



**HAL**  
open science

## Targeted Vaccination Strategies for an Infinite-dimensional SIS Model

Jean-François Delmas, Dylan Dronnier, Pierre-André Zitt

► **To cite this version:**

Jean-François Delmas, Dylan Dronnier, Pierre-André Zitt. Targeted Vaccination Strategies for an Infinite-dimensional SIS Model. 2021. hal-03171312v2

**HAL Id: hal-03171312**

**<https://hal.science/hal-03171312v2>**

Preprint submitted on 25 Aug 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# 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 overall proportion of infected individuals in the endemic state. We prove the existence of Pareto optimal strategies for both loss functions.

We also show that vaccinating according to the profile of the endemic state is a critical allocation, in the sense that, if the initial reproduction number is larger than 1, then this vaccination strategy yields an effective reproduction number equal to 1.

## 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 [19]. 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 on stochastic block models. In companion papers, 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 recalling a few classical results in mathematical epidemiology; we refer to Keeling and Rohani’s monograph [30] 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. 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.

---

*Date:* August 25, 2021.

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

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

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

Assuming  $R_0 > 1$  in an homogeneous population, suppose now that only a proportion  $\eta^{\text{uni}}$  of the population can catch the disease, the rest 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 [42] and Dietz in 1975 [13] of a simple threshold theorem: the incidence of an infection declines if the proportion of non-immune individuals is reduced below  $\eta_{\text{crit}}^{\text{uni}} = 1/R_0$ . This effect is called *herd immunity*, and the corresponding percentage  $1 - \eta_{\text{crit}}^{\text{uni}}$  of people that have to be vaccinated is called *herd immunity threshold*; see for instance [43, 44].

It is of course unrealistic to depict human populations as homogeneous, and many generalizations of the homogeneous model have been studied; see [30, 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 [12]. After a vaccination campaign, let the vaccination strategy  $\eta$  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 vaccination strategy  $\eta$  is *critical* if  $R_e(\eta) = 1$ . The possible choices of  $\eta$  naturally raises a question that may be expressed as the following informal optimization problem:

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

If the quantity of available vaccine is limited, then one is also interested in:

$$(2) \quad \begin{cases} \textbf{Minimize:} & \text{the effective reproduction number } R_e \\ \textbf{subject to:} & \text{a given quantity of available vaccine.} \end{cases}$$

Interestingly enough, the strategy  $\eta_{\text{crit}}^{\text{uni}}$ , which consists in delivering the vaccine *uniformly* to the population, without taking inhomogeneity into account, leaves a proportion  $\eta_{\text{crit}}^{\text{uni}} = 1/R_0$  of the population unprotected, and is therefore critical since  $R_e(\eta_{\text{crit}}^{\text{uni}}) = 1$ . In particular it is admissible for the optimization problem (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 [30]. For example, the discussion of vaccination control of gonorrhea in [24, Section 4.5] suggests that it may be better to prioritize the vaccination of people that have already caught the disease: this lead us to consider a vaccination strategy guided by the equilibrium state. This strategy denoted by  $\eta^{\text{equi}}$  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: prioritizing the individuals that are more likely to get infected in a S(E)IR epidemic may be more efficient than distributing uniformly the vaccine in the population.

The main goal of this paper is two-fold: formalize the optimization problems (1) and (2) for a particular infinite dimensional SIS model, recasting them more generally as a bi-objective optimization problem; and give existence and properties of solutions to this bi-objective problem. We will also consider a closely related problem, where one wishes to minimize the size of the epidemic rather than the reproduction number. We will in passing provide insight on the efficiency of classical vaccination strategies such as  $\eta_{\text{crit}}^{\text{uni}}$  or  $\eta^{\text{equi}}$ .

**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) and (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 [25, 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) has been studied in this setting by Hill and Longini [25]. 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 derive in [38] an analytical formula for the solutions of Problem (2), for populations divided in two groups. Many papers also contain numerical studies of the optimization problems (1) and (2) on real-world data using gradient techniques or similar methods; see for example [14, 17, 18, 21, 47].

Finally, the effective reproduction number is not the only reasonable way of quantifying a population's vulnerability to an infection. For an SIR infection for example, the proportion of individuals that eventually catch (and recover from) the disease, often referred to as the *attack rate*, is broadly used. We refer to [14, 15] 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, 39]. Since Van Mieghem *et al.* proved in [46] 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 [40] and references therein.

**1.4. Main results.** The differential equations governing the epidemic dynamics in meta-population SIS models were developed by Lajmanovich and Yorke in their pioneer paper [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  $\Omega$ , where  $x \in \Omega$  represents a feature and the probability measure  $\mu(dx)$  represents the fraction of the population with feature  $x$ .

**1.4.1. Regularity of the effective reproduction function  $R_e$ .** We consider the effective reproduction function in a general operator framework which we call the *kernel model*. This model is characterized by a probability space  $(\Omega, \mathcal{F}, \mu)$  and a measurable non-negative kernel  $k : \Omega \times \Omega \rightarrow \mathbb{R}_+$ .

Let  $T_k$  be the corresponding integral operator defined by:

$$T_k(h)(x) = \int_{\Omega} k(x, y)h(y) \mu(dy).$$

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  $\gamma$  by the product  $k(x, y) = k(x, y)/\gamma(y)$ ; and the reproduction number  $R_0$  is then the spectral radius  $\rho(T_k)$  of  $T_k$ .

Following [7, Section 5], we represent a vaccination strategy by a function  $\eta : \Omega \rightarrow [0, 1]$ , where  $\eta(x)$  represents the fraction of **non-vaccinated** individuals with feature  $x$ ; the effective reproduction number associated to  $\eta$  is then given by

$$(3) \quad R_e(\eta) = \rho(T_{k\eta}),$$

where  $\rho$  stands for the spectral radius and  $k\eta$  stands for the kernel  $(k\eta)(x, y) = k(x, y)\eta(y)$ . If  $R_0 \geq 1$ , then a vaccination strategy  $\eta$  is called *critical* if it achieves precisely the herd immunity threshold, that is  $R_e(\eta) = 1$ .

In particular, the “strategy” that consists in vaccinating no one corresponds to  $\eta \equiv \mathbb{1}$ , and of course  $R_e(\mathbb{1}) = R_0$ . 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:

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

is critical, as  $R_e(\eta_{\text{crit}}^{\text{uni}}) = 1$ . This is consistent with results obtained in the homogeneous model given in Section 1.2.

Let  $\Delta$  be the set of strategies, that is the set of  $[0, 1]$ -valued functions defined on  $\Omega$ . The usual technique to obtain the existence of solutions to optimization problems like (1) or (2) is to prove that the function  $R_e$  is continuous with respect to a topology for which the set of strategies  $\Delta$  is compact. It is natural to try and prove this continuity by writing  $R_e$  as the composition of the spectral radius  $\rho$  and the map  $\eta \mapsto T_{k\eta}$ . The spectral radius is indeed continuous at compact operators (and  $T_{k\eta}$  is in fact compact under a technical integrability assumption on the kernel  $k$  formalized on page 10 as Assumption 1), if we endow the set of bounded operators with the operator norm topology; see [5, 37]. However, this would require choosing the uniform topology on  $\Delta$ , which then is not compact.

We instead endow  $\Delta$  with the weak topology, see Section 3.1, 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 [29, p. 431]. However, the family of operators  $(T_{k\eta}, \eta \in \Delta)$  is *collectively compact* which enables us to recover continuity, using a serie of results obtained by Anselone [1]. This leads to the following result, proved in Theorem 4.2 below. We recall that Assumption 1, formulated on page 10, 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  $R_e : \Delta \rightarrow \mathbb{R}_+$  is continuous with respect to the weak topology on  $\Delta$ .*

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  $R_e[k]$  to stress the dependence of the function  $R_e$  in the kernel  $k$ . In Proposition 4.3, we prove the stability of  $R_e$ , by giving natural sufficient conditions on a sequence of kernels  $(k_n, n \in \mathbb{N})$  converging to  $k$  which imply that  $R_e[k_n]$  converges uniformly towards  $R_e[k]$ . 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.

1.4.2. *On the maximal endemic equilibrium in the SIS model.* We consider the SIS model from [7]. This model is characterized by a probability space  $(\Omega, \mathcal{F}, \mu)$ , the transmission kernel  $k : \Omega \times \Omega \rightarrow \mathbb{R}_+$  and the recovery rate  $\gamma : \Omega \rightarrow \mathbb{R}_+^*$ . We suppose in the following that the technical Assumption 2, formulated on page 11, holds, so that the SIS dynamical evolution is well defined.

This evolution is encoded as  $u = (u_t, t \in \mathbb{R}_+)$ , where  $u_t \in \Delta$  for all  $t$  and  $u_t(x)$  represents the probability of an individual with feature  $x \in \Omega$  to be infected at time  $t \geq 0$ , and follows the equation:

$$(4) \quad \partial_t u_t = F(u_t) \quad \text{for } t \in \mathbb{R}_+, \quad \text{where} \quad F(g) = (1 - g)\mathcal{T}_k(g) - \gamma g \quad \text{for } g \in \Delta,$$

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, see (16). It is proved in [7] that such a solution  $u$  exists and is unique under Assumption 2. An *equilibrium* of (4) is a function  $g \in \Delta$  such that  $F(g) = 0$ . According to [7], there exists a maximal equilibrium  $\mathbf{g}$ , *i.e.*, an equilibrium such that all other equilibria  $h \in \Delta$  are dominated by  $\mathbf{g}$ :  $h \leq \mathbf{g}$ . Furthermore, we have  $R_0 \leq 1$  if and only if  $\mathbf{g} = 0$ . In the connected case (for example if  $k > 0$ ), then 0 and  $\mathbf{g}$  are the only equilibria; besides  $\mathbf{g}$  is the long-time distribution of infected individuals in the population:  $\lim_{t \rightarrow +\infty} u_t = \mathbf{g}$  as soon as the initial condition is non-zero; see [7, Theorem 4.14].

As hinted in [24, Section 4.5] for vaccination control of gonorrhoea, it is interesting to consider vaccinating people with feature  $x$  with probability  $\mathbf{g}(x)$ ; this corresponds to the strategy based on the maximal equilibrium:

$$\eta^{\text{equi}} = 1 - \mathbf{g}.$$

The following result entails that this strategy is critical and thus achieves the herd immunity threshold. Recall that Assumption 2, formulated page 11, provides technical conditions on the parameters  $k$  and  $\gamma$  of the SIS model. The effective reproduction number of the SIS model is the function  $R_e$  defined in (3) with the kernel  $k = k/\gamma$ .

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

This result will be proved below as a part of Proposition 8.2. Let us finally describe informally another consequence of this Proposition. 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 8.2 gives spectral information on the formal linearization of the dynamics (4) near any equilibrium  $h$ ; in particular if  $h \neq \mathbf{g}$  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  $\eta$  is given by (4), where  $F$  is replaced by  $F_\eta$  defined by:

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

and  $u_t$  now describes the proportion of infected *among the non-vaccinated population*. We denote by  $\mathbf{g}_\eta$  the corresponding maximal equilibrium (thus considering  $\eta \equiv 1$  gives  $\mathbf{g} = \mathbf{g}_1$ ), so that  $F_\eta(\mathbf{g}_\eta) = 0$ . Since the probability for an individual  $x$  to be infected in the stationary regime is  $\mathbf{g}_\eta(x)\eta(x)$ , the *fraction of infected individuals at equilibrium*,  $\mathfrak{I}(\eta)$ , is thus given by:

$$(5) \quad \mathfrak{I}(\eta) = \int_{\Omega} \mathbf{g}_\eta \eta \, d\mu = \int_{\Omega} \mathbf{g}_\eta(x) \eta(x) \mu(dx).$$

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 [14, 15]. 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  $\mathfrak{J}$  that are very similar to the ones on  $R_e$ . Recall that Assumption 2 on page 11 ensures that the infinite-dimensional SIS model, given by equation (4), is well defined. The next theorem corresponds to Theorem 4.6.

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

In Proposition 4.7, we prove the stability of  $\mathfrak{J}$ , by giving natural sufficient condition on a sequence of kernels and functions  $((k_n, \gamma_n), n \in \mathbb{N})$  converging to  $(k, \gamma)$  which imply that  $\mathfrak{J}[k_n, \gamma_n]$  converges uniformly towards  $\mathfrak{J}[k, \gamma]$ . We also prove that the loss functions  $L = R_e$  and  $L = \mathfrak{J}$  are both non-decreasing ( $\eta \leq \eta'$  implies  $L(\eta) \leq L(\eta')$ ), and sub-homogeneous ( $L(\lambda\eta) \leq \lambda L(\eta)$  for all  $\lambda \in [0, 1]$ ); see Propositions 4.1 and 4.5.

1.4.4. *Optimizing the protection of the population.* Consider a cost function  $C : \Delta \rightarrow [0, 1]$  which measures the cost for the society of a vaccination strategy (production and diffusion). Since the vaccination strategy  $\eta$  represents the non-vaccinated population, the cost function  $C$  should be decreasing (roughly speaking  $\eta < \eta'$  implies  $C(\eta) > C(\eta')$ ; see Definition 5.1). We shall also assume that  $C$  is continuous with respect to the weak topology on  $\Delta$ , and that doing nothing costs nothing, that is,  $C(\mathbf{1}) = 0$ . A simple and natural choice is the uniform cost  $C_{\text{uni}}$  given by the overall proportion of vaccinated individuals:

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

See Remark 5.2 for comments on other examples of cost functions.

Our problem may now be seen as a bi-objective minimization problem: we wish to minimize both the loss  $L(\eta)$  and the cost  $C(\eta)$ , subject to  $\eta \in \Delta$ , with the loss function  $L$  being either  $R_e$  or  $\mathfrak{J}$ . Following classical terminology for multi-objective optimisation problems [36], we call a strategy  $\eta_*$  *Pareto optimal* if no other strategy is strictly better:

$$C(\eta) < C(\eta_*) \implies L(\eta) > L(\eta_*) \quad \text{and} \quad L(\eta) < L(\eta_*) \implies C(\eta) > C(\eta_*).$$

The set of Pareto optimal strategies will be denoted by  $\mathcal{P}_L$ , and we define the *Pareto frontier* as the set of Pareto optimal outcomes:

$$\mathcal{F}_L = \{(C(\eta_*), L(\eta_*)) : \eta_* \in \mathcal{P}_L\}.$$

Notice that, with this definition, the Pareto frontier is empty when there is no Pareto optimal strategy.

For any strategy  $\eta$ , the cost and loss of  $\eta$  vary between the following bounds:

$$\begin{aligned} 0 = C(\mathbf{1}) \leq C(\eta) \leq C(0) = c_{\max} &= \text{cost of vaccinating the whole population,} \\ 0 = L(0) \leq L(\eta) \leq L(\mathbf{1}) = \ell_{\max} &= \text{loss incurred in the absence of vaccination.} \end{aligned}$$

Let  $L_*$  be the *optimal loss function* and  $C_{*,L}$  the *optimal cost function* defined by:

$$\begin{aligned} L_*(c) &= \inf \{L(\eta) : \eta \in \Delta, C(\eta) \leq c\} \quad \text{for } c \in [0, c_{\max}], \\ C_{*,L}(\ell) &= \inf \{C(\eta) : \eta \in \Delta, L(\eta) \leq \ell\} \quad \text{for } \ell \in [0, \ell_{\max}]. \end{aligned}$$

We simply write  $C_*$  for  $C_{*,L}$  when no confusion on the loss function can arise. Proposition 5.5 (in a more general framework in particular for the cost function) and Lemma 5.6 states that the Pareto frontier is non empty and has a continuous parametrization for the cost  $C = C_{\text{uni}}$  and the loss  $L = R_e$  or  $L = \mathfrak{J}$ ; see Figure 1(B) below for a visualization of the Pareto frontier.

**Theorem 1.4** (Properties of the Pareto frontier). *For the kernel model with loss function  $L = R_e$  or the SIS model with  $L \in \{R_e, \mathfrak{J}\}$ , and the uniform cost function  $C = C_{\text{uni}}$ , the function  $C_{*,L}$*

is continuous and decreasing on  $[0, \ell_{\max}]$ , the function  $L_\star$  is continuous on  $[0, c_{\max}]$  decreasing on  $[0, C_{\star, L}(0)]$  and zero on  $[C_{\star, L}(0), c_{\max}]$ ; furthermore the Pareto frontier is connected and:

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

We also establish that  $\mathcal{P}_L$  is compact in  $\Delta$  for the weak topology in Corollary 5.7; that the set of outcomes or feasible region  $\mathbf{F} = \{(C(\eta), L(\eta)), \eta \in \Delta\}$  has no holes in Proposition 6.1; and that the Pareto frontier is convex if  $C$  and  $L$  are convex in Proposition 6.6. We study in Proposition 6.2 the stability of the Pareto frontier and the set of Pareto optima when the parameters vary.

In a sense the Pareto optimal strategies are intuitively the “best” strategies. Similarly, we also study the “worst” strategies, which we call anti-Pareto optimal strategies, and describe the corresponding anti-Pareto frontier. Understanding the “worst strategies” also helps to avoid pitfalls when one has to consider sub-optimal strategies: for example, we prove in [8] that disconnecting strategies are not the “worst” strategies, and we provide in [10, Section 4] an elementary example where the same strategies can be “best” or “worst” according to model parameters values. Surprisingly, proving properties of the anti-Pareto frontier sometimes necessitates stronger assumptions than in the Pareto case: for example, the connectedness of the anti-Pareto frontier is only proved under a *quasi-irreducibility* assumption on the kernel, see Lemmas 5.11 and 5.12.

*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  $R_e(\eta) > 1$ . Consider the uniform cost  $C = C_{\text{uni}}$ . 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  $\eta$  is Pareto optimal for the objectives  $(\mathfrak{I}, C)$  and satisfies  $\mathfrak{I}(\eta) = 0$ :

$$(6) \quad \eta \in \mathcal{P}_{R_e} \text{ and } R_e(\eta) = 1 \iff \eta \in \mathcal{P}_{\mathfrak{I}} \text{ and } \mathfrak{I}(\eta) = 0.$$

*Remark 1.6* (Minimal cost of eradication). Assume  $R_0 > 1$  and the uniform cost  $C = C_{\text{uni}}$ . The equivalence (6) implies directly that:

$$C_{\star, R_e}(\mathbf{1}) = C_{\star, \mathfrak{I}}(0).$$

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

$$C_{\star, R_e}(\mathbf{1}) = C_{\star, \mathfrak{I}}(0) \leq \min \left( 1 - \frac{1}{R_0}, \int_{\Omega} \mathbf{g} \, d\mu \right).$$

1.4.5. *Equivalence of models.* Our last results address a natural question stemming from our choice of a very general framework to modelize the infection. Since our models are infinite dimensional and depend on the choices of the probability space  $(\Omega, \mathcal{F}, \mu)$ , the kernel  $\mathbf{k}$  (for the kernel model) and the kernel  $k$  and recovery rate  $\gamma$  (for the SIS model), they are different equivalent ways to model the same situation. We study in Section 7 a way to ensure that, even if the parameters are different, we end up with the same Pareto frontiers. This situation is similar to random variables having the same law in probability theory, or to equivalent graphons in graphon theory. In particular it allows us to treat the same meta-population model in either a discrete or a continuous setting, see Figure 4 for an illustration and Example 1.7.



1.4.6. *An illustrative example: the multipartite graphon.* Let us illustrate some of our results on an example, which will be discussed in details in a forthcoming companion paper [10].

*Example 1.7* (Multipartite graphon). Graphs that can be colored with  $\ell$  colors, so that no two endpoints of an edge have the same color are known as  $\ell$ -partite graphs. In a biological setting, this corresponds to a population of  $\ell$  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). Using the equivalence of models from Section 7, we can represent this model by using a continuous state space  $\Omega = [0, 1]$ , endowed with  $\mu$  the Lebesgue measure on  $\Omega$ , 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  $k = \kappa(1 - \sum_{n \in \mathbb{N}^*} \mathbb{1}_{I_n \times I_n})$ ; it is represented in Figure 1(A).

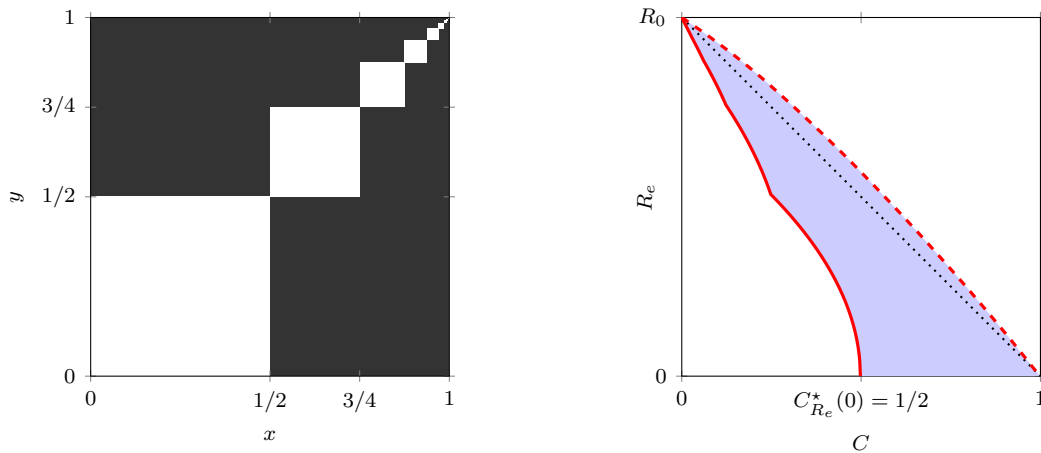
Consider the loss  $L = R_e$  and the cost  $C = C_{\text{uni}}$  giving the overall proportion of vaccinated individuals. Based on the results of [16, 45], we prove in [10] that the vaccination strategies  $\mathbb{1}_{[0, 1-c]}$ , with cost  $C(\mathbb{1}_{[0, 1-c]}) = c \in [0, 1/2]$ , are Pareto optimal. Remembering that the natural definition of the degree in a continuous graph is given by  $\text{deg}(x) = \int_{\Omega} k(x, y) \mu(dy)$ , 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 (*i.e.*, the outcome of the “best” vaccination strategies) is drawn as the solid red line; the blue-colored zone corresponds to the feasible region that is, all the possible values of  $(C(\eta), R_e(\eta))$ , where  $\eta$  ranges over  $\Delta$ ; 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 red dashed curve corresponds to the anti-Pareto frontier (*i.e.*, the outcome of the “worst” vaccination strategies), which for this model correspond to the uniform vaccination of the nodes with the updated lower degree; see [10]. Notice that the path  $(\mathbb{1}_{[0, 1-c]}, c \in [0, 1/2])$  is an increasing continuous (for the topology of the simple convergence and thus the  $L^1(\mu)$  topology) path of Pareto optima which gives a complete parametrization of the Pareto frontier. The latter has been computed numerically using the power iteration method. In particular, we obtained the following value:  $R_0 \simeq 0.697\kappa$ .

**1.5. On the companion papers.** We detail some developments in forthcoming papers where only the uniform cost  $C = C_{\text{uni}}$  is considered. In [8], motivated by the conjecture formulated by Hill and Longini in finite dimension [25, Conjecture 8.1], we investigate the convexity and concavity of the effective reproduction function  $R_e$ . We also prove that a disconnecting strategy is better than the worst, *i.e.*, is not anti-Pareto optimal.

In [11], under monotonicity properties of the kernel, satisfied for example by the configuration model, it is proven that vaccinating the individuals with the highest (resp. lowest) number of contacts is Pareto (resp. anti-Pareto) optimal. In this case the greedy algorithm, which performs infinitesimal locally optimal steps, is optimal as it browses continuously the set of Pareto (resp. anti-Pareto) optimal strategies, providing an increasing parametrization of the Pareto (resp. anti-Pareto) frontier. In this setting, we provide some examples of SIS models where the set of Pareto optimal strategies coincide for the losses  $R_e$  and  $\mathfrak{J}$ :

$$(7) \quad \mathcal{P}_{\mathfrak{J}} = \mathcal{P}_{R_e} \cap \{\eta \in \Delta : R_e(\eta) \geq 1\}.$$

In [10], which includes a detailed study of the multipartite kernel of Example 1.7, we study the optimal vaccination when the individuals have the same number of contacts. This provides examples where the uniform vaccination is Pareto optimal, or anti-Pareto optimal, or not optimal for either problem. We also provide an example where the set  $\mathcal{P}_{R_e}$  has a countable



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

FIGURE 1. Example of optimization with  $L = R_e$ .

number of connected components (and is thus not connected). This implies in particular that the greedy algorithm is not optimal in this case.

In [9], we give a comprehensive treatment of the two groups model,  $\Omega = \{1, 2\}$ , for  $L = R_e$ , and some partial results for  $L = \mathcal{J}$ . Despite its apparent simplicity, the derivation of formulae for the Pareto optimal strategies is non trivial, see also [38]. In addition, this model is rich enough to give examples of various interesting behaviours:

- *On the critical strategies  $\eta_{\text{crit}}^{\text{uni}}$  and  $\eta^{\text{equi}}$ .* Depending on the parameters, the strategies  $\eta_{\text{crit}}^{\text{uni}}$  and/or  $\eta^{\text{equi}}$  may or may not be Pareto optimal, and the cost  $C(\eta_{\text{crit}}^{\text{uni}})$  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.
- *Dependence on the choice of the loss function.* For examples where  $R_0 > 1$ , the optimal strategies for the losses  $\mathcal{J}$  and  $R_e$  may coincide, so that (7) holds, or not at all, so that  $\mathcal{P}_{\mathcal{J}} \cap \mathcal{P}_{R_e} \cap \{\eta \in \Delta : 1 < R_e(\eta) < R_0\} = \emptyset$ , depending on the parameters.

**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  $\mathcal{J}$ . After recalling a few topological facts in Section 3, we study the regularity properties of  $R_e$  and  $\mathcal{J}$  in Section 4. We present the multi-objective optimization problem in Section 5 under general condition on the loss function  $L$  and cost function  $C$  and prove the results on the Pareto frontier. This is completed in Section 6 with miscellaneous properties of the Pareto frontier. In Section 7, we discuss the equivalent representation of models with different parameters. Proofs of a few technical results are gathered in Section 8.

## 2. SETTING AND NOTATION

**2.1. Spaces, operators, spectra.** All metric spaces  $(S, d)$  are endowed with their Borel  $\sigma$ -field denoted by  $\mathcal{B}(S)$ . The set  $\mathcal{K}$  of compact subsets of  $\mathbb{C}$  endowed with the Hausdorff distance  $d_H$  is a metric space, and the function  $\text{rad}$  from  $\mathcal{K}$  to  $\mathbb{R}_+$  defined by  $\text{rad}(K) = \max\{|\lambda|, \lambda \in K\}$  is Lipschitz continuous from  $(\mathcal{K}, d_H)$  to  $\mathbb{R}$  endowed with its usual Euclidean distance.

Let  $(\Omega, \mathcal{F}, \mu)$  be a probability space. We denote by  $\mathcal{L}^\infty$ , the Banach spaces of bounded real-valued measurable functions defined on  $\Omega$  equipped with the sup-norm,  $\mathcal{L}_+^\infty$  the subset of  $\mathcal{L}^\infty$  of non-negative function, and  $\Delta = \{f \in \mathcal{L}^\infty : 0 \leq f \leq 1\}$  the subset of non-negative functions bounded by 1. For  $f$  and  $g$  real-valued functions defined on  $\Omega$ , we may write  $\langle f, g \rangle$  or  $\int_\Omega fg \, d\mu$  for  $\int_\Omega f(x)g(x) \, \mu(dx)$  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 on  $\Omega$  such that  $\|g\|_p = (\int |g|^p \, d\mu)^{1/p}$  (with the convention that  $\|g\|_\infty$  is the  $\mu$ -essential supremum of  $|g|$ ) is finite, where functions which agree  $\mu$ -almost surely are identified. We denote by  $L_+^p$  the subset of  $L^p$  of non-negative functions.

Let  $(E, \|\cdot\|)$  be a Banach space. We denote by  $\|\cdot\|_E$  the operator norm on  $\mathcal{L}(E)$  the Banach algebra of bounded operators. The spectrum  $\text{Spec}(T)$  of  $T \in \mathcal{L}(E)$  is the set of  $\lambda \in \mathbb{C}$  such that  $T - \lambda \text{Id}$  does not have a bounded inverse operator, where  $\text{Id}$  is the identity operator on  $E$ . Recall that  $\text{Spec}(T)$  is a compact subset of  $\mathbb{C}$ , and that the spectral radius of  $T$  is given by:

$$(8) \quad \rho(T) = \text{rad}(\text{Spec}(T)) = \lim_{n \rightarrow \infty} \|T^n\|_E^{1/n}.$$

The element  $\lambda \in \text{Spec}(T)$  is an eigenvalue if there exists  $x \in E$  such that  $Tx = \lambda x$  and  $x \neq 0$ .

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

**2.2. Kernel operators.** We define a *kernel* (resp. *signed kernel*) on  $\Omega$  as a  $\mathbb{R}_+$ -valued (resp.  $\mathbb{R}$ -valued) measurable function defined on  $(\Omega^2, \mathcal{F}^{\otimes 2})$ . For  $f, g$  two non-negative measurable functions defined on  $\Omega$  and  $k$  a kernel on  $\Omega$ , we denote by  $fk g$  the kernel defined by:

$$(9) \quad fkg : (x, y) \mapsto f(x)k(x, y)g(y).$$

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

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

$$(10) \quad \|k\|_{p,q} = \left( \int_\Omega \left( \int_\Omega |k(x, y)|^q \, \mu(dy) \right)^{p/q} \mu(dx) \right)^{1/p} \quad \text{with } q \text{ given by } \frac{1}{p} + \frac{1}{q} = 1.$$

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

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

$$(11) \quad T_k(g)(x) = \int_\Omega k(x, y)g(y) \, \mu(dy) \quad \text{for } g \in L^p \text{ and } x \in \Omega.$$

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

$$(12) \quad \|T_k\|_{L^p} \leq \|k\|_{p,q}.$$

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

$$(13) \quad \text{Spec}[k](\eta) = \text{Spec}(T_{k\eta}),$$

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

$$(14) \quad R_e[k](\eta) = \text{rad}(\text{Spec}(T_{k\eta})) = \rho(T_{k\eta}),$$

and the corresponding *reproduction number*:

$$(15) \quad R_0[k] = R_e[k](\mathbf{1}) = \rho(T_k).$$

When there is no ambiguity, we simply write  $R_e$  for  $R_e[k]$  and  $R_0$  for  $R_0[k]$ . We say a vaccination strategy  $\eta \in \Delta$  is *critical* if  $R_e(\eta) = 1$ .

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

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

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

$$(16) \quad \mathcal{T}_k(g)(x) = \int_{\Omega} k(x, y)g(y) \mu(dy) \quad \text{for } g \in \mathcal{L}^{\infty} \text{ and } x \in \Omega.$$

Notice that the integral operators  $\mathcal{T}_k$  and  $T_k$  corresponds respectively to the operators  $T_k$  and  $\hat{T}_k$  in [7]. According to [7, Lemma 3.7], the operator  $\mathcal{T}_k^2$  on  $\mathcal{L}^{\infty}$  is compact and  $\mathcal{T}_k$  has the same spectral radius as  $T_k$ :

$$(17) \quad \rho(\mathcal{T}_k) = \rho(T_k).$$

**2.3. Dynamics for the SIS model and equilibria.** In accordance with [7], we consider the following assumption. Recall that  $k/\gamma = k\gamma^{-1}$ .

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

Assumption 2 implies Assumption 1 for the kernel  $k = k/\gamma$ . Under Assumption 2, we also consider the bounded operators  $\mathcal{T}_{k/\gamma}$  on  $\mathcal{L}^{\infty}$ , as well as  $T_{k/\gamma}$  on  $L^p$ , which are the so called *next-generation operator*. The SIS dynamics considered in [7] (under Assumption 2) follows the vector field  $F$  defined on  $\mathcal{L}^{\infty}$  by:

$$(18) \quad F(g) = (1 - g)\mathcal{T}_k(g) - \gamma g.$$

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

$$(19) \quad \partial_t u_t = F(u_t) \quad \text{for } t \in \mathbb{R}_+,$$

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 (19) is a function  $g \in \Delta$  such that  $F(g) = 0$ . According to [7], there exists a maximal equilibrium  $\mathbf{g}$ , *i.e.*, an equilibrium such that all other equilibria  $h \in \Delta$  are dominated by  $\mathbf{g}$ :  $h \leq \mathbf{g}$ . The *reproduction number*  $R_0$  associated to the SIS model given by (19) is the spectral radius of the next-generation operator, so that using the definition of the effective reproduction number (14), (15) and (17), this amounts to:

$$(20) \quad R_0 = \rho(\mathcal{T}_{k/\gamma}) = R_0[k/\gamma] = R_e[k/\gamma](\mathbf{1}).$$

If  $R_0 \leq 1$  (sub-critical and critical case), then  $u_t$  converges pointwise to 0 when  $t \rightarrow \infty$ . In particular, the maximal equilibrium  $\mathbf{g}$  is equal to 0 everywhere. If  $R_0 > 1$  (super-critical case), then 0 is still an equilibrium but different from the maximal equilibrium  $\mathbf{g}$ , as  $\int_{\Omega} \mathbf{g} d\mu > 0$ .

**2.4. Vaccination strategies.** A *vaccination strategy*  $\eta$  of a vaccine with perfect efficiency is an element of  $\Delta$ , where  $\eta(x)$  represents the proportion of *non-vaccinated* individuals with feature  $x$ . Notice that  $\eta d\mu$  corresponds in a sense to the effective population.

Recall the definition of the kernel  $fkg$  from (9). 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  $\mathcal{L}^\infty$ . According to [7, Section 5.3.], the SIS equation with vaccination strategy  $\eta$  is given by (19), where  $F$  is replaced by  $F_\eta$  defined by:

$$(21) \quad F_\eta(g) = (1 - g)\mathcal{T}_{k\eta}(g) - \gamma g.$$

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 a non-vaccinated 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 (14) as  $\rho(\mathcal{T}_{k\eta/\gamma}) = \rho(T_{k\eta/\gamma}) = R_e[k/\gamma](\eta)$  with  $k = k/\gamma$ . We denote by  $\mathfrak{g}_\eta$  the corresponding maximal equilibrium (so that  $\mathfrak{g} = \mathfrak{g}_1$ ). In particular, we have:

$$(22) \quad F_\eta(\mathfrak{g}_\eta) = 0.$$

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:

$$(23) \quad \mathfrak{I}(\eta) = \int_{\Omega} \mathfrak{g}_\eta \eta d\mu = \int_{\Omega} \mathfrak{g}_\eta(x)\eta(x)\mu(dx).$$

We deduce from (21) and (22) that  $\mathfrak{g}_\eta\eta = 0$   $\mu$ -almost surely is equivalent to  $\mathfrak{g}_\eta = 0$ . Applying the results of [7] to the kernel  $k\eta$ , we deduce that:

$$(24) \quad \mathfrak{I}(\eta) > 0 \iff R_e[k/\gamma](\eta) > 1.$$

We conclude this section with a result on the maximal equilibrium  $\mathfrak{g}$  which is a direct consequence of Proposition 8.2 proved in Section 8.1. This result 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 and write  $R_e$  for  $R_e[k/\gamma]$ .*

- (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  $R_e(1 - \mathfrak{g}) = 1$ .*

### 3. PRELIMINARY TOPOLOGICAL RESULTS

**3.1. On the weak topology.** We first recall briefly some properties we shall use frequently. We can see  $\Delta$  as a subset of  $L^1$ , and consider the corresponding *weak topology*: a sequence  $(g_n, n \in \mathbb{N})$  of elements of  $\Delta$  converges weakly to  $g$  if for all  $h \in L^\infty$  we have:

$$(25) \quad \lim_{n \rightarrow \infty} \int_{\Omega} h g_n d\mu = \int_{\Omega} h g d\mu.$$

Notice that (25) can easily be extended to any function  $h \in L^q$  for any  $q \in (1, +\infty)$ ; so that the weak-topology on  $\Delta$ , seen as a subset of  $L^p$  with  $1/p + 1/q = 1$ , can be seen as the trace on  $\Delta$  of the weak topology on  $L^p$ . The main advantage of this topology is the following compactness result.

**Lemma 3.1** (Topological properties of  $\Delta$ ). *We have that:*

- (i) *The set  $\Delta$  endowed with the weak topology is compact and sequentially compact.*
- (ii) *A function from  $\Delta$  (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  $\Delta$  as the trace on  $\Delta$  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  $\Delta$  is closed and convex, therefore it is weakly closed; see [6, Corollary V.1.5]. Thus,  $\Delta$  is weakly compact as a weakly closed subset of the weakly compact unit ball. By the Eberlein–Šmulian theorem [6, Theorem V.13.1],  $\Delta$  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.  $\square$

**3.2. Invariance and continuity of the spectrum for compact operators.** We recall a few facts on operators. Let  $(E, \|\cdot\|)$  be a Banach space. Let  $A \in \mathcal{L}(E)$ . We denote by  $A^\top$  the adjoint of  $A$ . A sequence  $(A_n, n \in \mathbb{N})$  of elements of  $\mathcal{L}(E)$  converges strongly to  $A \in \mathcal{L}(E)$  if  $\lim_{n \rightarrow \infty} \|A_n x - Ax\| = 0$  for all  $x \in E$ . Following [1], a set of operators  $\mathcal{A} \subset \mathcal{L}(E)$  is *collectively compact* if the set  $\{Ax : A \in \mathcal{A}, \|x\| \leq 1\}$  is relatively compact.

We collect some known results on the spectrum of compact operators. Recall that the spectrum of a compact operator is finite or countable and has at most one accumulation point, which is 0. Furthermore, 0 belongs to the spectrum of compact operators in infinite dimension.

**Lemma 3.2.** *Let  $A, B$  be elements of  $\mathcal{L}(E)$ .*

(i) *If  $A, B$  and  $A - B$  are positive operators, then we have:*

$$(26) \quad \rho(A) \geq \rho(B).$$

(ii) *If  $A$  is compact, then we have:*

$$(27) \quad \text{Spec}(A) = \text{Spec}(A^\top)$$

$$(28) \quad \text{Spec}(AB) = \text{Spec}(BA)$$

*and in particular:*

$$(29) \quad \rho(AB) = \rho(BA).$$

(iii) *Let  $(E', \|\cdot\|')$  be a Banach space such that  $E'$  is continuously and densely embedded in  $E$ . Assume that  $A(E') \subset E'$ , and denote by  $A'$  the restriction of  $A$  to  $E'$  seen as an operator on  $E'$ . If  $A$  and  $A'$  are compact, then we have:*

$$(30) \quad \text{Spec}(A) = \text{Spec}(A').$$

(iv) *Let  $(A_n, n \in \mathbb{N})$  be a collectively compact sequence which converges strongly to  $A$ . Then, we have  $\lim_{n \rightarrow \infty} \text{Spec}(A_n) = \text{Spec}(A)$  in  $(\mathcal{K}, d_H)$ , and  $\lim_{n \rightarrow \infty} \rho(T_n) = \rho(T)$ .*

*Proof.* Property (i) can be found in [35, Theorem 4.2]. Equation (27) from Property (ii) can be deduced from the [32, Theorem page 20]. Using the [32, Proposition page 25], we get that  $\text{Spec}(AB) \cap \mathbb{C}^* = \text{Spec}(BA) \cap \mathbb{C}^*$ , and thus (29). As  $A$  is compact we get that  $AB$  and  $BA$  are compact, thus 0 belongs to their spectrum in infinite dimension. Whereas in finite dimension, as  $\det(AB) = \det(A)\det(B) = \det(BA)$  (where  $A$  and  $B$  denote also the matrix of the corresponding operator in a given base), we get that 0 belongs to the spectrum of  $AB$  if and only if it belongs to the spectrum of  $BA$ . This gives (28).

Property (iii) follows from [23, Corollary 1 and Section 6]. We eventually check Property (iv). We deduce from [1, Theorems 4.8 and 4.16] (see also (d) and (e) in [2, Section 3]) that  $\lim_{n \rightarrow \infty} \text{Spec}(T_n) = \text{Spec}(T)$ . Then use that the function  $\text{rad}$  is continuous to deduce the convergence of the spectral radius from the convergence of the spectra (see also (f) in [2, Section 3]).  $\square$

4. FIRST PROPERTIES OF THE FUNCTIONS  $R_e$  AND  $\mathfrak{J}$ 

**4.1. The effective reproduction number  $R_e$ .** We consider the kernel model  $[(\Omega, \mathcal{F}, \mu), k]$  under Assumption 1, so that  $k$  is a kernel on  $\Omega$  with finite double norm. Recall the effective reproduction number function  $R_e[k]$  defined on  $\Delta$  by (14):  $R_e[k](\eta) = \rho(T_k M_\eta)$  and the reproduction number  $R_0[k] = \rho(T_k)$ . We simply write  $R_e$  and  $R_0$  for  $R_e[k]$  and  $R_0[k]$  respectively when no confusion on the kernel can arise.

**Proposition 4.1** (Basic properties of  $R_e$ ). *Suppose Assumption 1 holds. Let  $\eta, \eta_1, \eta_2 \in \Delta$ . The function  $R_e = R_e[k]$  satisfies the following properties:*

- (i)  $R_e(\eta_1) = R_e(\eta_2)$  if  $\eta_1 = \eta_2$   $\mu$ -almost surely.
- (ii)  $R_e(0) = 0$  and  $R_e(\mathbf{1}) = R_0$ .
- (iii)  $R_e(\eta_1) \leq R_e(\eta_2)$  if  $\eta_1 \leq \eta_2$   $\mu$ -almost surely.
- (iv)  $R_e(\lambda\eta) = \lambda R_e(\eta)$  for all  $\lambda \in [0, 1]$ .

*Proof.* If  $\eta_1 = \eta_2$   $\mu$ -almost surely, then we have that  $T_{k\eta_1} = T_{k\eta_2}$ , 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  $\lambda A$  is equal to  $\{\lambda s, s \in \text{Spec}(A)\}$ , Point (iv) is clear. Finally, note that if  $\eta_1 \leq \eta_2$   $\mu$ -almost everywhere, then the operator  $T_{k\eta_2} - T_{k\eta_1}$  is positive. According to (26), we get that  $\rho(T_{k\eta_1}) \leq \rho(T_{k\eta_2})$ . This concludes the proof of Point (iii).  $\square$

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

**Theorem 4.2** (Continuity of  $R_e[k]$  and  $\text{Spec}[k]$ ). *Suppose Assumption 1 holds. Then, the functions  $\text{Spec}[k]$  and  $R_e[k]$  are continuous functions from  $\Delta$  (endowed with the weak-topology) respectively to  $\mathcal{K}$  (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  $T_k(B)$  is relatively compact. For all  $\eta \in \Delta$ , set  $\eta B = \{\eta g : g \in B\}$ . As  $\eta B \subset B$ , we deduce that  $T_{k\eta}(B) = T_k(\eta B) \subset T_k(B)$ . This implies that the family  $(T_{k\eta}, \eta \in \Delta)$  is collectively compact.

Let  $(\eta_n, n \in \mathbb{N})$  be a sequence in  $\Delta$  converging weakly to some  $\eta \in \Delta$ . Let  $g \in L^p$ . The weak convergence of  $\eta_n$  to  $\eta$  implies that  $(T_{k\eta_n}(g), n \in \mathbb{N})$  converges  $\mu$ -almost surely to  $T_{k\eta}(g)$ . Consider the function:

$$K(x) = \left( \int_{\Omega} k(x, y)^q \mu(dy) \right)^{1/q},$$

which belongs to  $L^p$ , thanks to (10). Since for all  $x$ ,

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

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

$$(31) \quad \lim_{n \rightarrow \infty} \|T_{k\eta_n}(g) - T_{k\eta}(g)\|_p = 0,$$

so that  $T_{k\eta_n}$  converges strongly to  $T_{k\eta}$ . Using Lemma 3.2 (iv) (with  $T_n = T_{k\eta_n}$  and  $T = T_{k\eta}$ ) on the continuity of the spectrum, we get that  $\lim_{n \rightarrow \infty} \text{Spec}[k](\eta_n) = \text{Spec}[k](\eta)$ . The function  $\text{Spec}[k]$  is thus sequentially continuous, and, thanks to Lemma 3.1, it is continuous from  $\Delta$  endowed with the weak topology to the metric space  $\mathcal{K}$  endowed with the Hausdorff distance. The continuity of  $R_e[k]$  then follows from its definition (8) as the composition of the continuous functions  $\text{rad}$  and  $\text{Spec}[k]$ .  $\square$

We give a stability property of the spectrum and spectral radius with respect to the kernel  $k$ .

**Proposition 4.3** (Stability of  $R_e[k]$  and  $\text{Spec}[k]$ ). *Let  $p \in (1, +\infty)$ . Let  $(k_n, n \in \mathbb{N})$  and  $k$  be kernels on  $\Omega$  with finite double norms on  $L^p$ . If  $\lim_{n \rightarrow \infty} \|k_n - k\|_{p,q} = 0$ , then we have:*

$$(32) \quad \lim_{n \rightarrow \infty} \sup_{\eta \in \Delta} \left| R_e[k_n](\eta) - R_e[k](\eta) \right| = 0 \quad \text{and} \quad \lim_{n \rightarrow \infty} \sup_{\eta \in \Delta} d_{\text{H}} \left( \text{Spec}[k_n](\eta), \text{Spec}[k](\eta) \right) = 0.$$

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

The operators  $\mathcal{A} = \{T_k\} \cup \{T_{k_n} : n \in \mathbb{N}\}$  are compact, and we deduce from (12) that:

$$\lim_{n \rightarrow \infty} \|T_{k_n} - T_k\|_{L^p} = 0.$$

The family  $\mathcal{A}$  is then easily seen to be collectively compact. (Indeed, let  $(y_n = T_{k_{i_n}}(x_n), n \in \mathbb{N})$  be a sequence with  $\|x_n\| \leq 1$ ,  $i_n \in \mathbb{N} \cup \{\infty\}$  and the convention  $k_{i_n} = k$  if  $i_n = \infty$ . Up to taking a sub-sequence, we can assume that either the sequence  $(i_n, n \in \mathbb{N})$  is constant and  $(T_{k_{i_0}}(x_n), n \in \mathbb{N})$  is convergent (as  $T_{k_{i_0}}$  is compact) or that the sequence  $(i_n, n \in \mathbb{N})$  is increasing and the sequence  $(T_k(x_n), n \in \mathbb{N})$  is convergent (as  $T_k$  is compact) towards a limit, say  $y$ . In the former case, clearly the sequence  $(y_n, n \in \mathbb{N})$  converges. In the latter case, we have:  $\|T_{k_{i_n}}(x_n) - y\|_p \leq \|T_{k_{i_n}} - T_k\|_{L^p} + \|T_k(x_n) - y\|_p$ , which readily implies that the sequence  $(y_n, n \in \mathbb{N})$  converges towards  $y$ . This proves that the family  $\mathcal{A}$  is collectively compact.) This implies, see [1, Proposition 4.1(2)] for details, that the family  $\mathcal{A}' = \{T' M_\eta : T' \in \mathcal{A} \text{ and } \eta \in \Delta\}$  is collectively compact. We deduce that the sequence  $(T_n = T_{k_n \eta_n} = T_{k_n} M_{\eta_n}, n \in \mathbb{N})$  of elements of  $\mathcal{A}'$  is collectively compact and that  $T = T_{k\eta} = T_k M_\eta$  is compact.

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

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

Using  $\lim_{n \rightarrow \infty} \|T_{k_n} - T_k\|_{L^p} = 0$  and (31), we get that  $\lim_{n \rightarrow \infty} \|T_n(g) - T(g)\|_p = 0$ , thus  $(T_n, n \in \mathbb{N})$  converges strongly to  $T$ . With Lemma 3.2 (iv), we get that  $\lim_{n \rightarrow \infty} \text{Spec}(T_n) = \text{Spec}(T)$ , that is  $\lim_{n \rightarrow \infty} \text{Spec}[k_n](\eta_n) = \text{Spec}[k](\eta)$ .

Then, as the function  $\eta \mapsto d_{\text{H}} \left( \text{Spec}[k_n](\eta), \text{Spec}[k](\eta) \right)$  is continuous on the compact set  $\Delta$ , thanks to Theorem 4.2, it reaches its maximum say at  $\eta_n \in \Delta$  for  $n \in \mathbb{N}$ . As  $\Delta$  is compact, consider a sub-sequence which converges weakly to a limit say  $\eta$ . Since

$$\begin{aligned} & \sup_{\eta \in \Delta} d_{\text{H}} \left( \text{Spec}[k_n](\eta), \text{Spec}[k](\eta) \right) \\ &= d_{\text{H}} \left( \text{Spec}[k_n](\eta_n), \text{Spec}[k](\eta_n) \right) \\ &\leq d_{\text{H}} \left( \text{Spec}[k_n](\eta_n), \text{Spec}[k](\eta) \right) + d_{\text{H}} \left( \text{Spec}[k](\eta_n), \text{Spec}[k](\eta) \right), \end{aligned}$$

using the continuity of  $\text{Spec}[k]$ , we deduce that along this sub-sequence the right hand side converges to 0. Since this result holds for any converging sub-sequence, we get the second part of (32). The first part then follows from the definition (8) of  $R_e$  as a composition, and the Lipschitz continuity of the function  $\text{rad}$ .  $\square$

**4.2. The asymptotic proportion of infected individuals  $\mathfrak{J}$ .** We consider the SIS model  $[(\Omega, \mathcal{F}, \mu), k, \gamma]$  under Assumption 2. Recall from (23) that the asymptotic proportion of infected individuals  $\mathfrak{J}$  is given on  $\Delta$  by  $\mathfrak{J}(\eta) = \int_{\Omega} \mathfrak{g}_\eta \eta \, d\mu$ , where  $\mathfrak{g}_\eta$  is the maximal solution in  $\Delta$  of the equation  $F_\eta(h) = 0$ . We first give a preliminary result.

**Lemma 4.4.** *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  $g \leq \mathfrak{g}_\eta$ .  $\square$



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

**Proposition 4.5** (Basic properties of  $\mathfrak{J}$ ). *Suppose that Assumption 2 holds. Let  $\eta, \eta_1, \eta_2 \in \Delta$ . The function  $\mathfrak{J}$  has the following properties:*

- (i)  $\mathfrak{J}(\eta_1) = \mathfrak{J}(\eta_2)$  if  $\eta_1 = \eta_2$   $\mu$ -almost surely.
- (ii)  $\mathfrak{J}(\eta) = 0$  if and only if  $R_e[k/\gamma](\eta) \leq 1$ .
- (iii)  $\mathfrak{J}(\eta_1) \leq \mathfrak{J}(\eta_2)$  if  $\eta_1 \leq \eta_2$   $\mu$ -almost surely.
- (iv)  $\mathfrak{J}(\lambda\eta) \leq \lambda\mathfrak{J}(\eta)$  for all  $\lambda \in [0, 1]$ .

*Proof.* If  $\eta_1 = \eta_2$   $\mu$ -almost surely, then the operators  $\mathcal{T}_{k\eta_1}$  and  $\mathcal{T}_{k\eta_2}$  are equal. Thus, the equilibria  $\mathfrak{g}_{\eta_1}$  and  $\mathfrak{g}_{\eta_2}$  are also equal which in turns implies that  $\mathfrak{J}(\eta_1) = \mathfrak{J}(\eta_2)$ . Point (ii) is already stated in Equation (24).

To prove the monotonicity (Point (iii)), consider  $\eta_1 \leq \eta_2$ . Since  $\mathcal{T}_{k\eta_1} \leq \mathcal{T}_{k\eta_2}$ , we get  $F_{\eta_1}(g) \leq F_{\eta_2}(g)$  for all  $g \in \Delta$ . In particular, taking  $g = \mathfrak{g}_{\eta_1}$  and using (22), we get  $F_{\eta_2}(\mathfrak{g}_{\eta_1}) \geq 0$ . By Lemma 4.4 this implies  $\mathfrak{g}_{\eta_1} \leq \mathfrak{g}_{\eta_2}$ . To sum up, we get:

$$(33) \quad \eta_1 \leq \eta_2 \implies \mathfrak{g}_{\eta_1} \leq \mathfrak{g}_{\eta_2}.$$

This readily implies that  $\mathfrak{J}(\eta_1) = \int_{\Omega} \mathfrak{g}_{\eta_1} \eta_1 \, d\mu \leq \int_{\Omega} \mathfrak{g}_{\eta_2} \eta_2 \, d\mu = \mathfrak{J}(\eta_2)$ . We conclude using Point (i).

We now consider Point (iv). Since  $\lambda \in [0, 1]$ , we deduce from (33) that  $\mathfrak{g}_{\lambda\eta} \leq \mathfrak{g}_{\eta}$ . This implies that  $\mathfrak{J}(\lambda\eta) = \int_{\Omega} \mathfrak{g}_{\lambda\eta} \lambda\eta \, d\mu \leq \lambda \int_{\Omega} \mathfrak{g}_{\eta} \eta \, d\mu = \lambda\mathfrak{J}(\eta)$ .  $\square$

The proof of the following continuity results are both postponed to Section 8.1.

**Theorem 4.6** (Continuity of  $\mathfrak{J}$ ). *Suppose that Assumption 2 holds. The function  $\mathfrak{J}$  defined on  $\Delta$  is continuous with respect to the weak topology.*

We write  $\mathfrak{J}[k, \gamma]$  for  $\mathfrak{J}$  to stress the dependence on the parameters  $k, \gamma$  of the SIS model.

**Proposition 4.7** (Stability of  $\mathfrak{J}$ ). *Let  $((k_n, \gamma_n), n \in \mathbb{N})$  and  $(k, \gamma)$  be a sequence of kernels and functions satisfying Assumption 2. Assume furthermore that there exists  $p' \in (1, +\infty)$  such that  $k = \gamma^{-1}k$  and  $(k_n = \gamma_n^{-1}k_n, n \in \mathbb{N})$  have finite double norm in  $L^{p'}$  and that  $\lim_{n \rightarrow \infty} \|k_n - k\|_{p', q'} = 0$ . Then we have:*

$$(34) \quad \lim_{n \rightarrow \infty} \sup_{\eta \in \Delta} \left| \mathfrak{J}[k_n, \gamma_n](\eta) - \mathfrak{J}[k, \gamma](\eta) \right| = 0.$$

## 5. PARETO AND ANTI-PARETO FRONTIERS

5.1. **The setting.** To any vaccination strategy  $\eta \in \Delta$ , we associate a cost and a loss.

- **The cost function.** The cost  $C(\eta)$  measures all the costs of the vaccination strategy (production and diffusion). The cost is expected to be a decreasing function of  $\eta$ , since  $\eta$  encodes the non-vaccinated population. Since doing nothing costs nothing, we also expect  $C(\mathbf{1}) = 0$ , see Assumptions 3 below. We shall also consider natural hypothesis on  $C$ , see Assumptions 4 and 6. A simple cost model is the affine cost given by:

$$(35) \quad C_{\text{aff}}(\eta) = \int_{\Omega} (1 - \eta(x)) c_{\text{aff}}(x) \mu(dx),$$

where  $c_{\text{aff}}(x)$  is the cost of vaccination of population of feature  $x$ , with  $c_{\text{aff}} \in L^1$  positive. The particular case  $c_{\text{aff}} = 1$  is the uniform cost  $C = C_{\text{uni}}$ :

$$(36) \quad C_{\text{uni}}(\eta) = \int_{\Omega} (1 - \eta) \, d\mu.$$

The real cost of the vaccination may be a more complicated function  $\psi(C_{\text{aff}}(\eta))$  of the affine cost, for example if the marginal cost of producing a vaccine depends on the

quantity already produced. However, as long as  $\psi$  is strictly increasing, this will not affect the optimal strategies.

- **The loss function.** The loss  $L(\eta)$  measures the (non)-efficiency of the vaccination strategy  $\eta$ . Different choices are possible here. We prove in this section general results that only depend on a few natural hypothesis for  $L$ ; see Assumptions 3, 5 and 7. These hypothesis are in particular satisfied if the loss is the effective reproduction number  $R_e$  (kernel and SIS models), or the asymptotic proportion of infected individuals  $\mathfrak{I}$  (SIS model); more precisely see Lemmas 5.6, 5.11 and 5.12.

We shall consider cost and loss functions with some regularities.

**Definition 5.1.** *We say that a real-valued function  $H$  defined on  $\Delta$  endowed with the weak topology is:*

- **Continuous:** if  $H$  is continuous with respect to the weak topology on  $\Delta$ .
- **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  $\int_{\Omega} \eta_1 d\mu < \int_{\Omega} \eta_2 d\mu$ , we have  $H(\eta_1) > H(\eta_2)$ .
- **Sub-homogeneous:** if  $H(\lambda\eta) \leq \lambda H(\eta)$  for all  $\eta \in \Delta$  and  $\lambda \in [0, 1]$ .

The definition of non-increasing function and increasing function are similar.

**Assumption 3** (On the cost function and loss function). *The loss function  $L : \Delta \rightarrow \mathbb{R}$  is non-decreasing and continuous with  $L(0) = 0$ . The cost function  $C : \Delta \rightarrow \mathbb{R}$  is non-increasing and continuous with  $C(\mathbb{1}) = 0$ . We also have:*

$$\ell_{\max} := \max_{\Delta} L > 0 \quad \text{and} \quad c_{\max} := \max_{\Delta} C > 0.$$

Assumption 3 will always hold. In particular, the loss and the cost functions are non-negative and non-constant.

We will consider the multi-objective minimization and maximization problems:

$$(37) \quad \begin{cases} \text{Minimize:} & (C(\eta), L(\eta)) \\ \text{subject to:} & \eta \in \Delta \end{cases} \quad \text{and} \quad \begin{cases} \text{Maximize:} & (C(\eta), L(\eta)) \\ \text{subject to:} & \eta \in \Delta \end{cases}$$

Before going further, let us remark that for the reproduction number optimization in the vaccination context, one can without loss of generality consider the uniform cost instead of the affine cost.

*Remark 5.2.* Consider the kernel model  $\text{Param} = [(\Omega, \mathcal{F}, \mu), k]$  with the affine cost function  $C_{\text{aff}}$  and the loss  $R_e$ . Furthermore, if we assume that  $c_{\text{aff}}$  is bounded and bounded away from 0 (that is  $c_{\text{aff}}$  and  $1/c_{\text{aff}}$  belongs to  $\mathcal{L}_+^{\infty}$ ), and without loss of generality, that  $\int c_{\text{aff}} d\mu = 1$ , then we can consider the weighted kernel model  $\text{Param}_0 = [(\Omega, \mathcal{F}, \mu_0), k_0]$  with measure  $\mu_0(dx) = c_{\text{aff}}(x) \mu(dx)$  and kernel  $k_0 = k/c_{\text{aff}}$ . (Notice that if Assumption 2 holds for the model  $\text{Param}$ , then it also holds for the model  $\text{Param}_0$ .) Consider the loss  $L = R_e$ . Then for a strategy  $\eta \in \Delta$ , we get that  $(C_{\text{aff}}(\eta), L(\eta))$  for the model  $\text{Param}$  is equal to  $(C_{\text{uni}}(\eta), L(\eta))$  for the model  $\text{Param}_0$ . Therefore, for the loss function  $L = R_e$ , instead of the affine cost  $C_{\text{aff}}$ , one can consider without any real loss of generality the uniform cost. (This holds also for the SIS model.) However, this is no longer the case for the loss function  $L = \mathfrak{I}$  in the SIS model.

Multi-objective problems are in a sense ill-defined because in most cases, it is impossible to find a single solution that would be optimal to all objectives simultaneously. Hence, we recall the concept of Pareto optimality. Since the minimization problem is crucial for vaccination, we shall define Pareto optimality for the bi-objective minimization problem. A strategy  $\eta_{\star} \in \Delta$  is said to be *Pareto optimal* for the minimization problem in (37) if any improvement of one objective leads to a deterioration of the other, for  $\eta \in \Delta$ :

$$(38) \quad C(\eta) < C(\eta_{\star}) \implies L(\eta) > L(\eta_{\star}) \quad \text{and} \quad L(\eta) < L(\eta_{\star}) \implies C(\eta) > C(\eta_{\star}).$$

Similarly, a strategy  $\eta^* \in \Delta$  is *anti-Pareto optimal* if it is Pareto optimal for the bi-objective maximization problem in (37). Intuitively, the “best” vaccination strategies are the Pareto optima and the “worst” vaccination strategies are the anti-Pareto optima.

We define the *feasible region* as all possible outcomes:

$$\mathbf{F} = \{(C(\eta), L(\eta)), \eta \in \Delta\}.$$

Then, we first consider the minimization problem for the “best” strategies. The set of Pareto optimal strategies will be denoted by  $\mathcal{P}_L$ , and the Pareto frontier is defined as the set of Pareto optimal outcomes:

$$\mathcal{F}_L = \{(C(\eta), L(\eta)) : \eta \text{ Pareto optimal}\}.$$

We consider the minimization problems related to the “best” vaccination strategies, with  $\ell \in [0, \ell_{\max}]$  and  $c \in [0, c_{\max}]$ :

$$\begin{aligned} (39a) \quad & \mathbf{Minimize:} && L(\eta) \\ (39b) \quad & \mathbf{subject\ to:} && \eta \in \Delta, C(\eta) \leq c, \end{aligned}$$

as well as

$$\begin{aligned} (40a) \quad & \mathbf{Minimize:} && C(\eta) \\ (40b) \quad & \mathbf{subject\ to:} && \eta \in \Delta, L(\eta) \leq \ell. \end{aligned}$$

We denote the values of Problems (39) and (40) by:

$$\begin{aligned} L_\star(c) &= \inf\{L(\eta) : \eta \in \Delta \text{ and } C(\eta) \leq c\} \quad \text{for } c \in [0, c_{\max}], \\ C_\star(\ell) &= \inf\{C(\eta) : \eta \in \Delta \text{ and } L(\eta) \leq \ell\} \quad \text{for } \ell \in [0, \ell_{\max}]. \end{aligned}$$

We now consider the maximization problem related to the “worst” vaccination strategies, with  $\ell \in [0, \ell_{\max}]$  and  $c \in [0, c_{\max}]$ :

$$\begin{aligned} (41a) \quad & \mathbf{Maximize:} && L(\eta) \\ (41b) \quad & \mathbf{subject\ to:} && \eta \in \Delta, C(\eta) \geq c, \end{aligned}$$

as well as

$$\begin{aligned} (42a) \quad & \mathbf{Maximize:} && C(\eta) \\ (42b) \quad & \mathbf{subject\ to:} && \eta \in \Delta, L(\eta) \geq \ell. \end{aligned}$$

We denote the values of Problems (41) and (42) by:

$$\begin{aligned} L^\star(c) &= \sup\{L(\eta) : \eta \in \Delta \text{ and } C(\eta) \geq c\} \quad \text{for } c \in [0, c_{\max}], \\ C^\star(\ell) &= \sup\{C(\eta) : \eta \in \Delta \text{ and } L(\eta) \geq \ell\} \quad \text{for } \ell \in [0, \ell_{\max}]. \end{aligned}$$

We denote by  $\mathcal{P}_L^{\text{Anti}}$  the set of anti-Pareto optimal strategies, and by  $\mathcal{F}_L^{\text{Anti}}$  its frontier:

$$\mathcal{F}_L^{\text{Anti}} = \{(C(\eta), L(\eta)) : \eta \text{ anti-Pareto optimal}\}.$$

If necessary, we may write  $C_{\star, L}$  and  $C^{\star, L}$  to stress the dependence of the function  $C_\star$  and  $C^\star$  in the loss function  $L$ .

Under Assumption 3, as the loss and the cost functions are continuous on the compact set  $\Delta$ , the infima in the definitions of the value functions  $C_\star$  and  $L_\star$  are minima; and the suprema in the definition of the value functions  $C^\star$  and  $L^\star$  are maxima. Since  $\Delta$  is endowed with the weak topology, we will consider the set of Pareto and anti-Pareto optimal vaccination modulo  $\mu$ -almost sure equality.

See Figure 2 for a typical representation of the possible aspects of the feasible region  $\mathbf{F}$  (in light blue), the value functions and the Pareto and anti-Pareto frontiers under the general Assumption 3, and the connected Pareto and anti-Pareto frontiers under further regularity on the cost and loss functions (see Assumption 4-7 below) in Figure 2(D). In Figure 1(B), we

have plotted in solid red line the Pareto frontier and in dashed red line the anti-Pareto frontier from Example 1.7.

*Outline of the section.* It turns out that the anti-Pareto optimization problem can be recast as a Pareto optimization problem by changing signs and exchanging the cost and loss functions. In order to make use of this property for the kernel and SIS models, we study the Pareto problem under assumptions on the cost that are general enough to cover the choices  $C_{\text{uni}}$  and  $-L$ , and assumptions on the loss that cover the choices  $R_e$ ,  $\mathcal{J}$  and  $-C_{\text{uni}}$ .

The main result of this section states that all the solutions of the optimization Problems (39) or (40) are Pareto optimal, and gives a description of the Pareto frontier  $\mathcal{F}_L$  as a graph in Section 5.2, and similarly for the anti-Pareto frontier in Section 5.3. Surprisingly, the problem is not completely symmetric, compare Lemma 5.6 used for the Pareto frontier and Lemmas 5.11 and 5.12 used for the anti-Pareto frontier. In the latter lemmas, notice the kernel considered is quasi-irreducible, whereas this condition is not needed for the Pareto frontier.

**5.2. On the Pareto frontier.** We first check that Problems (39) and (40) have solutions.

**Proposition 5.3** (Optimal solutions for fixed cost or fixed loss). *Suppose that Assumption 3 holds. For any cost  $c \in [0, c_{\max}]$ , there exists a minimizer of the loss under the cost constraint  $C(\cdot) \leq c$ , that is, a solution to Problem (39). Similarly, for any loss  $\ell \in [0, \ell_{\max}]$ , there exists a minimizer of the cost under the loss constraint  $L(\cdot) \leq \ell$ , that is a solution to Problem (40).*

*Proof.* Let  $c \in [0, c_{\max}]$ . The set  $\{\eta \in \Delta : C(\eta) \leq c\}$  is non-empty as it contains  $\mathbf{1}$  since  $C(\mathbf{1}) = 0$ . It is also compact as  $C$  is continuous on the compact set  $\Delta$  (for the weak topology). Therefore, since the loss function  $L$  is continuous (for the weak topology), we get that  $L$  restricted to this compact set reaches its minimum. Thus, Problem (39) has a solution. The proof is similar for the existence of a solution to Problem (40).  $\square$

We start by a general result concerning the links between the three problems.

**Proposition 5.4** (Single-objective and bi-objective problems). *Suppose Assumption 3 holds.*

- (i) *If  $\eta_\star$  is Pareto optimal, then  $\eta_\star$  is a solution of (39) for the cost  $c = C(\eta_\star)$ , and a solution of (40) for the loss  $\ell = L(\eta_\star)$ . Conversely, if  $\eta_\star$  is a solution to both problems (39) and (40) for some values  $c$  and  $\ell$ , then  $\eta_\star$  is Pareto optimal.*
- (ii) *The Pareto frontier is the intersection of the graphs of  $C_\star$  and  $L_\star$ :*

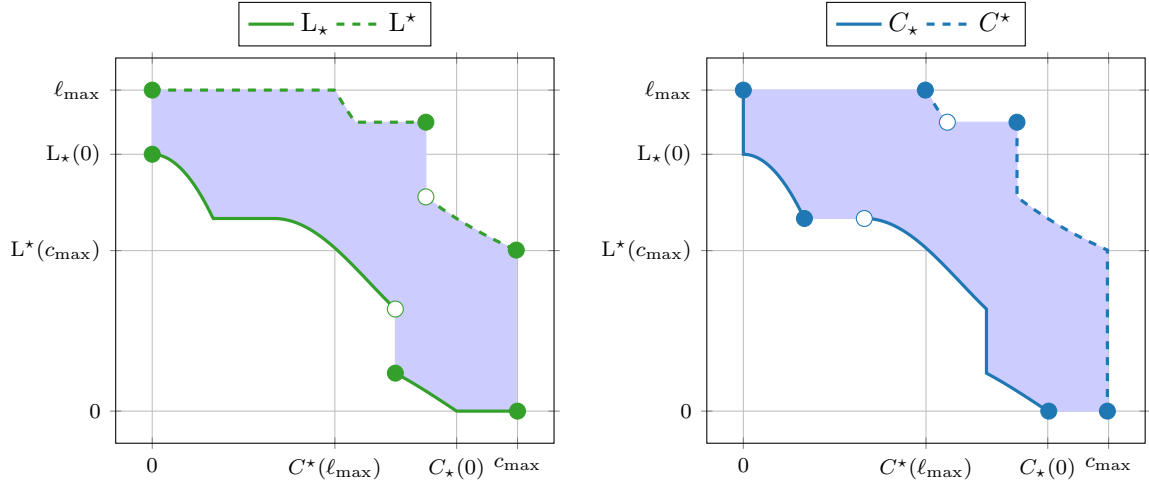
$$\mathcal{F}_L = \{(c, \ell) \in [0, c_{\max}] \times [0, \ell_{\max}] : c = C_\star(\ell) \text{ and } \ell = L_\star(c)\}.$$

- (iii) *The points  $(0, L_\star(0))$  and  $(C_\star(0), 0)$  both belong to the Pareto frontier, and we have  $C_\star(L_\star(0)) = L_\star(C_\star(0)) = 0$ . Moreover, we also have  $C_\star(\ell) = 0$  for  $\ell \in [L_\star(0), \ell_{\max}]$ , and  $L_\star(c) = 0$  for  $c \in [C_\star(0), c_{\max}]$ .*

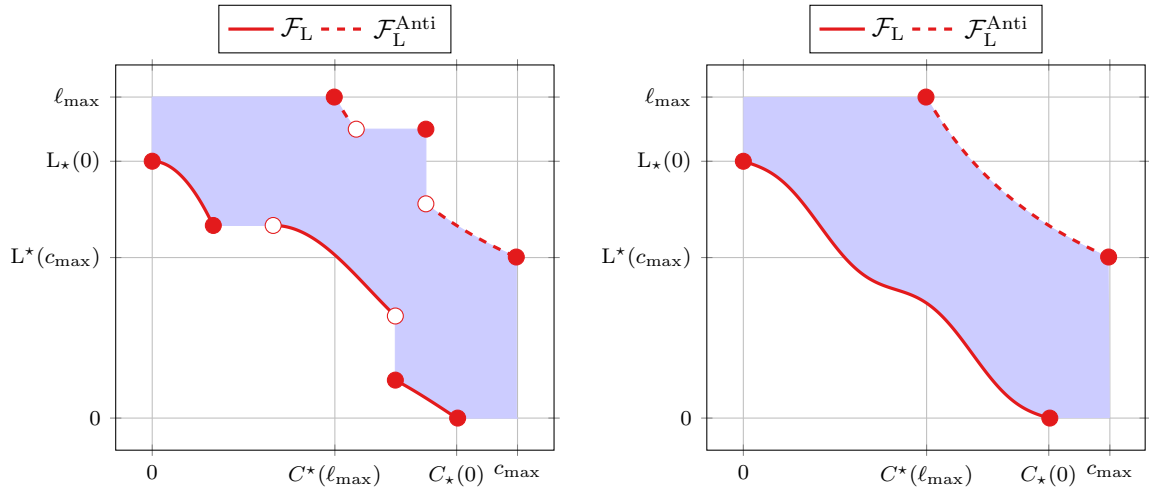
*Proof.* Let us prove (i). If  $\eta_\star$  is Pareto optimal, then for any strategy  $\eta$ , if  $C(\eta) \leq C(\eta_\star)$  then  $L(\eta) \geq L(\eta_\star)$  by taking the contraposition in (38), and  $\eta_\star$  is indeed a solution of Problem (39) with  $c = C(\eta_\star)$ . Similarly  $\eta_\star$  is a solution of Problem (40).

For the converse statement, let  $\eta_\star$  be a solution of (39) for some  $c$  and of (40) for some  $\ell$ . It is also a solution of (39) with  $c = C(\eta_\star)$ . In particular, we get that for  $\eta \in \Delta$ ,  $L(\eta) < L(\eta_\star)$  implies that  $C(\eta) > c = C(\eta_\star)$ , which is the second part of (38). Similarly, use that  $\eta_\star$  is a solution to (40), to get that the first part of (38) also holds. Thus the strategy  $\eta_\star$  is Pareto optimal.

To prove Point (ii), we first prove that  $\mathcal{F}_L$  is a subset of  $\{(c, \ell) : c = C_\star(\ell) \text{ and } \ell = L_\star(c)\}$ . A point in  $\mathcal{F}_L$  may be written as  $(C(\eta_\star), L(\eta_\star))$  for some Pareto optimal strategy  $\eta_\star$ . By Point (i),  $\eta_\star$  solves Problem (39) for the cost  $C(\eta_\star)$ , so  $L_\star(C(\eta_\star)) = L(\eta_\star)$ . Similarly, we have  $C_\star(L(\eta_\star)) = C(\eta_\star)$ , as claimed.



(A) Value functions for Problems (39) and (41). (B) Value functions for Problems (40) and (42).



(C) Pareto and anti-Pareto frontier.

(D) Pareto and anti-Pareto frontier under additional regularity Assumptions 4-7.

FIGURE 2. An example of the possible aspects of the feasible region  $\mathbf{F}$  (in light blue), the value functions  $L_*$ ,  $L^*$ ,  $C_*$ ,  $C^*$ , and the Pareto and anti-Pareto frontier (in red) under Assumption 3.

We now prove the reverse inclusion. Assume that  $c = C_*(\ell)$  and  $\ell = L_*(c)$ , and consider  $\eta$  a solution of Problem (40) for the loss  $\ell$ :  $L(\eta) \leq \ell$  and  $C(\eta) = C_*(\ell) = c$ . Then  $\eta$  is admissible for Problem (39) with cost  $c = C_*(\ell)$ , so  $L(\eta) \geq L_*(C_*(\ell)) = L_*(c) = \ell$ . Therefore, we get  $L(\eta) = L^*(c)$ , and  $\eta$  is also a solution of Problem (39). By Point (i),  $\eta$  is Pareto optimal, so  $(C(\eta), L(\eta)) = (c, \ell) \in \mathcal{F}_L$ , and the reverse inclusion is proved.

Finally we prove Point (iii). We have  $C_*(0) = \min\{C(\eta) : \eta \in \Delta \text{ and } L(\eta) = 0\} \in [0, c_{\max}]$ . Let  $\eta \in \Delta$  such that  $L(\eta) = 0$  and  $C(\eta) = C_*(0)$ . We deduce that  $L_*(C_*(0)) \leq L(\eta) = 0$  and thus  $L_*(C_*(0)) = 0$  as  $L$  is non-negative. We deduce from (ii) that  $(C_*(0), 0)$  belongs to  $\mathcal{F}_L$ . Since  $C_*$  is non-increasing, we also get that  $C_* = 0$  on  $[C_*(0), c_{\max}]$ . The other properties of (iii) are proved similarly.  $\square$

The next two hypotheses on  $C$  and  $L$  will imply that the Pareto frontier is connected.

**Assumption 4.** *If the cost  $C$  has a local minimum (for the weak topology) at  $\eta$ , then  $C(\eta) = 0$  and  $\eta$  is a global minimum of  $C$ .*

**Assumption 5.** *If the loss  $L$  has a local minimum (for the weak topology) at  $\eta$ , then  $L(\eta) = 0$  and  $\eta$  is a global minimum of  $L$ .*

Under these hypotheses, the picture becomes much nicer, see Figure 2(D), where the only flat parts of the graphs of  $C_\star$  and  $L_\star$  occur at zero cost or zero loss.

**Proposition 5.5.** *Under Assumption 3 and 4 the following properties hold:*

(i) *The optimal cost  $C_\star$  is decreasing on  $[0, L_\star(0)]$ .*

(ii) *If  $\eta$  solves Problem (40) for the loss  $\ell \in [0, L_\star(0)]$ , then  $L(\eta) = \ell$  (that is, the constraint is binding). Moreover  $\eta$  is Pareto optimal, and:*

$$(43) \quad L_\star(C_\star(\ell)) = \ell.$$

(iii) *The Pareto frontier is the graph of  $C_\star$ :*

$$(44) \quad \mathcal{F}_L = \{(C_\star(\ell), \ell) : \ell \in [0, L_\star(0)]\}.$$

Similarly, under Assumptions 3 and 5, the following properties hold:

(iv) *The optimal loss  $L_\star$  is decreasing on  $[0, C_\star(0)]$ .*

(v) *If  $\eta$  solves Problem (39) for the cost  $c \in [0, C_\star(0)]$ , then  $C(\eta) = c$ . Moreover  $\eta$  is Pareto optimal, and  $C_\star(L_\star(c)) = c$ .*

(vi) *The Pareto frontier is the graph of  $L_\star$ :*

$$(45) \quad \mathcal{F}_L = \{(c, L_\star(c)) : c \in [0, C_\star(0)]\}.$$

Finally, if Assumptions 4 and 5 hold, then  $L_\star$  is a continuous decreasing bijection of  $[0, C_\star(0)]$  onto  $[0, L_\star(0)]$  and  $C_\star$  is the inverse bijection, and the Pareto frontier is compact and connected.

*Proof.* We prove (i). Let  $0 \leq \ell < \ell' \leq L_\star(0)$ , and let  $\eta_\star$  be a solution of Problem (40):

$$(46) \quad C(\eta_\star) = C_\star(\ell) \quad \text{and} \quad L(\eta_\star) \leq \ell.$$

The set  $\mathcal{O} = \{\eta : L(\eta) < \ell'\}$  is open and contains  $\eta_\star$ . Since  $L(\eta_\star) < L_\star(0)$ , we get  $C(\eta_\star) > 0$ , so  $\eta_\star$  is not a global minimum for  $C$ . By Assumption 4, it cannot be a local minimum for  $C$ , so  $\mathcal{O}$  contains at least one point  $\eta'$  for which  $C(\eta') < C(\eta_\star)$ . Since  $\eta' \in \mathcal{O}$ , we get  $L(\eta') \leq \ell'$ , so that  $C_\star(\ell') \leq C(\eta') < C(\eta_\star) = C_\star(\ell)$ . Since  $\ell < \ell'$  are arbitrary,  $C_\star$  is decreasing on  $[0, L_\star(0)]$ .

We now prove (ii). If the inequality in (46) was strict, that is  $L(\eta_\star) < \ell$ , then we would get a contradiction as  $C(\eta_\star) \geq C_\star(L(\eta_\star)) > C_\star(\ell) = C(\eta_\star)$ . Therefore any solution  $\eta_\star$  of (40) satisfies  $L(\eta_\star) = \ell$ , and in particular  $C_\star(L(\eta_\star)) = C_\star(\ell) = C(\eta_\star)$ . This implies in turn that  $\eta_\star$  also solves (39): if  $\eta$  satisfies  $L(\eta) < L(\eta_\star)$ , then using the definition of  $C_\star$ , the fact that it decreases, and the definition of  $\eta_\star$ , we get:

$$C(\eta) \geq C_\star(L(\eta)) > C_\star(L(\eta_\star)) = C(\eta_\star).$$

By contraposition, we have  $L(\eta) \geq L(\eta_\star)$  for any  $\eta$  such that  $C(\eta) \leq C(\eta_\star)$ , proving that  $\eta_\star$  is also a solution of (39) with  $c = C(\eta_\star)$ . By Point (i) of Proposition 5.4,  $\eta_\star$  is Pareto optimal. Therefore  $(C(\eta_\star), L(\eta_\star)) = (C_\star(\ell), \ell)$  belongs to the Pareto frontier. Using Point (ii) of Proposition 5.4, we deduce that  $\ell = L_\star(C_\star(\ell))$ .

To prove Point (iii), note that Equation (43) shows that, if  $c = C_\star(\ell)$  for  $\ell \in [0, L_\star(0)]$ , then  $\ell = L_\star(c)$ . Use Point (ii) and (iii) of Proposition 5.4, to get that  $\mathcal{F}_L = \{(c, \ell) : c = C_\star(\ell), \ell \in [0, L_\star(0)]\}$ .

The claims (iv), (v) and (vi) are proved in the same way, exchanging the roles of  $L$  and  $C$ .

To conclude the proof, it remains to check that  $C_\star$  and  $L_\star$  are continuous under Assumptions 3, 4 and 5. We deduce from Point (ii) and Proposition 5.3 that  $[0, L_\star(0)]$  is in the range of  $L_\star$ .

Since  $L_\star$  is decreasing, thanks to Point (iv) and  $L_\star(C_\star(0)) = 0$ , see Proposition 5.4 (iii), we get that  $L_\star$  is continuous and decreasing on  $[0, L_\star(0)]$ , and thus one-to-one from  $[0, C_\star(0)]$  onto  $[0, L_\star(0)]$ . Then use (43) to get that  $C_\star$  is its inverse bijection. The continuity of  $L_\star$  and (45) implies that  $\mathcal{F}_L$  is compact and connected.  $\square$

Finally, let us check that Assumptions 4 and 5 hold under very simple assumptions, which are in particular satisfied by the cost functions  $C_{\text{uni}}$  and  $C_{\text{aff}}$  and the loss functions  $R_e$  and  $\mathfrak{J}$  (recall from Propositions 4.1 and 4.5 that  $R_e$  and  $\mathfrak{J}$  are sub-homogeneous).

**Lemma 5.6.** *Suppose Assumption 3 holds. If the cost function  $C$  is decreasing, then Assumption 4 holds and  $L_\star(0) = \ell_{\max}$ . If the loss function  $L$  is sub-homogeneous, then Assumption 5 holds.*

*Proof.* Let  $\eta \in \Delta$ . If  $C$  has a local minimum at  $\eta$ , then, as  $C$  is non-increasing, for  $\varepsilon > 0$  small enough, we get that  $C(\eta) \geq C(\eta + \varepsilon(1 - \eta)) \geq C(\eta)$ . If  $C$  is decreasing, this is only possible if  $\eta = 1$  almost surely, so that  $\eta$  is a global minimum of  $C$ . This also gives  $L_\star(0) = \ell_{\max}$ . Similarly if  $L$  has a local minimum at  $\eta$ , then for  $\varepsilon > 0$  small enough  $L(\eta) \leq L((1 - \varepsilon)\eta) \leq (1 - \varepsilon)L(\eta)$ , so  $L(\eta) = 0$  and  $\eta$  is a global minimum of  $L$ .  $\square$

**Corollary 5.7.** *Suppose that Assumptions 3, 4 and 5 hold. The set of Pareto optimal strategies  $\mathcal{P}_L$  is compact (for the weak topology).*

*Proof.* Since  $L_\star$  is continuous thanks to Proposition 5.5, we deduce that  $\mathcal{F}_L$ , which is given by (45), is compact and thus closed. Since  $\mathcal{P}_L = f^{-1}(\mathcal{F}_L)$ , where the function  $f = (C, L)$  defined on  $\Delta$  is continuous, we deduce that  $\mathcal{P}_L$  is closed and thus compact as  $\Delta$  is compact.  $\square$

### 5.3. On the anti-Pareto frontier.

5.3.1. *The general setting.* Letting  $C'(\eta) = \ell_{\max} - L(\eta)$  and  $L'(\eta) = c_{\max} - C(\eta)$ , it is easy to see that:

$$C'_\star(c) = \ell_{\max} - L^\star(c_{\max} - c) \quad \text{and} \quad L'_\star(\ell) = c_{\max} - C^\star(\ell_{\max} - \ell),$$

so that Proposition 5.5 may be applied to the cost function  $C'$  and the loss function  $L'$  to yield the following result.

**Proposition 5.8** (Single-objective and bi-objective problems for the anti-Pareto strategies). *Suppose Assumption 3 holds.*

- (i) *If  $\eta^\star$  is anti-Pareto optimal, then  $\eta^\star$  is a solution of (41) for the cost  $c = C(\eta^\star)$ , and a solution of (42) for the loss  $\ell = L(\eta^\star)$ . Conversely, if  $\eta^\star$  is a solution to both problems (41) and (42) for some values  $c$  and  $\ell$ , then  $\eta^\star$  is anti-Pareto optimal.*
- (ii) *The anti-Pareto frontier is the intersection of the graphs of  $C^\star$  and  $L^\star$ :*

$$\mathcal{F}_L^{\text{Anti}} = \{(c, \ell) \in [0, c_{\max}] \times [0, \ell_{\max}] : c = C^\star(\ell) \text{ and } \ell = L^\star(c)\}.$$

- (iii) *The points  $(C^\star(\ell_{\max}), \ell_{\max})$  and  $(c_{\max}, L^\star(c_{\max}))$  both belong to the anti-Pareto frontier, and we have  $C^\star(L^\star(c_{\max})) = c_{\max}$  and  $L^\star(C^\star(\ell_{\max})) = \ell_{\max}$ . Moreover, we also have  $C^\star(\ell) = c_{\max}$  for  $\ell \in [0, L^\star(c_{\max})]$ , and  $L^\star(c) = \ell_{\max}$  for  $c \in [0, C^\star(\ell_{\max})]$ .*

The following additional hypotheses rule out the occurrence of flat parts in the anti-Pareto frontier.

**Assumption 6.** *If the cost  $C$  has a local maximum at  $\eta$  (for the weak topology), then  $C(\eta) = c_{\max}$  and  $\eta$  is a global maximum of  $C$ .*

**Assumption 7.** *If the loss  $L$  has a local maximum at  $\eta$  (for the weak topology), then  $L(\eta) = \ell_{\max}$  and  $\eta$  is a global maximum of  $L$ .*

The following result is now a consequence of Proposition 5.5 and Corollary 5.7 applied to the loss function  $L'$  and cost function  $C'$ .

**Proposition 5.9.** *Under Assumption 3 and 6 the following properties hold:*

- (i) *The optimal cost  $C^*$  is decreasing on  $[C^*(\ell_{\max}), c_{\max}]$ .*
- (ii) *If  $\eta$  solves Problem (42) for the loss  $\ell \in [L^*(c_{\max}), \ell_{\max}]$ , then  $L(\eta) = \ell$  (that is, the constraint is binding). Moreover  $\eta$  is anti-Pareto optimal, and  $L^*(C^*(\ell)) = \ell$ .*
- (iii) *The anti-Pareto frontier is the graph of  $C^*$ :*

$$(47) \quad \mathcal{F}_L^{\text{Anti}} = \{(C^*(\ell), \ell) : \ell \in [L^*(c_{\max}), \ell_{\max}]\}.$$

Similarly, under Assumptions 3 and 7, the following properties hold:

- (iv) *The optimal loss  $L^*$  is decreasing on  $[L^*(c_{\max}), \ell_{\max}]$ .*
- (v) *If  $\eta$  solves Problem (41) for the cost  $c \in [C^*(\ell_{\max}), c_{\max}]$ , then  $C(\eta) = c$ . Moreover  $\eta$  is anti-Pareto optimal, and  $C^*(L^*(c)) = c$ .*
- (vi) *The anti-Pareto frontier is the graph of  $L^*$ :*

$$(48) \quad \mathcal{F}_L^{\text{Anti}} = \{(c, L^*(c)) : c \in [C^*(\ell_{\max}), c_{\max}]\}.$$

Finally, if Assumptions 3, 6 and 7 hold, then  $L^*$  is a continuous decreasing bijection of  $[C^*(\ell_{\max}), c_{\max}]$  onto  $[L^*(c_{\max}), \ell_{\max}]$ ,  $C^*$  is the inverse bijection, and the anti-Pareto frontier is compact and connected. Furthermore, the set of anti-Pareto optimal strategies  $\mathcal{P}_L^{\text{Anti}}$  is compact (for the weak topology).

The following result is similar to the first part of Lemma 5.6.

**Lemma 5.10.** *Suppose Assumption 3 holds. If the cost function  $C$  is decreasing, then Assumption 6 holds and  $L^*(c_{\max}) = 0$ .*

*Proof.* Let  $\eta \in \Delta$  and  $\varepsilon \in (0, 1)$ . Since  $C$  is decreasing,  $C((1 - \varepsilon)\eta) \geq C(\eta)$ , with equality if and only if  $\eta = 0$   $\mu$ -almost surely. Therefore the only local maximum of  $C$  is  $\eta = 0$ , and it is a global maximum. Since  $C(\eta) = c_{\max}$  implies that  $\eta = 0$   $\mu$ -almost surely, we also get that  $L^*(c_{\max}) = L(0) = 0$ .  $\square$

5.3.2. *The particular case of the kernel and SIS models.* We show that, under an irreducibility hypothesis on the kernel, Assumption 7 holds for the loss functions  $R_e$  and  $\mathfrak{J}$ . The reducible case is more delicate and it is studied in more details in [8] for the loss function  $L = R_e$ ; in particular Assumption 7 may not hold and the anti-Pareto frontier may not be connected.

Let us recall some notation. Let  $k$  be a kernel with finite double norm. For  $A, B \in \mathcal{F}$ , we write  $A \subset B$  a.s. if  $\mu(B \cap A^c) = 0$  and  $A = B$  a.s. if  $A \subset B$  a.s. and  $B \subset A$  a.s. For  $A, B \in \mathcal{F}$ ,  $x \in \Omega$  and a kernel  $k$ , we simply write  $k(x, A) = \int_A k(x, y) \mu(dy)$ ,  $k(B, x) = \int_B k(z, x) \mu(dz)$  and:

$$k(B, A) = \int_{B \times A} k(z, y) \mu(dz) \mu(dy).$$

A set  $A \subset \mathcal{F}$  is  $k$ -invariant if  $k(A^c, A) = 0$ . (Notice that if  $A$  is  $k$ -invariant, then  $L^p(A, \mu)$  is an invariant closed subspace for  $T_k$ , seen as an operator on  $L^p(\Omega, \mu)$ .) A kernel  $k$  is *irreducible* (or *connected*) if any  $k$ -invariant set  $A$  is such that a.s.  $A = \emptyset$  or a.s.  $A = \Omega$ . Define  $\{k \equiv 0\}$  as  $\{x \in \Omega : k(x, \Omega) + k(\Omega, x) = 0\}$ , so that  $k(A, \Omega) + k(\Omega, A) = 0$  implies that a.s.  $A \subset \{k \equiv 0\}$ . A kernel  $k$  is *quasi-irreducible* if the restriction of  $k$  to  $\{k \equiv 0\}^c$  is irreducible, that is if any  $k$ -invariant set  $A$  is such that  $A \subset \{k \equiv 0\}$  a.s. or  $A^c \subset \{k \equiv 0\}$  a.s. Notice the definition of the quasi-irreducibility from [3, Definition 2.11] is slightly stronger as it uses a topology on  $\Omega$ .

**Lemma 5.11.** *Consider the kernel model  $\text{Param} = [(\Omega, \mathcal{F}, \mu), k]$  under Assumption 1. If  $k$  is quasi-irreducible, then Assumption 7 holds for  $L = R_e[k]$ , and  $C^*(\ell_{\max}) = C(\mathbb{1}_{\{k \equiv 0\}^c})$  (which is 0 if  $k$  is irreducible).*

*Proof.* The quasi-irreducible case can easily be deduced from the irreducible case, so we assume that  $k$  is irreducible. In particular, we have  $k(\Omega, y) > 0$  almost surely. Let  $\eta \in \Delta$  be a local maximum of  $R_e$  on  $\Delta$ ; we want to show that it is also a global maximum.



Suppose first that  $\inf \eta > 0$ . Then  $k\eta$  is irreducible with finite double norm. According [41, Theorem V.6.6 and Example V.6.5.b], the eigenspace of  $T_{k\eta}$  associated to  $R_e(\eta)$  is one-dimensional and it is spanned by a vector  $v_d$  such that  $v_d > 0$  almost surely, and the corresponding left eigenvector associated to  $R_e(\eta)$ , say  $v_g$ , can be chosen such that  $\langle v_g, v_d \rangle = 1$  and  $v_g > 0$  almost surely. According to [31, Theorem 2.6], applied to  $L_0 = T_{k\eta}$  and  $L = T_{k(\eta+\varepsilon(1-\eta))}$ , we have, using that  $\|L_0 - L\| = O(\varepsilon)$  thanks to (12):

$$R_e(\eta + \varepsilon(1 - \eta)) = R_e(\eta) + \varepsilon \langle v_g, T_{k(1-\eta)} v_d \rangle + O(\varepsilon^2).$$

Since  $R_e$  has a local maximum at  $\eta$ , the first order term on the right hand side vanishes, so  $v_g(x)k(x, y)(1 - \eta(y))v_d(y) = 0$  for  $\mu$  almost all  $x$  and  $y$ . Since  $v_g$  and  $v_d$  are positive almost surely and  $k$  is irreducible, we get that  $k(\Omega, y)(1 - \eta(y)) = 0$  almost surely and thus  $\eta(y) = 1$  almost surely. Therefore  $\eta = \mathbf{1}$ , which is a global maximum for  $R_e$ .

Finally, suppose that  $\inf \eta = 0$ . Let  $\mathcal{O}$  be an open subset of  $\Delta$  on which  $R_e \leq R_e(\eta)$  and with  $\eta \in \mathcal{O}$ . For  $\varepsilon > 0$  small enough, the strategy  $\eta_\varepsilon = \eta + \varepsilon(1 - \eta)$  belongs to  $\mathcal{O}$  and satisfies  $R_e(\eta) \leq R_e(\eta_\varepsilon) \leq R_e(\eta)$  (where the first inequality comes from the fact that  $R_e$  is non-decreasing). Therefore  $\eta_\varepsilon$  is a local maximum, and thus  $\eta_\varepsilon = \mathbf{1}$  almost surely. This readily implies that  $\eta = \mathbf{1}$  almost surely.

We deduce that if  $\eta$  is a local maximum, then  $\eta = 1$  almost surely. Thus  $\eta$  is a global maximum and  $C^*(\ell_{\max}) = C(\mathbf{1}) = 0$ . This ends the proof.  $\square$

**Lemma 5.12.** *Consider the SIS model  $\text{Param} = [(\Omega, \mathcal{F}, \mu), k, \gamma]$  under Assumption 2. If  $k$  is quasi-irreducible, then Assumption 7 holds for  $L = \mathfrak{J}$  and  $C^*(\ell_{\max}) = C(\mathbf{1}_{\{k=0\}^c})$  (which is 0 if  $k$  is irreducible).*

*Proof.* The quasi-irreducible case can easily be deduced from the irreducible case, so we assume that  $k$  is irreducible.

Set  $k = k/\gamma$ . Suppose that  $\mathfrak{J}$  has a local maximum at some  $\eta \in \Delta$ . For  $\varepsilon \in (0, 1)$ , the kernel  $k\eta_\varepsilon$ , with  $\eta_\varepsilon = \eta + \varepsilon(1 - \eta)$ , is irreducible (with finite double norm) since  $k$  is irreducible and  $\gamma$  is positive and bounded. We have that for  $\varepsilon > 0$  small enough:

$$\mathfrak{J}(\eta) \geq \mathfrak{J}(\eta_\varepsilon) = \int_{\Omega} \mathfrak{g}_{\eta_\varepsilon} \eta_\varepsilon d\mu \geq \int_{\Omega} \mathfrak{g}_{\eta_\varepsilon} \eta d\mu \geq \int_{\Omega} \mathfrak{g}_\eta \eta d\mu = \mathfrak{J}(\eta),$$

where we used that  $\eta \leq \eta_\varepsilon$  and  $0 \leq \mathfrak{g}_\eta \leq \mathfrak{g}_{\eta_\varepsilon}$ , see (33). Therefore all these quantities are equal. Since the equilibrium  $\mathfrak{g}_{\eta_\varepsilon}$  is  $\mu$ -a.e. positive thanks to [7, Remark 4.11] as  $k\eta_\varepsilon$  is irreducible, we must have  $\eta_\varepsilon = \eta$  a.s., which is only possible if  $\eta = 1$  almost surely.

Since  $\mathfrak{J}(\mathbf{1}) > \mathfrak{J}(\eta)$  for any  $\eta \neq \mathbf{1}$ , we also get  $C^*(\ell_{\max}) = 0$ , with  $\ell_{\max} = \mathfrak{J}(\mathbf{1}) = \int_{\Omega} \mathfrak{g} d\mu$ .  $\square$

## 6. MISCELLANEOUS PROPERTIES FOR SET OF OUTCOMES AND THE PARETO FRONTIER

We prove results concerning the feasible region, the stability of the Pareto frontier and its geometry.

**6.1. No holes in the feasible region.** We check there is no hole in the feasible region.

**Proposition 6.1.** *Suppose that Assumption 3 holds. The feasible region  $\mathbf{F}$  is compact, path connected, and its complement is connected in  $\mathbb{R}^2$ . It is the whole region between the graphs of the one-dimensional value functions:*

$$\begin{aligned} \mathbf{F} &= \{(c, \ell) \in \mathbb{R}^2 : 0 \leq c \leq c_{\max}, L_\star(c) \leq \ell \leq L^\star(c)\} \\ &= \{(c, \ell) \in \mathbb{R}^2 : 0 \leq \ell \leq \ell_{\max}, C_\star(\ell) \leq c \leq C^\star(\ell)\}. \end{aligned}$$

*Proof.* The region  $\mathbf{F}$  is compact and path-connected as a continuous image by  $(C, L)$  of the compact, path-connected set  $\Delta$ .

By symmetry, it is enough to prove that  $\mathbf{F}$  is equal to  $F_1 = \{(c, \ell) \in \mathbb{R}^2 : 0 \leq c \leq c_{\max}, L_*(c) \leq \ell \leq L^*(c)\}$ . Let  $(c, \ell) \in \mathbf{F}$  and  $\eta \in \Delta$  be such that  $(c, \ell) = (C(\eta), L(\eta))$ . By definition of  $L_*$  and  $L^*$ , we have:  $L_*(c) = L_*(C(\eta)) \leq L(\eta) \leq L^*(C(\eta)) = L^*(c)$ . We deduce that  $(c, \ell) \in F_1$ .

Let us now prove that  $F_1 \subset \mathbf{F}$ . Let us first consider a point of the form  $(c, L_*(c))$ , where  $0 \leq c \leq c_{\max}$ . By definition, there exists  $\eta$  such that  $C(\eta) \leq c$  and  $L(\eta) = L_*(c)$ . Let  $\eta_t = t\eta$ . The map  $t \mapsto C(\eta_t)$  is continuous from  $[0, 1]$  to  $[C(\eta), c_{\max}]$ , and  $c \in [C(\eta), c_{\max}]$ , so there exists  $s$  such that  $C(\eta_s) = c$ . Since  $L$  is non-decreasing,  $L(\eta_s) \leq L(\eta)$ . By definition of  $L_*(c)$ ,  $L(\eta_s) \geq L_*(c)$ . Therefore  $(c, L_*(c)) = (C(\eta_s), L(\eta_s))$  belongs to  $\mathbf{F}$ . Similarly the graphs of  $C_*$ ,  $C^*$  and  $L^*$  are also included in  $\mathbf{F}$ .

So, it is enough to check that, if  $A = (c, \ell)$  is in  $F_1$ , with  $c \in (0, c_{\max})$  and  $\ell \in (L_*(c), L^*(c))$ , then  $A$  belongs to  $\mathbf{F}$ . We shall assume that  $A \notin \mathbf{F}$  and derive a contradiction by building a loop in  $\mathbf{F}$  that encloses  $A$  and which can be continuously contracted into a point in  $\mathbf{F}$ .

Since  $L_*(c) < \ell < L^*(c)$ , there exist  $\eta_{SO}$  and  $\eta_{NE}$  such that:

$$C(\eta_{SO}) \leq c, \quad L(\eta_{SO}) < \ell, \quad C(\eta_{NE}) \geq c \quad \text{and} \quad L(\eta_{NE}) > \ell.$$

We concatenate the four paths defined for  $u \in [0, 1]$ :

$$u \mapsto u\eta_{SO}, \quad u \mapsto (1-u)\eta_{SO} + u, \quad u \mapsto (1-u) + u\eta_{NE} \quad \text{and} \quad u \mapsto (1-u)\eta_{NE},$$

to obtain a continuous loop  $(\eta_t, t \in [0, 4])$  from  $[0, 4]$  to  $\Delta$ , such that:

$$\eta_0 = \eta_4 = 0, \quad \eta_1 = \eta_{SO}, \quad \eta_2 = 1 \quad \text{and} \quad \eta_3 = \eta_{NE}.$$

We now define a continuous family of loops  $(\gamma_s, s \in [0, 1])$  in  $\mathbb{R}^2$  by

$$\gamma_s(t) = (C(s\eta_t), L(s\eta_t), t \in [0, 4]).$$

By definition, for all  $s \in [0, 1]$ ,  $\gamma_s$  is a continuous loop in  $F$ . Since  $A = (c, \ell) \notin \mathbf{F}$ , the loops  $\gamma_s$  do not contain  $A$ , so the winding number  $W(\gamma_s, A)$  is well-defined (see for example [26, Definition 6.1]). As  $A \notin \mathbf{F}$ , we get that  $\gamma_s$  is a continuous deformation in  $\mathbb{R}^2 \setminus \{A\}$  from  $\gamma_1$  to  $\gamma_0$ . Thanks to [26, Theorem 6.5], this implies that  $W(\gamma_s, A)$  does not depend on  $s \in [0, 1]$ .

For  $s = 0$ , the loop degenerates to the single point  $(C(0), 0)$  so the winding number is 0. For  $s = 1$ , let us check that the winding number is 1, which will provide the contradiction. To do this, we compare  $\gamma_1$  with a simpler loop  $\delta$  defined by:

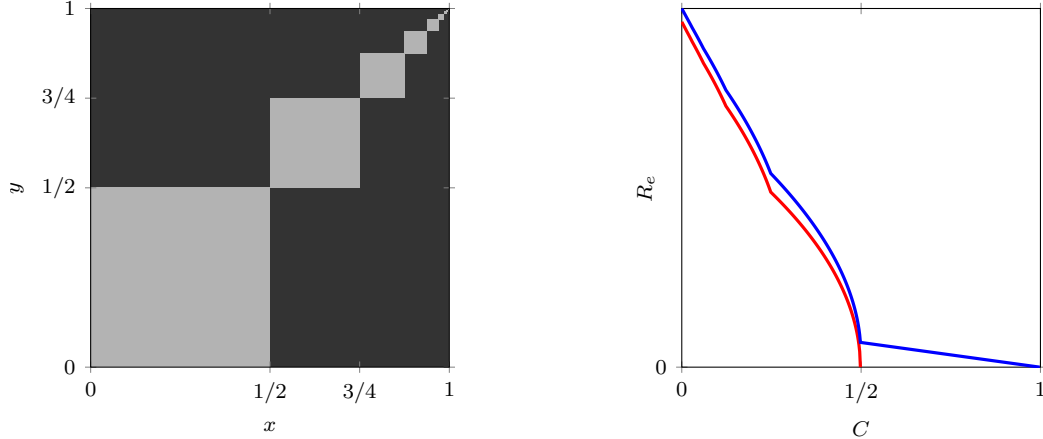
$$\delta(0) = \delta(4) = (c_{\max}, 0), \quad \delta(1) = (0, 0), \quad \delta(2) = (0, \ell_{\max}) \quad \text{and} \quad \delta(3) = (c_{\max}, \ell_{\max}),$$

and by linear interpolation for non integer values of  $t$ : in other words,  $\delta$  runs around the perimeter of the axis-aligned rectangle with corners  $(0, 0)$  and  $(c_{\max}, \ell_{\max})$ . Clearly, we have  $W(\delta, A) = 1$ .

Let  $M_t, N_t$  denote  $\gamma_1(t)$  and  $\delta(t)$  respectively. For  $t \in [0, 1]$ , we have  $N_t = ((1-t)c_{\max}, 0)$ , so the second coordinate of  $\overrightarrow{AN_t}$  is non-positive. On the other hand  $L(t\eta_{SO}) \leq L(\eta_{SO}) < \ell$ , so the second coordinate of  $\overrightarrow{AM_t}$  is negative. Therefore the two vectors  $\overrightarrow{AN_t}$  and  $\overrightarrow{AM_t}$  cannot point in opposite directions. Similar considerations for the other values of  $t \in [1, 4]$  show that  $\overrightarrow{AN_t}$  and  $\overrightarrow{AM_t}$  never point in opposite directions. By [26, Theorem 6.1], the winding numbers  $W(\gamma_1, A)$  and  $W(\delta, A)$  are equal, and thus  $W(\gamma_1, A) = 1$ .

This gives that  $A \in \mathbf{F}$  by contradiction, and thus  $F_1 \subset \mathbf{F}$ .

Finally, it is easy to check that  $F_1$  has a connected complement, because  $F_1$  is bounded, and all the points in  $F_1^c$  can reach infinity by a straight line: for example, if  $\ell > L^*(c)$ , then the half-line  $\{(c, \ell'), \ell' \geq \ell\}$  is in  $F_1^c$ .  $\square$



(A) Grayplot of the kernel  $k_\varepsilon$ , with  $\Omega = [0, 1]$  and  $\mu$  the Lebesgue measure ( $k_\varepsilon$  is equal to the constant  $\kappa > 0$  on the black zone and to  $\varepsilon$  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  $k_\varepsilon$  in blue.

FIGURE 3. On the stability of the Pareto frontier

**6.2. Stability.** We can consider the stability of the Pareto frontier and the set of Pareto optima. Recall that, thanks to (45), the graph  $\{(c, L_\star(c)) : c \in [0, c_{\max}]\}$  of  $L_\star$  is the union of the Pareto frontier and the straight line joining  $(0, C_\star(0))$  to  $(0, c_{\max})$  and can thus be seen as an extended Pareto frontier. The proof of the following proposition is immediate. It implies in particular the convergence of the extended Pareto frontier. This result can also easily be adapted to the anti-Pareto frontier.

**Proposition 6.2.** *Let  $C$  be a cost function and  $(L^{(n)}, n \in \mathbb{N})$  a sequence of loss functions converging uniformly on  $\Delta$  to a loss function  $L$ . Assume that Assumptions 3, 4 and 5 hold for the cost  $C$  and the loss functions  $L^{(n)}, n \in \mathbb{N}$ , and  $L$ . Then  $L_\star^{(n)}$  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  $\eta_n \in \mathcal{P}_{L^{(n)}}$  for all  $n \in \mathbb{N}$ . If  $C(\eta) \leq C_\star(0)$ , then we have  $\eta \in \mathcal{P}_L$ .*

*Remark 6.3 (On the continuity of the Pareto Frontier).* It might happen that some elements of  $\mathcal{P}_L$  are not weak limit of sequence of elements of  $\mathcal{P}_{L^{(n)}}$ ; see [9] for such discontinuity. It might also happen that a sequence  $(\eta_n, n \in \mathbb{N})$  such that  $\eta_n \in \mathcal{P}_{L^{(n)}}$  and  $L^{(n)}(\eta_n) > 0$  converges to some  $\eta$  that does not belong to  $\mathcal{P}_L$  if  $L(\eta) = 0$ . In particular, in this case,  $C_{\star, L^{(n)}}(0)$  does not converge to  $C_{\star, L}(0)$ , where  $C_{\star, L'}$  is the value function  $C_\star$  associated to the loss  $L'$ . This situation is represented in Figure 3. In Figure 3(A), we have plotted a perturbation  $k_\varepsilon = k + \varepsilon \sum_{n \in \mathbb{N}^*} \mathbb{1}_{I_n \times I_n}$  of the multipartite kernel  $k$  defined in Example 1.7 for  $\varepsilon > 0$  small. According to Proposition 4.3,  $R_e[k_\varepsilon]$  converges uniformly to  $R_e[k]$  when  $\varepsilon$  vanishes. However, the Pareto optimal strategies for  $k_\varepsilon$  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  $k_\varepsilon$  (in blue) corresponding to costs larger than  $1/2$  does not have a counterpart in the Pareto frontier of  $k$  (in red).

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

**Lemma 6.4.** *Suppose that Assumption 3 holds, the cost function is affine (i.e.,  $C = C_{\text{aff}}$  given by (35)) and the loss function  $L$  is sub-homogeneous. Then, we have  $L_\star(\theta c + (1-\theta)c_{\max}) \leq \theta L_\star(c)$  for all  $c \in [0, c_{\max}]$  and  $\theta \in [0, 1]$ .*

*Remark 6.5.* Geometrically, Lemma 6.4 means that the graph of the loss  $L_\star : [0, c_{\max}] \rightarrow [0, \ell_{\max}]$  is below its chords with end point  $(1, L_\star(\mathbb{1})) = (1, 0)$ . See Figures 1(B) for a typical representation of the Pareto frontier (red solid line).

*Proof.* Let  $c \in [0, c_{\max}]$  and  $\theta \in [0, 1]$ . Thanks to Lemma 5.6, Assumption 5 holds. Thus, thanks to Proposition 5.5 (iv), there exists  $\eta \in \mathcal{P}_L$  with cost  $C(\eta) = c$  and thus  $L(\eta) = L_\star(c)$ . Since  $C$  is affine, we have:

$$C(\theta\eta) = \theta C(\eta) + (1 - \theta)c_{\max} \leq \theta c + (1 - \theta)c_{\max}.$$

Therefore,  $\theta\eta$  is admissible for Problem (39) with cost constraint  $C(\cdot) \leq \theta c + (1 - \theta)c_{\max}$ . This implies that  $L_\star(\theta c + (1 - \theta)c_{\max}) \leq L(\theta\eta) \leq \theta L_\star(c)$ , thanks to the sub-homogeneity of the loss function  $L$ .  $\square$

In some case, we shall prove that the considered loss function is convex (which in turn implies Assumption 5). In this case, choosing a convex cost function implies that Assumption 4 holds and the Pareto frontier is convex. A similar result holds in the concave case. We provide a short proof of this result.

**Proposition 6.6.** *Suppose that Assumption 3 holds. If the cost function  $C$  and the loss function  $L$  are convex, then the functions  $C_\star$  and  $L_\star$  are convex. If the cost function  $C$  and the loss function  $L$  are concave, then the functions  $C^\star$  and  $L^\star$  are convex.*

*Proof.* Let  $\ell_0, \ell_1 \in [0, \ell_{\max}]$ . By Proposition 5.3, there exist  $\eta_0, \eta_1$  such that  $L(\eta_i) \leq \ell_i$  and  $C(\eta_i) = C_\star(\ell_i)$  for  $i \in \{0, 1\}$ . For  $\theta \in [0, 1]$ , let  $\ell = (1 - \theta)\ell_0 + \theta\ell_1$ . Since  $C$  and  $L$  are assumed to be convex,  $\eta = (1 - \theta)\eta_0 + \theta\eta_1$  satisfies:

$$C(\eta) \leq (1 - \theta)C_\star(\ell_0) + \theta C_\star(\ell_1) \quad \text{and} \quad L(\eta) \leq (1 - \theta)\ell_0 + \theta\ell_1.$$

Therefore, we get that  $C_\star((1 - \theta)\ell_0 + \theta\ell_1) \leq C(\eta) \leq (1 - \theta)C_\star(\ell_0) + \theta C_\star(\ell_1)$ , and  $C_\star$  is convex. The proof of the convexity of  $L_\star$  is similar. The concave case is also similar.  $\square$

## 7. EQUIVALENCE OF MODELS BY COUPLING

Even if in full generality, the cost function could also be treated as a parameter, we shall for simplicity consider only the uniform cost  $C_{\text{uni}}$  given by (36) in this section. (The interested reader can use Remark 5.2 for a first generalization to the affine cost function given by (35).)

**7.1. Motivation.** The aim of this section is to provide examples of different set of parameters for which two kernel or SIS models are “equivalent”, in the intuitive sense that their Pareto frontiers are the same (as subsets of  $\mathbb{R}_+^2$ ), and it is possible to map nicely the Pareto optima from one model to the another. In Section 7.4, we present an example where discrete models can be represented as a continuous models and an example based on measure preserving transformation in the spirit of the graphon theory. We shall consider the two families of models:

- **the kernel model** characterized by  $\text{Param} = [(\Omega, \mathcal{F}, \mu), k]$ , with Assumption 1 fulfilled, and loss function  $L = R_e$ ;
- **the SIS model** characterized by  $\text{Param} = [(\Omega, \mathcal{F}, \mu), k, \gamma]$ , with Assumption 2 fulfilled, and loss function  $L \in \{R_e, \mathfrak{J}\}$ ;

where  $(\Omega, \mathcal{F}, \mu)$  is a probability space,  $k$  and  $k$  are non-negative kernels on  $\Omega$  and  $\gamma$  is a non-negative function on  $\Omega$ .

In order to emphasize the dependence of a quantity  $H$  on the parameters  $\text{Param}$  of the model, we shall write  $H[\text{Param}]$  for  $H$ . For example we write:  $\Delta[\text{Param}]$  for the set of functions  $\{\eta \in \mathcal{L}^\infty(\Omega, \mathcal{F}) : 1 \geq \eta \geq 0\}$ , which clearly depends on the parameters  $\text{Param}$ ; and the effective reproduction function  $R_e[\text{Param}]$ . For example, under Assumption 2, we have the equality of the following functions:  $R_e[(\Omega, \mathcal{F}, \mu), k, \gamma] = R_e[(\Omega, \mathcal{F}, \mu), k/\gamma, 1] = R_e[(\Omega, \mathcal{F}, \mu), k/\gamma]$ ,

where for the last equality the left hand-side refers to the SIS model and the right hand-side refers to the kernel model (where Assumption 1 holds as a consequence of Assumption 2). Using (29) if  $\inf \gamma > 0$  (see [8, Section 3] for details and more general results), then we also have  $R_e[(\Omega, \mathcal{F}, \mu), k/\gamma] = R_e[(\Omega, \mathcal{F}, \mu), \gamma^{-1}k]$ .

**7.2. On measurability.** Let us recall some well-known facts on measurability. Let  $(E, \mathcal{E})$  and  $(E', \mathcal{E}')$  be two measurable spaces. If  $E' = \mathbb{R}$ , then we take  $\mathcal{E}' = \mathcal{B}(\mathbb{R})$  the Borel  $\sigma$ -field. Let  $f$  be a function from  $E$  to  $E'$ . We denote by  $\sigma(f) = \{f^{-1}(A) : A \in \mathcal{E}'\}$  the  $\sigma$ -field generated by  $f$ . In particular  $f$  is measurable from  $(E, \mathcal{E})$  to  $(E', \mathcal{E}')$  if and only if  $\sigma(f) \subset \mathcal{E}$ . Let  $\varphi$  be a measurable function from  $(E, \mathcal{E})$  to  $(E', \mathcal{E}')$ . For  $\nu$  a measure on  $(E, \mathcal{E})$ , we write  $\varphi_{\#}\nu$  for the push-forward measure on  $(E', \mathcal{E}')$  of the measure  $\nu$  by the function  $\varphi$  (that is  $\varphi_{\#}\nu(A) = \nu(\varphi^{-1}(A))$  for all  $A \in \mathcal{E}'$ ). By definition of  $\varphi_{\#}\nu$ , for a non-negative measurable function  $g$  defined from  $(E', \mathcal{E}')$  to  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ , we have:

$$(49) \quad \int_{E'} g d\varphi_{\#}\nu = \int_E g \circ \varphi d\nu.$$

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

$$(50) \quad \sigma(f) \subset \sigma(\varphi) \implies f = g \circ \varphi,$$

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

The random variables we consider are defined on a probability space, say  $(\Omega_0, \mathcal{F}_0, \mathbb{P})$ .

**7.3. Coupled models.** We refer the reader to [27] for a similar development in the graphon setting. We first define coupled models in the next definition and state in Proposition 7.3 that coupled models have related (anti-)Pareto optima and the same (anti-)Pareto frontiers.

In the kernel model, we consider the models  $\text{Param}_i = [(\Omega_i, \mathcal{F}_i, \mu_i), k_i]$  for  $i \in \{1, 2\}$ , where Assumption 1 holds for each model; in the SIS model, we consider the models  $\text{Param}_i = [(\Omega_i, \mathcal{F}_i, \mu_i), k_i, \gamma_i]$  for  $i \in \{1, 2\}$ , where Assumption 2 holds for each model. In what follows, we simply write  $\Delta_i$  the set of functions  $\Delta$  for the model  $\text{Param}_i$ .

A measure  $\pi$  on  $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2)$  is a *coupling* if its marginals are  $\mu_1$  and  $\mu_2$ .

**Definition 7.1** (Coupled models). *The models  $\text{Param}_1$  and  $\text{Param}_2$  are coupled if there exists two independent  $\Omega_1 \times \Omega_2$ -valued random vectors  $(X_1, X_2)$  and  $(Y_1, Y_2)$  (defined on a probability space  $(\Omega_0, \mathcal{F}_0, \mathbb{P})$ ) with the same distribution given by a coupling (i.e.  $X_i$  and  $Y_i$  have distribution  $\mu_i$ ) such that,  $\mathbb{P}$ -almost surely:*

$$\text{Kernel model: } k_1(X_1, Y_1) = k_2(X_2, Y_2),$$

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

*In this case, two real-valued measurable functions  $v_1$  and  $v_2$  defined respectively on  $\Omega_1$  and  $\Omega_2$  are coupled (through  $V$ ) if there exists a real-valued  $\sigma(X_1, X_2)$ -measurable integrable random variable  $V$  such that  $\mathbb{P}$ -almost surely:*

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

*Remark 7.2.* We keep notation from Definition 7.1

- (i) Since  $V$  is real-valued and  $\sigma(X_1, X_2)$ -measurable, we deduce from (50) that there exists a measurable function  $v$  defined on  $\Omega_1 \times \Omega_2$  such that  $V = v(X_1, X_2)$ , thus the following equality holds  $\mathbb{P}$ -almost surely:

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

- (ii) If  $W$  is a real-valued integrable  $\sigma(X_1) \cap \sigma(X_2)$ -measurable random variable, then setting  $v_i(X_i) = \mathbb{E}[W | X_i] = W$ , the equality  $v_1(X_1) = v_2(X_2)$  holds almost surely, and we get that  $v_1$  and  $v_2$  are coupled (through  $W$ ).

- (iii) Let  $\eta_1 \in \Delta_1$ . According to (50), there exists  $\eta_2 \in \Delta_2$  such that  $\mathbb{E}[\eta_1(X_1)|X_2] = \eta_2(X_2)$ . Thus, by definition  $\eta_1$  and  $\eta_2$  are coupled (through  $V = \eta_1(X_1)$ ).

The main result of this section, whose proof is given in Section 8.2, states that coupled models have coupled Pareto optimal strategies, and thus the same (anti-)Pareto frontier.

**Proposition 7.3** (Coupling and Pareto optimality). *Let  $\text{Param}_1$  and  $\text{Param}_2$  be two coupled (kernel or SIS) models with the uniform cost function  $C = C_{\text{uni}}$  and loss function  $L$  (with  $L = R_e$  in the kernel model and  $L \in \{R_e, \mathcal{J}\}$  in the SIS model). If the functions  $\eta_1 \in \Delta_1$  and  $\eta_2 \in \Delta_2$  are coupled, then:*

$$\eta_1 \text{ is Pareto optimal (for Param}_1) \iff \eta_2 \text{ is Pareto optimal (for Param}_2).$$

Furthermore, if  $\eta_1 \in \Delta_1$  is Pareto optimal (for  $\text{Param}_1$ ), then there exists a Pareto optimal (for  $\text{Param}_2$ ) strategy  $\eta_2 \in \Delta_2$  such that  $\eta_1$  and  $\eta_2$  are coupled. In particular, the (anti-)Pareto frontiers are the same for the two models  $\text{Param}_1$  and  $\text{Param}_2$ .

The next Corollary is useful for model reduction, which corresponds to merging individuals with identical behavior, see the examples in Sections 7.4.1 and 7.4.3. Equation (51) below could also be stated for anti-Pareto optima; and the adaptation to the kernel model is immediate.

**Corollary 7.4.** *Let  $\text{Param} = [(\Omega, \mathcal{F}, \mathbb{P}), k, \gamma]$  be a SIS model with the uniform cost function  $C = C_{\text{uni}}$  and loss function  $L \in \{R_e, \mathcal{J}\}$ . Let  $\mathcal{G} \subset \mathcal{F}$  be a  $\sigma$ -field such that  $\gamma$  is  $\mathcal{G}$ -measurable and  $k$  is  $\mathcal{G} \otimes \mathcal{G}$ -measurable. Then, for any  $\eta \in \Delta[\text{Param}]$ , we have:*

$$(51) \quad \eta \text{ is Pareto optimal} \iff \mathbb{E}[\eta|\mathcal{G}] \text{ is Pareto optimal.}$$

*Proof.* Let  $\Omega_0 = \Omega^2$  endowed with the product  $\sigma$ -field and the product probability measure  $\mathbb{P}_0$ , and  $X$  (resp.  $Y$ ) be the projection on the first (resp. second) coordinate. Thus the random variables  $X$  and  $Y$  are independent,  $(\Omega, \mathcal{F})$ -valued with distribution  $\mathbb{P}$ . Write  $(X', Y')$  for  $(X, Y)$  when considered as  $(\Omega, \mathcal{G})$ -valued random variables. Notice that  $X'$  and  $Y'$  are by construction independent with distribution  $\mathbb{P}'$ , where  $\mathbb{P}'$  is the restriction of  $\mathbb{P}$  to  $\mathcal{G}$ . As  $\gamma$  is  $\mathcal{G}$ -measurable and  $k$  is  $\mathcal{G} \otimes \mathcal{G}$ -measurable, we can consider the model  $\text{Param}' = [(\Omega, \mathcal{G}, \mathbb{P}'), k, \gamma]$ . Then  $(X, X')$  and  $(Y, Y')$  are two trivial couplings such that  $k(X, Y) = k(X', Y')$  and  $\gamma(X) = \gamma(X')$ . Thus the models  $\text{Param}$  and  $\text{Param}'$  are coupled. We have that  $\eta \in \Delta$  and  $\eta' = \mathbb{E}[\eta|\mathcal{G}] \in \Delta'$  are coupled through  $\eta \circ X$  since  $\mathbb{E}_0[\eta \circ X | \sigma(X)] = \eta \circ X$  and  $\mathbb{E}_0[\eta \circ X | \sigma(X')] = \eta' \circ X'$  as  $\sigma(X') = X^{-1}(\mathcal{G})$  and  $X = X'$  can be seen as the identity map on  $\Omega$ . The conclusion then follows from Proposition 7.3.  $\square$

**7.4. Examples of couplings.** In this section, we consider the SIS model as the kernel model can be handled in the same way. We denote by  $\text{Leb}$  the Lebesgue measure.

**7.4.1. Discrete and continuous models.** We now formalize how finite population models can be seen as particular cases of models with a continuous population. Let  $\Omega_d \subset \mathbb{N}$ ,  $\mathcal{F}_d$  the set of subsets of  $\Omega_d$  and  $\mu_d$  a probability measure on  $\Omega_d$ . Without loss of generality, we can assume that  $\mu_d(\{\ell\}) > 0$  for all  $\ell \in \Omega_d$ . We set  $\Omega_c = [0, 1)$ , with  $\mathcal{F}_c$  its Borel  $\sigma$ -field and  $\mu_c = \text{Leb}$ . Let  $(B_\ell, \ell \in \Omega_d)$  be a partition of  $[0, 1)$  in measurable sets such that  $\text{Leb}(B_\ell) = \mu_d(\{\ell\})$  for all  $\ell \in \Omega_d$ . The measure  $\pi$  on  $\Omega_d \times \Omega_c$  uniquely defined by:

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

for all measurable  $A \subset [0, 1)$  and  $\ell \in \Omega_d$  is clearly a coupling of  $\mu_d$  and  $\mu_c$ . If the kernels  $k_d$  on  $\Omega_d$  and  $k_c$  on  $\Omega_c$  and the functions  $\gamma_d$  and  $\gamma_c$  are related through the formula:

$$\gamma_c(x) = \gamma_d(\ell) \quad \text{and} \quad k_c(x, y) = k_d(\ell, j), \quad \text{for } x \in B_\ell, y \in B_j \text{ and } \ell, j \in \Omega_d,$$

then the discrete model  $\text{Param}_d = [(\Omega_d, \mathcal{F}_d, \mu_d), k_d, \gamma_d]$  and the continuous model  $\text{Param}_c = [([0, 1), \mathcal{F}_c, \text{Leb}), k_c, \gamma_c]$  are coupled. Roughly speaking, we can blow up the atomic part of the

measure  $\mu_d$  into a continuous part, or, conversely, merge all points that behave similarly for  $k_c$  and  $\gamma_c$  into an atom, without altering the Pareto frontier.

*Example 7.5.* We consider the so called stochastic block model, with 2 populations for simplicity, in the setting of the SIS model, and give in this elementary case the corresponding discrete and continuous models. Then, we explicit the relation with the formalism of the same model developed in [33] by Lajmanovich and Yorke.

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

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

Notice  $p$  is the relative size of population 1. The corresponding discrete model is  $\text{Param}_d = [(\{1, 2\}, \mathcal{F}_d, \mu_d), k_d, \gamma_d]$ ; see Figure 4(B).

The continuous model is defined on the state space  $\Omega_c = [0, 1]$  is endowed with its Borel  $\sigma$ -field,  $\mathcal{F}_c$ , and the Lebesgue measure  $\mu_c = \text{Leb}$ . 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  $\gamma_c$  are given by:

$$k_c(x, y) = k_{ij} \quad \text{and} \quad \gamma_c(x) = \gamma_i \quad \text{for } x \in B_i, y \in B_j, \text{ and } i, j \in \{1, 2\}.$$

The corresponding continuous model is  $\text{Param}_c = [([0, 1], \mathcal{F}_c, \text{Leb}), k_c, \gamma_c]$ ; see Figure 4(A). By the general discussion above, the discrete and continuous models are coupled, and in particular they have the same Pareto and anti-Pareto frontiers.

Furthermore, in this simple example, it is easily checked that a discrete vaccination  $\eta_d = (\eta_1, \eta_2)$  and a continuous vaccination  $\eta_c = (\eta_c(x), x \in [0, 1])$  are coupled if and only if there exists a function  $\eta$  defined on  $\Omega_c \times \Omega_d = [0, 1] \times \{1, 2\}$  such that:

$$\begin{cases} \eta_i = \frac{1}{\text{Leb}(B_i)} \int_{B_i} \eta(x, i) dx, & i \in \{1, 2\}, \\ \eta_c(x) = \eta(x, 1)\mathbb{1}_{B_1}(x) + \eta(x, 2)\mathbb{1}_{B_2}(x), & \text{Leb-a.s.}, \end{cases}$$

which occurs if and only if:

$$\eta_i = \frac{1}{\text{Leb}(B_i)} \int \eta_c(x)\mathbb{1}_{B_i}(x) dx, \quad i \in \{1, 2\}.$$

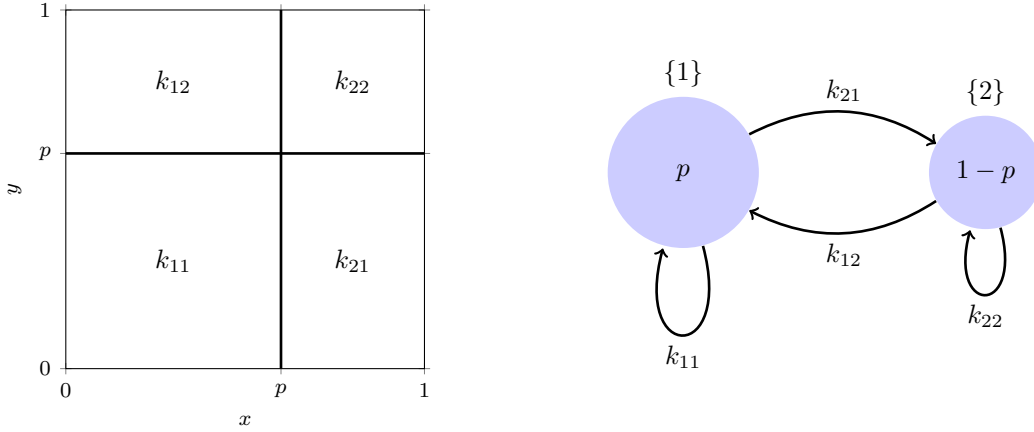
Therefore, in this case, the optimal strategies of the continuous model are easily deduced from the optimal strategies of the discrete model.

To conclude this example, we rewrite, using the formalism of the discrete model  $\text{Param}_d$ , the next-generation matrix  $K$  in the setting of [33], and the effective next-generation matrix  $K_e(\eta)$  when the vaccination strategy  $\eta$  is in force (recall  $\eta_i$  is the proportion of population with feature  $i$  which is not vaccinated):

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

with  $p = \mu_d(\{1\})$ ,  $1-p = \mu_d(\{2\})$  and  $k_d = k_d/\gamma_d$ , that is:

$$k_d = \begin{pmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{pmatrix} = \begin{pmatrix} k_{11}/\gamma_1 & k_{12}/\gamma_2 \\ k_{21}/\gamma_1 & k_{22}/\gamma_2 \end{pmatrix}.$$



(A) Continuous model: kernel  $k_c$  on  $\Omega_c = [0, 1)$  with the Lebesgue measure.

(B) Discrete model: kernel  $k_d$  on  $\Omega_d = \{1, 2\}$  with the measure  $p\delta_1 + (1-p)\delta_2$ .

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

**7.4.2. Measure preserving function.** This section is motivated by the theory of graphons, which are indistinguishable by measure preserving transformation, see [34, Sections 7.3 and 10.7]. Let  $(\Omega, \mathcal{F}, \mu)$  be a measurable space. We say a measurable function  $\varphi$  from  $(\Omega, \mathcal{F})$  to itself is *measure preserving* if  $\mu = \varphi_{\#}\mu$ . For example the function  $\varphi : x \mapsto 2x \pmod{1}$  defined on the probability space  $([0, 1], \mathcal{B}([0, 1]), \text{Leb})$  is measure preserving.

Let  $\varphi$  be measure preserving function on  $\Omega$ . Let  $k_1$  be a kernel and  $\gamma_1$  a function on  $\Omega$  such that the model  $\text{Param}_1 = [(\Omega, \mathcal{F}, \mu), k_1, \gamma_1]$  satisfies Assumption 2. Let  $X_1$  be a random variable with probability distribution  $\mu$  and let  $X_2 = \varphi(X_1)$ , so that  $(X_1, X_2)$  is a coupling of  $(\Omega, \mathcal{F}, \mu)$  with itself. Then for the kernel  $k_2$  and the function  $\gamma_2$  defined by:

$$k_2(x, y) = k_1(\varphi(x), \varphi(y)) \quad \text{and} \quad \gamma_2(x) = \gamma_1(\varphi(x)),$$

the models  $\text{Param}_1$  and  $\text{Param}_2 = [(\Omega, \mathcal{F}, \mu), k_2, \gamma_2]$  are coupled. Roughly speaking, we can give different labels to the features of the population without altering the Pareto and anti-Pareto frontiers.

**7.4.3. Model reduction using deterministic coupling.** This example is in the spirit of Section 7.4.1, where one merges individual with identical behavior. We consider a SIS model  $\text{Param}_1 = [(\Omega_1, \mathcal{F}_1, \mu_1), k_1, \gamma_1]$ . Let  $\varphi$  be a measurable function from  $(\Omega_1, \mathcal{F}_1)$  to  $(\Omega_2, \mathcal{F}_2)$ . Assume that:

$$\sigma(\gamma_1) \subset \sigma(\varphi) \quad \text{and} \quad \sigma(k_1) \subset \sigma(\varphi) \otimes \sigma(\varphi).$$

We can then build an elementary coupling. Let  $X_1$  and  $Y_1$  be independent  $\mu_1$  distributed random elements of  $\Omega_1$ , and set  $(X_2, Y_2) = (\varphi(X_1), \varphi(Y_1))$ . Since  $\sigma(\gamma_1) \subset \sigma(\varphi)$  and  $\sigma(k_1) \subset \sigma(\varphi) \otimes \sigma(\varphi)$ , we get that  $\gamma_1(X_1)$  is  $\sigma(X_2)$ -measurable and  $k_1(X_1, Y_1)$  is  $\sigma(X_2, Y_2)$ -measurable. According to (50), there exists two measurable functions  $\gamma_2 : \Omega_2 \rightarrow \mathbb{R}$  and  $k_2 : \Omega_2 \times \Omega_2 \rightarrow \mathbb{R}$  such that  $\gamma_1 = \gamma_2 \circ \varphi$  and  $k_1 = k_2(\varphi \otimes \varphi)$  that is almost surely:

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

Let  $\mu_2 = \varphi_{\#}\mu_1$  be the push-forward measure of  $\mu_1$  by  $\varphi$ . Using (49) it is easy to check that the integrability condition from Assumption 2 is fulfilled, so we can consider the reduced model  $\text{Param}_2 = [(\Omega_2, \mathcal{F}_2, \mu_2), k_2, \gamma_2]$ . By Definition 7.1,  $\text{Param}_1$  is coupled with  $\text{Param}_2$  through the (deterministic) coupling  $\pi$  given by the distribution of  $(X_1, \varphi(X_1))$ .



Eventually, we get from Corollary 7.4 with  $\mathcal{G} = \sigma(\varphi)$ , that  $\eta_1 \in \Delta_1$  is Pareto optimal if and only if  $\mathbb{E}_1[\eta_1 | \varphi]$  is Pareto optimal (for the model  $\text{Param}_1$ ), where  $\mathbb{E}_1$  correspond to the expectation with respect to the probability measure  $\mu_1$  on  $(\Omega_1, \mathcal{F}_1)$ .

## 8. TECHNICAL PROOFS

**8.1. The SIS model: properties of  $\mathfrak{J}$  and of the maximal equilibrium.** We prove here Theorem 4.6 and Proposition 4.7, 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  $A$ , we denote by  $A^\top$  its adjoint. We first give a preliminary lemma.

**Lemma 8.1.** *Suppose Assumption 2 holds, and consider the positive bounded linear integral operator  $\mathcal{T}_{k/\gamma}$  on  $\mathcal{L}^\infty$ . If there exists  $g \in \mathcal{L}_+^\infty$ , with  $\int_\Omega g \, d\mu > 0$  and  $\lambda > 0$  satisfying:*

$$\mathcal{T}_{k/\gamma}(g)(x) > \lambda g(x), \quad \text{for all } x \text{ such that } g(x) > 0,$$

then we have  $\rho(\mathcal{T}_{k/\gamma}) > \lambda$ .

*Proof.* Set  $\mathcal{T} = \mathcal{T}_{k/\gamma}$ . 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  $\mathcal{L}^\infty$ , 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  $\int_A v \, d\mu > 0$  and  $\int_\Omega v g \, d\mu > 0$ . 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 (26) that  $\rho(\mathcal{T}) \geq \rho(\mathcal{T}') > \lambda$ .  $\square$

We now state an interesting result on the characterization of the maximal equilibrium  $\mathfrak{g}$ . We keep notations from Sections 2.3 and 2.4 and write  $R_e$  for  $R_e[k/\gamma]$ . Recall that  $R_0 = R_e(\mathbb{1})$  and  $F$  defined by (18). Let  $DF[h]$  denote the bounded linear operator on  $\mathcal{L}^\infty$  of the derivative of the map  $f \mapsto F(f)$  defined on  $\mathcal{L}^\infty$  at point  $h$ :

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

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

**Proposition 8.2.** *Suppose Assumption 2 holds and write  $R_e$  for  $R_e[k/\gamma]$ . Let  $h$  in  $\Delta$  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:  $\mathfrak{g} = 0 \iff R_0 \leq 1$ ; as well as:  $\mathfrak{g} \neq 0 \implies R_e(1 - \mathfrak{g}) = 1$ .

*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:

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

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

$$(52) \quad k(x, y) = (1 - h(x)) \frac{k(x, y)(1 - h(y))}{\gamma(y)}$$

and thus  $\mathcal{T}_k = M_{1-h} \mathcal{T}_{k/\gamma} M_{1-h}$ , where  $M_f$  is the multiplication operator by  $f$ . Recall the definition (14) of  $R_e$ . According to (29), we have:

$$(53) \quad \rho(\mathcal{T}_k) = \rho(\mathcal{T}_{k/\gamma} M_{(1-h)^2}) = R_e((1-h)^2).$$

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

We prove that (i) implies (iv). Suppose that  $R_e(1-h) > 1$ . Thanks to (29), 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  $\mathcal{T}_{(1-h)k/\gamma}^\top(v) = R_e(1-h)v$ . 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  $\langle v, \mathbf{1}_A \rangle = 0$ , where  $A = \{h > 0\}$  denote the support of the function  $h$ . Since  $\mathcal{T}_{(1-h)k/\gamma}^\top(v) = R_e(1-h)v$  and setting  $v' = (1-h)v$  (so that  $v' = v$   $\mu$ -almost surely on  $A^c$ ), we deduce that:

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

where  $k' = \mathbf{1}_{A^c} k \mathbf{1}_{A^c}$ . 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 (26) 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  $w \in \mathcal{L}_+^\infty \setminus \{0\}$  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  $F(h) = 0$ ,  $\mathcal{T}_k(w) = \mathcal{T}_{k'}(w) = (\gamma + \lambda)w$  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 4.4 to deduce that  $h + \varepsilon w \leq \mathfrak{g}$  and thus  $h \neq \mathfrak{g}$ .

To see that (iv) implies (iii), notice that  $(1-h)^2 \leq (1-h)$ , and then deduce from Proposition 4.1 (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  $\gamma(\mathfrak{g} - h)/(1-h) \in \mathcal{L}_+^\infty$ . An elementary computation, using  $F(h) = F(\mathfrak{g}) = 0$  and  $k$  defined in (52), gives:

$$\mathcal{T}_k \left( \gamma \frac{\mathfrak{g} - h}{1-h} \right) = (1-h)\mathcal{T}_k(\mathfrak{g} - h) = \gamma \frac{\mathfrak{g} - h}{1-\mathfrak{g}} = \frac{1-h}{1-\mathfrak{g}} \gamma \frac{\mathfrak{g} - h}{1-h}.$$

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 8.1 (with  $k$  replaced by  $k\gamma$ ) that  $\rho(\mathcal{T}_k) > 1$ . Then use (53) 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  $R_0 = R_e(\mathbf{1})$ .

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  $\mu$ -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  $F_\eta$ ,  $R_e$  or  $\mathfrak{g}_\eta$ , in the parameters  $k$  and  $\gamma$  (which satisfy Assumption 2) of the model, we shall write  $H[k, \gamma]$ . Recall that if  $k$  and  $\gamma$  satisfy Assumption 2, then the kernel  $k/\gamma$  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 8.3.** *Let  $((k_n, \gamma_n), n \in \mathbb{N})$  and  $(k, \gamma)$  be kernels and functions satisfying Assumption 2 and  $(\eta_n, n \in \mathbb{N})$  be a sequence of elements of  $\Delta$  converging weakly to  $\eta$ .*

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

*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  $\Delta$  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 (22), we have:

$$(54) \quad g_n = \frac{\mathcal{T}_{k_n}(\eta_n g_n)}{1 + \mathcal{T}_{k_n}(\eta_n g_n)} = \frac{\mathcal{T}_{k_n}(h_n)}{1 + \mathcal{T}_{k_n}(h_n)}.$$

We set  $g = \mathcal{T}_k(h)/(1 + \mathcal{T}_k(h))$ . Notice that  $\mathcal{T}_{k_n}(h_n) = (\mathcal{T}_{k_n} - \mathcal{T}_k)(h_n) + \mathcal{T}_k(h_n)$ . We have  $\lim_{n \rightarrow \infty} \mathcal{T}_k(h_n) = \mathcal{T}_k(h)$  pointwise. Since  $\|(\mathcal{T}_{k_n} - \mathcal{T}_k)(h_n)\|_{p'} \leq \|k_n - k\|_{p', q'}$ , up to taking a sub-sequence, we deduce that  $\lim_{n \rightarrow \infty} (\mathcal{T}_{k_n} - \mathcal{T}_k)(h_n) = 0$  almost surely. (Notice the previous step is not used in the proof of (i) as  $k_n = k$  and  $\lim_{n \rightarrow \infty} \mathcal{T}_k(h_n) = \mathcal{T}_k(h)$  pointwise.) This implies that  $g_n$  converges almost surely to  $g$ . By the dominated convergence theorem, we deduce that  $g_n$  converges also in  $L^p$  to  $g$ . This proves that  $h = \eta g$  almost surely. We get  $g = \mathcal{T}_k(\eta g)/(1 + \mathcal{T}_k(\eta g))$  and thus  $F_\eta[k, \gamma](g) = 0$ :  $g$  is an equilibrium for  $F_\eta[k, \gamma]$ . We recall from [8, Section 3] the functional equality  $R_e[k'h] = R_e[hhk']$ , where  $k'$  is a kernel,  $h$  a non-negative functions such that the kernels  $k'h$  and  $hhk'$  have some finite double norm. We deduce from the weak-continuity and the stability of  $R_e$ , see Theorem 4.2 and Proposition 4.3, that:

$$\begin{aligned} R_e[k/\gamma](\eta(1-g)) &= R_e[k](\eta(1-g)) = \lim_{n \rightarrow \infty} R_e[k_n](\eta_n(1-g_n)) \\ &= \lim_{n \rightarrow \infty} R_e[k_n/\gamma_n](\eta_n(1-g_n)) \\ &\leq 1. \end{aligned}$$

(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 8.2 holds with  $k$  replaced by  $k\eta$ , and thus property (i) therein implies that  $g = \mathfrak{g}_\eta[k, \gamma]$ .  $\square$

*Proofs of Theorem 4.6 and Proposition 4.7.* Under the assumptions of Lemma 8.3, taking the pair  $(k_n, \gamma_n)$  equal to  $(k, \gamma)$  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:

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

Taking  $(k_n, \gamma_n) = (k, \gamma)$  provides the continuity of  $\mathfrak{J}[k, \gamma]$  and thus Theorem 4.6. Then, arguing as in the end of the proof of Proposition 4.3, we get Proposition 4.7.  $\square$

**8.2. Coupling and Pareto optimality.** We prove here Proposition 7.3. We only consider the SIS model  $\text{Param} = [(\Omega, \mathcal{F}, \mu), k, \gamma]$ , as the kernel model can be handled similarly. We suppose throughout this section that Assumption 2 holds.

The random variables we consider, are defined on a probability space, say  $(\Omega_0, \mathcal{F}_0, \mathbb{P})$ . We recall an elementary result on conditional independence. Let  $\mathcal{A}$ ,  $\mathcal{B}$  and  $\mathcal{H}$  be  $\sigma$ -fields subsets of  $\mathcal{F}_0$ , such that  $\mathcal{H} \subset \mathcal{A} \cap \mathcal{B}$ . Then, according to [28, Theorem 8.9], we have that for any integrable real-valued random variable  $X$  which is  $\mathcal{B}$ -measurable:

$$(55) \quad \mathcal{A} \text{ and } \mathcal{B} \text{ are conditionally independent given } \mathcal{H} \implies \mathbb{E}[X | \mathcal{A}] = \mathbb{E}[X | \mathcal{H}].$$

We now state two technical lemmas.

**Lemma 8.4** (Measurability). *Let  $\text{Param}_1$  and  $\text{Param}_2$  be coupled models with independent coupling  $(X_1, X_2)$  and  $(Y_1, Y_2)$ . Then the random variable  $\gamma_1(X_1)$  is  $\sigma(X_1) \cap \sigma(X_2)$ -measurable. For any measurable function  $v : \Omega_1 \times \Omega_2 \rightarrow \mathbb{R}$ , such that  $k_1(X_1, Y_1)v(Y_1, Y_2)$  is integrable, the random variable  $\mathbb{E}[k_1(X_1, Y_1)v(Y_1, Y_2)|X_1]$  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  $\mathbb{E}[k(X_1, Y_1)v(Y_1, Y_2)|X_1]$  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  $\sigma$ -fields  $\mathcal{A} = \sigma(X_1, X_2)$  and  $\mathcal{B} = \sigma(X_1, Y_1, Y_2)$  are conditionally independent given  $\mathcal{H} = \sigma(X_1)$ . Using (55), we deduce that:

$$\mathbb{E}[k_1(X_1, Y_1)v(Y_1, Y_2)|X_1] = \mathbb{E}[k_1(X_1, Y_1)v(Y_1, Y_2)|X_1, X_2].$$

Since  $k_1(X_1, Y_1) = k_2(X_2, Y_2)$   $\mathbb{P}$ -almost surely, we get:

$$\begin{aligned} \mathbb{E}[k_1(X_1, Y_1)v(Y_1, Y_2)|X_1] &= \mathbb{E}[k_2(X_2, Y_2)v(Y_1, Y_2)|X_1, X_2] \\ &= \mathbb{E}[k_2(X_2, Y_2)v(Y_1, Y_2)|X_2], \end{aligned}$$

where the last equality follows from another application of (55) with  $\mathcal{A} = \sigma(X_1, X_2)$ ,  $\mathcal{B} = \sigma(X_2, Y_1, Y_2)$  which are conditionally independent given  $\mathcal{H} = \sigma(X_2)$ . The last expression is  $\sigma(X_2)$  measurable, so the proof is complete.  $\square$

In the following key lemma, we simply write  $H_i$  for  $H[\text{Param}_i]$  for  $H$  the loss functions  $R_e$  and  $\mathfrak{J}$ , the cost function  $C = C_{\text{uni}}$  and the spectrum  $\text{Spec}$ .

**Lemma 8.5.** *If  $\text{Param}_1$  and  $\text{Param}_2$  are coupled models, and if the functions  $\eta_1 \in \Delta_1$  and  $\eta_2 \in \Delta_2$  are coupled, then  $\text{Spec}_1(\eta_1) \cup \{0\} = \text{Spec}_2(\eta_2) \cup \{0\}$  and for  $H$  any one of the mappings  $C_{\text{uni}}$ ,  $R_e$  or  $\mathfrak{J}$ :*

$$(56) \quad H_1(\eta_1) = H_2(\eta_2).$$

*Proof.* Let  $(X_1, X_2)$  and  $(Y_1, Y_2)$  be two independent couplings, and assume that  $\eta_1$  and  $\eta_2$  are coupled through the function  $\eta$ , see Remark 7.2 (i):

$$(57) \quad \mathbb{E}[\eta(X_1, X_2)|X_i] = \eta_i(X_i) \quad \text{for } i \in \{1, 2\}.$$

*Step 1:* The cost function ( $H = C_{\text{uni}}$ ). We directly have:

$$C_1(\eta_1) = 1 - \mathbb{E}[\eta_1(X_1)] = 1 - \mathbb{E}[\eta(X_1, X_2)] = 1 - \mathbb{E}[\eta_2(X_2)] = C_2(\eta_2).$$

*Step 2:* The spectrum and the effective reproduction function ( $H = R_e$ ). Set  $k_i = k_i/\gamma_i$  for  $i \in \{1, 2\}$ . Let  $\lambda$  be a non-zero eigenvalue of  $T_{k_1\eta_1}$  associated with an eigenvector  $v_1$ . Notice that  $k(X_1, Y_1)\eta_1(Y_1)v(Y_1)$  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:

$$\lambda^{-1}\mathbb{E}[k_1(X_1, Y_1)\eta_1(Y_1)v_1(Y_1)|X_1].$$

By Lemma 8.4 applied to the function  $v(y_1, y_2) = (v_1\eta_1/\gamma_1)(y_1)$ , the real-valued random variable  $v_1(X_1)$  is  $\sigma(X_1) \cap \sigma(X_2)$ -measurable and thus  $\sigma(X_2)$ -measurable. Thanks to (50), there exists  $v_2$  such that  $v_2(X_2) = v_1(X_1)$  almost surely. Since  $(Y_1, Y_2)$  is distributed as  $(X_1, X_2)$ , we deduce that (57) holds also with  $(X_1, X_2)$  replaced by  $(Y_1, Y_2)$  and that  $v_2(Y_2) = v_1(Y_1)$  almost surely. Recall that  $k_i = k_i/\gamma_i$ , so that  $k_1(X_1, Y_1) = k_2(X_2, Y_2)$  almost surely. We may now

compute:

$$\begin{aligned}
\lambda v_2(X_2) &= \lambda v_1(X_1) \\
&= \mathbb{E} [k_1(X_1, Y_1) \eta_1(Y_1) v_1(Y_1) | X_1] \\
&= \mathbb{E} [k_1(X_1, Y_1) \eta(Y_1, Y_2) v_1(Y_1) | X_1] \quad (\text{de-conditioning on } (Y_1, X_1)) \\
(58) \quad &= \mathbb{E} [k_1(X_1, Y_1) \eta(Y_1, Y_2) v_1(Y_1) | X_2] \quad (\text{Lemma 8.4}) \\
&= \mathbb{E} [k_2(X_2, Y_2) \eta(Y_1, Y_2) v_2(Y_2) | X_2] \quad (\text{a.s. equality}) \\
&= \mathbb{E} [k_2(X_2, Y_2) \eta_2(Y_2) v_2(Y_2) | X_2] \quad (\text{conditioning on } (Y_2, X_2)) \\
&= T_{k_2 \eta_2} v_2(X_2).
\end{aligned}$$

Since the distribution of  $X_2$  is  $\mu_2$ , we have  $\lambda v_2 = T_{k_2 \eta_2} v_2$   $\mu_2$ -almost surely. Therefore  $\lambda$  is also an eigenvalue for  $T_{k_2 \eta_2}$ . By symmetry we deduce that the spectrum up to  $\{0\}$  of  $T_{k_1 \eta_1}$  and  $T_{k_2 \eta_2}$  coincide, that is  $\text{Spec}_1(\eta_1) \cup \{0\} = \text{Spec}_2(\eta_2) \cup \{0\}$ , and in particular the spectral radius coincide.

*Step 3:* The total proportion of infected population function ( $H = \mathfrak{J}$ ). 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 (56) 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<sub>1</sub>. Since  $F_{\eta_1}(g_1) = 0$ , see (22), we have:

$$(59) \quad g_1 = \frac{\mathcal{T}_{k_1}(\eta_1 g_1)}{\gamma_1 + \mathcal{T}_{k_1}(\eta_1 g_1)}.$$

By Lemma 8.4, this implies that  $g_1(X_1)$  is  $\sigma(X_1) \cap \sigma(X_2)$  measurable. Thus, there exists  $g'_2$  such that  $g'_2(X_2) = g_1(X_1)$   $\mathbb{P}$ -almost surely.. Therefore, by the same computation as in (58):

$$\mathcal{T}_{k_1}(\eta_1 g_1)(X_1) = \mathcal{T}_{k_2}(\eta_2 g'_2)(X_2) \quad \mathbb{P} - \text{a.s.}$$

We set:

$$(60) \quad g_2 = \frac{\mathcal{T}_{k_2}(\eta_2 g'_2)}{\gamma_2 + \mathcal{T}_{k_2}(\eta_2 g'_2)}.$$

Then, we deduce from (59) that  $g_2(X_2) = g'_2(X_2)$   $\mathbb{P}$ -almost surely, that is  $g_2 = g'_2$   $\mu_2$ -almost surely. Thus (60) holds with  $g'_2$  replaced by  $g_2$ . In other words,  $g_2$  satisfies (22): it is an equilibrium for the model given by Param<sub>2</sub>.

Let us now prove that  $g_2$  is in fact the maximal equilibrium. Since  $g_2(X_2) = g_1(X_1)$   $\mathbb{P}$ -almost surely and  $g_1(X_1)$  is  $\sigma(X_1) \cap \sigma(X_2)$ -measurable, we deduce from Remark 7.2 (ii), that  $(1 - g_1)$  and  $(1 - g_2)$  are coupled, so  $R_e[\text{Param}_1](1 - g_1) = R_e[\text{Param}_2](1 - g_2)$ , by Property (56) applied to  $H = R_e$ . Since  $R_0 > 1$  and  $g_1$  is the maximal equilibrium for Param<sub>1</sub>, we deduce from Proposition 8.2 that  $R_e[\text{Param}_1](1 - g_1) = 1$ . Using again Proposition 8.2, this gives that  $g_2$  is the maximal equilibrium for Param<sub>2</sub>.

We may now compute:

$$\begin{aligned}
\mathfrak{J}_1(\eta_1) &= \mathbb{E} [\eta_1(X_1) g_1(X_1)] \\
&= \mathbb{E} [\eta(X_1, X_2) g_1(X_1)] \quad (\text{deconditioning on } X_1) \\
&= \mathbb{E} [\eta(X_1, X_2) g_2(X_2)] \quad (\text{a.s. equality}) \\
&= \mathbb{E} [\eta_2(X_2) g_2(X_2)] \quad (\text{conditioning on } X_2) \\
&= \mathfrak{J}_2(\eta_2),
\end{aligned}$$

thus (56) holds for  $H = \mathfrak{J}$ , and the proof is complete.  $\square$

We now give the proof of Proposition 7.3. Its first part is an elementary consequence of the Lemma 8.5; and the second part is a direct consequence of Remark 7.2 (iii).

## REFERENCES

- [1] P. M. Anselone. *Collectively compact operator approximation theory and applications to integral equations*. Prentice-Hall, 1971.
- [2] P. M. Anselone and J. W. Lee. “Spectral properties of integral operators with nonnegative kernels”. *Linear Algebra and its Applications* 9 (1974), pp. 67–87.
- [3] B. Bollobás, S. Janson, and O. Riordan. “The phase transition in inhomogeneous random graphs”. *Random Structures Algorithms* 31.1 (2007), pp. 3–122.
- [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* 369.6505 (2020), pp. 846–849.
- [5] L. Burlando. “Continuity of spectrum and spectral radius in Banach algebras”. *Banach Center Publications* 30 (1994), pp. 53–100.
- [6] J. B. Conway. *A course in functional analysis*. 2nd ed. Vol. 96. Graduate Texts in Mathematics. Springer-Verlag, 1990.
- [7] J.-F. Delmas, D. Dronnier, and P.-A. Zitt. *An Infinite-Dimensional SIS Model*. Version 1. 2020. arXiv: 2006.08241 [math.DS].
- [8] J.-F. Delmas, D. Dronnier, and P.-A. Zitt. “Effective reproduction number: convexity, invariance and cordons sanitaires”. 2021.
- [9] J.-F. Delmas, D. Dronnier, and P.-A. Zitt. “Optimal vaccination for a 2 sub-populations SIS model”. 2021.
- [10] J.-F. Delmas, D. Dronnier, and P.-A. Zitt. “Optimal vaccination: various (counter) intuitive examples”. 2021.
- [11] J.-F. Delmas, D. Dronnier, and P.-A. Zitt. “Vaccinating highly connected people is (sometimes) optimal”. 2021.
- [12] 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 (1990), pp. 365–382.
- [13] K. Dietz. “Transmission and control of arbovirus diseases”. *Epidemiology* 104 (1975), pp. 104–121.
- [14] 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 (2016), pp. 68–81.
- [15] L. E. Duijzer, W. L. van Jaarsveld, J. Wallinga, and R. Dekker. “Dose-optimal vaccine allocation over multiple populations”. *Production and Operations Management* 27.1 (2018), pp. 143–159.
- [16] F. Esser and F. Harary. “On the spectrum of a complete multipartite graph”. *European Journal of Combinatorics* 1.3 (1980), pp. 211–218.
- [17] 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 (2017), pp. 93–104.
- [18] 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 (2015), pp. 177–187.
- [19] P. Fine, K. Eames, and D. L. Heymann. “Herd Immunity: A Rough Guide”. *Clinical Infectious Diseases* 52.7 (2011), pp. 911–916.
- [20] A. Ganesh, L. Massoulie, and D. Towsley. “The effect of network topology on the spread of epidemics”. *Proceedings IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies*. (Miami, FL, USA). 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* 7.46 (2010), pp. 755–764.
- [22] J. J. Grobler. “Compactness conditions for integral operators in Banach function spaces”. *Indagationes Mathematicae (Proceedings)* 32 (1970), pp. 287–294.
- [23] C. J. A. Halberg Jr. and A. E. Taylor. “On the spectra of linked operators”. *Pacific Journal of Mathematics* 6 (1956), pp. 283–290.
- [24] H. W. Hethcote and J. A. Yorke. *Gonorrhea Transmission Dynamics and Control*. Ed. by S. Levin. Vol. 56. Lecture Notes in Biomathematics. Springer-Verlag, 1984.
- [25] A. N. Hill and I. M. Longini Jr. “The critical vaccination fraction for heterogeneous epidemic models”. *Mathematical Biosciences* 181.1 (2003), pp. 85–106.
- [26] S. Huggett and D. Jordan. *A topological aperitif*. 2nd ed. Springer-Verlag, 2009.
- [27] S. Janson. *Graphons, cut norm and distance, couplings and rearrangements*. NYJM Monographs. New York Journal of Mathematics, 2013.
- [28] O. Kallenberg. *Foundations of modern probability*. 3rd ed. Vol. 99. Probability Theory and Stochastic Modelling. Springer-Verlag, 2021, p. 946.
- [29] T. Kato. *Perturbation theory for linear operators*. Vol. 132. Springer-Verlag, 2013.
- [30] M. J. Keeling and P. Rohani. *Modeling infectious diseases in humans and animals*. Princeton University Press, 2008.
- [31] B. R. Kloeckner. “Effective perturbation theory for simple isolated eigenvalues of linear operators”. *Journal of Operator Theory* 81.1 (2019), pp. 175–194.
- [32] H. König. *Eigenvalue distribution of compact operators*. Vol. 16. Operator Theory: Advances and Applications. Birkhäuser Verlag, 1986.
- [33] A. Lajmanovich and J. A. Yorke. “A deterministic model for gonorrhea in a nonhomogeneous population”. *Mathematical Biosciences* 28.3 (1976), pp. 221–236.
- [34] L. Lovász. *Large networks and graph limits*. American Mathematical Society colloquium publications 60. American Mathematical Society, 2012.
- [35] I. Marek. “Frobenius theory of positive operators: comparison theorems and applications”. *SIAM Journal on Applied Mathematics* 19.3 (1970), pp. 607–628.
- [36] K. Miettinen. *Nonlinear multiobjective optimization*. Springer-Verlag, 1998.
- [37] J. D. Newburgh. “The variation of spectra”. *Duke Mathematical Journal* 18.1 (1951), pp. 165–176.
- [38] 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 (2018), pp. 1795–1831.
- [39] J. G. Restrepo, E. Ott, and B. R. Hunt. “Characterizing the Dynamical Importance of Network Nodes and Links”. *Physical Review Letters* 97.9 (2006).
- [40] S. Saha, A. Adiga, B. A. Prakash, and A. K. S. Vullikanti. “Approximation Algorithms for Reducing the Spectral Radius to Control Epidemic Spread”. *Proceedings of the 2015 SIAM International Conference on Data Mining*. Society for Industrial and Applied Mathematics, 2015.
- [41] H. H. Schaefer. *Banach lattices and positive operators*. Vol. 215. Grundlehren der mathematischen Wissenschaften. Springer-Verlag, 1974.
- [42] C. E. G. Smith. “Prospects for the control of infectious disease”. *Proceedings of the Royal Society of Medicine* 63.11 Pt 2 (1970), pp. 1181–1190.
- [43] P. G. Smith. “Concepts of herd protection and immunity”. *Procedia in Vaccinology* 2.2 (2010), pp. 134–139.
- [44] M. Somerville, K. Kumaran, and R. Anderson. *Public health and epidemiology at a glance*. John Wiley & Sons, 2016.
- [45] D. Stevanović, I. Gutman, and M. U. Rehman. “On spectral radius and energy of complete multipartite graphs”. *Ars Mathematica Contemporanea* 9.1 (2015), pp. 109–113.

- [46] 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* 84.1 (2011).
- [47] H. Zhao and Z. Feng. “Identifying optimal vaccination strategies via economic and epidemiological modeling”. *Journal of Biological Systems* 27.4 (2019), pp. 423–446.

JEAN-FRANÇOIS DELMAS, CERMICS, ÉCOLE DES PONTS, FRANCE  
*Email address:* `jean-francois.delmas@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`