr{F}, \mu), \gamma^{-1}k]$.

The aim of this section is to provide examples of different set of parameters for which the models are "equivalent" in the intuitive sense that is: the Pareto frontiers (as subset of $\mathbb{R}_+^2$) are the same and it is possible to map nicely the Pareto optimum from one model to the another.

5.2. **On measurability.**

We 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 the push-forward measure, for a measurable function g from $(E', \mathscr{E}')$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, we have :

$$\int_{E'} g \, \mathrm{d}\varphi_\# \nu = \int_E g \circ \varphi \, \mathrm{d}\nu. \tag{39}$$

Let f be a measurable function from $(E, \mathscr{E})$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. We have:

$$\sigma(f) \subset \sigma(\varphi) \implies f = g \circ \varphi, \tag{40}$$

for some measurable function g from $(E', \mathscr{E}')$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

The random variables we shall consider, will be defined on a probability space, say $(\Omega_0, \mathscr{F}_0, \mathbb{P})$. 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 [30, Theorem 8.9], we have for any integrable real-valued random variable X which is $\mathscr{B}$-measurable:

$$\mathscr{A} \text{ and } \mathscr{B} \text{ are conditionally independent given } \mathscr{H} \implies \mathbb{E}\left[X \,|\, \mathscr{A}\right] = \mathbb{E}\left[X \,|\, \mathscr{H}\right]. \tag{41}$$

5.3. **Coupled models.** In this section, we only consider the SIS model; the adaptation the kernel model is immediate. We refer the reader to [29] for a similar developpment in the graphon setting. Let $\mathrm{Param}_i = [(\Omega_i, \mathscr{F}_i, \mu_i), k_i, \gamma_i]$ for $i \in \{1, 2\}$ be two sets of parameters for the SIS model. In particular, Assumption 2 holds for each model. In what follows, we simply write Δ_i the set of functions Δ for the model with parameters 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 5.1 (Coupled models). *Two sets of parameters* Param_1 *and* Param_2 *are* coupled *if there exists two independent 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 such that,* $\mathbb{P}$*-a.s.:*

$$\gamma_1(X_1) = \gamma_2(X_2) \quad \textit{and} \quad k_1(X_1, Y_1) = k_2(X_2, Y_2).$$

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* $\mathbb{P}$*-a.s.:*

$$\mathbb{E}\left[V \,|\, X_i\right] = v_i(X_i) \quad \textit{for } i \in \{1, 2\}.$$

Remark 5.2. In the previous definition, since V is real-valued and $\sigma(X_1, X_2)$-measurable, we deduce from (40) there exits a measurable function v defined on $\Omega_1 \times \Omega_2$ such that $V = v(X_1, X_2)$. And, thus, we get that $\mathbb{P}$-a.s.:

$$\mathbb{E}\left[v(Y_1, Y_2) \,|\, Y_i\right] = v_i(Y_i) \quad \text{for } i \in \{1, 2\}.$$

Remark 5.3. If V is a real-valued integrable $\sigma(X_1) \cap \sigma(X_2)$-measurable random variable, then setting $v_i(X_i) = \mathbb{E}\left[V | X_i\right] = V$, we get that a.s. $v_1(X_1) = v_2(X_2)$, so that v_1 and v_2 are coupled (through V).

Remark 5.4. Let $\eta_1 \in \Delta_1$. According to (40) 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 $V = \eta_1(X_1)$).

Lemma 5.5 (Measurability). *Let* Param_1 *and* Param_2 *be coupled parameters set 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. And for any measurable function* $v : \Omega_1 \to \mathbb{R}$*, such that* $k(X_1, Y_1) v(Y_1)$ *is integrable, the random variable* $\mathbb{E}\left[k(X_1, Y_1) v(Y_1) | X_1\right]$ *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 $\mathbb{E}[k(X_1, Y_1)v(Y_1)|X_1]$ 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 $\mathscr{B} = \sigma(X_1, Y_1)$ are conditionally independent with respect to $\mathscr{H} = \sigma(X_1)$. Using (41), we deduce that:

$$\mathbb{E}[k(X_1, Y_1)v(Y_1)|X_1] = \mathbb{E}[k(X_1, Y_1)v(Y_1)|X_1, X_2].$$

Since $k(X_1, Y_1) = k_2(X_2, Y_2)$ $\mathbb{P}$-a.s., we get:

$$\begin{aligned}\mathbb{E}[k(X_1, Y_1)v(Y_1)|X_1] &= \mathbb{E}[k(X_2, Y_2)v(Y_1)|X_1, X_2] \\ &= \mathbb{E}[k(X_2, Y_2)v(Y_1)|X_2],\end{aligned}$$

where the last equality follows from another application of (41) 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. ☐

The main result of this section state that coupled models have coupled Pareto optimal strategies.

Proposition 5.6 (Coupling and Pareto optimality)**.** *If* Param_1 *and* Param_2 *are coupled and 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.*

Before going into the proof of Proposition 5.6, we derive an important example.

Example 5.7 (Deterministic coupling, model reduction). Let $\mathrm{Param}_1 = [(\Omega_1, \mathscr{F}_1, \mu_1), k_1, \gamma_1]$. Let φ be a measurable function from $(\Omega_1, \mathscr{F}_1)$ to $(\Omega_2, \mathscr{F}_2)$, let X_1 and Y_1 be independent μ_1 distributed random elements of Ω_1, and set $(X_2, Y_2) = (\varphi(X_1), \varphi(Y_1))$. Notice that $\sigma(X_2) \subset \sigma(X_1)$ so that $\sigma(X_1) \cap \sigma(X_2) = \sigma(X_2)$.

Assume that $\sigma(\gamma_1) \subset \sigma(\varphi)$ and $\sigma(k_1) \subset \sigma(\varphi) \otimes \sigma(\varphi)$. This implies 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 (40) 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 $k_1(x, y) = k_2 \circ (\varphi \otimes \varphi)$ that is a.s.:

$$\gamma_1(X_1) = \gamma_2(X_2) \quad \text{and} \quad k_1(X_1, Y_1) = k_2(X_2, Y_2).$$

Let $\mu_2 = \varphi_{\#}\mu_1$ be the push-forward measure of μ_1 by φ. Using (39) it is easy to check that the integrability condition from Assumption 2 is fulfilled, so we can consider the parameter set $\mathrm{Param}_2 = [(\Omega_2, \mathscr{F}_2, \mu_2), k_2, \gamma_2]$. By Definition 5.1, Param_1 is coupled with Param_2 through the (deterministic) coupling π given by the distribution of $(X_1, \varphi(X_1))$.

Let $\eta_1 \in \Delta_1$ be a vaccination strategy. According to Remark 5.4, there exists $\eta_2 \in \Delta_2$ such that η_1 and η_2 are coupled, and:

$$\eta_2(X_2) = \mathbb{E}[\eta_1(X_1)|X_2] = \eta_2(X_2).$$

Notice that $\tilde{\eta}_1 = \eta_2 \circ \varphi = \mathbb{E}[\eta|\,\varphi]$ and η_2 are also coupled, thanks to Remark 5.3 (take $V = \tilde{\eta}_1(X_1) = \eta_2(X_2)$). We deduce from Proposition 5.6, that for $\eta \in \Delta_1$:

$$\eta \text{ is Pareto optimal} \iff \mathbb{E}[\eta|\,\varphi] \text{ is Pareto optimal}\,, \tag{42}$$

where the Pareto optimum are for the parameter set Param_1. Notice we only require that that $\sigma(\gamma_1) \subset \sigma(\varphi)$ and $\sigma(k_1) \subset \sigma(\varphi) \otimes \sigma(\varphi)$.

Remark 5.8. As a consequence of the previous example, if one consider the smallest σ-field $\mathscr{G} \subset \mathscr{F}$ such that γ is $\mathscr{G}$-measurable and k is $\mathscr{G} \otimes \mathscr{G}$ measurable, then $\eta \in \Delta$ for the model $[(\Omega, \mathscr{F}, \mu), k, \gamma]$ is Pareto optimal if and only if $\mathbb{E}[\eta|\mathscr{G}]$ is also Pareto optimal, with $\mathbb{E}$ the expectation corresponding to the probability measure $\mathbb{P} = \mu$ on $(\Omega, \mathscr{F})$. This could be formalized through a coupling between the initial model $[(\Omega, \mathscr{F}, \mu), k, \gamma]$ and the reduced model $[(\Omega, \mathscr{G}, \mu), k', \gamma']$, where $\gamma' = \mathbb{E}[\gamma|\mathscr{G}]$ and $k' = \mathbb{E}[k|\mathscr{G} \otimes \mathscr{G}]$.

The first part of Proposition 5.6 is an elementary consequence of the following key lemma; and the second part is a direct consequence of Remark 5.4. Their proofs are left to the reader. In what follows, we simply write H_i for $H[\mathrm{Param}_i]$ for H the loss functions R_e and $\mathfrak{I}$, the cost function C and the spectrum Spec.

Lemma 5.9. *If* Param_1 *and* Param_2 *are coupled parameters, 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}*,* R_e *or* $\mathfrak{I}$*:*

$$H_1(\eta_1) = H_2(\eta_2). \tag{43}$$

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 5.2:

$$\mathbb{E}\left[\eta(X_1, X_2)|\, X_i\right] = \eta_i(X_i) \quad \text{for } i \in \{1, 2\}. \tag{44}$$

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 ($H = R_e$). Set $\mathsf{k}_i = k_i/\gamma_i$ for $i \in \{1, 2\}$. Let λ be a non-zero eigenvalue of $T_{\mathsf{k}_1\eta_1}$ associated with an eigenvector v_1. Notice that $\mathsf{k}(X_1, Y_1)\eta_1(Y_1)v(Y_1)$ is integrable thanks to the integrability condition from Assumption 2. By definition of eigenvectors, $v_1(X_1)$ is a version of the conditional expectation

$$\lambda^{-1}\mathbb{E}\left[\mathsf{k}_1(X_1, Y_1)\, \eta_1(Y_1)v_1(Y_1)|X_1\right].$$

By Lemma 5.5 applied to the function $v = v_1\eta_1$, 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 (40), there exists v_2 such that a.s. $v_2(X_2) = v_1(X_1)$. Since (Y_1, Y_2) is distributed as (X_1, X_2), we deduce that (44) holds also with (X_1, X_2) replaced by (Y_1, Y_2) and that a.s. $v_2(Y_2) = v_1(Y_1)$. We may now compute:

$$\begin{aligned}
\lambda v_2(X_2) &= \lambda v_1(X_1) \\
&= \mathbb{E}\left[\mathsf{k}_1(X_1, Y_1)\, \eta_1(Y_1)v_1(Y_1)|X_1\right] \\
&= \mathbb{E}\left[\mathsf{k}_1(X_1, Y_1)\, \eta_1(Y_1)v_1(Y_1)|X_2\right] && \text{(Lemma 5.5)} \\
&= \mathbb{E}\left[\mathsf{k}_1(X_1, Y_1)\, \eta(Y_1, Y_2)v_1(Y_1)|X_2\right] && \text{(de-conditioning on } Y_1) \\
&= \mathbb{E}\left[k_2(X_2, Y_2)\, \eta(Y_1, Y_2)v_2(Y_2)|X_2\right] && \text{(a.s. equality)} \\
&= \mathbb{E}\left[k_2(X_2, Y_2)\, \eta_2(Y_2)v_2(Y_2)|X_2\right] && \text{(conditioning on } Y_2) \\
&= T_{\mathsf{k}_2\eta_2}v_2(X_2).
\end{aligned} \tag{45}$$

Since the distribution of X_2 is μ_2, we have μ_2-a.s. that $\lambda v_2(x_2) = T_{\mathsf{k}_2\eta_2} v_2(x_2)$. Therefore λ is also an eigenvalue for $T_{\mathsf{k}_2\eta_2}$. By symmetry we deduce that the spectrum up to $\{0\}$ of $T_{\mathsf{k}_1\eta_1}$ and $T_{\mathsf{k}_2\eta_2}$ 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 (43) with $H = R_e$ and $\eta_1 = \eta_2 = 1$. Let $g_1 = \mathfrak{g}_{\eta_1}$ be the maximal equilibrium for the model Param_1. Since $F_{\eta_1}(g_1) = 0$, see (17), we have:

$$g_1 = \frac{\mathcal{T}_{k_1}(\eta_1 g_1)}{\gamma_1 + \mathcal{T}_{k_1}(\eta_1 g_1)}. \tag{46}$$

By Lemma 5.5, this implies that $g_1(X_1)$ is $\sigma(X_1)\cap\sigma(X_2)$ measurable. Thus, there exists g_2' such that $\mathbb{P}$-a.s. $g_2'(X_2) = g_1(X_1)$. Therefore, by the same computation as in (45), $\mathbb{P}$-a.s.:

$$\mathcal{T}_{k_1}(\eta_1 g_1)(X_1) = \mathcal{T}_{k_2}(\eta_2 g_2')(X_2).$$

We set:

$$g_2 = \frac{\mathcal{T}_{k_2}(\eta_2 g_2')}{\gamma_2 + \mathcal{T}_{k_2}(\eta_2 g_2')}. \tag{47}$$

Then, we deduce from (46) that $\mathbb{P}$-a.s. $g_2(X_2) = g_2'(X_2)$, that is, μ_2-a.s., $g_2 = g_2'$. Thus (47) holds with g_2' replaced by g_2. In other words, g_2 satisfies (17): 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 $\mathbb{P}$-a.s. $g_2(X_2) = g_1(X_1)$ and $g_1(X_1)$ is $\sigma(X_1)\cap\sigma(X_2)$-measurable, we deduce from Remark 5.3, that $(1-g_1)$ and $(1-g_2)$ are coupled, so $R_e[\mathrm{Param}_1](1-g_1) = R_e[\mathrm{Param}_2](1-g_2)$, by Property (43) applied to $H = R_e$. Since $R_0 > 1$ and g_1 is the maximal equilibrium for Param_1, we deduce from Proposition 7.2 that $R_e[\mathrm{Param}_1](1-g_1) = 1$. Using again Proposition 7.2, 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 (43) holds for $H = \mathfrak{I}$, and the proof is complete. □

5.4. **Examples of couplings.** In this section, we consider the kernel model when describing the parameters to fix the notations, but the SIS can be handled in the same way. We denote by Leb the Lebesgue measure.

5.4.1. *Measure preserving function.* This section is motivated by the theory of graphon, which are indistinguishable by measure preserving transformation, see [35, 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 $([0,1], \mathscr{B}([0,1], \mathrm{Leb})$ is measure preserving.

Let φ be measure preserving, and k be an arbitrary kernel with a finite double norm. Let X_1 be with probability distribution μ and let $X_2 = \varphi(X_1)$, so that (X_1, X_2) is a coupling of $(\Omega, \mathscr{F}, \mu)$ and itself. Then for

$$\mathsf{k}_1(x,y) = \mathsf{k}(x,y) \quad \text{and} \quad \mathsf{k}_2(x,y) = \mathsf{k}(\varphi(x), \varphi(y)),$$

the models $\mathrm{Param}_1 = [(\Omega, \mathscr{F}, \mu), \mathsf{k}_1]$ and $\mathrm{Param}_2 = [(\Omega, \mathscr{F}, \mu), \mathsf{k}_2]$ are coupled. Roughly speaking, we can give different label to the features of the population without altering the Pareto frontier.

5.4.2. *Discrete and continuous models.* We now formalize how finite population models can be seen as particular cases of models with a continuous population. Let $\Omega_1 \subset \mathbb{N}$, $\mathscr{F}_1$ the set of subsets of Ω_1 and μ_1 a probability measure on Ω_1. Without loss of generality, we can assume that $\mu_1(\{\ell\}) > 0$ for all $\ell \in \Omega_1$. We set $\Omega_2 = [0,1)$, with $\mathscr{F}_2$ its Borel σ-field and $\mu_2 = \mathrm{Leb}$. Let $(B_\ell, \ell \in \Omega_1)$ be a partition of $[0,1)$ in measurable sets such that $\mathrm{Leb}(B_\ell) = \mu_1(\{\ell\})$ for all $\ell \in \Omega_1$. The measure π on $\Omega_1 \times \Omega_2$ uniquely defined by:

$$\pi(\{\ell\} \times A) = \mathrm{Leb}(B_\ell \cap A)$$

for all measurable $A \subset [0,1)$ is clearly a coupling of μ_1 and μ_2. If the kernels k_1 on Ω_1 and k_2 on Ω_2 are related through the formula

$$\mathsf{k}_2(x,y) = \mathsf{k}_1(\ell, j), \quad \text{for } x \in B_\ell,\ y \in B_j,$$

then $\mathrm{Param}_1 = [(\Omega_1, \mathscr{F}_1, \mu_1), \mathsf{k}_1]$ and $\mathrm{Param}_2 = [([0,1), \mathscr{F}_2, \mathrm{Leb}), \mathsf{k}_2]$ are coupled. Roughly speaking, we can blow up the atomic part of the measure μ_1 into a continuous part, or, conversely, merge all points that behave similarly for k_2 into an atom, without altering the Pareto frontier.

Example 5.10. We consider the so called stochastic block model with 2 populations in the setting of the SIS model in order to stick to the discrete model developed in [33] by Lajmanovich and Yorke.

The discrete SIS model is defined on $\Omega_\mathrm{d} = \{1,2\}$ with the probability measure μ_d defined by $\mu_\mathrm{d}(1) = 1 - \mu_\mathrm{d}(2) = p$ with $p \in (0,1)$, and a kernel k_d and recovery function γ_d given by the matrix and the vector:

$$k_\mathrm{d} = \begin{pmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{pmatrix} \quad \text{and} \quad \gamma_\mathrm{d} = \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix}.$$

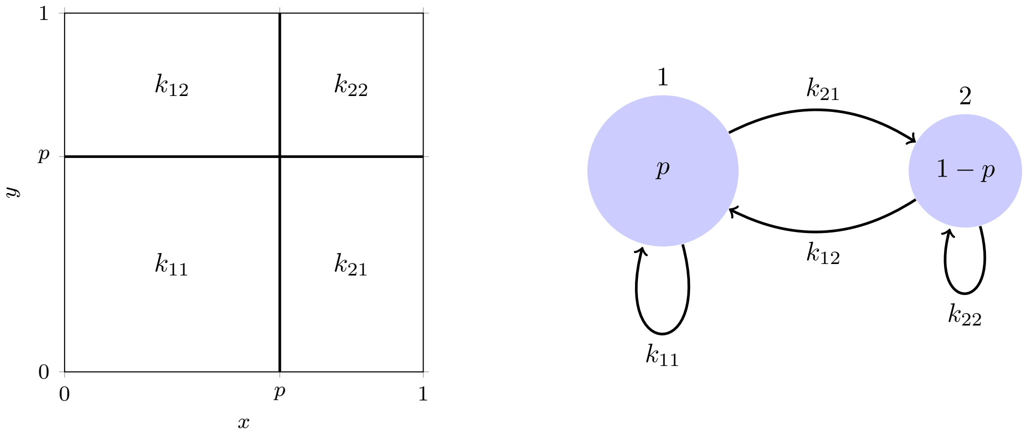
The corresponding model is $\mathrm{Param}_\mathrm{d} = [(\{1,2\}, \mathscr{F}_\mathrm{d}, \mu_\mathrm{d}), k_\mathrm{d}, \gamma_\mathrm{d}]$; see Figure 4(B).

The continuous model is defined on the state space $\Omega_\mathrm{c} = [0,1)$ is endowed with its Borel σ-field, $\mathscr{F}_\mathrm{c}$, and the Lebesgue measure $\mu_\mathrm{c} = \mathrm{Leb}$. 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:

$$k_\mathrm{c}(x,y) = k_{ij} \quad \text{and} \quad \gamma_\mathrm{c}(x) = \gamma_i \quad \text{for } x \in B_i,\ y \in B_j.$$

The corresponding model is $\mathrm{Param}_\mathrm{c} = [([0,1), \mathscr{F}_\mathrm{c}, \mathrm{Leb}), k_\mathrm{c}, \gamma_\mathrm{c}]$; see Figure 4(A). By the general discussion above, the discrete and continuous models are coupled, and in particular they have the same Pareto frontier.

Furthermore, in this simple example, it is easily checked that a discrete vaccination $\eta_\mathrm{d} = (\eta_1, \eta_2)$ and a continuous vaccination $\eta_\mathrm{c} = (\eta_\mathrm{c}(x), x \in [0,1))$ are coupled if and only if



(A) Continuous model: kernel k_{c} on $\Omega_{\mathrm{c}} = [0, 1]$ with the Lebesgue measure

(B) Discrete model: kernel k_{d} on $\Omega_{\mathrm{d}} = \{1, 2\}$ with the measure $p\delta_1 + (1-p)\delta_2$

FIGURE 4. Coupled continuous model (left) and discrete model (right).

there exists a function η defined on $\Omega_{\mathrm{c}} \times \Omega_{\mathrm{d}} = [0, 1) \times \{1, 2\}$ 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 Pareto optimal strategies of the continuous model are easily deduced from the optimal strategies of the discrete model.

To conclude this example, we also give the next-generation matrix K in the setting of [33] which is defined on the discrete model Ω_{d}, 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):

$$K = \begin{pmatrix} \mathsf{k}_{11}\, p & \mathsf{k}_{12}\, (1-p) \\ \mathsf{k}_{21}\, p & \mathsf{k}_{22}\, (1-p) \end{pmatrix} \quad \text{and} \quad K_e(\eta) = \begin{pmatrix} \mathsf{k}_{11}\, p\, \eta_1 & \mathsf{k}_{12}\, (1-p)\, \eta_2 \\ \mathsf{k}_{21}\, p\, \eta_1 & \mathsf{k}_{22}\, (1-p)\, \eta_2 \end{pmatrix},$$

with $\mathsf{k}_{\mathrm{d}} = k_{\mathrm{d}}/\gamma_{\mathrm{d}}$, that is:

$$\mathsf{k}_{\mathrm{d}} = \begin{pmatrix} \mathsf{k}_{11} & \mathsf{k}_{12} \\ \mathsf{k}_{21} & \mathsf{k}_{22} \end{pmatrix} = \begin{pmatrix} k_{11}/\gamma_1 & k_{12}/\gamma_2 \\ k_{21}/\gamma_1 & k_{22}/\gamma_2 \end{pmatrix}.$$

5.5. **Diagonal similarity and transposition.**

5.5.1. *The operator case.* In this section, we consider the kernel model with a given probability state space $(\Omega, \mathscr{F}, \mu)$, and we discuss two operations on the kernel k that leave the functions $\mathrm{Spec}[\mathsf{k}]$ and $R_e[\mathsf{k}]$ defined on Δ, and thus the set of Pareto optima unchanged. Recall the convention (4) for the kernel $f\mathsf{k}g$ defined from the kernel k and the non-negative functions f and g.

Lemma 5.11. *Let k be a kernel on Ω and h be a non-negative measurable function on Ω.*

(i) If $h\mathsf{k}$ and $\mathsf{k}h$ have finite double norm (with possibly different p), then we have:

$$\mathrm{Spec}[h\mathsf{k}] = \mathrm{Spec}[h\mathsf{k}\mathbb{1}_{\{h>0\}}] = \mathrm{Spec}[\mathbb{1}_{\{h>0\}}\mathsf{k}h] = \mathrm{Spec}[\mathsf{k}h]$$
$$R_e[h\mathsf{k}] = R_e[h\mathsf{k}\mathbb{1}_{\{h>0\}}] = R_e[\mathbb{1}_{\{h>0\}}\mathsf{k}h] = R_e[\mathsf{k}h].$$

(ii) If h is positive and if k and $h\mathsf{k}/h$ have finite double norm (with possibly different p), then we have:

$$\mathrm{Spec}[\mathsf{k}] = \mathrm{Spec}[h\mathsf{k}/h] \quad \textit{and} \quad R_e[\mathsf{k}] = R_e[h\mathsf{k}/h].$$

(iii) If k has finite double norm (with $p \in (1, +\infty)$), then its transpose $\mathsf{k}^\top : (x, y) \mapsto \mathsf{k}(y, x)$ has finite double norm (with p replaced by its conjugate q), and we have:

$$\mathrm{Spec}[\mathsf{k}] = \mathrm{Spec}[\mathsf{k}^\top] \quad \textit{and} \quad R_e[\mathsf{k}] = R_e[\mathsf{k}^\top].$$

Even if (ii) is a consequence of (i), we state it separately, since (ii) and (iii) describe two modifications of k that leave the function R_e invariant.

Proof. We give the detailed proof of (ii), in the spirit of the proof of Lemma 3.9, 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 $T_{\mathsf{k}\eta}$ and $T_{h\mathsf{k}\eta/h}$ 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 $T_{\mathsf{k}\eta/h} = T_\mathsf{k}\, M_{\eta/h}$ is compact and (22) for the second equality:

$$\mathrm{Spec}(T_{\mathsf{k}\eta}) = \mathrm{Spec}(T_{\mathsf{k}\eta/h}\, M_h) = \mathrm{Spec}(M_h T_{\mathsf{k}\eta/h}) = \mathrm{Spec}(T_{h\mathsf{k}\eta/h}).$$

Since $\eta \in \Delta$ is arbitrary, this gives that $\mathrm{Spec}[\mathsf{k}] = \mathrm{Spec}[h\mathsf{k}/h]$.

In the general case, we use an approximation scheme. Define the kernel $\mathsf{k}_n = (v_n \mathsf{k} v_n) \wedge n$ with $v_n = \mathbb{1}_{\{h \geq 1/n\}}$ 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 $\mathrm{Spec}[\mathsf{k}_n] = \mathrm{Spec}[\mathsf{k}'_n]$, with $\mathsf{k}'_n = h_n \mathsf{k}_n / h_n$. Since $\|\mathsf{k}\|_{p,q}$ is finite for some $p \in (1, +\infty)$, we get by dominated convergence that $\lim_{n\to\infty} \|\mathsf{k} - \mathsf{k}_n\|_{p,q} = 0$, and we deduce from Proposition 3.6 that $\lim_{n\to\infty} \mathrm{Spec}[\mathsf{k}_n] = \mathrm{Spec}[\mathsf{k}]$. Similarly, setting $\mathsf{k}' = h\mathsf{k}/h$, the norm $\|\mathsf{k}'\|_{p',q'}$ is finite for some $p' \in (1, +\infty)$, and thus $\lim_{n\to\infty} \|\mathsf{k}' - \mathsf{k}'_n\|_{p',q'} = 0$, so that $\lim_{n\to\infty} \mathrm{Spec}[\mathsf{k}'_n] = \mathrm{Spec}[\mathsf{k}']$. This proves that $\mathrm{Spec}[\mathsf{k}] = \mathrm{Spec}[\mathsf{k}']$, and thus (ii).

We now prove (iii). For any $\eta \in \Delta$, the kernel $\mathsf{k}^\top\eta$ defines a bounded integral operator in L^q, whose adjoint is $T_{\eta\mathsf{k}}$. Since the spectrum of an operator and its adjoint are the same, we get $\mathrm{Spec}[\mathsf{k}^\top](\eta) = \mathrm{Spec}(T_{\mathsf{k}^\top\eta}) = \mathrm{Spec}(T_{\eta\mathsf{k}}) = \mathrm{Spec}(M_\eta T_\mathsf{k}) = \mathrm{Spec}(T_\mathsf{k} M_h) = \mathrm{Spec}[\mathsf{k}](\eta)$, where the fourth equality follows once more from (22). Since this is true for any $\eta \in \Delta$, this gives $\mathrm{Spec}[\mathsf{k}^\top] = \mathrm{Spec}[\mathsf{k}]$. □

Remark 5.12. A case of particular interest is the SIS model defined under Assumption 2 where $\inf \gamma > 0$. Indeed, taking $h = 1/\gamma$ in Lemma 5.11 (i), we get the models with parameters $[(\Omega, \mathscr{F}, \mu), k, \gamma]$, $[(\Omega, \mathscr{F}, \mu), k/\gamma, 1]$ and $[(\Omega, \mathscr{F}, \mu), \gamma^{-1}k, 1]$ are equivalent, as far as the optimization of the loss R_e is concerned, as they have the same Pareto optima and Pareto frontier. Notice the operator $T_{\gamma^{-1}k}$, which is the next generation operator for the model $[\gamma^{-1}k, 1]$ appears naturally in the definition of the maximal equilibrium, see (13).

5.5.2. *The matrix case.* It is then natural to ask if the invariance properties stated in Lemma 5.11 describe all possible cases. In other words, does $\mathrm{Spec}[\mathsf{k}] = \mathrm{Spec}[\tilde{\mathsf{k}}]$ or even the weaker condition $R_e[\mathsf{k}] = R_e[\tilde{\mathsf{k}}]$ imply that k and $\tilde{\mathsf{k}}$, or k and $\tilde{\mathsf{k}}^\top$, are diagonally similar? Building on results from [25, 34], 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, let $\Delta = [0,1]^n$. For $\eta \in \Delta$, let $K\eta$ denote the square matrix $K \cdot \mathrm{Diag}(\eta)$, defined by $(K\eta)_{ij} = K_{ij}\eta_j$. We define two maps:

$$\mathrm{Spec}[K] : \Delta \to \mathscr{K} \quad \text{and} \quad R_e[K] : \Delta \to \mathbb{R}_+,$$

where for all $\eta \in \Delta$, $\mathrm{Spec}[K](\eta)$ (resp. $R_e[K](\eta)$) 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 $\{1, ..., n\}$ we denote by $K[\alpha, \beta]$ the submatrix 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. It is elementary to check that the characteristic polynomial of K may be written as:

$$\chi_K(t) = \sum_{k=0}^{n} (-1)^k c_{n-k} t^k, \tag{48}$$

where $c_0 = 1$ and, for $j \geq 1$, c_j is the sum of all principal minors of size j of K.

Definition 5.13. *Let K be a square matrix of size n. A (non-empty) subset α of $\{1, ..., n\}$ is a* clan *if $K[\alpha, \alpha^c]$ and $K[\alpha^c, \alpha]$ have rank at most 1. The matrix K is* clan-free *if there exists no clan.*

Assume that $\alpha = \{1, ..., m\}$ is a clan for K. 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:

$$K = \begin{pmatrix} A & vb^\top \\ cw^\top & D \end{pmatrix} \tag{49}$$

Let us then say that:

$$\tilde{K} = \begin{pmatrix} A^\top & wb^\top \\ cv^\top & D \end{pmatrix} \tag{50}$$

is a *partial transpose* of K (note that the partial transpose is not unique).

Remark 5.14. Such transformations have been considered in the special case where $v = w$ in [34, Lemma 5]; see also [3] 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 $K = D\tilde{K}D^{-1}$. The matrix K is *irreducible* if $K[\alpha, \beta] \neq 0$ for all non-empty subsets α and β. 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: consider the oriented graph $G = (V, E)$ with $V = \{1, \ldots, n\}$ and $ij \in E$, that is ij is an oriented edge of G, if and only if $K(i,j) > 0$. Then K is irreducible if there is an oriented path from i to j for any choice of vertices $i, j \in V$; 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 5.15 (Matrix case)**.** *Let K and $\tilde{K}$ be square matrices of the same size with non-negative entries.*

(i) Assume K is symmetric and $R_e[K] = R_e[\tilde{K}]$. If $\tilde{K}$ is competely reducible then $\tilde{K}$ is diagonally similar to K; if $\tilde{K}$ is symmetric then $\tilde{K} = K$.
(ii) Assume that K is irreducible, clan-free and $R_e[K] = R_e[\tilde{K}]$. Then $\tilde{K}$ is diagonally similar to K or to $K^\top$.
(iii) If K is not clan-free, then $R_e[K] = R_e[\tilde{K}]$ 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 R_e coincide.

Lemma 5.16. *Let K and $\tilde{K}$ be square matrices of the same size with non-negative entries. The following are equivalent:*

(i) The functions $R_e[K]$ and $R_e[\tilde{K}]$ coincide on Δ.
(ii) The functions $R_e[K]$ and $R_e[\tilde{K}]$ 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.

Remark 5.17. Property (ii) from Lemma 5.16 does not imply (i) nor (v) if the entries of the matrices are signed. 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 $R_e[K]$ and $R_e[\tilde{K}]$ clearly coincide on $\mathcal{E}(\Delta) = \{(1,1), (1,0), (0,1)\,(0,0)\}$ even if $R_e[K] \neq R_e[\tilde{K}]$.

According to the the proof of Lemma 5.16, we have that (v) implies (i). Mimicking the proof by induction of (ii) $\Longrightarrow$ (v) from Lemma 5.16, it is not difficult to check that if K and $\tilde{K}$ are square matrices of the same size with at least one non-zero diagonal entry, then (i) implies (v). This result is however not true if all the diagonal entries are zero. Indeed, consider the following two matrices:

$$K = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad \tilde{K} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

We have $R_e[K] = R_e[\tilde{K}]$ on $\Delta = [0,1]^2$, but, even if all the principal minors of size 1 coincide, the principal minor of size two is different.

Proof of Lemma 5.16. 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 multilinearity 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 (48) this implies that $K\eta$ and $\tilde{K}\eta$ have the same spectrum. Thus, (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 n. Let K and $\tilde{K}$ be two matrices of dimension $n+1$, and assume that $R_e[K]$ and $R_e[\tilde{K}]$ coincide on $\mathcal{E}(\Delta)$. For any non-empty $\alpha \subset \{1, ..., n+1\}$, let η_α be the column vector $(\mathbb{1}_\alpha(i), 1 \leq i \leq n+1)$. Notice that for any matrix K':

$$R_e[K'](\eta_\alpha) = \rho(K'\eta_\alpha) = \rho(K'[\alpha]).$$

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:

$$R_e[K'[\alpha]](\tilde{\eta}_\beta) = \rho(K'[\alpha]\tilde{\eta}_\beta) = \rho(K'\eta_\alpha\eta_\beta) = \rho(K'\eta_\beta) = R_e[K'](\eta_\beta). \tag{51}$$

Since the vector η_β is extremal in Δ, we get $R_e[K](\eta_\beta) = R_e[\tilde{K}](\eta_\beta)$ for all $\beta \subset \alpha$. We deduce from (51) that $R_e[K[\alpha]] = R_e[\tilde{K}[\alpha]]$ on the extremal points. By the induction hypothesis the principal minors of $K[\alpha]$ and $\tilde{K}[\alpha]$ are equal, and in particular they have the same determinant. Therefore, 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 (48) that:

$$\chi_K(t) - \det(K) = \chi_{\tilde{K}}(t) - \det(\tilde{K}),$$

Since K and K' have non-negative entries, by Perron-Frobenius theorem, their spectral radius is also an eigenvalue, and thus a root of their characteristic polynomial. We deduce that $\det(K) = \det(\tilde{K})$. This concludes the proof of the induction step. □

Proof of Proposition 5.15. To prove the first two points (i) and (ii), note that thanks to Lemma 5.16, the principal minors of K and $\tilde{K}$ coincide. The results then follow directly from [14, Theorem 3.5], for the symmetric case, and [34, Theorem 1] for the clan-free case.

To prove the last 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 (49) and (50). For any $\lambda \notin \mathrm{Spec}(D)$, using a classical formula for determinants of block matrices,

$$\begin{aligned}
\det(K - \lambda I) &= \det(A - \lambda I - vb^\top(D - \lambda I)^{-1}cw^\top)\det(D - \lambda I)\\
\det(\tilde{K} - \lambda I) &= \det(A^\top - \lambda I - wb^\top(D - \lambda I)^{-1}cv^\top)\det(D - \lambda I)\\
&= \det(A - \lambda I - vc^\top((D - \lambda I)^{-1})^\top bw^\top)\det(D - \lambda I).
\end{aligned}$$

Since $b^\top(D - \lambda I)^{-1}c$ 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 $\lambda \notin \mathrm{Spec}(D)$, 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 $R_e[K]$ and $R_e[\tilde{K}]$ coincide. □

6. A characterization of $C^\star_{R_e}(0)$ when the support of k is symmetric

We have seen in Remark 4.5 that the vaccination strategy $\eta = 1$ (*i.e.* nobody is vaccinated) is Pareto optimal. This is the unique strategy (modulo μ-a.e. equality) that minimizes C on Δ. When the kernel k has a symmetric support, the Pareto optimal strategies which minimize R_e can also be characterized.

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 [28], we generalize this definition to kernels.

Definition 6.1 (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 $\mathsf{k} = 0$ $\mu^{\otimes 2}$-a.s. on $A \times A$. The independence number $\alpha(\mathsf{k})$ of the kernel k is:*

$$\alpha(\mathsf{k}) = \sup\{\mu(A) : \ A \text{ is an independent set of } \mathsf{k}\}.$$

A compactness argument will show that the supremum defining α is reached.

Proposition 6.2 (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 $\mu(A) = \alpha(\mathsf{k})$.*

Proof. First, notice that the independent sets and maximal independent sets of a kernel k depends only on the support $\{\, \mathsf{k} > 0 \,\}$ of k. Therefore, the maximal independent sets of the kernel k and of the kernel $\mathbb{1}_{\{\, \mathsf{k}>0 \,\}}$ are the same. In particular, we can assume without loss of generality that k is bounded.

Let $(A_n,\, n \in \mathbb{N})$ be a sequence of independent sets for k such that:

$$\lim_{n\to\infty} \mu(A_n) = \alpha(\mathsf{k}).$$

Since Δ is compact for the weak topology, 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 operator $\mathcal{T}_\mathsf{k}$ is well defined. We deduce that $\mathcal{T}_\mathsf{k}(\mathbb{1}_{A_n})$ belongs to Δ and converges pointwise towards $\mathcal{T}_\mathsf{k}(g)$. This implies that $\mathbb{1}_{A_n}\mathcal{T}_\mathsf{k}(\mathbb{1}_{A_n})$ converges weakly towards $g\mathcal{T}_\mathsf{k}(g)$. We deduce that:

$$\int_\Omega g\mathcal{T}_\mathsf{k}(g)\,\mathrm{d}\mu = \lim_{n\to\infty}\int_\Omega \mathbb{1}_{A_n}\mathcal{T}_\mathsf{k}(\mathbb{1}_{A_n})\,\mathrm{d}\mu = 0.$$

As $g \in \Delta$, this implies that $\{\, g > 0 \,\}$ is an independent set of k and thus $\mu\,(g > 0) \leq \alpha(\mathsf{k})$. Besides, since $(\mathbb{1}_{A_n}, n \in \mathbb{N})$ converges weakly to g, we get:

$$\int_\Omega g\,\mathrm{d}\mu = \lim_{n\to\infty} \mu(A_n) = \alpha(\mathsf{k}).$$

This implies that $\mu\,(g > 0) \geq \int_\Omega g\,\mathrm{d}\mu = \alpha(\mathsf{k})$. We deduce that $\mu\,(g > 0) = \alpha(\mathsf{k})$, 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 R_e and the cost function C_{uni} given by (31) corresponding to the cost $C^\star_{R_e}(0)$. See Example 1.7 for an example where $C^\star_{R_e}(0) < 1$, that is, where it is possible to prevent infections without vaccinating the whole population.

Proposition 6.3. *Suppose Assumption 1 holds and that* $\{\mathsf{k} > 0\}$*, the support of the kernel* k*, is a symmetric subset of* Ω^2*. We consider the cost* $C = C_{\text{uni}}$ *given by* (31)*. For any maximal independent set* A *of* k*, the strategy* $\mathbb{1}_A$ *is Pareto optimal for the loss* $\mathrm{L} = R_e[\mathsf{k}]$ *and we have:*

$$C^\star_{R_e[\mathsf{k}]}(0) = C(\mathbb{1}_A) = 1 - \alpha(\mathsf{k}). \tag{52}$$

Proof. The existence of a maximum independent set A is given by Proposition 6.2. The basic reproduction number obviously vanishes for the strategy $\mathbb{1}_A$ with cost $1 - \alpha(\mathsf{k})$ as $(T_{\mathsf{k}\mathbb{1}_A})^2 = T_\mathsf{k}\, T_{\mathbb{1}_A \mathsf{k} \mathbb{1}_A} = 0$. Now, let $\eta \in \Delta$ be a strategy such that $R_e[\mathsf{k}](\eta) = 0$. To complete the proof of the proposition, it is enough to prove that $C(\eta) \geq 1 - \alpha(\mathsf{k})$.

Since $R_e[\mathsf{k}](\eta) = 0$, the spectral radius of $T_{\mathsf{k}\eta}$ is equal to 0. Let $\varepsilon > 0$ and consider the kernel k_ε defined on Ω by:

$$\mathsf{k}_\varepsilon(x, y) = \mathbb{1}_{\{\, \mathsf{k}(x,y) > \varepsilon \,\}}.$$

Since $T_{\mathsf{k}\eta} - \varepsilon T_{\mathsf{k}_\varepsilon \eta}$ is a positive operator, we deduce from (23) that $\varepsilon \rho(T_{\mathsf{k}_\varepsilon \eta}) = \rho(\varepsilon T_{\mathsf{k}_\varepsilon \eta}) \leq \rho(T_{\mathsf{k}\eta}) = 0$ and thus $\rho(T_{\mathsf{k}_\varepsilon \eta}) = 0$. Set $\mathsf{k}' = \mathbb{1}_{\{\, \mathsf{k} > 0 \,\}}$. Since $\lim_{\varepsilon \to 0+} \|\mathsf{k}_\varepsilon - \mathsf{k}'\|_{p,q} = 0$, we deduce from Proposition 3.6 on the stability of R_e that $\rho(T_{\mathsf{k}'\eta}) = R_e[\mathsf{k}'](\eta) = \lim_{\varepsilon \to 0+} R_e[\mathsf{k}_\varepsilon](\eta) = \lim_{\varepsilon \to 0+} \rho(T_{\mathsf{k}_\varepsilon \eta}) = 0$. As the support of k is symmetric, we deduce that the kernel k' is symmetric. According to (21), we have:

$$\rho(T_{\mathsf{k}''}) = \rho(T_{\mathsf{k}'\eta}) = 0,$$

with $\mathsf{k}'' = \sqrt{\eta}\, \mathsf{k}' \sqrt{\eta} = \sqrt{\eta}\, \mathbb{1}_{\{\, \mathsf{k} > 0 \,\}} \sqrt{\eta}$. Since the kernel k'' is symmetric, non-negative and bounded by 1, this implies that $\mathsf{k}'' = 0$ $\mathrm{d}\mu^{\otimes 2}$-a.s., and thus $\{\, \eta > 0 \,\}$ is an independent set for k. This gives $\mu\,(\eta > 0) \leq \alpha(\mathsf{k})$. Therefore, we have the following lower bound for the cost $C(\eta)$:

$$C(\eta) = 1 - \int_\Omega \eta \, \mathrm{d}\mu \geq 1 - \mu\,(\eta > 0) \geq 1 - \alpha(\mathsf{k}).$$

This ends the proof of the proposition. □

7. Proofs of Theorem 3.12 and Proposition 3.13, and properties of the maximal equilibrium

For the convenience of the reader, we only use references to the results recalled in [7] for positive operators on Banach spaces. For an operator $\mathcal{T}$, we denote by $\mathcal{T}^\top$ its adjoint. We first give a preliminary lemma.

Lemma 7.1. *Let* $\mathcal{T}$ *be a positive bounded operator on* $\mathscr{L}^\infty$*. If there exists* $g \in \mathscr{L}^\infty_+$*, with* $\int_\Omega g \, \mathrm{d}\mu > 0$ *and* $\lambda > 0$ *satisfying:*

$$\mathcal{T}(g)(x) > \lambda g(x), \quad \textit{for all} \quad x \textit{ such that } g(x) > 0,$$

then we have $\rho(\mathcal{T}) > \lambda$.

Proof. Let $A = \{\, g > 0 \,\}$ be the support of the function g. Let $\mathcal{T}'$ be the bounded operator defined by $\mathcal{T}'(f) = \mathbb{1}_A \mathcal{T}(\mathbb{1}_A f)$. Since $\mathcal{T}'(g) = \mathbb{1}_A \mathcal{T}(\mathbb{1}_A g) = \mathbb{1}_A \mathcal{T}(g) > \lambda g$, we deduce from the Collatz-Wielandt formula, see [7, Proposition 3.6], that $\rho(\mathcal{T}') \geq \lambda > 0$. According to [7, Lemma 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 $\mathcal{T}'$, that is such that $(\mathcal{T}')^\top(v) = \rho(\mathcal{T}')v$. 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:

$$(\rho(\mathcal{T}') - \lambda) \, \langle v, g \rangle = \langle v, \mathcal{T}'(g) - \lambda g \rangle > 0.$$

This implies that $\rho(\mathcal{T}') > \lambda$. Since $\mathcal{T} - \mathcal{T}'$ is a positive operator, we deduce from (23) that $\rho(\mathcal{T}) \geq \rho(\mathcal{T}') > \lambda$. □

We now state an interesting result on the characterization of the maximal equilibrium. We keep notations from Sections 2.3 and 2.4 and write R_e for $R_e[k/\gamma]$. Recall that $R_0 = R_e(1)$. 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:

$$DF[h](g) = (1-h)\mathcal{T}_k(g) - (\gamma + \mathcal{T}_k(h))g \quad \text{for } h, g \in \mathscr{L}^\infty.$$

Recall $s(A)$ denotes the spectral bound of the bounded operator A, see (33) in [7].

Proposition 7.2. *Suppose Assumption 2 holds and write R_e for $R_e[k/\gamma]$. 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)* $R_e((1-h)^2) \leq 1$.
- *(iv)* $R_e(1-h) \leq 1$.

We also have: $\mathfrak{g} = 0 \Longleftrightarrow R_0 \leq 1$; *as well as:* $\mathfrak{g} \neq 0 \Longrightarrow R_e(1-\mathfrak{g}) = 1$.

Proof. Let $h \in \Delta$ be an equilibrium, that is $F(h) = 0$.

Let us show the equivalence between (ii) and (iii). According to [7, Proposition 4.2], $s(DF[h]) \leq 0$ if and only if:

$$\rho(\mathcal{T}_{\mathsf{k}}) \leq 1 \quad \text{with} \quad \mathsf{k}(x,y) = (1-h(x))\frac{k(x,y)}{\gamma(y) + \mathcal{T}_k(h)(y)}.$$

Since $F(h) = 0$, we have $(1-h)/\gamma = 1/(\gamma + \mathcal{T}_k(h))$. This gives:

$$\mathsf{k}(x,y) = (1-h(x))\frac{k(x,y)(1-h(y))}{\gamma(y)} \tag{53}$$

and thus $\mathcal{T}_{\mathsf{k}} = M_{1-h}\,\mathcal{T}_{k/\gamma}\,M_{1-h}$, where M_f the multiplication operator by f. Recall the definition (9) of R_e. According to (21), we have:

$$\rho(\mathcal{T}_{\mathsf{k}}) = \rho\left(\mathcal{T}_{k/\gamma}M_{(1-h)^2}\right) = R_e((1-h)^2). \tag{54}$$

This gives the equivalence between (ii) and (iii).

We prove that (i) implies (iv). Suppose that $R_e(1-h) > 1$. Thanks to (21), we have $\rho(M_{1-h}\mathcal{T}_{k/\gamma}) = \rho(\mathcal{T}_{k/\gamma}M_{1-h}) = R_e(1-h) > 1$. According to [7, Lemma 3.7 (v)], there exists $v \in L^q_+ \setminus \{0\}$ a left Perron eigenfunction of $\mathcal{T}_{(1-h)k/\gamma}$, that is $\mathcal{T}^\top_{(1-h)k/\gamma}(v) = R_e(1-h)v$. Using $F(h) = 0$, and thus $(1-h)\mathcal{T}_k(h) = \gamma h$, for the last equality, we have:

$$R_e(1-h)\langle v, \gamma h\rangle = \langle v, (1-h)\mathcal{T}_{k/\gamma}(\gamma h)\rangle = \langle v, \gamma h\rangle.$$

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 $\mathcal{T}^\top_{(1-h)k/\gamma}(v) = R_e(1-h)v$ and setting $v' = (1-h)v$ (so that $v' = v$ μ-a.s.), we deduce that:

$$\mathcal{T}^\top_{k'/\gamma}(v') = R_e(1-h)v',$$

where $k' = \mathbb{1}_{A^c}\,k\,\mathbb{1}_{A^c}$. This implies that $\rho(\mathcal{T}_{k'/\gamma}) \geq R_e(1-h)$. Since $k' = (1-h)k'$ and $\mathcal{T}_{k/\gamma} - \mathcal{T}_{k'/\gamma}$ is a positive operator as $k - k' \geq 0$, we get, using (23) for the inequality, that $\rho(\mathcal{T}_{k'/\gamma}) = \rho(M_{1-h}\mathcal{T}_{k'/\gamma}) \leq \rho(M_{1-h}\mathcal{T}_{k/\gamma}) = R_e(1-h)$. Thus, the spectral radius of $\mathcal{T}_{k'/\gamma}$

is equal to $R_e(1-h)$. According to [7, Proposition 4.2], since $\rho(\mathcal{T}_{k'/\gamma}) > 1$, there exists $w \in \mathscr{L}^\infty_+ \setminus \{0\}$ and $\lambda > 0$ such that:

$$\mathcal{T}_{k'}(w) - \gamma w = \lambda w.$$

This also implies that $w = 0$ on $A = \{ h > 0 \}$, that is $wh = 0$ and thus $w\mathcal{T}_k(h) = 0$ as $\mathcal{T}_k(h) = \gamma h/(1-h)$. Using that $F(h) = 0$, $\mathcal{T}_{k'}(w) = \mathcal{T}_k(w)$ and $h\mathcal{T}_k(w) = 0$, we obtain:

$$F(h+w) = w(\lambda - \mathcal{T}_k(w)).$$

Taking $\varepsilon > 0$ small enough so that $\varepsilon\mathcal{T}_k(w) \leq \lambda/2$ and $\varepsilon w \leq 1$, we get $h + \varepsilon w \in \Delta$ and $F(h+\varepsilon w) \geq 0$. Then use Lemma 3.10 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 3.4 (iii) that $R_e((1-h)^2) \leq R_e(1-h)$.

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 $\gamma/(1-h) = \gamma + \mathcal{T}_k(h)$, so that $\gamma(\mathfrak{g}-h)/(1-h) \in \mathscr{L}^\infty_+$. An elementary computation, using $F(h) = F(\mathfrak{g}) = 0$ and (53), gives:

$$\mathcal{T}_{\mathsf{k}}\left(\gamma\frac{\mathfrak{g}-h}{1-h}\right) = (1-h)\mathcal{T}_k(\mathfrak{g}-h) = \gamma\frac{\mathfrak{g}-h}{1-\mathfrak{g}} = \frac{1-h}{1-\mathfrak{g}}\gamma\frac{\mathfrak{g}-h}{1-h}\cdot$$

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 7.1 that $\rho(\mathcal{T}_{\mathsf{k}}) > 1$. Then use (54) to conclude.

To conclude notice that $\mathfrak{g} = 0 \iff R_0 \leq 1$ is a consequence of the equivalence between (i) and (iv) with $h = 0$ and $R_0 = R_e(1)$.

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 $R_e(1-\mathfrak{g}) \geq 1$. Then use (iv) to deduce that $R_e(1-\mathfrak{g}) = 1$ if $\mathfrak{g} \neq 0$. $\square$

In the SIS model, in order to stress, if necessary, the dependence of a quantity H, such as F_η, R_e or $\mathfrak{g}_\eta$, in the parameters k and γ (which satisfy Assumption 2) of the model, we shall write $H[k,\gamma]$. Recall that if k and γ satisfy Assumption 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 7.3. *Let $((k_n,\gamma_n), n \in \mathbb{N})$ and (k,γ) be kernels and functions satisfying Assumption 2 and $(\eta_n,\, n \in \mathbb{N})$ be a sequence of elements of Δ converging weakly to η.*

- *(i) We have μ-a.s. $\lim_{n\to\infty} \mathfrak{g}_{\eta_n}[k,\gamma] = \mathfrak{g}_\eta[k,\gamma]$.*
- *(ii) Assume furthermore there exists $p' \in (1,+\infty)$ such that $\mathsf{k} = \gamma^{-1}k$ and $(\mathsf{k}_n = \gamma_n^{-1}k_n, n \in \mathbb{N})$ have finite double norm on $L^{p'}$ and that $\lim_{n\to\infty}\|\mathsf{k}_n - \mathsf{k}\|_{p',q'} = 0$. Then, we have μ-a.s. $\lim_{n\to\infty} \mathfrak{g}_{\eta_n}[k_n,\gamma_n] = \mathfrak{g}_\eta[k,\gamma]$.*

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 (17), we have:

$$g_n = \frac{\mathcal{T}_{\mathsf{k}_n}(\eta_n g_n)}{1+\mathcal{T}_{\mathsf{k}_n}(\eta_n g_n)} = \frac{\mathcal{T}_{\mathsf{k}_n}(h_n)}{1+\mathcal{T}_{\mathsf{k}_n}(h_n)}\cdot \tag{55}$$

We set $g = \mathcal{T}_{\mathsf{k}}(h)/(1 + \mathcal{T}_{\mathsf{k}}(h))$. Notice that $\mathcal{T}_{\mathsf{k}_n}(h_n) = (\mathcal{T}_{\mathsf{k}_n} - \mathcal{T}_{\mathsf{k}})(h_n) + \mathcal{T}_{\mathsf{k}}(h_n)$. We have $\lim_{n\to\infty} \mathcal{T}_{\mathsf{k}}(h_n) = \mathcal{T}_{\mathsf{k}}(h)$ pointwise. Since $\|(\mathcal{T}_{\mathsf{k}_n} - \mathcal{T}_{\mathsf{k}})(h_n)\|_{p'} \leq \|\mathsf{k}_n - \mathsf{k}\|_{p',q'}$, up to taking a sub-sequence, we deduce that a.s. $\lim_{n\to\infty}(\mathcal{T}_{\mathsf{k}_n} - \mathcal{T}_{\mathsf{k}})(h_n) = 0$. (Notice the previous step is not used in the proof of (i) as $\mathsf{k}_n = \mathsf{k}$ and $\lim_{n\to\infty} \mathcal{T}_k(h_n) = \mathcal{T}_k(h)$ pointwise.) This implies that g_n converges a.s. to g. By the dominated convergence theorem, we deduce that g_n converges also in L^p to g. This proves that h is actually equal a.s. to ηg. This gives $g = \mathcal{T}_{\mathsf{k}}(\eta g)/(1 + \mathcal{T}_{\mathsf{k}}(\eta g))$ and thus $F_\eta[k,\gamma](g) = 0$: g is an equilibrium for $F_\eta[k,\gamma]$. We deduce from the weak-continuity and the stability of R_e, see Theorem 3.5 and Proposition 3.6, that $R_e[\mathsf{k}](\eta(1-g)) = \lim_{n\to\infty} R_e[\mathsf{k}_n](\eta_n(1-g_n)) \leq 1$. Using Lemma 5.11 (i) with $h = 1/\gamma$, we get that $R_e[k/\gamma](\eta(1-g) = R_e[\mathsf{k}](\eta(1-g)) \leq 1$. (Only the weak-continuity of $\eta' \mapsto R_e[k/\gamma](\eta')$ is used in the proof of (i) to get $R_e[k/\gamma](\eta(1-g)) \leq 1$.) We deduce that property (iv) of Proposition 7.2 holds with k replaced by $k\eta$, and thus property (i) implies that $g = \mathfrak{g}_\eta[k,\gamma]$. ☐

Proofs of Theorem 3.12 and Proposition 3.13. Under the assumption of Lemma 7.3, taking $(k_n,\gamma_n) = (k,n)$ 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:

$$\lim_{n\to\infty} \mathfrak{I}[k_n,\gamma_n](\eta_n) = \lim_{n\to\infty} \int_\Omega \eta_n\, \mathfrak{g}_\eta[k_n,\gamma_n]\, \mathrm{d}\mu = \int_\Omega \eta_n\, \mathfrak{g}_\eta[k,\gamma]\, \mathrm{d}\mu = \mathfrak{I}[k,\gamma](\eta).$$

Taking $(k_n,\gamma_n) = (k,\gamma)$ provides the continuity of $\mathfrak{I}[k,\gamma]$ and thus Theorem 3.12. Then, arguing as in the end of the proof of Proposition 3.6, we get Proposition 3.13. ☐

References

[1] P. M. Anselone. *Collectively compact operator approximation theory and applications to integral equations*. Prentice-Hall, Inc., Englewood Cliffs, N. J., 1971. With an appendix by Joel Davis, Prentice-Hall Series in Automatic Computation.

[2] P. M. Anselone and J. W. Lee. Spectral properties of integral operators with nonnegative kernels. *Linear Algebra and its Applications*, 9:67–87, 1974.

[3] A. Boussaïri and B. Chergui. A transformation that preserves principal minors of skew-symmetric matrices. *Electron. J. Linear Algebra*, 32:131–137, 2017.

[4] T. Britton, F. Ball, and P. Trapman. A mathematical model reveals the influence of population heterogeneity on herd immunity to sars-cov-2. *Science*, 369:846–849, 8 2020.

[5] L. Burlando. Continuity of spectrum and spectral radius in Banach algebras. In *Functional analysis and operator theory (Warsaw, 1992)*, volume 30 of *Banach Center Publ.*, pages 53–100. Polish Acad. Sci. Inst. Math., Warsaw, 1994.

[6] J. B. Conway. *A course in functional analysis*, volume 96 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1990.

[7] J.-F. Delmas, D. Dronnier, and P.-A. Zitt. An Infinite-Dimensional SIS Model. *arXiv:2006.08241 [math]*, June 2020. arXiv: 2006.08241.

[8] J.-F. Delmas, D. Dronnier, and P.-A. Zitt. Examples of optimal vaccination for SIS model. *Work in progress*, 2021.

[9] J.-F. Delmas, D. Dronnier, and P.-A. Zitt. Optimal vaccination for a 2 population SIS model. *Work in progress*, 2021.

[10] 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. *Journal of Mathematical Biology*, 28(4):365–382, Jun 1990.

[11] K. Dietz. Transmission and control of arbovirus diseases. *Epidemiology*, 104:104–121, 1975.

[12] E. Duijzer, W. van Jaarsveld, J. Wallinga, and R. Dekker. The most efficient critical vaccination coverage and its equivalence with maximizing the herd effect. *Mathematical Biosciences*, 282:68–81, 12 2016.

[13] L. E. Duijzer, W. L. van Jaarsveld, J. Wallinga, and R. Dekker. Dose-optimal vaccine allocation over multiple populations. *Production and Operations Management*, 27:143–159, 1 2018.

[14] G. M. Engel and H. Schneider. Matrices diagonally similar to a symmetric matrix. *Linear Algebra Appl.*, 29:131–138, 1980.

[15] F. Esser and F. Harary. On the spectrum of a complete multipartite graph. *European Journal of Combinatorics*, 1(3):211–218, September 1980.

[16] Z. Feng, A. N. Hill, A. T. Curns, and J. W. Glasser. Evaluating targeted interventions via meta-population models with multi-level mixing. *Mathematical Biosciences*, 287:93–104, May 2017.

[17] 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. *Journal of Theoretical Biology*, 386:177–187, 12 2015.

[18] P. Fine, K. Eames, and D. L. Heymann. "Herd immunity": A rough guide. *Clinical Infectious Diseases*, 52:911–916, 4 2011.

[19] S. Friedland. Convex spectral functions. *Linear and Multilinear Algebra*, 9(4):299–316, 1980/81.

[20] 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.* IEEE, 2005.

[21] 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. *Journal of The Royal Society Interface*, 7:755–764, 5 2010.

[22] J. J. Grobler. Compactness conditions for integral operators in Banach function spaces. *Nederl. Akad. Wetensch. Proc. Ser. A 73=Indag. Math.*, 32:287–294, 1970.

[23] C. J. A. Halberg, Jr. and A. E. Taylor. On the spectra of linked operators. *Pacific J. Math.*, 6:283–290, 1956.

[24] D. J. Hartfiel and R. Loewy. On matrices having equal corresponding principal minors. *Linear Algebra Appl.*, 58:147–167, 1984.

[25] D. J. Hartfiel and R. Loewy. On matrices having equal corresponding principal minors. *Linear Algebra Appl.*, 58:147–167, 1984.

[26] H. W. Hethcote and J. A. Yorke. *Gonorrhea transmission dynamics and control*, volume 56 of *Lecture Notes in Biomathematics*. Springer-Verlag, Berlin, 1984. With a foreword by Paul J. Wiesner and Willard Cates, Jr.

[27] A. N. Hill and I. M. Longini Jr. The critical vaccination fraction for heterogeneous epidemic models. *Math. Biosci.*, 181(1):85–106, 2003.

[28] J. Hladký and I. Rocha. Independent sets, cliques, and colorings in graphons. *European Journal of Combinatorics*, 88:103108, August 2020.

[29] S. Janson. Graphons, cut norm and distance, couplings and rearrangements. *arXiv preprint arXiv:1009.2376*, 2010.

[30] O. Kallenberg. *Foundations of modern probability*, volume 99 of *Probability Theory and Stochastic Modelling*. Springer, Cham, third edition, [2021] ©2021.

[31] T. Kato. *Perturbation theory for linear operators*, volume 132. Springer-Verlag, 2013.

[32] M. J. Keeling and P. Rohani. *Modeling infectious diseases in humans and animals*. Princeton University Press, Princeton, 2008.

[33] A. Lajmanovich and J. A. Yorke. A deterministic model for gonorrhea in a nonhomogeneous population. *Mathematical Biosciences*, 28(3):221–236, 01 1976.

[34] R. Loewy. Principal minors and diagonal similarity of matrices. *Linear Algebra Appl.*, 78:23–64, 1986.

[35] L. Lovász. *Large networks and graph limits*. Number 60 in American Mathematical Society colloquium publications. American Mathematical Society, Providence, RI, 2012.

[36] I. Marek. Frobenius theory of positive operators: comparison theorems and applications. *SIAM Journal on Applied Mathematics*, 19(3):607–628, 1970.

[37] K. Miettinen. *Nonlinear multiobjective optimization*. Springer US, 1998.

[38] J. D. Newburgh. The variation of spectra. *Duke Math. J.*, 18(1):165–176, 03 1951.

[39] G. K. Pedersen. *C^*-algebras and their automorphism groups*. Pure and Applied Mathematics. Academic Press, an imprint of Elsevier, London San Diego Cambridge Oxford, second edition edition, 2018.

[40] G. Poghotanyan, Z. Feng, J. W. Glasser, and A. N. Hill. Constrained minimization problems for the reproduction number in meta-population models. *Journal of Mathematical Biology*, 77(6):1795–1831, 12 2018.

[41] J. G. Restrepo, E. Ott, and B. R. Hunt. Characterizing the dynamical importance of network nodes and links. *Physical Review Letters*, 97, 9 2006.

[42] 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, 6 2015.

[43] C. E. G. Smith. Prospects for the control of infectious disease. *Proceedings of the Royal Society of Medicine*, 63(11 Pt 2):1181–1190, November 1970.

[44] P. G. Smith. Concepts of herd protection and immunity. *Procedia in Vaccinology*, 2(2):134–139, January 2010.

[45] M. Somerville, K. Kumaran, and R. Anderson. *Public health and epidemiology at a glance*. John Wiley & Sons, August 2016.

[46] D. Stevanović, I. Gutman, and M. U. Rehman. On spectral radius and energy of complete multipartite graphs. *Ars Math. Contemp.*, 9(1):109–113, 2015.

[47] 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. *Physical Review E*, 84, 7 2011.

[48] A. C. Zaanen. *Linear analysis: measure and integral, Banach and Hilbert space, linear integral equations*. Bibl. Matematica. North-Holland, Amsterdam, 1956.

[49] H. Zhao and Z. Feng. Identifying optimal vaccination strategies via economic and epidemiological modeling. *Journal of Biological Systems*, 27:423–446, 12 2019.

Jean-François Delmas, CERMICS, École des Ponts, France
Email address: `jean-francois.delmas@cermics.enpc.fr`

Dylan Dronnier, CERMICS, École des Ponts, France
Email address: `dylan.dronnier@enpc.fr`

Pierre-André Zitt, LAMA, Université Gustave Eiffel, France
Email address: `pierre-andre.zitt@univ-eiffel.fr`