

Equivalence by coupling for heterogeneous SIS models*

Jean-François Delmas[†] Dylan Dronnier[†] Pierre-André Zitt[‡]

Abstract

We consider the optimal allocation of (perfect) vaccine in an heterogeneous SIS model. Using a coupling approach, we explain how different models for the heterogeneity of the population lead to the same Pareto frontier in the cost/loss valuation of the vaccinations strategies. This covers in particular the elementary continuous representation of discrete models and the measure preserving transformation which appears in graphon theory.

Keywords: Coupling; SIS Model; vaccination strategy; effective reproduction number; multi-objective optimization; Pareto frontier.

MSC2020 subject classifications: 92D30;58E17.

1 Introduction

We consider the effect of vaccination in an heterogeneous SIS model (with S=Susceptible and I=Infectious), in the framework introduced in [1]. The model, which will be recalled in detail below, is parametrized by four elements: a feature space, denoted by \mathcal{X} ; two real-valued functions γ and c on \mathcal{X} , representing the feature-dependent recovery rate and vaccination cost; a real-valued function k on \mathcal{X}^2 , encoding the infection rate between individuals of different features. We focus on optimizing feature-dependent vaccination strategies, as discussed in [2].

In classical probability theory, the same random experiment may be represented by two different probability spaces and random variables, with the same distribution. Unsurprisingly, the same situation occurs here in the choice of the trait space and the associated parameters. The goal of this article is to describe precisely a notion of equivalence between models via a coupling, and to compare equilibria and optimal vaccination strategies between equivalent models.

We address the three following questions, see Theorem 3.6 and Corollary 3.7:

1. Do equivalent models lead to comparable optimal vaccination strategies? Is knowing the optima for one model enough to find the optima in equivalent models?
2. If the feature space is “too rich”, and encodes features that are not relevant to the propagation of the epidemic, is it possible to reduce the model by “forgetting” irrelevant features?
3. Do equivalent models evolve in the same way, and in particular can we compare their equilibria?

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

[†]CERMICS, École nationale des ponts et chaussées, 77455 Marne-la-Vallée, France. E-mail: jean-francois.delmas, dylan.dronnier@enpc.fr

[‡]LAMA, Université Gustave Eiffel, 77420 Champs-sur-Marne, France. E-mail: pierre-andre.zitt@univ-eiffel.fr

In the next section, we introduce the necessary notation, borrowing heavily from the presentation of [7]. The main result is stated in Section 3 and gives positive answers to the three questions; the proofs are postponed to Section 5. Detailed examples are discussed in Section 4.

2 Framework and notation

2.1 The heterogeneous SIS model

We recall the differential equations governing the epidemic dynamics in meta-population SIS models introduced in [1], to which we refer for additional context and details.

Let $(\mathcal{X}, \mathcal{F}, \mu)$ be a probability space, where $x \in \mathcal{X}$ represents a feature and the probability measure $\mu(dx)$ represents the fraction of the population with feature x . The parameters of the SIS model are given by a *recovery rate function* γ , which is a positive bounded measurable function defined on \mathcal{X} , and a *transmission rate kernel* k , where a kernel is a nonnegative measurable function defined on \mathcal{X}^2 . In accordance with [1], we consider for a kernel k on \mathcal{X} and $q \in (1, +\infty)$ its norm: $\|k\|_{\infty, q} = \sup_{x \in \mathcal{X}} \left(\int_{\mathcal{X}} k(x, y)^q \mu(dy) \right)^{1/q}$. For a kernel k on \mathcal{X} such that $\|k\|_{\infty, q}$ is finite for some $q \in (1, +\infty)$, we define the integral operator \mathcal{T}_k on the set \mathcal{L}^∞ of bounded measurable real-valued function on \mathcal{X} by:

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

By convention, for f, g two nonnegative measurable functions defined on \mathcal{X} and k a kernel on \mathcal{X} , we denote by fkg the kernel on \mathcal{X} defined by:

$$fkg : (x, y) \mapsto f(x)k(x, y)g(y). \quad (2.1)$$

We shall consider the kernel $\mathbf{k} = k\gamma^{-1}$, which is thus defined by:

$$\mathbf{k}(x, y) = k(x, y)\gamma(y)^{-1}.$$

We assume that:

$$\|\mathbf{k}\|_{\infty, q} < \infty \quad \text{for some } q \in (1, +\infty). \quad (2.2)$$

The integral operator $\mathcal{T}_{\mathbf{k}}$ is the so called *next-generation operator*.

Let $\Delta = \{f \in \mathcal{L}^\infty : 0 \leq f \leq 1\}$ be the subset of nonnegative functions bounded by 1, and let $0, 1 \in \Delta$ be the constant functions equal respectively to 0 and to 1. The SIS dynamics considered in [1] follows the vector field F defined on Δ by:

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

More precisely, we consider $u = (u_t, t \in \mathbb{R})$, where $u_t \in \Delta$ for all $t \in \mathbb{R}_+$, and u solves in \mathcal{L}^∞ :

$$\partial_t u_t = F(u_t) \quad \text{for } t \in \mathbb{R}_+, \quad (2.3)$$

with initial condition $u_0 \in \Delta$. The value $u_t(x) = u(t, x)$ models the probability that an individual of feature x is infected at time t ; it is proved in [1] that such a solution u exists and is unique.

An *equilibrium* of (2.3) is a function $g \in \Delta$ such that $F(g) = 0$. According to [1], there exists a *maximal equilibrium* \mathbf{g} , that is, an equilibrium such that all other equilibria $h \in \Delta$ are dominated by \mathbf{g} : $h \leq \mathbf{g}$. This maximal equilibrium is obtained as the long time pointwise limit of the SIS model started with its whole population infected: $\lim_{t \rightarrow \infty} u_t = \mathbf{g}$ where $u_0 = 1$. The *fraction of infected individuals at equilibrium*, \mathfrak{I}_0 , is thus given by:

$$\mathfrak{I}_0 = \int_{\mathcal{X}} \mathbf{g} d\mu.$$

For T a bounded operator on \mathcal{L}^∞ endowed with its usual supremum norm, we denote by $\|T\|_{\mathcal{L}^\infty}$ its operator norm. The spectral radius of T is then given by $\rho(T) = \lim_{n \rightarrow \infty} \|T^n\|_{\mathcal{L}^\infty}^{1/n}$. The *reproduction number* R_0 associated to the SIS model given by (2.3) is the spectral radius of the next-generation operator:

$$R_0 = \rho(\mathcal{T}_{\mathbf{k}}). \quad (2.4)$$

If $R_0 \leq 1$ (sub-critical and critical case), then u_t converges pointwise to $\mathbb{0}$ when $t \rightarrow \infty$. In particular, the maximal equilibrium \mathbf{g} is equal to $\mathbb{0}$ and $\mathfrak{F}_0 = 0$. If $R_0 > 1$ (super-critical case), then $\mathbb{0}$ is still an equilibrium but different from the maximal equilibrium \mathbf{g} , as $\mathfrak{F}_0 = \int_{\mathcal{X}} \mathbf{g} \, d\mu > 0$.

2.2 Vaccination strategies

A *vaccination strategy* η of a vaccine with perfect efficiency is an element of Δ , where $\eta(x)$ represents the proportion of **non-vaccinated** individuals with feature x . Notice that $\eta \, d\mu$ corresponds in a sense to the effective population. In particular, the “strategy” that consists in vaccinating no one corresponds to $\eta = \mathbb{1}$, the constant function equal to 1, while $\eta = \mathbb{0}$, the constant function equal to 0, corresponds to vaccinating everybody.

Recall the definition of the kernel fkg from (2.1). For $\eta \in \Delta$, the kernel $\mathbf{k}\eta = k\eta/\gamma$ has finite norm $\|\cdot\|_{\infty, q}$, so we can consider the bounded positive operators $\mathcal{T}_{\mathbf{k}\eta}$ and $\mathcal{T}_{k\eta}$ on \mathcal{L}^∞ . According to [1, Section 5.3.], the SIS equation with vaccination strategy η is given by $u^\eta = (u_t^\eta, t \geq 0)$ solution to (2.3) with F is replaced by F_η defined by:

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

The quantity $u_t^\eta(x) = u^\eta(t, x)$ then represents the probability for a non-vaccinated individual of feature x to be infected at time t ; so at time t among the population of feature x , a fraction $1 - \eta(x)$ is vaccinated, a fraction $\eta(x) u_t^\eta(x)$ is not vaccinated and infected, and a fraction $\eta(x) (1 - u_t^\eta(x))$ is not vaccinated and not infected.

We define the *effective reproduction number* $R_e(\eta)$ associated to the vaccination strategy η as the spectral radius of the effective next-generation operator $\mathcal{T}_{\mathbf{k}\eta}$:

$$R_e(\eta) = \rho(\mathcal{T}_{\mathbf{k}\eta}). \quad (2.5)$$

For example, for the trivial vaccination strategies we get $R_e(\mathbb{1}) = R_0$ and $R_e(\mathbb{0}) = 0$. We also denote by \mathbf{g}_η the corresponding maximal equilibrium and, using that $\eta \, d\mu$ is the effective population, we define the *effective fraction of infected individuals at equilibrium* as:

$$\mathfrak{F}(\eta) = \int_{\mathcal{X}} \mathbf{g}_\eta \, \eta \, d\mu. \quad (2.6)$$

For example, we have $\mathfrak{F}(\mathbb{1}) = \mathfrak{F}_0$ and $\mathfrak{F}(\eta) = 0$ for all $\eta \in \Delta$ such that $R_e(\eta) \leq 1$.

2.3 Optimal strategies

For a vaccination strategy $\eta \in \Delta$, we consider its loss $L(\eta)$, given either by the effective reproduction number ($L = R_e$) or by the effective fraction of infected individuals at equilibrium ($L = \mathfrak{F}$). Following [2], we measure the cost for the society of a vaccination strategy (production, diffusion, ...) by a nonnegative function C defined on Δ . We shall concentrate on the affine case:

$$C(\eta) = \int_{\mathcal{X}} (\mathbb{1} - \eta) \, c \, d\mu$$

where the nonnegative function $c \in L^1$ represents the feature-dependent cost of vaccinating individuals. Notice that doing nothing costs nothing, that is, $C(\mathbb{1}) = 0$. A simple and natural choice is the uniform cost C_{uni} corresponding to $c = \mathbb{1}$.

Let us note that if H is any of the three functionals R_e , \mathfrak{F} or C , and if $\eta_1 = \eta_2$ μ -a.s., then $H(\eta_1) = H(\eta_2)$. Following [2] we therefore consider the set of vaccination strategies as a subset of L^∞ :

$$\Delta = \{\eta \in L^\infty : 0 \leq \eta \leq 1 \quad \mu - \text{a.s.}\}. \quad (2.7)$$

In [2, Section 4], we formalized and study the problem of optimal allocation strategies for a perfect vaccine in the SIS model. This question may be viewed as a bi-objective minimization problem, where one tries to minimize simultaneously the cost of the vaccination and its corresponding loss:

$$\min_{\Delta} (C, L).$$

We call a strategy η_\star *Pareto optimal* if no other strategy is strictly better:

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

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

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

We call a strategy η^\star *anti-Pareto optimal* if no other strategy is strictly worse, that is, $C(\eta) > C(\eta^\star) \implies L(\eta) < L(\eta^\star)$ and $L(\eta) > L(\eta^\star) \implies C(\eta) < C(\eta^\star)$. The set of anti-Pareto optimal strategies will be denoted by $\mathcal{P}^{\text{Anti}} \subset \Delta$, and we define the *anti-Pareto frontier* as the set of anti-Pareto optimal outcomes $\mathcal{F}^{\text{Anti}} = \{(C(\eta^\star), L(\eta^\star)) : \eta^\star \in \mathcal{P}^{\text{Anti}}\}$. We refer to [2] for an extensive study and alternate characterizations of the Pareto and anti-Pareto frontiers; let us simply mention that under our assumptions both frontiers are non-trivial.

2.4 Parameters of the SIS model in a nutshell

Let us summarize the setup. The SIS model is given by a probability space $(\mathcal{X}, \mathcal{F}, \mu)$, a positive recovery rate function $\gamma \in \mathcal{L}^\infty(\mathcal{X}, \mathcal{F})$, a transmission rate kernel k (that is, a measurable nonnegative function defined on \mathcal{X}^2) such that $\|k/\gamma\|_{\infty, q} < \infty$ for some $q \in (1, +\infty)$, see (2.2), and an affine cost function with a nonnegative density $\mathfrak{c} \in L^1(\mathcal{X}, \mathcal{F}, \mu)$. We denote the parameters of the SIS model by:

$$\text{Param} = [(\mathcal{X}, \mathcal{F}, \mu), (k, \gamma), \mathfrak{c}].$$

Finally, we write $H[\text{Param}]$ to emphasize the dependence of any quantity H on the parameters: for example $R_e[\text{Param}](\eta)$ is the effective reproduction number associated to the vaccination strategy η in the model defined by Param .

3 Equivalence of models by coupling

We now define our main tool: the *coupling* of two SIS models, which gives rise to a notion of *conjugation* between functions defined on the first and the second model. This tool is then used to state our main results. All proofs are postponed to Section 5.

Remark 3.1 (Graphons and weak isometry). In Section 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 refer the reader to [12] for similar developments in the graphon setting.

3.1 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 σ -field. Let f be a function from E to E' .

We denote by $\sigma(f) = \{f^{-1}(A) : A \in \mathcal{E}'\}$ the σ -field generated by f . In particular the function f is measurable from (E, \mathcal{E}) to (E', \mathcal{E}') if and only if $\sigma(f) \subset \mathcal{E}$. Let φ be a measurable function from (E, \mathcal{E}) to (E', \mathcal{E}') . For ν a measure on (E, \mathcal{E}) , we write $\nu' = \varphi_{\#}\nu$ for the push-forward measure on (E', \mathcal{E}') of the measure ν by the function φ ; by definition of ν' , for a nonnegative measurable function g defined from (E', \mathcal{E}') to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, we have:

$$\int_{E'} g d\nu' = \int_E g \circ \varphi d\nu. \quad (3.1)$$

In particular, if f, g are measurable functions defined from (E', \mathcal{E}') to some measurable space, then we have that ν' -a.e. $f = g$ if and only if ν -a.e. $g \circ \varphi = f \circ \varphi$. Thus, if g belongs to $L^p(E', \mathcal{E}', \nu')$, then $g \circ \varphi$ is well defined as an element of $L^p(E, \mathcal{E}, \nu)$.

Let f be a measurable function from (E, \mathcal{E}) to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. We recall (see for example [13, Lemma 1.14]) that:

$$\sigma(f) \subset \sigma(\varphi) \implies f = g \circ \varphi, \quad (3.2)$$

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

In what follows the random variables are defined on some probability space $(\Omega_0, \mathcal{F}_0, \mathbb{P})$.

3.2 Coupling and conjugate functions

Let $(E_1, \mathcal{E}_1, \mu_1)$ and $(E_2, \mathcal{E}_2, \mu_2)$ be measurable spaces. A *coupling* is a measure π on $(E_1 \times E_2, \mathcal{E}_1 \otimes \mathcal{E}_2)$ with marginals μ_1 and μ_2 . By abuse of notation we also call coupling a random variable $Z = (Z_1, Z_2)$ with distribution π , and also say that E_1 and E_2 are coupled through Z .

We introduce a notion of conjugacy whose basic properties are similar to convex conjugation.

Definition 3.2 (Conjugate functions). *Let $(E_1, \mathcal{E}_1, \mu_1)$ and $(E_2, \mathcal{E}_2, \mu_2)$ be coupled through (Z_1, Z_2) . Let $f_i \in L^1(E_i, \mu_i)$ for $i = 1, 2$. The conjugate f_1^* of f_1 is the element of $L^1(E_2)$ defined by:*

$$f_1^*(Z_2) = \mathbb{E}[f_1(Z_1) | \mathcal{C}] \quad \text{with} \quad \mathcal{C} = \sigma(Z_1) \cap \sigma(Z_2);$$

its existence is justified by (3.2). Similarly $f_2^ \in L^1(E_1)$ is defined by $f_2^*(Z_1) = \mathbb{E}[f_2(Z_2) | \mathcal{C}]$.*

The pair (f_1, f_2) is called conjugate if $f_1 = f_2^$ and $f_2 = f_1^*$; it is called pre-conjugate if the pair (f_2^*, f_1^*) is conjugate (that is, $f_2^* = f_1^{**}$ and $f_1^* = f_2^{**}$).*

Notice that a conjugate pair is also pre-conjugate, but the converse is false in general.

We shall see below that if the transmission kernels, recovery functions and the density of the cost functions of two SIS model are conjugate, then any vaccinations strategies which are pre-conjugate have the same loss and cost, and thus are (anti-)Pareto optima simultaneously.

We first give another characterization of the conjugation.

Lemma 3.3 (Characterization of conjugation). *Let $(E_1, \mathcal{E}_1, \mu_1)$ and $(E_2, \mathcal{E}_2, \mu_2)$ be coupled through (Z_1, Z_2) . Let $f_i \in L^1(E_i)$ for $i = 1, 2$. We have:*

$$(f_1, f_2) \text{ is conjugate} \iff f_1(Z_1) = f_2(Z_2) \quad \pi\text{-a.s.}$$

If the pair (f_1, f_2) is conjugate, then $f_i(Z_i)$ is \mathcal{C} -measurable for $i = 1, 2$, with $\mathcal{C} = \sigma(Z_1) \cap \sigma(Z_2)$.

Proof. The proof is immediate as, for X and Y integrable random variables and a sub- σ -field \mathcal{C} , the equalities $\mathbb{E}[X | \mathcal{C}] = Y$ and $\mathbb{E}[Y | \mathcal{C}] = X$ imply that a.s. $X = Y$. \square

We shall complete the next result with other properties in Section 5.

Lemma 3.4 (Properties of conjugation). *Let $(E_1, \mathcal{E}_1, \mu_1)$ and $(E_2, \mathcal{E}_2, \mu_2)$ be coupled through $Z = (Z_1, Z_2)$. Let $f \in L^1(E_1)$.*

- (i) *The pair (f, f^*) is pre-conjugate and the pair (f^{**}, f^*) is conjugate.*

(ii) Set $\mathcal{C} = \sigma(Z_1) \cap \sigma(Z_2)$. We have:

$$f(Z_1) \text{ is } \mathcal{C}\text{-measurable} \iff f = f^{**} \iff (f, f^*) \text{ is conjugate.}$$

Proof. By definition, we have $f^{**}(Z_1) = \mathbb{E}[f^*(Z_2)|\mathcal{C}] = \mathbb{E}[\mathbb{E}[f(Z_1)|\mathcal{C}]|\mathcal{C}]$, which yields that $f^{**}(Z_1) = f^*(Z_2)$. By Lemma 3.3 this implies that (f^{**}, f^*) is conjugate and thus that (f, f^*) is pre-conjugate. This gives (i).

We now prove (ii). Notice first that $f^{**} = f$ is equivalent to the pair (f, f^*) being conjugate. Secondly, if (f, f^*) is conjugate, then by Lemma 3.3, we get that $f(Z_1)$ is \mathcal{C} -measurable. Conversely, if $f(Z_1)$ is \mathcal{C} -measurable, we deduce that $f^*(Z_2) = f(Z_1)$ and thus $f^{**} = f$. \square

Let two spaces E_1 and E_2 be coupled through π . The product spaces $\mathbf{E}_1 = E_1 \times E_1$ and $\mathbf{E}_2 = E_2 \times E_2$ may always be coupled through the random variable $(\mathbf{Z}_1, \mathbf{Z}_2) = ((X_1, Y_1), (X_2, Y_2))$, where the two vectors (X_1, X_2) and (Y_1, Y_2) are independent and follow the distribution π . We denote the distribution of $(\mathbf{Z}_1, \mathbf{Z}_2)$ by π and call it the *extended coupling*. Conjugates are preserved by extension in the following sense; the proof is given in Section 5.2.

Lemma 3.5 (Extended coupling and conjugacy). *If the measurable function $g : \mathbf{E}_1 \rightarrow \mathbb{R}$ only depends on its first argument, $g(x_1, y_1) = f(x_1)$, then $g^*(X_2, Y_2) = f^*(X_2)$ (where g^* is the conjugate through π and f^* the conjugate through π).*

3.3 Coupled models

We consider the SIS models $\text{Param}_i = [(\mathcal{X}_i, \mathcal{F}_i, \mu_i), (k_i, \gamma_i), c_i]$ for $i = 1, 2$. In what follows, we simply write Δ_i for the set of functions Δ , see (2.7), in the model Param_i .

Theorem 3.6 (Coupling, equilibria and optimal vaccinations). *Consider two SIS models Param_1 and Param_2 , with a coupling between $(\mathcal{X}_1, \mathcal{F}_1, \mu_1)$ and $(\mathcal{X}_2, \mathcal{F}_2, \mu_2)$. Let $\eta_i \in \Delta_i$ be a vaccination strategies for the SIS model $i = 1, 2$.*

(i) *If the pair $(k_1/\gamma_1, k_2/\gamma_2)$ is conjugate (for the extended coupling), then*

$$(\eta_1, \eta_2) \text{ is pre-conjugate} \implies R_e[\text{Param}_1](\eta_1) = R_e[\text{Param}_2](\eta_2).$$

(ii) *If both pairs (k_1, k_2) and (γ_1, γ_2) are conjugate, then the equilibria are (pre-)conjugate: if g_1 is an equilibrium of Param_1 , then there exists an equilibrium g_2 of Param_2 such that the pair (g_1, g_2) is conjugate. We also have:*

$$(\eta_1, \eta_2) \text{ is pre-conjugate} \implies \mathfrak{F}[\text{Param}_1](\eta_1) = \mathfrak{F}[\text{Param}_2](\eta_2).$$

(iii) *Suppose the assumptions of item (i), for $L = R_e$, or of item (ii), for $L = \mathfrak{F}$, hold. Assume also that the pair (c_1, c_2) is conjugate. If the pair (η_1, η_2) is pre-conjugate, then::*

$$\eta_1 \text{ is (anti-)Pareto optimal for } \text{Param}_1 \iff \eta_2 \text{ is (anti-)Pareto optimal for } \text{Param}_2. \quad (3.3)$$

For $\eta \in \Delta_1$, we have $\eta^ \in \Delta_2$ and $H[\text{Param}_1](\eta) = H[\text{Param}_2](\eta^*)$ for H equal to the loss L or the cost C ; in particular, if η is (anti-)Pareto optimal for Param_1 , then its conjugate η^* is (anti-)Pareto optimal for Param_2 .*

As a direct consequence, we get the following result, where the set of outcomes is defined as $\mathbf{F} = \{(C(\eta), L(\eta)), \eta \in \Delta\}$.

Corollary 3.7 (Coupling and frontiers). *Let Param_1 and Param_2 be coupled SIS models, with conjugate parameters γ , c and k . For any of the two choices $L \in \{R_e, \mathfrak{F}\}$, the models Param_1 and Param_2 have the same set of outcomes \mathbf{F} and the same (anti-)Pareto frontiers \mathcal{F} and $\mathcal{F}^{\text{Anti}}$.*

Remark 3.8 (Obvious couplings). If the costs are uniform in both models Param_1 and Param_2 , then the pair (c_1, c_2) is trivially conjugate as both functions are a.s. constant equal to 1.

Using a trivial coupling, one sees that the recovery rate and transmission kernel in the SIS model could have been defined only almost everywhere without affecting the set of outcomes and the (anti-)Pareto frontiers.

The coupling hypotheses are strong and give strong results, allowing to compare equilibria and vaccinations between models. Let us note that other, weaker ways of comparing models exist, and may yield interesting results.

Remark 3.9 (Life without coupling — normalizing γ and c). If we are only interested in the loss function $L = R_e$, various invariance properties of the spectral radius may be used to normalize models. Indeed, consider a SIS model $\text{Param} = [(\mathcal{X}, \mathcal{F}, \mu), (k, \gamma), c]$ for which both γ and c are bounded away from zero, and assume without loss of generality that $\int_{\mathcal{X}} c \, d\mu = 1$. Define another model by $\text{Param}_0 = [(\mathcal{X}, \mathcal{F}, \mu_0), (k_0, \gamma_0), c_0]$, where:

$$\mu_0(dx) = c(x) \mu(dx), \quad k_0 = k/(c\gamma), \quad \gamma_0 = c_0 = 1.$$

Notice that as (2.2) holds for the model Param , then it also holds for the model Param_0 as we assumed c to be bounded away from 0.

We trivially have $\Delta(\text{Param}) = \Delta(\text{Param}_0)$. Clearly, we have $C(\eta)$ for the model Param is equal to $C_{\text{uni}}(\eta)$ for the model Param_0 . Using also that $L^p(\mu)$ and $L^p(\mu_0)$ are compatible (see [5, Lemma 2.2]) and the corresponding integral operators are consistent (see [5, Section 2.2] and Lemma 2.1(iii)), we get that $R_e(\eta)$ for the model Param is equal to $R_e(\eta)$ for the model Param_0 , for all strategies $\eta \in \Delta$. In particular the (anti-)Pareto optimal strategies and the (anti-)Pareto frontiers are the same for the two models. Therefore we may focus on Param_0 and assume without loss of generality that the only dependence on the features is in the transmission kernel, while both the vaccination cost and the recovery rate are uniform.

4 Examples of couplings

We discuss three examples, all of which are built on the following special case of coupling, each one taking a slightly different point of view.

Lemma 4.1 (Deterministic coupling). *Let $(E_1, \mathcal{E}_1, \mu_1)$ and $(E_2, \mathcal{E}_2, \mu_2)$ be two probability spaces and assume that $\phi : E_1 \rightarrow E_2$ is measurable and pushes μ_1 forward to μ_2 . Then E_1 and E_2 are coupled through $(X_1, \phi(X_1))$, with $X_1 \sim \mu_1$, and for any two functions $f_i \in L^1(E_i)$, $i = 1, 2$ we have, with \mathbb{E}_1 the expectation w.r.t. μ_1 :*

1. $f_2^* = f_2 \circ \phi$, $f_1^* \circ \phi = \mathbb{E}_1[f_1 \mid \sigma(\phi)]$ and $f_2^{**} = f_2$;
2. The pair (f_1, f_2) is conjugate if and only if $f_1 = f_2 \circ \phi$;
3. The pair (f_1, f_2) is pre-conjugate if and only if $f_2 = f_1^*$;
4. The pair of kernels (k_1, k_2) (respectively on \mathbf{E}_1 and \mathbf{E}_2) is conjugate (through the extended coupling) if and only if $\mu_1(dx_1) \otimes \mu_1(dy_1)$ -a.e. $k_1(x_1, y_1) = k_2(\phi(x_1), \phi(y_1))$.

The proof is elementary and left to the reader.

4.1 Starting from E_1 : model reduction using deterministic coupling

We consider a SIS model $\text{Param}_1 = [(\mathcal{X}_1, \mathcal{F}_1, \mu_1), (k_1, \gamma_1), c_1]$. Let ϕ be a measurable function from $(\mathcal{X}_1, \mathcal{F}_1)$ to $(\mathcal{X}_2, \mathcal{F}_2)$, let μ_2 be the push-forward $\phi_{\#}\mu_1$, and consider the coupling given by $(X_1, \phi(X_1))$ where $X_1 \sim \mu_1$. By Lemma 4.1, the functions c_1, γ_1 and k_1 will be part of conjugate pairs for this coupling if and only if they all factor through ϕ , in the sense that for some functions c_2, γ_2 on \mathcal{X}_2 and k_2 on $\mathcal{X}_2 \times \mathcal{X}_2$:

$$c_1 = c_2 \circ \phi, \quad \gamma_1 = \gamma_2 \circ \phi \quad \text{and} \quad k_1(\cdot, \cdot) = k_2(\phi(\cdot), \phi(\cdot)). \quad (4.1)$$

If that is the case, then by Theorem 3.6 and Lemma 4.1, the vaccination strategy $\eta_1 \in \Delta_1$ is (anti-)Pareto optimal for Param_1 if and only if its conjugate η_1^* defined by $\eta_1^* \circ \phi = \mathbb{E}[\eta_1 | \sigma(\phi)]$ is (anti-)Pareto optimal for the simplified model Param_2 . In words, the behaviour of an individual x only depends on $\phi(x)$, and in the trait space \mathcal{X}_2 , individuals with identical behavior are merged.

We may deduce the following result.

Corollary 4.2 (Model reduction). *Let $\text{Param} = [(\mathcal{X}, \mathcal{F}, \mu), (k, \gamma), \mathbf{c}]$ be a SIS model with loss function $L \in \{R_e, \mathfrak{F}\}$. Let $\mathcal{G} \subset \mathcal{F}$ be a σ -field such that γ and \mathbf{c} are \mathcal{G} -measurable and k is $\mathcal{G} \otimes \mathcal{G}$ -measurable. Then, for any $\eta \in \Delta$, we have, with \mathbb{E}_μ the expectation w.r.t. μ :*

$$\eta \text{ is (anti-)Pareto optimal} \iff \mathbb{E}_\mu[\eta | \mathcal{G}] \text{ is (anti-)Pareto optimal.} \quad (4.2)$$

Proof. Denote with a subscript 1 the parameters of the original model (e.g., set $\mathcal{X}_1 = \mathcal{X}$). Let $\text{Param}_2 = [(\mathcal{X}_2, \mathcal{F}_2, \mu_2), (k_2, \gamma_2), \mathbf{c}_2]$ where most parameters are the same: $\mathcal{X}_2 = \mathcal{X}$, $k_2 = k$, $\gamma_2 = \gamma$, $\mathbf{c}_2 = \mathbf{c}$, but we equip \mathcal{X}_2 with $\mathcal{F}_2 = \mathcal{G}$, and the measure $\mu_2 = (\mu_1)_{|\mathcal{G}}$. Note that this is legitimate, in the sense that the measurability hypotheses on (k, γ, \mathbf{c}) , imply that γ_2, \mathbf{c}_2 are measurable from $(\mathcal{X}_2, \mathcal{F}_2)$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ and k_2 is measurable on the product space $(\mathcal{X}_2 \times \mathcal{X}_2, \mathcal{F}_2 \otimes \mathcal{F}_2)$.

Now we define $\phi : \mathcal{X}_1 \rightarrow \mathcal{X}_2$ by $\phi(x) = x$; since $\mathcal{G} \subset \mathcal{F}$, ϕ is measurable from $(\mathcal{X}_1, \mathcal{F}_1) = (\mathcal{X}, \mathcal{F})$ to $(\mathcal{X}_2, \mathcal{F}_2) = (\mathcal{X}, \mathcal{G})$. This function defines a deterministic coupling between the two spaces. Since ϕ is the identity if we forget the measure structure, it is clear that $\gamma_1 = \gamma_2 \circ \phi$, $\mathbf{c}_1 = \mathbf{c}_2 \circ \phi$ and $k_1(\cdot, \cdot) = k_2(\phi(\cdot), \phi(\cdot))$, so that all three pairs of functions are conjugate, by Lemma 4.1. Applying Theorem 3.6 twice, we get that $\eta \in \Delta$ is Pareto-optimal for Param_1 if and only if η^* is Pareto-optimal for Param_2 , if and only if η^{**} is Pareto-optimal for Param_1 .

Let us finally identify η^{**} . Let X be μ_1 -distributed. The coupling is (Z_1, Z_2) where $Z_1 = X$ and $Z_2 = \phi(Z_1)$, so $\mathcal{C} = \sigma(Z_1) \cap \sigma(Z_2) = \sigma(Z_2) = X^{-1}(\phi^{-1}(\mathcal{G})) = X^{-1}(\mathcal{G})$. By definition we have $\eta^{**}(X) = \mathbb{E}[\eta(X) | X^{-1}(\mathcal{G})]$. We deduce that $\eta^{**} = \mathbb{E}_\mu[\eta | \mathcal{G}]$ as for any $B \in \mathcal{G}$ and $A = X^{-1}(B)$:

$$\begin{aligned} \mathbb{E}[\eta(X) \mathbb{1}_A] &= \mathbb{E}[\eta(X) \mathbb{1}_B(X)] = \int_{\mathcal{X}} \eta(x) \mathbb{1}_B(x) \mu_1(dx) = \int_{\mathcal{X}} \mathbb{E}_\mu[\eta | \mathcal{G}](x) \mathbb{1}_B(x) \mu_1(dx) \\ &= \mathbb{E}[\mathbb{E}_\mu[\eta | \mathcal{G}](X) \mathbb{1}_A]. \end{aligned} \quad \square$$

4.2 Linking E_1 and E_2 : discrete and continuous models

We now consider a particular case, and formalize how finite population models can be seen as images of models with a continuous population. We denote by $\mathcal{B}([0, 1])$ and by Leb the Borel σ -field and the Lebesgue measure on $[0, 1]$.

Let $\mathcal{X}_d \subset \mathbb{N}$, \mathcal{F}_d the set of subsets of \mathcal{X}_d and μ_d a probability measure on \mathcal{X}_d . Without loss of generality, we can assume that $\mu_d(\{\ell\}) > 0$ for all $\ell \in \mathcal{X}_d$. We set $\mathcal{X}_c = [0, 1]$, $\mathcal{F}_c = \mathcal{B}([0, 1])$ and let μ_c be a probability measure on $(\mathcal{X}_c, \mathcal{F}_c)$ without atoms (for example one can take the Lebesgue measure Leb). Let $(B_\ell, \ell \in \mathcal{X}_d)$ be a partition of $[0, 1]$ in measurable sets such that $\mu_c(B_\ell) = \mu_d(\{\ell\})$ for all $\ell \in \mathcal{X}_d$. The map $\phi : \mathcal{X}_c \rightarrow \mathcal{X}_d$ defined by $\phi(x) = \sum \ell \mathbb{1}_{B_\ell}(x)$ clearly defines a deterministic coupling between μ_c and μ_d . If the kernels k_d on \mathcal{X}_d and k_c on \mathcal{X}_c and the functions (γ_d, \mathbf{c}_d) and (γ_c, \mathbf{c}_c) are related through the formula:

$$\gamma_c(x) = \gamma_d(\ell), \quad \mathbf{c}_c(x) = \mathbf{c}_d(\ell) \quad \text{and} \quad k_c(x, y) = k_d(\ell, j) \quad \text{for } x \in B_\ell, y \in B_j \text{ and } \ell, j \in \mathcal{X}_d,$$

then all pairs are conjugate, and all the hypotheses of Theorem 3.6 and Corollary 3.7 are satisfied.

Roughly speaking, we can blow up the atomic part of the measure μ_d into a continuous part, or, conversely, merge all points that behave similarly for k_c , γ_c and \mathbf{c}_c into an atom, without altering the Pareto frontier.

Example 4.3 (The stochastic block model). To be more concrete, we consider the so called stochastic block model, with 2 populations for simplicity 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 [14] by Lajmanovich and Yorke. For simplicity, we assume the cost is uniform (that is, $\mathfrak{c} = \mathbb{1}$), so that the conjugation condition for the costs is trivially satisfied.

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

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

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

The continuous SIS model is defined on the probability space $(\mathcal{X}_c = [0, 1], \mathcal{F}_c = \mathcal{B}([0, 1]), \mu_c = \text{Leb})$. The segment $\mathcal{X}_c = [0, 1]$ is partitioned into two intervals $B_1 = [0, p)$ and $B_2 = [p, 1]$, the transmission kernel k_c and recovery rate γ_c are given by:

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

The corresponding continuous model is $\text{Param}_c = [(\mathcal{X}_c, \mathcal{F}_c, \mu_c), (k_c, \gamma_c), \mathfrak{c}_c = \mathbb{1}]$; see Figure 1b. By the general discussion above, these two models have the same (anti-)Pareto frontiers, and their equilibria and optimal vaccinations may be transferred to one another by conjugation. Let us note that, in this example, by Lemma 4.1 a function f_d on $\mathcal{X}_d = \{1, 2\}$ and f_c on \mathcal{X}_c are :

- pre-conjugate if and only if $\frac{1}{\mu_c(B_i)} \int_{B_i} f_c d\mu_c = f_d(i)$, for $i = 1, 2$;
- conjugate if and only if $f_c(x) = f_d(i)$, a.e. for $x \in B_i$ and $i = 1, 2$.

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, using the formalism of the discrete model Param_d , the next-generation matrix K in the setting of [14], and the effective next-generation matrix $K_e(\eta)$ when the vaccination strategy η is in force (recall η_i is the proportion of population with feature i which is not vaccinated), are given by:

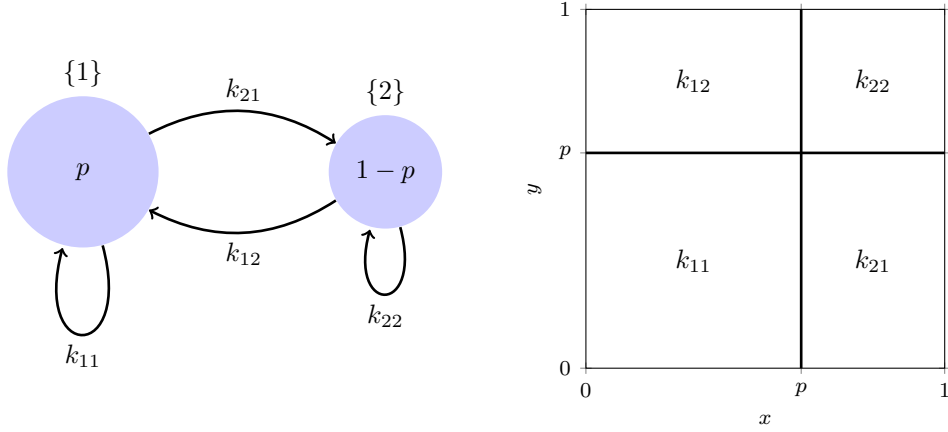
$$K = \begin{pmatrix} \mathbf{k}_{11} p & \mathbf{k}_{12} (1 - p) \\ \mathbf{k}_{21} p & \mathbf{k}_{22} (1 - p) \end{pmatrix} \quad \text{and} \quad K_e(\eta) = \begin{pmatrix} \mathbf{k}_{11} p \eta_1 & \mathbf{k}_{12} (1 - p) \eta_2 \\ \mathbf{k}_{21} p \eta_1 & \mathbf{k}_{22} (1 - p) \eta_2 \end{pmatrix} \quad \text{with} \quad \mathbf{k}_{ij} = k_{ij} / \gamma_j.$$

4.3 Starting from E_2 : measure preserving function

Finally, let us briefly discuss an example motivated by the theory of graphons, which are indistinguishable by measure preserving transformation, see [15, Sections 7.3 and 10.7].

Let $(\mathcal{X}, \mathcal{F}, \mu)$ be a probability space. We say a measurable function $\varphi : \mathcal{X} \rightarrow \mathcal{X}$ is *measure preserving* if $\mu = \varphi_{\#} \mu$. For example the function $\varphi : x \mapsto 2x \mod (1)$ defined on the probability space $([0, 1], \mathcal{B}([0, 1]), \text{Leb})$ is measure preserving. Note it is not one-to-one in general.

Now consider a SIS model with parameters $\text{Param}_2 = [(\mathcal{X}, \mathcal{F}, \mu), (k, \gamma), \mathfrak{c}]$ and a measure preserving function ϕ . Define $\gamma_1 = \gamma \circ \phi$, $\mathfrak{c}_1 = \mathfrak{c} \circ \phi$ and $k_1(\cdot, \cdot) = k_2(\phi(\cdot), \phi(\cdot))$. Then the models $\text{Param}_1 = [(\mathcal{X}, \mathcal{F}, \mu), (k_1, \gamma_1), \mathfrak{c}_1]$ and Param_2 are coupled and all consequences of Theorem 3.6 and Corollary 3.7 hold. Roughly speaking, we can give different labels to the features of the population without altering the (anti-)Pareto frontiers.



(a) Discrete model: kernel k_d on $\mathcal{X}_d = \{1, 2\}$ with the measure $\mu_d = p\delta_1 + (1-p)\delta_2$. (b) Continuous model: kernel k_c on $\mathcal{X}_c = [0, 1]$ with the Lebesgue measure μ_c .

Figure 1: Coupled discrete model (left) and continuous model (right).

5 Proofs

5.1 Elementary properties of conjugation

We give further technical properties of the conjugation.

Lemma 5.1 (Other properties of conjugation). *Let $(E_1, \mathcal{E}_1, \mu_1)$ and $(E_2, \mathcal{E}_2, \mu_2)$ be coupled through $Z = (Z_1, Z_2)$. Let $f \in L^1(E_1)$ and $f_i \in L^1(E_i)$, $i = 1, 2$.*

- (i) *We have $f^{***} = f^*$.*
- (ii) *Let $g \in L^\infty(E_1)$. If $f^{**} = f$, then we have $(fg)^* = f^*g^*$.*
- (iii) *If the pair (f_1, f_2) is pre-conjugate, then $\int_{E_1} f_1 d\mu_1 = \int_{E_2} f_2 d\mu_2$.*
- (iv) *Let $g_i \in L^\infty(E_i)$, $i = 1, 2$. If the pair (f_1, f_2) is conjugate and the pair (g_1, g_2) is pre-conjugate, then the pair (f_1g_1, f_2g_2) is pre-conjugate.*

Proof. Since (f^{**}, f^*) is conjugate by Lemma 3.4, we deduce that $f^{***} = f^*$ by definition of conjugation. To prove (ii) note that $(fg)^*(Z_2) = \mathbb{E}[f(Z_1)g(Z_1)|\mathcal{C}]$, but $f(Z_1)$ is \mathcal{C} -measurable as $f^{**} = f$, so we may pull it out. Since $f(Z_1) = f^*(Z_2)$ and $\mathbb{E}[g(Z_1)|\mathcal{C}] = g^*(Z_2)$, the result follows.

If (f_1, f_2) is pre-conjugate, we have $\mathbb{E}[f_1(Z_1)|\mathcal{C}] = f_1^*(Z_2) = f_2^*(Z_1) = \mathbb{E}[f_2(Z_2)|\mathcal{C}]$; then take the expectation to get (iii). Point (iv) is a direct consequence of Point (ii) and Lemma 3.3. \square

5.2 Proof of Lemma 3.5 and a key lemma

Let us first recall an elementary result on conditional independence. The random variables we consider are defined on a probability space, say $(\Omega_0, \mathcal{F}_0, \mathbb{P})$. Let \mathcal{A} , \mathcal{B} and \mathcal{I} be sub- σ -fields of \mathcal{F}_0 . We recall that \mathcal{A} and \mathcal{B} are conditionally independent given \mathcal{I} , denoted by $\mathcal{A} \perp_{\mathcal{I}} \mathcal{B}$, if $\mathbb{P}[A \cap B|\mathcal{I}] = \mathbb{P}[A|\mathcal{I}]\mathbb{P}[B|\mathcal{I}]$ for all $A \in \mathcal{A}$ and $B \in \mathcal{B}$. According to [13, Theorem 8.9], if $\mathcal{I} \subset \mathcal{A} \cap \mathcal{B}$, the conditional independence $\mathcal{A} \perp_{\mathcal{I}} \mathcal{B}$ holds if and only if

$$\mathbb{E}[W|\mathcal{B}] = \mathbb{E}[W|\mathcal{I}] \text{ for any nonnegative } \mathcal{A}\text{-measurable variable } W. \quad (5.1)$$

We start by a probabilistic result.

Lemma 5.2. *Let E_1 and E_2 be coupled, and \mathbf{E}_1 and \mathbf{E}_2 coupled through the extended coupling $(\mathbf{Z}_1, \mathbf{Z}_2) = ((X_1, Y_1), (X_2, Y_2))$. Let $\mathcal{C} = \sigma(\mathbf{Z}_1) \cap \sigma(\mathbf{Z}_2)$, $\mathcal{C}_X = \sigma(X_1) \cap \sigma(X_2)$ and $\mathcal{C}_Y = \sigma(Y_1) \cap \sigma(Y_2)$.*

- (i) *The following conditional independence holds: $\sigma(X_1, X_2) \perp_{\mathcal{C}_X} \mathcal{C}$ and $\sigma(Y_1, Y_2) \perp_{\mathcal{C}_Y} \mathcal{C}$.*
- (ii) *Let K, V be nonnegative random variables. If K is \mathcal{C} -measurable and V is $\sigma(Y_1, Y_2)$ -measurable, then we have:*

$$\mathbb{E}[KV|X_i] = \mathbb{E}[K \mathbb{E}[V|\mathcal{C}]|\mathcal{C}_X] = \mathbb{E}[K \mathbb{E}[V|\mathcal{C}]|X_i], \quad i = 1, 2. \quad (5.2)$$

Proof. By (5.1) the first independence in Point (i) holds if, for any \mathcal{C} -measurable nonnegative random variable W , $\mathbb{E}[W|\mathcal{C}_X] = \mathbb{E}[W|X_1, X_2]$. Let W be \mathcal{C} -measurable and nonnegative; so $W = \phi(X_1, Y_1)$ for some function ϕ . Let $W' = \mathbb{E}[W|X_1, X_2]$. Since $\sigma(X_1, X_2) \perp_{\sigma(X_1)} \sigma(X_1, Y_1)$,

$$W' = \mathbb{E}[\phi(X_1, Y_1)|X_1, X_2] = \mathbb{E}[\phi(X_1, Y_1)|X_1] = \mathbb{E}[W|X_1].$$

Therefore the random variable W' is $\sigma(X_1)$ -measurable. By symmetry, it is also $\sigma(X_2)$ -measurable, so it is in fact \mathcal{C}_X -measurable. Therefore, by the tower property, we get:

$$\mathbb{E}[W|X_1, X_2] = W' = \mathbb{E}[W'|\mathcal{C}_X] = \mathbb{E}[W|\mathcal{C}_X].$$

This proves the first point.

Since K is \mathcal{C} -measurable we may write it as $K = k(X_1, Y_1)$; similarly $V = v(Y_1, Y_2)$. Since $\sigma(X_1, Y_1, Y_2) \perp_{\sigma(X_1)} \sigma(X_1, X_2)$, and since KV is $\sigma(X_1, Y_1, Y_2)$ -measurable, we get:

$$\mathbb{E}[KV|X_1] = \mathbb{E}[KV|X_1, X_2].$$

Let W denote this random variable. The same argument applied with the conditional independence $\sigma(X_2, Y_1, Y_2) \perp_{\sigma(X_2)} \sigma(X_1, X_2)$ yields symmetrically $\mathbb{E}[KV|X_2] = \mathbb{E}[KV|X_1, X_2] = W$. In particular W is measurable with respect to both X_1 and X_2 , so W is \mathcal{C}_X -measurable. Using the tower property of conditional expectations with $\mathcal{C}_X \subset \sigma(X_1, X_2)$ and $\mathcal{C}_X \subset \mathcal{C}$, and the fact that K is \mathcal{C} -measurable, we get:

$$W = \mathbb{E}[W|\mathcal{C}_X] = \mathbb{E}[KV|\mathcal{C}_X] = \mathbb{E}[\mathbb{E}[KV|\mathcal{C}]|\mathcal{C}_X] = \mathbb{E}[K \mathbb{E}[V|\mathcal{C}]|\mathcal{C}_X].$$

This proves the first equality of (5.2) for $i = 1$ and then for $i = 2$ by symmetry. Set $V' = \mathbb{E}[V|\mathcal{C}]$ which is \mathcal{C}_Y -measurable and then $\sigma(Y_1, Y_2)$ -measurable. Then apply the first equality of (5.2) with V replaced by V' to get the second equality of (5.2). The proof is then complete. \square

The fact that conjugacy behaves well on extended spaces is now easy to establish.

Proof of Lemma 3.5. Let $\phi(X_1, Y_1) = f_1(X_1)$. Since $\sigma(X_1) \perp_{\mathcal{C}_X} \mathcal{C}$ by the first point of Lemma 5.2, we get by (5.1) that $\mathbb{E}[\phi(X_1, Y_1)|\mathcal{C}] = \mathbb{E}[f_1(X_1)|\mathcal{C}] = \mathbb{E}[f_1(X_1)|\mathcal{C}_X]$. \square

The next lemma is the key to all our main results. For a probability space (E, \mathcal{E}, μ) , say that a kernel k on E is *nice* if $k \in L^1(E^2)$ and satisfies $\int_E k(\cdot, y) \mu(dy) \in L^\infty(E)$. For a nice kernel k we define the bounded operator T_k on $L^\infty(E)$ by $T_k(g) = \int_E k(\cdot, y) g(y) \mu(dy)$.

Lemma 5.3 (Operator defined by conjugated kernels). *Let two spaces E_1 and E_2 be coupled through π . If the nice kernel k on E_1 satisfies $k = k^{**}$ (for the extended coupling) and if $v \in L^\infty(E_1)$, then k^* is a nice kernel on E_2 , $v^* \in L^\infty(E_2)$ and:*

$$T_k(v) = T_k(v)^{**} = T_k(v^{**}) \quad \text{and} \quad T_k(v)^* = T_{k^*}(v^*). \quad (5.3)$$

Proof. Let $(X_1, Y_1, X_2, Y_2) \sim \pi$ denote the extended coupling. Let k be a nice kernel on E_1 such that $k = k^{**}$. As $k = k^{**}$, we deduce from Lemma 3.4 (ii) that (k, k^*) is conjugate and by Lemma 3.3 that a.s. $k(X_1, Y_1) = k^*(X_2, Y_2)$ and that this random variable is \mathcal{C} -measurable.

Let $v \in L^\infty(E_1)$. The function $T_k(v)$ admits the probabilistic representation:

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

We apply Lemma 5.2 (ii) with $K = k(X_1, Y_1)$ and $V = v(Y_1)$ to get that $T_k(v)(X_1)$ is \mathcal{C}_X -measurable, and by Lemma 3.4 (ii) that $T_k(v) = T_k(v)^{**}$. This gives the first equality of (5.3).

It is obvious that if $v^* \in L^\infty(E_2)$. Using the definition of the conjugate, and then $\mathbb{E}[V|\mathcal{C}] = v^*(Y_2)$ from Lemma 3.5 and Equation (5.2), we obtain:

$$T_k(v)^*(X_2) = \mathbb{E}[k(X_1, Y_1)v(Y_1)|\mathcal{C}_X] = \mathbb{E}[k^*(X_2, Y_2)v^*(Y_2)|X_2] = T_{k^*}(v^*)(X_2).$$

Taking $v = \mathbb{1}$, we deduce that $T_{k^*}(\mathbb{1}) = T_k(\mathbb{1})^*$ belongs to $L^\infty(E_2)$, thus k^* is a nice kernel on E_2 and T_{k^*} is a bounded operator on $L^\infty(E_2)$. We have also proven that $T_k(v)^* = T_{k^*}(v^*)$ which is the last equality of (5.3). Using this equality again with k and v replaced by k^* and v^* , we obtain that $T_k(v)^{**} = T_{k^*}(v^*)^* = T_{k^{**}}(v^{**}) = T_k(v^{**})$, which is the second equality of (5.3). \square

5.3 Proof of the main result, Theorem 3.6

The spectrum and effective reproduction number. We prove the first item of Theorem 3.6. Recall the spectral radius of a bounded operator is the maximal modulus of its complex eigenvalues. Set $k_i = k_i/\gamma_i$ for $i = 1, 2$. Notice the bounded operators T_{k_i} on $L^\infty(E_i)$ and \mathcal{T}_{k_i} on $\mathcal{L}^\infty(E_i)$ have the same spectrum and thus the same spectral radius and more generally $R_e(\eta_i) = \rho(T_{k_i\eta_i})$ for $\eta_i \in \Delta_i$. For simplicity, write $k = k_1$ and thus, as (k_1, k_2) is a conjugate pair, $k^* = k_2$ and $k^{**} = k$.

Let $\eta \in \Delta_1$ and λ be a non-zero eigenvalue of $T_{k\eta}$ associated with an eigenvector $v \in L^1(E_1)$. By definition, we have:

$$\lambda v = T_{k\eta}(v) = T_k(\eta v).$$

Thanks to the first two equalities in (5.3) of Lemma 5.3, the function λv is equal to its biconjugate (that is, the pair (v, v^*) is conjugate) and $\lambda v = T_k((\eta v)^{**})$.

Assume the pair (η, η_2) is pre-conjugate. By Lemma 5.1 (iv), the pair $(\eta v, \eta_2 v^*)$ is pre-conjugate, and thus $(\eta_2 v^*)^* = (\eta v)^{**}$. Then, using Lemmas 5.3 and 5.1 (ii), we get:

$$T_{k^*}(\eta_2 v^*) = T_{k^*}((\eta_2 v^*)^{**}) = T_k((\eta_2 v^*)^*)^* = T_k((\eta v)^{**})^* = \lambda v^*.$$

Since $v^{**} = v \neq 0$, the function v^* is non-zero and it is therefore an eigenvector of $T_{k^*\eta_2}$, associated to the eigenvalue λ . By symmetry we deduce that the spectrum up to $\{0\}$ of $T_{k\eta}$ and $T_{k^*\eta_2}$ coincide, and thus their spectral radius are equal. This proves Point (i).

The equilibria. Let us now prove the first part of Point (ii) on the equilibria are conjugate. Let $g \in \mathcal{L}^\infty(E_1)$ be an equilibrium of the model Param_1 . Since $F_\eta(g) = 0$, we have:

$$g = \frac{\mathcal{T}_{k_1}(\eta g)}{\gamma_1 + \mathcal{T}_{k_1}(\eta g)}.$$

By Lemma 5.3, seeing g as an element of $L^\infty(E_1)$, we get that $T_{k_1}(\eta g)$ is equal to its biconjugate. Since μ_1 -a.e. $\gamma_1^{**} = \gamma_1$, we easily deduce using Lemma 5.1 (ii) that:

$$g^* = \frac{T_{k_1}(\eta g)^*}{\gamma_1^* + T_{k_1}(\eta g)^*} \quad \text{and then} \quad g^{**} = \frac{T_{k_1}(\eta g)}{\gamma_1 + T_{k_1}(\eta g)},$$

that is, μ_1 -a.e. $g^{**} = g$. So (g, g^*) is conjugate. By Lemma 5.1 (iv), the pair $(\eta g, \eta_2 g^*)$ is pre-conjugate, and thus $(\eta g)^* = (\eta_2 g^*)^{**}$. We get, using Lemma 5.3 for the first and last equalities:

$$T_{k_1}(\eta g)^* = T_{k_2}((\eta g)^*) = T_{k_2}((\eta_2 g^*)^{**}) = T_{k_2}(\eta_2 g^*).$$

Notice that if μ_2 -a.s. $f = h$ then $\mathcal{T}_{k_2}(f) = \mathcal{T}_{k_2}(g)$, so that $\mathcal{T}_{k_2}(\eta_2 g^*)$ is a well defined element of $\mathcal{L}^\infty(E_2)$. Thus defining $g_2 \in \mathcal{L}^\infty(E_2)$ by:

$$g_2 = \frac{\mathcal{T}_{k_2}(\eta_2 g^*)}{\gamma_2 + \mathcal{T}_{k_2}(\eta_2 g^*)},$$

we get that μ_2 -a.e. $g_2 = g^*$ and that $F_{\eta_2}(g_2) = 0$. In other words, g_2 is an equilibrium for the model given by Param_2 when using the vaccination strategy η_2 , and, seeing g_i as an element of $L^1(E_i)$, the pair (g_1, g_2) is conjugate. This proves the first part of Point (ii).

The fraction of infected individuals \mathfrak{F} . We now prove that $\mathfrak{F}[\text{Param}_1](\eta_1) = \mathfrak{F}[\text{Param}_2](\eta_2)$ whenever the pair (η_1, η_2) is preconjugate. We assume without loss of generality that $R_0[\text{Param}_1] = R_e[\text{Param}_1](1) > 1$ which is equivalent to $R_0[\text{Param}_2] = R_e[\text{Param}_2](1) > 1$, thanks to Theorem 3.6 (i) as the pair $(1, 1)$ is conjugate and thus pre-conjugate. Let $g_1 = g_{\eta_1}$ be the maximal equilibrium for the model Param_1 when using the vaccination strategy η_1 . By the previous result there exists an equilibrium g_2 for SIS model Param_2 such that μ_2 -a.s. $g_2 = g_1^*$. Let us now prove that it is the maximal one. Since $(1 - g_2) = (1 - g_1)^*$ in $L^1(E_2)$, we get $R_e[\text{Param}_1](1 - g_1) = R_e[\text{Param}_2](1 - g_2)$, again by Theorem 3.6 (i). Since $R_0[\text{Param}_1] > 1$ and g_1 is the maximal equilibrium for Param_1 , we deduce from [7, Proposition 5.5] that the vaccination strategy associated to g_1 is critical, that is, $R_e[\text{Param}_1](1 - g_1) = 1$. Since g_2 is an equilibrium for Param_2 satisfying $R_e[\text{Param}_2](1 - g_2) = 1$, we deduce using again [7, Proposition 5.5] that g_2 is also the maximal equilibrium for Param_2 . Using Point (iv) of Lemma 5.1, we deduce that the pair $(g_1 \eta_1, g_2 \eta_2)$ is pre-conjugate and then from Point (iii) therein that $\mathfrak{F}_1(\eta_1) = \int_{E_1} \eta_1 g_1 d\mu_1 = \int_{E_2} \eta_2 g_2 d\mu_2 = \mathfrak{F}_2(\eta_2)$. This ends the proof of Point (ii).

Proof of Point (iii). Thanks to Points (i) and (ii), it is enough to check that $C[\text{Param}_1](\eta_1) = C[\text{Param}_2](\eta_2)$ whenever the pair (η_1, η_2) is pre-conjugate. Since the pair (c_1, c_2) is conjugate, this is a direct consequence of Points (iii) and (iv) from Lemma 5.1.

References

- [1] J.-F. Delmas, D. Dronnier, and P.-A. Zitt. “An infinite-dimensional metapopulation SIS model”. *Journal of Differential Equations* 313 (2022), pp. 1–53.
- [2] J.-F. Delmas, D. Dronnier, and P.-A. Zitt. “Optimal vaccination strategies for an heterogeneous SIS model”. 2025. arXiv: 2503.15960 [math.OC].
- [3] J.-F. Delmas, D. Dronnier, and P.-A. Zitt. “Optimal vaccination: various (counter) intuitive examples”. *Journal of Mathematical Biology* 86.26 (2023), pp. 1–57.
- [4] J.-F. Delmas, D. Dronnier, and P.-A. Zitt. “Optimal vaccinations: cordons sanitaires, reducible population and optimal rays”. 2022. arXiv: 2209.07381.
- [5] J.-F. Delmas, D. Dronnier, and P.-A. Zitt. “The effective reproduction number: convexity, concavity and invariance”. *Journal of the European Mathematical Society* (2024).
- [6] J.-F. Delmas, D. Dronnier, and P.-A. Zitt. “Topological Properties of the Effective Reproduction Number in an Heterogeneous SIS Model”. 2025. arXiv: 2504.00055 [math.FA].
- [7] J.-F. Delmas, D. Dronnier, and P.-A. Zitt. “Vaccinating according to the maximal endemic equilibrium achieves herd immunity”. *SIAM Journal on Applied Mathematics* 85.2 (2025), pp. 806–820.
- [8] J.-F. Delmas, P. Frasca, F. Garin, V. C. Tran, A. Velleret, and P.-A. Zitt. “Individual-based SIS models on (not so) dense large random networks”. *ALEA Lat. Am. J. Probab. Math. Stat.* 21.2 (2024), pp. 1375–1405.

- [9] 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.
- [10] 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.
- [11] H. Heesterbeek. “The law of mass-action in epidemiology: a historical perspective”. *Ecological paradigms lost: routes of theory change* (2005), pp. 81–104.
- [12] S. Janson. *Graphons, cut norm and distance, couplings and rearrangements*. NYJM Monographs. New York Journal of Mathematics, 2013.
- [13] O. Kallenberg. *Foundations of modern probability*. 3rd ed. Vol. 99. Probability Theory and Stochastic Modelling. Springer-Verlag, 2021.
- [14] A. Lajmanovich and J. A. Yorke. “A deterministic model for gonorrhea in a nonhomogeneous population”. *Mathematical Biosciences* 28.3 (1976), pp. 221–236.
- [15] L. Lovász. *Large networks and graph limits*. American Mathematical Society colloquium publications 60. Providence, Rhode Island: American Mathematical Society, 2012.
- [16] H. McCallum, N. Barlow, and J. Hone. “How should pathogen transmission be modelled?” *Trends in ecology & evolution* 16.6 (2001), pp. 295–300.
- [17] K. Miettinen. *Nonlinear multiobjective optimization*. Springer-Verlag, 1998.