Flow Redirection for Epidemic Reaction-DIffusion Control

Pierre-Yves Massé, Quentin Laborde, Maria Cherifa, Jules Olayé, Laurent Oudre

Flow Redirection for Epidemic Reaction-Diffusion Control

Pierre-Yves Massé[†‡*], Quentin Laborde[†*], Maria Cherifa[†], Jules Olayé[†] and Laurent Oudre[†]
[†] Centre Borelli [‡] Corresponding author [*] Equal contribution

***Abstract*—We show we can control an epidemic SEIR reaction-diffusion process on a directed, and heterogeneous, network by redirecting the flows, thanks to the optimisation of well-designed loss functions, in particular the basic reproduction number of the model. We provide a final size relation linking the basic reproduction number to the epidemic final sizes, for diffusions around a reference diffusion with basic reproduction number less than 1. Experimentally, we show control is possible for different topologies, network heterogeneity levels, and speeds of diffusion. Our numerical simulations results highlight the relevance of the basic reproduction number loss, compared to more common losses.**

***Keywords*—Networked Control Systems, Optimization, Biological Networks, Reaction-Diffusion**

I. Introduction

NETWORKS are important infrastructures, whether they are transportation [1], telecommunication [2] [3], or supply [4] networks. However, they may be invaded by undesirable process, such as diseases [5] or virus malwares [6]. A typical modeling of these processes is based on systems of coupled ordinary differential equations (ODE), where the equations associated with each node (representing populations such as cities or countries) are coupled in some way by the network [7]. The ODEs used are deterministic and were introduced at the beginning of the 20$^{\text{th}}$ century, following notably Kermack and McKendrick [8]. This formalization described originally how individuals transitioned from state to state — healthy, infected, recovered, for instance — when confronted with a disease [9].

Two major classes of coupling have emerged. On the one hand, the interactions between populations may be described by a static contact structure [7], [10], [11], where individuals in a node remain in the same one, and may be infected by others in the neighboring nodes. On the other hand, in the epidemic reaction-diffusion models, also known as metapopulation models with explicit movement [12], individuals can only be infected by others in the same node, but are allowed to move to neighboring nodes. Following their appearance in ecology [13], these models have sparked considerable interest in mathematical epidemiology [12], [14]–[20].

To contain the spread of the disease, a first option consists in acting on the disease parameters, such as the infection rate and the curing rate [21]–[24]. For instance, treatments may reduce the likelihood for an individual to get infected, or speed up recovery. Another option is to modify the network structure. Notably, Preciado and Zargham [25] reduce the flows between cities, in a static contact model, in order to limit the spread of an epidemic. Recently, Umar B. Niazi *et al.* [26] have shown how to lower the flows of a model with diffusive dynamics in order to optimally control the epidemic. In our work, we propose to merely redirect the flows of the diffusion dynamics: indeed, redirection involves a smaller alteration of the network structure, and can help it keep functioning as normally as possible, which is desirable. Our goal is to reduce the final size of the epidemic, that is the final number of individuals who have been infected. We use the basic reproduction number of the system [27]–[29] as a criterion to redirect the flows. We aim at minimizing it by gradient descent, with respect to some relevant parameterization of the diffusion dynamics.

We start by discussing related works and presenting our contributions, in Section II. We then define our model, state the problem we study, and describe our methodology, in Section III. Section IV gathers our theoretical results, the main one being a final size relation linking the basic reproduction number, and the final size of the epidemic. We present the results of our numerical simulations in Section V[1]. The bulk of the proofs is deferred to Section VII for ease of reading.

II. Related Works and Contributions

We first present related works on epidemic control (Section II-A) and final size relations (Section II-B). Then, we state our contributions (Section II-C), before discussing our work in light of some closely related references (Section II-D).

A. Network Deterministic Epidemic Control

Control of deterministic epidemic processes on networks is an alive direction of research: see for instance Nowzari *et al.* [7] and Zino and Cao [30]. A usual aim is the stabilization of the disease-free-equilibrium (DFE) of the system (the point at which no disease is present in the population). The control represents a preventive intervention undertaken by some regulator (e.g. the distribution of prophylactic treatments). To assess the amount of intervention needed, one common strategy consists in tracking its effect on the reduction of the maximum real part of the eigenvalues of the Jacobian of the system at the DFE, known as the spectral abscissa. Indeed, once the abscissa is

This work was partially supported by the French Agence de l'innovation de défense as part of the project Onadap.

[1]The code for the simulations, written in Python, is available on the git repository https://reine.cmla.ens-cachan.fr/masse/flow_redirection.

negative, the DFE becomes exponentially stable. The Jacobian typically writes $J = \boldsymbol{\beta} A - \boldsymbol{\delta}$, where A is the adjacency matrix of the network, $\boldsymbol{\beta}$ contains the infectivity parameters and $\boldsymbol{\delta}$ contains the curing parameters. Two types of interventions are possible. First, one may act on the network structure, that is on A. For instance, Preciado and Jadbabaie [31] design well-suited values for the connectivity radius of a random geometric graph, ensuring the stability of the DFE, while Preciado and Zargham [25] reduce the flows between cities. Second, the control may be exercised on the epidemic parameters. In Preciado *et al.* [22] and Preciado *et al.* [23], the infectivity and curing parameters are optimised under two objectives. First, the authors seek to minimize as much as possible the spectral abscissa, under a given budget. Conversely, they look for the minimal budget ensuring the abscissa is taken below some threshold. In particular, Nowzari *et al.* [24] show that these optimization problems may be cast as geometric programs, and can be solved with standard solvers. Rather than the Jacobian, Knipl [32] uses as a criterion the spectral radius of some next-generation like matrices [29]. She leverages the fact these spectral radius act as thresholds for the spectral abscissa of the Jacobian at the DFE. She identifies a subset of entries of the next-generation matrices, and the target values they must have, in order to stabilize the DFE. Then, she reverse-engineers the alteration of the parameters of the model needed to reach these target values.

Let us discuss works focusing on the network as a control parameter. Umar B. Niazi *et al.* [26] propose an optimal control of solutions to mitigate an outbreak while maintaining as much as possible economic activity, in a metapopulation model with diffusive coupling. Ottaviano *et al.* [33] leverage the structure of the network to optimally distribute curing treatments on the network. Notably, their use of equitable partitions means the optimization problem they need to solve is lower-dimensional than the system they aim at controlling. Paré *et al.* [34] act on the network structure of a time-varying system to reduce the spread of the disease. Meanwhile, some works act on both network and disease-related parameters, like Knipl [32] mentioned earlier: for instance, Cenedese *et al.* [35] solve a Nonlinear Model Predictive Control problem to design the most cost-efficient way to both reduce infection and contact rates specified by the network adjacency matrix. Finally, time-varying networks exhibit specific epidemic behaviors compared to static [36], therefore require tailor-made procedures in order to control epidemics [37].

B. Final Size Relations

The final size of an epidemic is the asymptotic number of individuals confronted to the disease. It is an important outcome of the epidemic: estimating it and studying its dependency on the model's parameters has been given plenty of attention: see notably Junling and David [38]. Kermack and McKendrick [8] established the following well-known equation for a scalar, deterministic model: the final size $r(\infty)$ (the lowercase emphasizes it is a real number) and the basic reproduction number $\mathcal{R}_0$ are linked by

$$\mathcal{R}_0 r(\infty) + \log\left(1 - r(\infty)\right) = 0. \tag{1}$$

The basic reproduction number — see for instance Macdonald [27] — is a key statistic governing the behavior of epidemic models. It is often interpreted as the number of secondary infections caused by an initial infected individual in a fully susceptible population, though this needs qualifying (see discussion in Section II-D). One major property of Equation (1) is the monotonous link it shows between $\mathcal{R}_0$, which concerns the onset of the epidemic, and its final size, which concerns its outcome. Several studies have since been devoted to extending this relation to more evolved models [39]. Magal *et al.* [40] and Magal *et al.* [41] study the final size of a multi-group SIR epidemic model. Andreasen [42] proves the unicity of the solution of the final size equation in the open unit cube, when $\mathcal{R}_0 > 1$, in a network model with static contacts. See also Brauer [43] about the final size for models on networks with static contacts, also known as mixing models. Final size relations have also been derivated in continuous-space settings [44]. Finally, Gao [45] study the influence of the speed of diffusion on the total infection size of the endemic state of an SIS-metapopulation model. They show in particular that the basic reproduction number diminishes when the speed of diffusion increases, but the infection size may not be monotonic. However, up to our best knowledge, there is no general result expressing the final size of an epidemic reaction-diffusion process as a function of its basic reproduction number.

C. Contributions

We show we can control an epidemic running on a directed, heterogeneous network, by redirecting flows of individuals. Theoretically, we provide a final size relation linking the basic reproduction number and the final size of the epidemic (Theorem 3). It applies to reaction-diffusion processes with diffusion matrices close to a reference diffusion matrix, whose reproduction number is strictly less than 1. To obtain this result, we prove a uniform stability of the Disease Free Equilibrium (Corollary 6) which extends uniformly, in some neighborhood of the reference diffusion matrix, the standard stability criterion is given by the next-generation matrix.

Then, we present our methodology based on the control of the basic reproduction number to control the epidemic spread. We design a parameterization of the diffusion allowing us to redirect the flows, and define several loss functions that we compare to the base loss in reproduction number. The losses are differentiable so that even though the optimization problem is nonlinear and nonsymmetric, it can be solved by gradient descent.

Finally, we validate our approach with numerical simulations, using synthetic data presenting different topologies, different levels of network heterogeneity, and a range of diffusion speeds. The procedure works for general reaction models although, for the sake of clarity, we only work with a SEIR model. We use SEIR because it allows us to show the procedure works in a slightly involved setting, while not overburdening the analysis with technical details.

D. Comparison with some Related Works

Let us highlight our differences with some closely related works, first as far as control is concerned, and second as far as the final size relation is concerned.

Contrary to Knipl [32], we do not identify a target next-generation matrix, but directly optimize the basic reproduction number. In contrast to Umar B. Niazi *et al.* [26], we do not aim at lowering the flows, but redirect them. Moreover, we do not cast the problem as an optimal control one, but optimize the basic reproduction number. Extending our methodology to an optimal control framework is beyond the scope of our work, but would represent an interesting direction of future research.

Unlike Andreasen [42], we study a model with a diffusive coupling, and do not analyze the structure of a final size equation, but bound the final size. Contrary to Gao [45], we do not consider changes in the speed of the diffusion dynamics, but redirection of it. Moreover, we show numerically, and argue theoretically, that even though $\mathcal{R}_0$ and the final size may not be monotonically related, lowering $\mathcal{R}_0$ enough ends up diminishing the epidemic final size. Note that we consider a system with no possibility of reinfection, and therefore no endemic state, so that the final size we study is different from the infection size at the endemic state in Gao [45].

Our approach uses the basic reproduction number $\mathcal{R}_0$ as a stability criterion. Note that it is not uniquely defined [46], and its common interpretation is often incorrect. We retain its definition as the spectral radius of the next-generation matrix and use it as a threshold for the eigenvalues of the Jacobian. We thus are not liable to the issues identified in Li *et al.* [46], and can benefit from the following advantages. First, as it acts as a threshold, taking it below 1 is equivalent to ensuring the eigenvalues of the Jacobian of the system of Equation (2) at the DFE have negative real parts. Second, as the spectral radius of the next-generation matrix, it is obtained through the study of a matrix of order $|\mathcal{N}|$, whereas the Jacobian is of order $4|\mathcal{N}|$. Finally, the next-generation matrix is positive, so that its spectral radius is differentiable, and an analytical formula exists for its derivative (see Section IV-B).

III. Model, Problem Statement and Methodology

We first define our model, together with our notations, in Section III-A. The main notations are gathered in Table I. Second, we give our problem statement, and introduce our methodology, in Section III-B.

A. Background Material, Definitions and Notations on Metapopulation Models with Diffusion

Let $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ be a strongly connected, directed graph, with node set $\mathcal{N}$, and edge set $\mathcal{E}$. For each node $n \in \mathcal{N}$, we write β_n, δ_n and γ_n the positive infection, incubation and curing rates respectively, of a scalar SEIR model [9]. We write $\boldsymbol{\beta} = \operatorname{diag}\left(\beta_1, \dots, \beta_{|\mathcal{N}|}\right)$ the corresponding diagonal matrix, and likewise for the other coefficients. Capital letters like S, E, I or R are vectors of size $|\mathcal{N}|$, such that for instance, S_n is the numbers of individuals in compartment "S" of node $n \in \mathcal{N}$. Coupling between nodes is realised by a diffusion matrix, which definition we now recall.

Definition 1 (Diffusion Matrix)**.** *A diffusion matrix $\mathbf{M}$ on $\mathcal{G}$ has nonzero off-diagonal entries only for coordinates (i, j) such that the edge $i \rightsquigarrow j$ belongs to $\mathcal{E}$. $\mathbf{M}$ is Metzler, that is for $i, j \in \mathcal{N}$, $i \neq j$, we have $\mathbf{M}_{ij} \geq 0$. It is irreducible. Finally, the coordinates of each column sum to zero.*

Since $\mathcal{G}$ is strongly connected, such matrices do exist. Standard Perron-Frobenius theory guarantees that a diffusion matrix $\mathbf{M}$ admits a stationary distribution $\tilde{\mu}_{\mathbf{M}}$, that is a positive right eigenvector such that $\mathbf{M}\tilde{\mu}_{\mathbf{M}} = 0$, and which coordinates sum to 1. The reaction-diffusion extension of the standard SEIR system to a network evolves according to[2], for all $t \geq 0$,

$$\begin{cases} \dfrac{dS}{dt} = -\boldsymbol{\beta} S \odot I + \mathbf{M}S \\ \dfrac{dE}{dt} = \boldsymbol{\beta} S \odot I - \boldsymbol{\gamma} E + \mathbf{M}E \\ \dfrac{dI}{dt} = \boldsymbol{\gamma} E - \boldsymbol{\delta} I + \mathbf{M}I \\ \dfrac{dR}{dt} = \boldsymbol{\delta} I + \mathbf{M}R. \end{cases} \tag{2}$$

Standard results — see Arino [12] and references therein — guarantee that, for all nonnegative initial condition $\big(S(0), E(0), I(0), R(0)\big)$, the solution to Equation (2) is global, remains nonnegative, and converges to a fixed point of the form $\big(S(\infty), 0, 0, R(\infty)\big)$. Moreover, the total population is preserved: $\sum_{n\in\mathcal{N}} S_n(t)+E_n(t)+I_n(t)+R_n(t)$ is constant. In what follows, we assume it equals 1[3].

The "Disease Free Equilibrium" (DFE) $(\tilde{\mu}_{\mathbf{M}}, 0, 0, 0)$ is a fixed point of Equation (2), where there is no disease: all individuals are in the compartment S. We want it to be stable, and therefore we recall here a well-studied stability criterion, which we make extensive use of. The stability of the DFE is governed by the spectral radius of the next-generation matrix [28], [29], which is called the basic reproduction number, and written $\mathcal{R}_0$. The DFE is stable if, and only if, we have $\mathcal{R}_0 < 1$ [47]. The next-generation matrix in the sense of Diekmann and Heesterbeek [29] associated with the system of Equation (2) is $\mathbf{G}_{\mathbf{M}} := \boldsymbol{\beta} \operatorname{diag}(\tilde{\mu}_{\mathbf{M}}) (\mathbf{M} - \boldsymbol{\delta})^{-1} \boldsymbol{\gamma} (\mathbf{M} - \boldsymbol{\gamma})^{-1}$, where $\operatorname{diag}(\tilde{\mu}_{\mathbf{M}})$ is the diagonal matrix which diagonal coefficients are those of $\tilde{\mu}_{\mathbf{M}}$. The basic reproduction number depends on the diffusion matrix $\mathbf{M}$, and we write it $\mathcal{R}_0 = \mathcal{R}_0(\mathbf{M}) = \rho(\mathbf{G}_{\mathbf{M}})$, where ρ indicates the spectral radius.

Finally, we use the following notion of policy over the network.

Definition 2 (Policy over a Network)**.** *We call policy a stochastic matrix $\boldsymbol{\pi}$ of order $|\mathcal{N}|$ such that, for every node $n \in \mathcal{N}$, the row $(\boldsymbol{\pi}_{n,i})_{i\in\mathcal{N}}$ is a probability distribution over the outgoing neighbours of n*[4].

[2] For instance, for a node $n \in \mathcal{N}$, the equation on S_n reads: $dS_n/dt = -\beta_n S_n(t) I_n(t) + \sum_{i=1}^{|\mathcal{N}|} \mathbf{M}_{n,i} S_i(t)$.

[3] This is standard: see for instance Andreasen [42] or Gao [45].

[4] Therefore, for $i \in \mathcal{N}$, $\boldsymbol{\pi}_{n,i}$ is nonzero if, and only if, $\mathcal{E}$ contains an edge $n \rightsquigarrow i$.

For every $n \rightsquigarrow i \in \mathcal{E}$, the quantity $\boldsymbol{\pi}_{n,i}$ is the proportion of individuals who leave the node n through the edge $n \rightsquigarrow i$. We consider diffusion matrices of the form

$$\mathbf{M}(\theta) = \boldsymbol{f}\left(\boldsymbol{\pi}(\theta) - \mathrm{Id}_{|\mathcal{N}|}\right)^T,$$

where the outrate diagonal matrix $\boldsymbol{f}$, and the policy $\boldsymbol{\pi}(\theta)$, both in $\mathcal{M}_{|\mathcal{N}|}(\mathbb{R})$, are defined below. For every node $n \in \mathcal{N}$, we call outrate in n a positive real number f_n, which is the rate at which individuals leave the node. We write $\boldsymbol{f} = \mathrm{diag}(f_1, \dots, f_{|\mathcal{N}|})$. Let us define the parameter space

$$\Theta = \mathbb{R}^{|\mathcal{E}|}.$$

We index $\theta \in \Theta$ by the edges in $\mathcal{E}$: for instance, if $n, i \in \mathcal{N}$ are such that $n \rightsquigarrow i$ belongs to $\mathcal{E}$, we write $\theta_{n,i}$ the corresponding entry of θ. For every $\theta \in \Theta$, we define the policy $\boldsymbol{\pi}(\theta)$ over $\mathcal{N}$ such that, for every node n, $i \mapsto \boldsymbol{\pi}_{n,i}(\theta)$ is a softmax function over the outgoing neighbours of n[5]. As a result, for every θ, $\boldsymbol{\pi}(\theta)$ is indeed a policy, and the mapping $\theta \mapsto \boldsymbol{\pi}(\theta)$ is regular.

The matrices $\mathbf{M}(\theta)$ are diffusion matrices in the sense of Definition 1. Indeed, let us fix θ. For every node n, for every edge $n \rightsquigarrow i \in \mathcal{E}$, $\boldsymbol{\pi}_{n,i}(\theta)$ is positive by construction. Now, since $\mathcal{G}$ is strongly connected, we know $\boldsymbol{\pi}(\theta)$ is irreducible, therefore $\mathbf{M}(\theta)$ is as well, as $\boldsymbol{f}$ is positive. Finally, the columns of $\mathbf{M}(\theta)$ sum to 0, as $\boldsymbol{\pi}(\theta)$ is stochastic.

In what follows, we fix $\boldsymbol{f}$, as we only want to redirect the flows, by modifying $\boldsymbol{\pi}(\theta)$. Therefore, all quantities related to the diffusion dynamics are functions of θ, and we write accordingly $\mathbf{G}(\theta)$, $\mathcal{R}_0(\theta)$ and $\tilde{\mu}_\theta$ the associated next-generation matrix, basic reproduction number, and stationary distribution, respectively.

B. Problem Statement and Methodology

We want to minimize the final number of infected individuals, by optimising over the parameter θ[6]:

$$\min_{\theta \in \Theta} \sum_{n \in \mathcal{N}} R_n(\infty).$$

To do this, we define the three following loss functions on θ.

Epidemic loss. The main loss is the epidemic loss, defined by

$$\textsc{EpiLoss}(\theta) = \mathcal{R}_0(\theta).$$

The associated policy, $\boldsymbol{\pi}(\theta^*)$, with $\theta^* \in \arg\min_\theta \textsc{EpiLoss}(\theta)$, is called the epidemic policy. It aims at stabilising the DFE, taking $\mathcal{R}_0(\theta)$ below 1, and reducing the final size: see Section IV-A.

No diffusion loss. The second loss, or NoDiffLoss for "No Diffusion Loss", is defined by, for every θ,

$$\textsc{NoDiffLoss}(\theta) = \mathcal{S}_a\left(\boldsymbol{\beta}\boldsymbol{\delta}^{-1}\tilde{\mu}_\theta\right),$$

[5]Namely, for every edge $n \rightsquigarrow i \in \mathcal{E}$, we impose $\boldsymbol{\pi}_{n,i}(\theta) = e^{\theta_{n,i}} / \sum_{j, n \rightsquigarrow j \in \mathcal{E}} e^{\theta_{n,j}}$, while for every node i such that there is no edge $n \rightsquigarrow i$, we impose $\boldsymbol{\pi}_{n,i}(\theta) = 0$.

[6]The quantities $R_n(\infty)$ depend on the diffusion matrix $\mathbf{M}(\theta)$, which itself depends on θ.

Notation	Definition	Meaning
$\mathcal{G}$	—	Directed, strongly connected, graph
$\mathcal{N}$	—	Nodes set of the graph
$\mathcal{E}$	—	Edges set of the graph
n	integer, $1 \leq n \leq N$	Node of the graph
$\lvert\mathcal{N}\rvert$	positive integer	Cardinal of the nodes set
$\lvert\mathcal{E}\rvert$	positive integer	Cardinal of the edges set
$u \odot v$	$(u_n v_n)_{1 \leq n \leq \lvert\mathcal{N}\rvert} \in \mathbb{R}^{\lvert\mathcal{N}\rvert}$	Coordinate-wise product
$1_{\lvert\mathcal{N}\rvert}$	$(1, \dots, 1) \in \mathbb{R}^{\lvert\mathcal{N}\rvert}$	Unit vector
$\mathrm{Id}_{\lvert\mathcal{N}\rvert}$	$\mathrm{Id}_{\lvert\mathcal{N}\rvert} \in \mathcal{M}_{\lvert\mathcal{N}\rvert}(\mathbb{R})$	Identify matrix of size $\lvert\mathcal{N}\rvert$
β_n	$\beta_n > 0$	Infection rate of node n
γ_n	$\gamma_n > 0$	Incubation rate of node n
δ_n	$\delta_n > 0$	Curing rate of node n
$\boldsymbol{\beta}$	$\mathrm{diag}\left(\beta_1, \dots \beta_{\lvert\mathcal{N}\rvert}\right) \in \mathcal{M}_{\lvert\mathcal{N}\rvert}(\mathbb{R})$	Diagonal matrix
$\boldsymbol{\gamma}$	$\mathrm{diag}\left(\gamma_1, \dots \gamma_{\lvert\mathcal{N}\rvert}\right) \in \mathcal{M}_{\lvert\mathcal{N}\rvert}(\mathbb{R})$	Diagonal matrix
$\boldsymbol{\delta}$	$\mathrm{diag}\left(\delta_1, \dots \delta_{\lvert\mathcal{N}\rvert}\right) \in \mathcal{M}_{\lvert\mathcal{N}\rvert}(\mathbb{R})$	Diagonal matrix
S, I, R	$S, I, R \in \mathbb{R}_+^{\lvert\mathcal{N}\rvert}$	Vectors of numbers of individuals in compartments
$\mathcal{R}_0$	$\mathcal{R}_0 \geq 0$	Basic reproduction number
$\mathbf{M}$	$\mathbf{M} \in \mathcal{M}_{\lvert\mathcal{N}\rvert}(\mathbb{R})$, with constraints	Diffusion matrix, Definition 1
$\tilde{\mu}_\mathbf{M}$	$\tilde{\mu}_\mathbf{M} \in \mathbb{R}_+^{\lvert\mathcal{N}\rvert}$	Stationary distribution of $\mathbf{M}$
$\mathbf{G}$	$\mathbf{G} \in \mathcal{M}_{\lvert\mathcal{N}\rvert}(\mathbb{R})$	Next-generation matrix
$\boldsymbol{\pi}$	$\boldsymbol{\pi} \in \mathcal{M}_{\lvert\mathcal{N}\rvert}(\mathbb{R})$	Policy, Definition 2
f_n	$f > 0$	Outrate of node n
$\boldsymbol{f}$	$\mathrm{diag}\left(f_1, \dots f_{\lvert\mathcal{N}\rvert}\right) \in \mathcal{M}_{\lvert\mathcal{N}\rvert}(\mathbb{R})$	Matrix of outrates
θ	$\theta \in \mathcal{M}_{\lvert\mathcal{N}\rvert}(\mathbb{R})$, with constraints	Parameter to optimise over
Θ	$\Theta = \mathbb{R}^{\lvert\mathcal{E}\rvert}$	Parameter space

TABLE I: Main notations

where $\mathcal{S}_a : \mathbb{R}^{|\mathcal{N}|} \to \mathbb{R}$, is the a-smooth max function defined, for $a > 0$ and $u = (u_1, \dots, u_{|\mathcal{N}|})$, by

$$\mathcal{S}_a(u) = \frac{\sum_{i=1}^{|\mathcal{N}|} u_i \exp(a u_i)}{\sum_{i=1}^{|\mathcal{N}|} \exp(a u_i)}.$$

It is a smoothed maximum of the individual basic reproduction numbers $\beta_n \delta_n^{-1} \tilde{\mu}_\theta(n)$ of each node n, computed when there is no diffusion[7]. The maximum of these reproduction numbers is therefore the limit, when $\tau \to \infty$, of the basic reproduction number of the system of Equation (2), when the diffusion is replaced by $\mathbf{M}/\tau$. Here, τ acts as the typical time at which diffusion occurs.

Quick diffusion loss. Conversely, the third loss, QuickDiffLoss for "Quick Diffusion Loss", is the limit, when $\tau \to 0$, of the basic reproduction number — see Beaufort *et al.* [48] for a proof in the special case of a SIR reaction, but the same reasoning applies to SEIR. It is defined for every parameter θ by,

$$\textsc{QuickDiffLoss}(\theta) = \frac{\sum_{n \in \mathcal{N}} \beta_n \tilde{\mu}_\theta(n)^2}{\sum_{n \in \mathcal{N}} \delta_n \tilde{\mu}_\theta(n)}.$$

For each loss, the optimal parameter θ^* we obtain defines a control policy $\boldsymbol{\pi}(\theta^*)$. The policies obtained by minimising the three losses are called respectively EpiPol, NoDiffPol and QuickDiffPol.

Note that NoDiffLoss and QuickDiffLoss — which are simpler than EpiLoss — only consider the diffusion dynamics through its static part, the stationary distribution, while EpiLoss also takes into account its "dynamic" part, through its explicit dependency on $\mathbf{M}(\theta)$, which governs the transfers of population between the nodes during the epidemic. They

[7]Or an infinitely slow diffusive dynamics, with respect to the speed of the epidemic reaction.

both attempt at redistributing the population on the network, by sending it to nodes with low β_n, and high δ_n coefficients. Indeed, when β_n is small, and δ_n is high, the epidemic is less severe. Moreover, they balance it by taking into account the stationary distribution $\tilde{\mu}_{\mathbf{M}}$, as increasing the number of individuals in a node also increases its individual reproduction number $\beta_n \delta_n^{-1} \tilde{\mu}_\theta(n)$, which is bad for epidemic control. Second, the losses NoDiffLoss and QuickDiffLoss are limit cases, as far as the speed of the diffusive dynamics is concerned. Now, our numerical simulations — see Section V — show that results are better with EpiLoss, justifying the greater complexity of use it entails.

IV. Theoretical Results: Final Size and Optimisation Problem

We first present our final size relation in Section IV-A. Second, we prove the losses we study are differentiable in Section IV-B.

A. Final Size Relation

Consider some diffusion matrix $\mathbf{M}_{\text{ref}}$ such that $\mathcal{R}_0(\mathbf{M}_{\text{ref}}) < 1$. Then, for diffusion matrices $\mathbf{M}$ close enough to $\mathbf{M}_{\text{ref}}$, the ratio between the final size and the initial number of infected individuals is controlled by monotonic functions of the basic reproduction numbers $\mathcal{R}_0(\mathbf{M})$. First, we define two quantities needed to formally state this result. For every diffusion matrix $\mathbf{M}$, we write $v^{\mathbf{M}}$ a non-negative eigenvector (thus, not zero) summing up to 1, associated to the spectral radius of the next-generation matrix with large domain [29]. Vectors $v^{\mathbf{M}}$ exist because this matrix is non-negative (see the first result of Section 8.3 in Meyer [49]). Then, for every couple $\left(E^{\mathbf{M}}(0),\, I^{\mathbf{M}}(0)\right)$ of vectors, we write $\mathfrak{I}^{\mathbf{M}}(0) = \sum_{n\in\mathcal{N}} \left(E_n^{\mathbf{M}}(0) + I_n^{\mathbf{M}}(0)\right)$ the initial number of individuals either exposed or infected. We can now state our final size relation.

Theorem 3 (Final size relation)**.** *Let $\mathbf{M}_{\text{ref}}$ be a diffusion matrix, such that $\mathcal{R}_0(\mathbf{M}_{\text{ref}}) < 1$. Then, for every $\varepsilon > 0$ small enough, there exists a ball $\mathcal{B}$ of diffusion matrices around $\mathbf{M}_{\text{ref}}$, and $\eta > 0$ such that, for every $\mathbf{M} \in \mathcal{B}$, for every initial condition $\left\|\left(S^{\mathbf{M}}(0), E^{\mathbf{M}}(0), I^{\mathbf{M}}(0), R^{\mathbf{M}}(0)\right) - \left(\tilde{\mu}_{\mathbf{M}_{\text{ref}}}, 0, 0, 0\right)\right\| < \eta$, such that the relation $\left(E^{\mathbf{M}}(0),\, I^{\mathbf{M}}(0)\right) = \mathfrak{I}_0^{\mathbf{M}}(0) v^{\mathbf{M}}$ holds for some vector $v^{\mathbf{M}}$, we have*

$$\frac{\mathfrak{I}_0^{\mathbf{M}}(0)}{1-(1-\varepsilon)\,\mathcal{R}_0(\mathbf{M})} \leq \sum_{n\in\mathcal{N}} R_n^{\mathbf{M}}(\infty) \leq \frac{\mathfrak{I}_0^{\mathbf{M}}(0)}{1-(1+\varepsilon)\,\mathcal{R}_0(\mathbf{M})}, \tag{3}$$

where $R^{\mathbf{M}}(\infty)$ is the asymptotic vector of individuals in compartment R of the solution of Equation (2).

This result shows that, for diffusions $\mathbf{M}$ close enough to $\mathbf{M}_{\text{ref}}$, the final size is controlled by monotonic functions of the basic reproduction numbers $\mathcal{R}_0(\mathbf{M})$: the smaller the $\mathcal{R}_0(\mathbf{M})$, the closer the final size gets to the initial number of infected people, $\mathfrak{I}_0^{\mathbf{M}}(0)$, meaning the epidemic has not spread widely within the population. Note that the proof could be extended to other compartmental models, e.g SIR or SEPIR (see the supplementary materials for a description of the latter).

Our relation has two main limitations. First, it only applies for $\mathcal{R}_0(\mathbf{M}_{\text{ref}}) < 1$, that is when the DFE is already stable. Therefore, in this case, our control only aims at lowering the final size. However, even when the DFE is stable, the final size may be deemed unacceptably high, so that studying if it can be lowered remains relevant. Moreover, we show numerically (Figure 1) that, even when $\mathcal{R}_0 < 1$, lowering it still entails a reduction of the final size. Second, we restrict to initial conditions of the form $\mathfrak{I}^{\mathbf{M}}(0) v^{\mathbf{M}}$. However, by doing so, we only constrain their orientation, but not the absolute number of individuals $\mathfrak{I}_0^{\mathbf{M}}(0)$, which can be any sufficiently small value.

We have therefore addressed — albeit partially — the question of the extent to which the basic reproduction number influences the final size. Another question is the extent to which redirecting the flows modifies the basic reproduction number. We do not address this theoretically in this paper, but the subsequent sections show experimentally that by redirecting the flows, we manage to reduce the basic reproduction number, and the corresponding final size. Theorem 3 is proved in Section VII-C.

B. Differentiability of the Losses

The optimizations of the losses we use are nonlinear, and not symmetrical, optimisation problems, therefore we solve them by direct gradient descent (see Section V-A). This is possible because the losses are differentiable with respect to θ.

Proposition 4 (Differentiability of the Losses)**.** EpiLoss, NoDiffLoss *and* QuickDiffLoss *are differentiable with respect to θ on Θ.*

Since all three losses depend on a smooth way on the basic reproduction number, and the stationary distribution, we only need to prove these are differentiable, which we do in Lemma 5.

Lemma 5 (Differentiability of Relevant Quantities)**.** *The map $\theta \mapsto \tilde{\mu}_\theta$ is differentiable. Moreover, the map $\theta \mapsto \mathcal{R}_0(\theta)$ is differentiable and, for every $\theta \in \Theta$, its euclidean gradient is given by*

$$\frac{\partial \mathcal{R}_0}{\partial \theta}(\theta) = \left(\left\langle r(\theta)\, l(\theta), \frac{\partial}{\partial \theta_{i,j}} \mathbf{G}(\theta) \right\rangle \right)_{1\leq i,j\leq |\mathcal{N}|},$$

where $\langle \cdot, \cdot \rangle$ stands for the dot product on the space of matrices, and $l(\theta)$ and $r(\theta)$ are, respectively, a left and a right Perron-eigenvectors, (associated to the eigenvalue $\mathcal{R}_0(\theta)$) of the next-generation matrix $\mathbf{G}(\theta)$, such that $l(\theta) r(\theta) = 1$.

Proof. First, the stationary distribution is differentiable, as for δt small enough, it is that of the irreducible stochastic matrix $\mathrm{Id}_N + \delta t\, \mathbf{M}(\theta)^T$, which depends in a smooth way on the matrix [49], and $\theta \mapsto \mathbf{M}(\theta)$ is smooth. Second, $\theta \mapsto \mathbf{G}(\theta)$ is smooth. Indeed, from Section III-A we know that $\mathbf{G}(\theta) = \boldsymbol{\beta}\, \mathrm{diag}\left(\tilde{\mu}_\theta\right) \left(\mathbf{M}(\theta) - \boldsymbol{\delta}\right)^{-1} \boldsymbol{\gamma} \left(\mathbf{M}(\theta) - \boldsymbol{\gamma}\right)^{-1}$. Now, the maps $\theta \mapsto \tilde{\mu}_\theta$ and $\theta \mapsto \mathbf{M}(\theta)$ are smooth, and so is matrix inversion. Third, for every $\theta \in \Theta$, $\mathbf{G}(\theta)$ is positive.

We said in Section III-A that $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are diagonal positive matrices. Moreover, we also recalled that, since $\mathbf{M}(\theta)$ is a diffusion matrix, $\tilde{\mu}_\theta$ is a positive vector. Finally, we know that $-\left(\mathbf{M}(\theta)-\boldsymbol{\delta}\right)^{-1}$ and $-\left(\mathbf{M}(\theta)-\boldsymbol{\gamma}\right)^{-1}$ are positive: see for instance Lemma 1 of Arino [17]. Fourth, $\theta \mapsto \mathcal{R}_0(\theta)$ is differentiable. Indeed, thanks to Caswell [50], the map $\mathbf{G} \mapsto \rho(\mathbf{G})$ is differentiable on the set of non-negative, irreducible matrices. Since $\mathbf{G}(\theta)$ is positive, it is *a fortiori* irreducible. The formula then follows from Caswell [50] and the chain rule. □

V. Numerical Simulations

First, we describe the numerical simulations set-up in Section V-A, and discuss some algorithmic issues in Section V-B. Next, we study numerically the relation between the final size, and the basic reproduction number, in Section V-C. We compare the overall performances of the policies for various graphs sizes and topologies in V-D, before studying the effect of the network heterogeneity, and the rate of diffusion, in Sections V-E and V-F, respectively. Finally, we study a semi-real example in V-G.

A. Numerical Simulations Set-up

For each numerical simulation, we start by generating a graph from a random graph generator, specified below. Then, to obtain $\boldsymbol{f}$, we draw the outrates uniformly on $[0, 4\times10^{-1}]$. We draw uniformly in $[-10^{-1}, 10^{-1}]$ the coefficients of a parameter matrix θ_{ref}, with which we construct a reference softmax policy. In turn, this reference policy is used to construct a reference diffusion matrix, the associated trajectories of which provide the baseline against which we compare our results. Finally, when needed, we renormalise the diffusion matrix by the typical time of diffusion $\tau > 0$, which values we specify below.

The epidemiological coefficients are distributed according to the absolute values of normal variables, which parameters are available in the configuration files in the code. Having drawn the δ_n's, we compute the β_n's coefficients in such a way that the basic reproduction numbers of the nodes, $\mathcal{R}_0(n) = \beta_n \tilde{\mu}_{\mathbf{M}_{\text{ref}}}(n)/\delta_n$, are distributed around the threshold 1: some $\mathcal{R}_0(n)$'s are greater than 1, and some lesser. For each loss, we write θ^* the parameter obtained at the end of training. With it, we can compute the optimal policies $\boldsymbol{\pi}(\theta^*)$ associated with the different losses.

We then simulate the epidemic with the different policies on the time interval $[0, 1000]$. We use a uniform time discretisation step of $\Delta t = \min(1, \tau)$, where τ is the typical time at which diffusion occurs, and a time-discretisation scheme coinciding at first order with an Euler scheme, but preserving the positivity of the vectors S, E, I and R. For each setting, the population is initially distributed according to the reference stationary distribution $\tilde{\mu}_{\boldsymbol{\pi}_{\text{ref}}}$, and in 2 nodes chosen at random, 5% of the population is changed from susceptible to exposed.

Finally, we measure the worth of every policy by the relative final size of the epidemic with respect to the reference policy. Namely, if the parameter of the policy is θ^*, we compute $\sum_{n\in\mathcal{N}} R_n^{\mathbf{M}(\theta^*)}(\infty) / \sum_{n\in\mathcal{N}} R_n^{\mathbf{M}_{\text{ref}}}(\infty)$.

Parameter	Influence of size	Influence of graph	Influence of heterogeneity	Influence of diffusion rate
Number of nodes $\lvert\mathcal{N}\rvert$	10 to 50	40	25	25
Stepsize ζ_α	$1.25\cdot10^{-1}$	$1.25\cdot10^{-1}$	$1.25\cdot10^{-1}$	$1.25\cdot10^{-1}$
Momentum parameter ζ_β	10^{-1}	10^{-1}	10^{-1}	10^{-1}
Optimisation step	$4\cdot10^2$	$4\cdot10^2$	$4\cdot10^2$	$4\cdot10^2$
Mean $\mathcal{R}_0(n)$	1	1	1	1
S.d. $\mathcal{R}_0(n)$	$7\cdot10^{-1}$	$7\cdot10^{-1}$	$7\cdot10^{-2}$ to $7\cdot10^{-1}$	$7\cdot10^{-1}$
Mean γ_n, δ_n	$3\cdot10^{-1}$	$3\cdot10^{-1}$	$3\cdot10^{-1}$	$3\cdot10^{-1}$
S.d. γ_n	$4\cdot10^{-1}$	$4\cdot10^{-1}$	$4\cdot10^{-1}$	$4\cdot10^{-1}$
S.d. δ_n	$4\cdot10^{-1}$	$4\cdot10^{-1}$	$4\cdot10^{-2}$ to $4\cdot10^{-1}$	$4\cdot10^{-1}$
Typical time τ	1	1	1	$1\cdot10^{-2}$ to $1\cdot10^2$
Simulation length	$1\cdot10^3$	$1\cdot10^3$	$1\cdot10^3$ to $1\cdot10^5$	$1\cdot10^3$

TABLE II: Parameters used for the numerical simulations presented in Section V-D to V-F (S.d. : standard deviation).

B. Algorithmic Aspects

The parameter space Θ is of size at most $|\mathcal{N}|^2$ (in the case of a complete graph). Given the softmax parameterisation introduced above, the differential of the stationary distribution is of size $|\mathcal{N}|^3$, while that of the next-generation matrix, used to compute the differential of $\mathcal{R}_0(\theta)$, is of size $|\mathcal{N}|^4$. As a result, optimising EPILOSS is costlier than optimising the other two losses. An explicit formula exists for the differential of $\tilde{\mu}_\theta$ [51] but in our numerical simulations, we computed the differential through automatic differentiation, using the Python library TensorFlow.

Losses are optimised using a momentum gradient descent, iterated for 400 steps. Polyak's momentum, also known as the *heavy ball method* [52], introduces a "momentum" term $\zeta_\beta(\theta_k - \theta_{k-1})$ to the k-th optimisation step, where ζ_β is a hyperparameter — typically $\zeta_\beta \in [0, 1]$, although not limited to it. The full momentum update rule is given, denoting $\mathcal{L}$ the optimised loss, by:

$$\theta_{k+1} = \theta_k - \zeta_\alpha \nabla\mathcal{L}(\theta_k) + \zeta_\beta(\theta_k - \theta_{k-1}),$$

where ζ_α is the the step size. Concerning the convergence properties of this approach, it has been shown by Polyak that, for strongly convex and twice continuously differentiable loss functions, a good choice of the hyperparameters ζ_α and ζ_β produces trajectories that converge to the global minimum faster than the trajectories of the classical gradient descent method. Unfortunately, losses defined above — in particular the EPILOSS— are not convex. However, the convergence properties of the heavy ball algorithm in application to non-convex problems have been studied in the specific case of loss functions with L-Lipschitz continuous gradient, $L > 0$, on their domain [53] [54]. Future works should explicitly address this question by exploring the Lipschitz continuity of loss gradients involved in our approach.

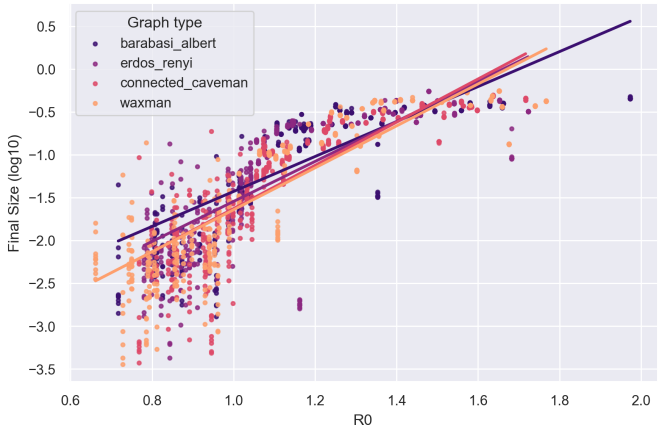


Fig. 1: Final size as a function of the $\mathcal{R}_0$, for various graphs of size $|\mathcal{N}| = 30$. The lines are the regression lines.

C. Basic Reproduction Number and Final Size

First, we study the relation between the basic reproduction number and the final size of the epidemic. We display in Figure 1 the logarithm of the final size as a function of the basic reproduction number. We run the simulation for four random graphs: the Erdős-Rényi, which is a standard model, the Waxman graph, which is a geometric graph, and the Relaxed Caveman graph and Barabási-Albert, which exhibit a somewhat more constrained structure. We see the relation is increasing (the bigger the $\mathcal{R}_0$, the bigger the final size), and that this applies to all types of graphs. That reducing the $\mathcal{R}_0$ would ultimately reduce the final size was expected, but we also knew the relation was not straightforward (Sections II-B and IV-A). We notably see the final size continues to diminish for $\mathcal{R}_0 < 1$. Therefore, even if $\mathcal{R}_0 < 1$ implies the DFE is stable, it is still possible to lower the final size by lowering $\mathcal{R}_0$, which is valuable. Therefore, these results validate our approach. The lines on the plot are the regression lines. We see the relation is closer to linear for the Barabási-Albert and Relaxed Caveman graphs (with sum of squared residuals $R^2 \leq 7$), than for the other graphs ($R^2 \geq 8$), which suggests a slightly better correlation between $\mathcal{R}_0$ and the final size for sparser graphs[8].

D. Policies Comparison

In this section, we compare the performances of the three policies obtained from the corresponding three losses. The parameters used to construct numerical simulations presented in this section — and the next two sections — are summarised in Table II. First, we consider standard Erdős-Rényi graphs of size ranging from 10 to 50 nodes, and show their relative final sizes. We see on Figure 2 EPIPOL performs best overall. On the other hand, QUICKDIFFPOL shows poor results for every graph size while NODIFFPOL performs better for large graphs. Second, in Figure 3, we report the performance of the losses, for different network topologies. We use the four random graphs already used in Section V-C. All graphs have size 40 nodes. Again, we show the median relative final sizes obtained from 15 runs. The performances remain quite good for every type of graph and every policy, except QUICKDIFFPOL which performs poorly. EPIPOL performs best in all cases, with the median consistently below the 50% level, and often much below. We think it is because the loss it comes from incorporates more information about the overall dynamics of the system, than the other two losses, which are derived as limit cases. In particular, EPILOSS has a direct dependency on the diffusion matrix $\mathbf{M}(\theta^*)$, contrary to the other losses. No discernible pattern is distinguishable between the topologies, suggesting performance is not too sensitive to it. These simulations prove the overall value of our approach for epidemic control.

[8]That the relationship between $\mathcal{R}_0$ and the topology is not clear-cut was also noted by Li *et al.* [46]: "the correspondences among $\mathcal{R}_0$ and the topological properties of the contact network are not one-to-one."

E. Influence of the Network Heterogeneity

Next, we study the influence of network heterogeneity. We call heterogeneity of the network the dispersion of the values of the individual basic reproduction numbers, and of the δ_n's. For several values of $x \in [0, 1]$, Erdős-Rényi graphs of size 25 were generated, and the individual basic reproduction numbers, and the δ_n's, were randomly sampled from normal distributions with standard deviations equal to $x\sigma_\delta$ and $x\sigma_{\mathcal{R}_0}$, respectively. The scale factor $x \in [0, 1]$ thus quantifies the heterogeneity of the network. (The γ_n's were left constant, so as not to advantage the epidemiological loss, which could take direct advantage of it, while the limit losses could only do so indirectly, through the stationary distribution.)

We show on Figure 4 the relative final size as a function of the heterogeneity factor x. We first see the final size decreases for all policies, as the scale parameter tends to 1: indeed, the more heterogeneous the network, the more leeway

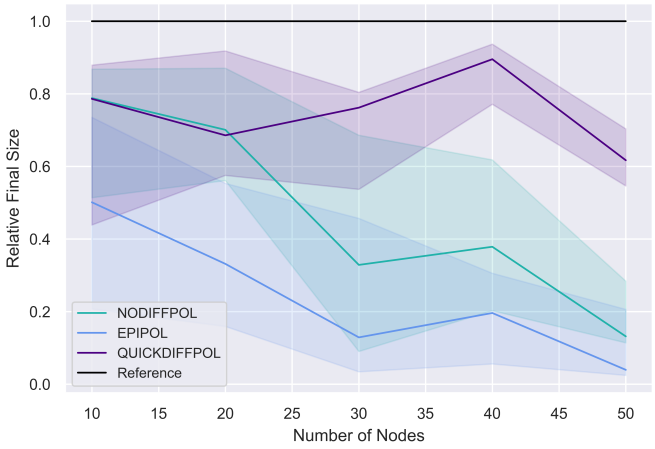


Fig. 2: Relative final size for the three policies, for various sizes of Erdős-Rényi graphs. The final sizes are those of the policies EPIPOL (blue), QUICKDIFFPOL (indigo) and NODIFFPOL (green), obtained from their respective losses. The Reference (black) refers to the final size without any control policy. The lines are the median values — obtained from 15 runs — and the shaded areas gather values in the $[30\%, 70\%]$ interval.

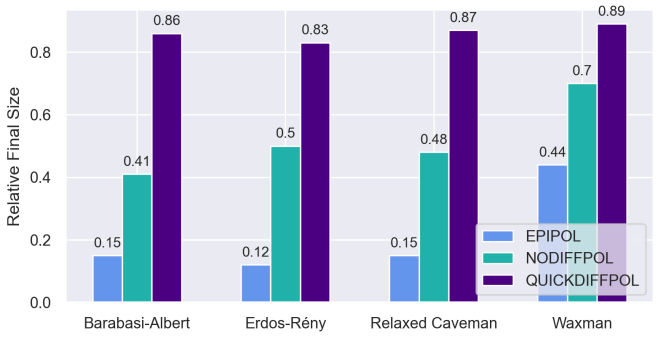


Fig. 3: Relative Final Size for Various Policies (obtained from the corresponding losses) and Various Graphs — median values obtained from 15 runs

.

there is for optimisation. Then, we see EPIPOL is consistently performing better than the policies derived from the limit losses: this suggests that EPILOSS is better able to exploit the heterogeneity, which we think is linked to the fact it incorporates knowledge about the dynamics, and not only the population distribution. These results show that, in order to control the epidemic by acting on the flows, there needs to be disparities in the reaction terms of the network: if all nodes share the same characteristics, the epidemic behaves the same no matter the population flows, so that flow redirection is not efficient.

F. Influence of the Rate of Diffusion

We now study the influence of the rate of diffusion. On Figure 5, we show the relative final sizes for a range of typical times τ of diffusion (equivalently, $1/\tau$ is the rate of diffusion). When $\tau \to 0$, the diffusion happens very quickly, while it happens slowly when $\tau \to \infty$. First, we see that for high typical times, EPIPOL and NODIFFPOL, give close results:

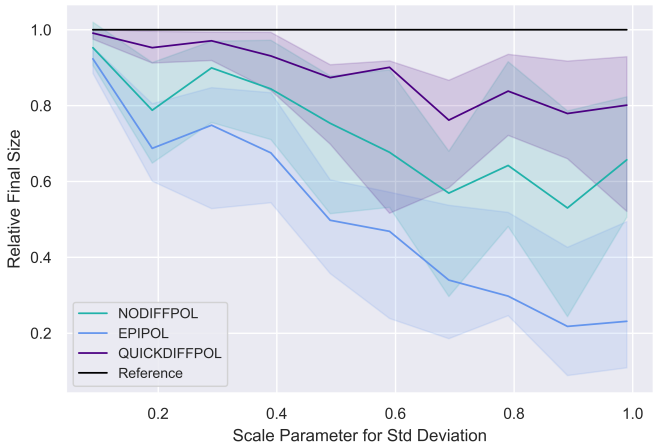


Fig. 4: Relative final size for different heterogeneity levels, for Erdős-Rényi graphs. The lines are the median values — obtained from 15 runs — and the shaded areas gather values in the $[30\%, 70\%]$ interval.

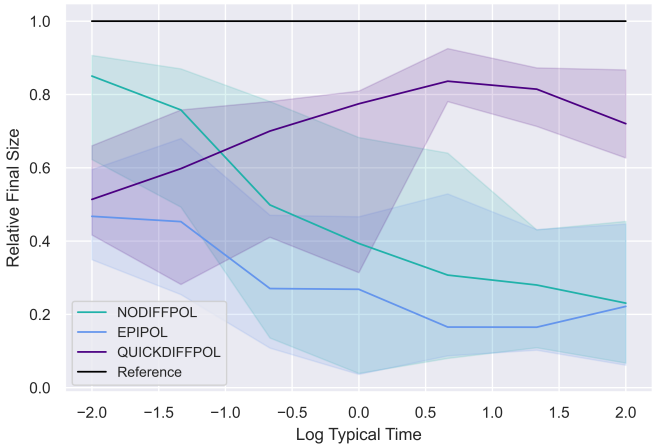
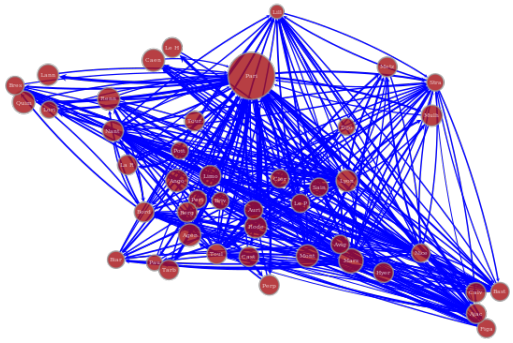


Fig. 5: Relative final size for different typical times, for Erdős-Rényi graphs. The lines are the median values — obtained from 15 runs — and the shaded areas gather values in the $[30\%, 70\%]$ interval.

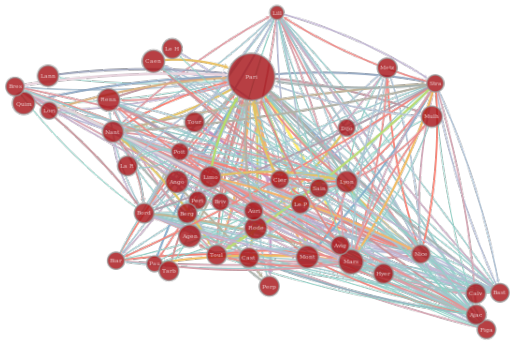
indeed, NODIFFLOSS is derived from the limit of the basic reproduction number when $\tau \to \infty$ — see Section III-B. Second, when τ diminishes, the results for NODIFFPOL deteriorate: NODIFFLOSS is no longer fit for these values. Third, the performance of QUICKDIFFPOL improves when $\tau \to 0$. Again, this was expected, as QUICKDIFFLOSS is tailored for the limit case of very quick diffusion. Fourth, and finally, we see that the EPILOSS produces the best results across the range of times of diffusion, emphasising its overall worth for epidemic reaction-diffusion control. Therefore, these results highlight the importance of the rate of diffusion regarding the performance of epidemic control with flow redirection.

G. Semi-real example

Finally, we conducted simulations on a semi-real example. First, we selected 45 cities with airports in France, and created the graph of the flights between them: each node is a city, and there are edges between nodes n_1 and n_2 if there is at least one flight leading from n_1 to n_2. The graph is strongly connected. Then, we computed the relative sizes of the populations of the cities. Finally, we used gradient descent to compute a diffusion matrix whose stationary distribution was as close as possible to the renormalised distribution of the population of these cities. The resulting network can be seen in Figure 6a — plotted with the graph-tool package [55]. The greater the value of the coordinate of the stationary distribution corresponding to a node, the bigger the node is. The greater the value of the diffusion matrix corresponding to an edge, the bigger the edge size is. Then, we conducted our simulations as described in Section V-A. The epidemic part of the example is synthetic. For each simulation, the relative final size for the EPIPOL policy was less than 20%, which shows a significant improvement. We displayed in Figure 6b the relative change between the reference diffusion matrix, and that obtained with the EPIPOL policy: for each non-diagonal entry, the relative change is $|\mathbf{M}_{i,j}(\theta_{\text{ref}}) - \mathbf{M}_{i,j}(\theta^*)| \,/\, (\mathbf{M}_{i,j}(\theta_{\text{ref}}) + \mathbf{M}_{i,j}(\theta^*))$.



(a) Original network



(b) Relative errors

Fig. 6: Original network, and relative changes for the reconstructed network where edge colours show the size of the relative changes: small (green), middle (pink), large (yellow).

The edge colours are linked to the size of the relative changes: the smallest changes are in green, then the middle changes are in pink, and the largest changes are in yellow. We see on Figure 6b that they are a significant number of changes of each size. Indeed, about 50% of edges have relative changes of size less than 0.1, while 10% of edges have relative changes of size more than 0.3. Moreover, we see on Figure 6b that edges in the south-east are predominantly green. This may be caused by cities in the margins of the network, which are not very likely to have an effect on the rest of it, or cities with already "good" epidemic parameters. However, the computations of the correlations between the relative changes of the edges, and the degree centrality (correlation: 0.34), the stationary distribution (correlation: 0.43), or the individual basic reproduction number (correlation: 0.07) of the nodes they go to, proved inconclusive[9]. The relationship between the amount of changes on the edges, and the topology of the network, is thus not straightforward, and requires work out of scope of our current study.

VI. Conclusion, Future Works

We have shown we can control an epidemic SEIR reaction-diffusion on a directed, and heterogeneous, network by redirecting the flows, thanks to the optimisation of well-designed loss functions, in particular the basic reproduction number of the model. We have provided a final size relation linking the basic reproduction number to the epidemic final sizes, for diffusions around a reference diffusion with basic reproduction number less than 1. Numerically, we have shown control is possible for different topologies, network heterogeneity levels, and speeds of diffusion. Moreover, our experimental results highlight the relevance of EPIPOL, the $\mathcal{R}_0$-based loss, compared to more straightforward losses. However, these improved performances should be balanced against the highest computational costs it entails with respect to the other losses.

Overall, we believe our results make the case for flow redirection as a relevant control tool of epidemic reaction-diffusion processes. Furthermore, we have identified key network parameters which may inform the optimisation design. In turn, this stresses the need to precisely measure them in real networks, which is a direction to further our work into.

One theoretical limitation of our work is the fact the final size relation only holds for $\mathcal{R}_0(\mathbf{M}_{\text{ref}}) < 1$: it would be interesting to extend it to the case $\mathcal{R}_0(\mathbf{M}_{\text{ref}}) > 1$. Then, optimisation reduces the final size, but modifies the network flow structure. Attempting to control the dynamics, while modifying as little as possible the usual behaviour of the network, or some expected output, in the spirit of Umar B. Niazi *et al.* [26], would represent an interesting direction of future research.

VII. Proofs of the final size relation, Theorem 3

We first give the outline of the proof in Section VII-A. It relies on some preliminary result, which we give in Section VII-B. Finally, we provide the proof in VII-C.

A. Outline of the Proof

Our proof consists in three stages: (i) a uniform first order Taylor expansion of the solution of the system Equation (2) at the vicinity of the DFE; (ii) a comparison between the true solution of Equation (2), and the solutions of the linearised system; (iii) the study of the linear system thanks to the next-generation matrix with large domain [29]. The first stage is obtained in Corollary 6 (see Section VII-B). The second stage is done in the first part of the final size relation proof, in Section VII-C (item a). The third stage consists first in integrating the comparison obtained (item b)). Second, it consists in making the next-generation matrix with large domain [29] appear for the linearised system, and using it to bound the $\boldsymbol{\delta} \int_0^\infty I^{\mathbf{M}}(s)\, ds$ term, which is the vector of the final size in each node (item c)). Finally, we conclude by summing along the coordinates of this vector (item d)).

B. Uniform stability at the DFE

To prove Theorem 3, we first need a strengthening of the standard stability result recalled in Section III-A. This strenghtening is obtained in Corollary 6, which establishes a uniform stability property for the Disease Free equilibrium.

Corollary 6. *Let $\mathcal{R}_0(\mathbf{M}_{\text{ref}}) < 1$. Then, for any $\varepsilon > 0$, there is a ball $\mathcal{B}$ around $\mathbf{M}_{\text{ref}}$, and $\eta > 0$ such that,*

[9] This was inconclusive as well for the same correlations, computed with the nodes the edges come from.

for every initial condition $\big(S(0),E(0),I(0),R(0)\big)$ *satisfying* $\big\|\big(S(0),E(0),I(0),R(0)\big)-\big(\tilde{\mu}_{\mathbf{M}_{\rm ref}},0,0,0\big)\big\|<\eta$, *for any diffusion matrix* $\mathbf{M}$ *in* $\mathcal{B}$, *for all* $t\geq 0$, *we have*

$$\big\|S^{\mathbf{M}}(t)-\tilde{\mu}_{\mathbf{M}}\big\|_{\infty}\leq\varepsilon\,\|\tilde{\mu}_{\mathbf{M}}\|_{\infty}.$$

To prove Corollary 6, we need several intermediary results.

Lemma 7 (Uniform exponential boundedness)**.** *Let* $\mathbf{Q}_0$ *be a, finite dimensional, matrix the eigenvalues of which have all negative real parts. Let* $\|\cdot\|$ *be any norm on the space of matrices. Then, we may find* $\lambda>0$, $\kappa\geq 0$, *and a ball* $\mathcal{B}$ *around* $\mathbf{Q}_0$ *such that, for every* $\mathbf{Q}\in\mathcal{B}$, *for all* $t\geq 0$, *we have* $\|\exp(t\mathbf{Q})\|\leq\kappa\,\exp(-\lambda t)$.

Proof. Let $\mathbf{P}$ be positive definite such that $\mathbf{P}\mathbf{Q}_0+\mathbf{Q}_0^T\mathbf{P}=-\mathrm{Id}$. Such a matrix exists because the eigenvalues of $\mathbf{Q}_0$ have negative real parts. For any vector X, we write $\|X\|_{\mathbf{P}}^2=X^T\mathbf{P}X$. Since all norms are equivalent, we can choose $\alpha>0$ such that $\alpha\|X\|_{\mathbf{P}}\leq\|X\|$ for all X. Let then $\mathcal{B}$ be a ball around $\mathbf{Q}_0$, such that, for every $\mathbf{Q}\in\mathcal{B}$, all the eigenvalues of $\mathbf{P}\mathbf{Q}+\mathbf{Q}^T\mathbf{P}$ are negative. This is possible because $\mathbf{Q}\mapsto\mathbf{P}\mathbf{Q}+\mathbf{Q}^T\mathbf{P}$ is continuous, and because the matrices $\mathbf{P}\mathbf{Q}+\mathbf{Q}^T\mathbf{P}$ are symmetric, so that all their eigenvalues are real. Let $-\lambda<0$ be any upper bound for the eigenvalues of the $\mathbf{P}\mathbf{Q}+\mathbf{Q}^T\mathbf{P}$'s for $\mathbf{Q}\in\mathcal{B}$. Let now $\mathbf{Q}\in\mathcal{B}$, X_0 be a vector, and X be the solution of $\frac{dX}{dt}=\mathbf{Q}\,X$, with initial condition $X(0)=X_0$. For all $t\geq 0$, we have

$$\begin{aligned}\frac{d}{dt}\|X\|_{\mathbf{P}}^2&=X^T\left(\mathbf{Q}^T\mathbf{P}+\mathbf{P}\mathbf{Q}\right)X\\&\leq-\lambda\,\|X\|^2\leq-\alpha\lambda\,\|X\|_{\mathbf{P}}^2,\end{aligned}$$

by construction of λ. As a consequence, for all $t\geq 0$, by comparison, we known that $\|X(t)\|_{\mathbf{P}}\leq\exp\left(-\frac{\alpha\lambda}{2}t\right)\|X_0\|_{\mathbf{P}}$, and therefore $\|\exp(\mathbf{Q}t)\,X_0\|_{\mathbf{P}}\leq\exp\left(-\frac{\alpha\lambda}{2}t\right)\|X_0\|_{\mathbf{P}}$. Since X_0 was any vector, the operator norm $\|\cdot\|_{\rm op}$ associated with $\|\cdot\|_{\mathbf{P}}$ at the start and finish satisfies, for all $t\geq 0$, $\|\exp(\mathbf{Q}t)\|_{\rm op}\leq\exp\left(-\frac{\alpha\lambda}{2}t\right)$. Being in finite dimension, all norms are equivalent so that, for some $\beta\geq 0$ independent of $\mathbf{Q}$, we have, for all $t\geq 0$, $\|\exp(\mathbf{Q}t)\|\leq\beta\,\|\exp(\mathbf{Q}t)\|_{\rm op}$. As result, for all $\mathbf{Q}\in\mathcal{B}$, for all $t\geq 0$, we have $\|\exp(\mathbf{Q}t)\|\leq\beta\exp\left(-\frac{\alpha\lambda}{2}t\right)$. □

Lemma 8 (Uniform closeness to the stationary distribution)**.** *Let* $\mathbf{M}_{\rm ref}$ *be a diffusion matrix. For all* $\varepsilon>0$, *we can find a ball* $\mathcal{B}$ *around* $\mathbf{M}_{\rm ref}$, *and* $\eta>0$, *such that, for all initial distribution* N_0 *verifying* $|N_0-\tilde{\mu}_{\mathbf{M}_{\rm ref}}|<\eta$, *for all diffusion matrix* $\mathbf{M}\in\mathcal{B}$, *for all* $t\geq 0$, *we have* $\|N_{\mathbf{M}}(t)-\tilde{\mu}_{\mathbf{M}}\|<\varepsilon$, *where* $N_{\mathbf{M}}(t)=\exp(t\mathbf{M})\,N_0$ *is the solution starting at* N_0 *of* $dN/dt=\mathbf{M}N$.

Proof. Let $\mathbf{M}$ be a diffusion matrix. We know that $\mathbb{R}^{|\mathcal{N}|}=\mathbb{R}\,\tilde{\mu}_{\mathbf{M}}\oplus\mathcal{H}$, where $\mathcal{H}=\left\{\nu\in\mathbb{R}^{|\mathcal{N}|}\,|\sum_n\nu_n=0\right\}$. Let N_0 be an initial distribution. We decompose $N_0=\alpha_{\mathbf{M}}\,\tilde{\mu}_{\mathbf{M}}+\nu$, with $\alpha_{\mathbf{M}}$ a real, and $\nu\in\mathcal{H}$. Indeed, $\alpha_{\mathbf{M}}=\sum_{n\in\mathcal{N}}N_{0,n}=1$ since N_0 is a distribution. For all $t\geq 0$, we have $N_{\mathbf{M}}(t)-\tilde{\mu}_{\mathbf{M}}=\exp(t\mathbf{M})\,N_0-\tilde{\mu}_{\mathbf{M}}=\exp(t\mathbf{M})\,(\tilde{\mu}_{\mathbf{M}}+\nu)-\tilde{\mu}_{\mathbf{M}}=\exp(t\mathbf{M})\,\tilde{\mu}_{\mathbf{M}}-\tilde{\mu}_{\mathbf{M}}+\exp(t\mathbf{M})\,\nu=\exp(t\mathbf{M})\,\nu$, since $\mathbf{M}\tilde{\mu}_{\mathbf{M}}=0$. Therefore, for all $t\geq 0$, we have $\|N_{\mathbf{M}}(t)-\tilde{\mu}_{\mathbf{M}}\|\leq\|\exp(t\mathbf{M})\|_{\rm op}\,\|\nu\|$. Now, we also have $\|\nu\|=\|N_0-\tilde{\mu}_{\mathbf{M}}\|=\|N_0-\tilde{\mu}_{\mathbf{M}}+\tilde{\mu}_{\mathbf{M}}-\tilde{\mu}_{\mathbf{M}}\|\leq\|N_0-\tilde{\mu}_{\mathbf{M}}\|+\|\tilde{\mu}_{\mathbf{M}}-\tilde{\mu}_{\mathbf{M}}\|$. All diffusion matrices have only eigenvalues with negative real parts in $\mathcal{H}$, so that it is true in particular for $\mathbf{M}_{\rm ref}$ and, thanks to the proof of Lemma 7, we may find a ball $\mathcal{B}$ of diffusion matrices around $\mathbf{M}_{\rm ref}$, and $\kappa\geq 0$ such that, for every $\mathbf{M}\in\mathcal{B}$, we have, for all $t\geq 0$, $\|\exp(t\mathbf{M})\|_{{\rm op},\mathcal{H}}<\kappa$, where $\|\cdot\|_{{\rm op},\mathcal{H}}$ is the operator norm for the restriction of matrices $\mathbf{M}$ to the space $\mathcal{H}$. We used the fact that restriction to $\mathcal{H}$ is a continuous function of the matrix.

Let then $\varepsilon>0$. Now, the stationary distribution of a diffusion matrix depends continuously on the matrix [51]. Therefore, upon diminishing $\mathcal{B}$, for all diffusion matrix $\mathbf{M}\in\mathcal{B}$, we may assume $\|\tilde{\mu}_{\mathbf{M}_{\rm ref}}-\tilde{\mu}_{\mathbf{M}}\|\leq\varepsilon/\kappa$. Choose N_0 such that $\|N_0-\tilde{\mu}_{\mathbf{M}_{\rm ref}}\|<\varepsilon/\kappa$. As a result, thanks to Lemma 7, for $\mathbf{M}\in\mathcal{B}$, for all $t\geq 0$, we obtain $\|N_{\mathbf{M}}(t)-\mu_{\mathbf{M}}\|\leq\|\exp(t\mathbf{M})\|_{{\rm op},\mathcal{H}}\,\frac{\varepsilon+\varepsilon}{\kappa}\leq 2\,\varepsilon$. □

We now introduce the following notations. For every diffusion matrix $\mathbf{M}$, and every initial condition, let us define (letting the dependency of E_+ and I_+ on $\mathbf{M}$ be implicit so as to simplify notations),

$$\begin{cases}\dfrac{dE_+}{dt}=\boldsymbol{\beta}N_{\mathbf{M}}\odot I_+-(\mathbf{M}-\boldsymbol{\gamma})\,E_+\\[2ex]\dfrac{dI_+}{dt}=\boldsymbol{\gamma}\,E_+-(\mathbf{M}-\boldsymbol{\delta})\,I_+.\end{cases}$$

For every $\mathbf{M}$, and every $t\geq 0$, define further

$$\mathbf{A}(\mathbf{M},t)=\begin{pmatrix}\mathbf{M}-\boldsymbol{\gamma}&\boldsymbol{\beta}\mathrm{Diag}\left(N_{\mathbf{M}}(t)\right)\\\boldsymbol{\gamma}&\mathbf{M}-\boldsymbol{\delta}\end{pmatrix}.$$

Then, in matrix notations, we have $\left(\frac{dE_+}{dt},\frac{dI_+}{dt}\right)=\mathbf{A}(\mathbf{M},t)\,(E_+(t),\,I_+(t))$.

Lemma 9 (Upper-bounding linear system)**.** *Let* $\mathbf{M}_{\rm ref}$ *be a diffusion matrix. Define* $\mathbf{A}=\mathbf{A}(\mathbf{M}_{\rm ref},\infty)$, *that is* $\mathbf{A}=\begin{pmatrix}\mathbf{M}_{\rm ref}-\boldsymbol{\gamma}&\boldsymbol{\beta}\mathrm{diag}\left(\tilde{\mu}_{\mathbf{M}_{\rm ref}}\right)\\\boldsymbol{\gamma}&\mathbf{M}_{\rm ref}-\boldsymbol{\delta}\end{pmatrix}$. *Assume* $\mathcal{R}_0(\mathbf{M}_{\rm ref})<1$. *Then, we can find a ball* $\mathcal{B}$ *around* $\mathbf{M}_{\rm ref}$, $\lambda>0$, $\eta>0$, *and* $\kappa\geq 0$ *such that, for all* $\|N_0-\tilde{\mu}_{\mathbf{M}_{\rm ref}}\|<\eta$, *for all diffusion matrices* $\mathbf{M}\in\mathcal{B}$, *for all* $t\geq 0$, *we have* $\|(E_+(t),\,I_+(t))\|\leq\kappa\,\exp(-\lambda t)\,\|(E_+(0),\,I_+(0))\|$.

Proof. Let us start by proving that the $\mathbf{A}(\mathbf{M},t)$'s are uniformly close to $\mathbf{A}$, subject to some conditions we now describe. Let $\varepsilon>0$. Thanks to Lemma 8, we can find a ball $\mathcal{B}$ around $\mathbf{M}_{\rm ref}$, and $\eta>0$, such that, for all diffusion matrix $\mathbf{M}\in\mathcal{B}$, for all $\|N_0-\tilde{\mu}_{\mathbf{M}_{\rm ref}}\|<\eta$, for all $t\geq 0$, we have $\|N_{\mathbf{M}}(t)-\tilde{\mu}_{\mathbf{M}}\|\leq\varepsilon$. Now, upon diminishing $\mathcal{B}$, we may also assume that, for all diffusion matrix $\mathbf{M}\in\mathcal{B}$, $\|\tilde{\mu}_{\mathbf{M}_{\rm ref}}-\tilde{\mu}_{\mathbf{M}}\|<\varepsilon$. As a result, for all such matrix $\mathbf{M}\in\mathcal{B}$, for all $\|N_0-\tilde{\mu}_{\mathbf{M}_{\rm ref}}\|<\eta$, for all $t\geq 0$, we have $\|N_{\mathbf{M}}(t)-\tilde{\mu}_{\mathbf{M}_{\rm ref}}\|\leq\|N_{\mathbf{M}}(t)-\tilde{\mu}_{\mathbf{M}}\|+\|\tilde{\mu}_{\mathbf{M}}-\tilde{\mu}_{\mathbf{M}_{\rm ref}}\|$ $\leq\varepsilon+\varepsilon\leq 2\,\varepsilon$. Now, all the other coefficients of $\mathbf{A}(\mathbf{M},t)$ depend continuously on $\mathbf{M}$ so that, upon diminishing $\mathcal{B}$ further, we may assume that, for all diffusion matrix $\mathbf{M}\in\mathcal{B}$, for all $\|N_0-\tilde{\mu}_{\mathbf{M}_{\rm ref}}\|<\eta$, for all $t\geq 0$, we have

$\|\mathbf{A}(\mathbf{M},t)-\mathbf{A}\|_\infty < 2\,\varepsilon$, where $\|\cdot\|_\infty$ is the infinity norm on the coefficients of the matrix, and we conclude by using the equivalence of norms.

Now, $\mathcal{R}_0(\mathbf{M}_{\rm ref})$ is strictly less than one so that, thanks to [29], all the eigenvalues of $\mathbf{A}$ have negative real parts. Proceeding as in the proof of Lemma 7, we may find $-1<-\lambda<0$, and a neighbourhood of $\mathbf{A}$ such that, for every matrix $\mathbf{B}$ inside it, all the eigenvalues of $\mathbf{PB}+\mathbf{B}^T\mathbf{P}$ are real and strictly less than $-\lambda$. By what precedes, upon choosing $\mathcal{B}$ and η small enough, we have that, for every diffusion matrix $\mathbf{M}\in\mathcal{B}$, for every $\|N_0-\tilde\mu_{\mathbf{M}_{\rm ref}}\|<\eta$, for every $t\geq 0$, all the eigenvalues of $\mathbf{P}A(\mathbf{M},t)-A(\mathbf{M},t)^T\mathbf{P}$ are less than $-\lambda$.

Fix now a diffusion matrix $\mathbf{M}\in\mathcal{B}$, and $\|N_0-\tilde\mu_{\mathbf{M}}\|<\eta$. Let then X_0 be a vector, and $X=(E_+,\,I_+)$ be the solution of $\frac{dX}{dt}=A(\mathbf{M},t)\,X$ with $X(0)=X_0$. Again, as in the proof of Lemma 7, we obtain some $\alpha\geq 0$ independent of $\mathbf{M}$ such that, for all $t\geq 0$, we have $\|X(t)\|_P\leq Y(t)^{1/2}\leq \exp\left(\frac{-\lambda\alpha}{2}t\right)\|X_0\|_P$, where Y is the solution, for $t\geq 0$, of $\frac{dY}{dt}=-\lambda\alpha Y$, with initial condition $Y(0)=\|X_0\|^2$.

Thanks to the equivalence of norms, we may find $\kappa\geq 0$ such that, for all $t\geq 0$, we have $\|X(t)\|\leq\kappa\exp\left(\frac{-\lambda\alpha}{2}t\right)\|X_0\|$. This stands for any diffusion matrix $\mathbf{M}\in\mathcal{B}$, and any $\|N_0-\tilde\mu_{\mathbf{M}_{\rm ref}}\|<\eta$, so that we have proven our claim. □

Lemma 10 (Comparison)**.** *Assume $\mathcal{R}_0(\mathbf{M}_{\rm ref})<1$. There exists a ball $\mathcal{B}$ around $\mathbf{M}_{\rm ref}$, $\lambda>0$, $\eta>0$ and $\kappa\geq 0$ such that, for any diffusion matrix $\mathbf{M}\in\mathcal{B}$, for all initial distribution $\|N_0-\tilde\mu_{\mathbf{M}_{\rm ref}}\|<\eta$, for all $t\geq 0$, we have $\left(E^{\mathbf{M}}(t),\,I^{\mathbf{M}}(t)\right)\leq\left(E^{\mathbf{M}}_+(t),\,I^{\mathbf{M}}_+(t)\right)$, and $\left\|\left(E^{\mathbf{M}}(t),\,I^{\mathbf{M}}(t)\right)\right\|\leq\kappa\exp(-\lambda t)\,\|(E(0),\,I(0))\|$, where $\left(E^{\mathbf{M}}(t),\,I^{\mathbf{M}}(t)\right)$ are the E and I coordinates of the system of Equation (2) when the diffusion matrix is $\mathbf{M}$, and the initial population is distributed according to $S(0)+E(0)+I(0)+R(0)=N_0$, and $\left(E^{\mathbf{M}}_+,\,I^{\mathbf{M}}_+\right)$ are introduced before Lemma 9, and have initial condition $(E(0),\,I(0))$.*

Proof. Choose $\mathcal{B}$, λ, η and κ as in Lemma 9. Fix an initial condition $(S(0),E(0),I(0),R(0))$ such that $\|N_0-\tilde\mu_{\mathbf{M}_{\rm ref}}\|<\eta$. Fix a diffusion matrix $\mathbf{M}\in\mathcal{B}$. We drop the "$\mathbf{M}$ exponents" to simplify the notations. By definition, we know that, for all $t\geq 0$, we have $\frac{dE}{dt}=\boldsymbol{\beta}S\odot I-\boldsymbol{\gamma}E+\mathbf{M}E$, and $\frac{dI}{dt}=\boldsymbol{\gamma}E-\boldsymbol{\delta}I+\mathbf{M}I$. Now, for all $t\geq 0$, we know that $S(t)\leq N_{\mathbf{M}}(t)=S(t)+E(t)+I(t)+R(t)$. As a result, for all $t\geq 0$, we have $\frac{dE}{dt}\leq\boldsymbol{\beta}N_{\mathbf{M}}(t)\odot I-\boldsymbol{\gamma}E+\mathbf{M}E$, so that $\left(\frac{dE}{dt},\,\frac{dI}{dt}\right)\leq\mathbf{A}(\mathbf{M},t)\,(E(t),\,I(t))$. Now, for all $t\geq 0$, $\mathbf{A}(\mathbf{M},t)$ is a Metzler matrix, so that, thanks to Section 5.5 of [56], $f(t,X)=\mathbf{A}(\mathbf{M},t)X$ is of type K. As a consequence, we may use the comparison Theorem B.1 of [57] to obtain that, for all $t\geq 0$, $(E(t),\,I(t))\leq(E_+(t),\,I_+(t))$. □

Lemma 11 (Uniform stability of the Disease Free Equilibrium)**.** *Let $\mathbf{M}_{\rm ref}$ be a diffusion matrix, and assume $\mathcal{R}_0(\mathbf{M}_{\rm ref})<1$. Then, for any $\varepsilon>0$, there is a ball $\mathcal{B}$ around $\mathbf{M}_{\rm ref}$, and $\eta>0$ such that, for every initial condition $\left(S(0),E(0),I(0),R(0)\right)$ satisfying $\left\|\left(S(0),E(0),I(0),R(0)\right)-\left(\tilde\mu_{\mathbf{M}_{\rm ref}},0,0,0\right)\right\|<\eta$, for any diffusion matrix $\mathbf{M}$ in $\mathcal{B}$, for all $t\geq 0$, we have*

$$\left\|\left(S^{\mathbf{M}}(t),E^{\mathbf{M}}(t),I^{\mathbf{M}}(t),R^{\mathbf{M}}(t)\right)-\left(\tilde\mu_{\mathbf{M}},0,0,0\right)\right\|\leq\varepsilon,$$

where $\left(S^{\mathbf{M}}(t),E^{\mathbf{M}}(t),I^{\mathbf{M}}(t),R^{\mathbf{M}}(t)\right)$ is the solution for $t\geq 0$ of Equation (2), with initial condition $\left(S(0),E(0),I(0),R(0)\right)$.

Proof. Thanks to Lemma 8 and Lemma 10, there exists a ball $\mathcal{B}$ around $\mathbf{M}_{\rm ref}$, $\lambda>0$, $\eta>0$ and $\kappa\geq 0$ such that, for any diffusion matrix $\mathbf{M}\in\mathcal{B}$, for all initial distribution $\|N_0-\tilde\mu_{\mathbf{M}_{\rm ref}}\|<\eta$, for all $t\geq 0$, we have $\|N_{\mathbf{M}}(t)-\tilde\mu_{\mathbf{M}}\|<\varepsilon$, and $\left\|\left(E^{\mathbf{M}}(t),\,I^{\mathbf{M}}(t)\right)\right\|\leq\kappa\exp(-\lambda t)\,\|(E(0),\,I(0))\|$. Fix a diffusion matrix $\mathbf{M}\in\mathcal{B}$ and some initial condition satisfying the requirements above. Note that by assumption $\|(E(0),\,I(0))\|\leq\eta$. Hence, for all $t\geq 0$, we have $\left\|\left(E^{\mathbf{M}}(t),\,I^{\mathbf{M}}(t)\right)\right\|\leq\kappa\,\eta\exp(-\lambda t)$. As a result, for all $t\geq 0$, we have

$$\begin{aligned}
&\left\|\left(S^{\mathbf{M}}(t),E^{\mathbf{M}}(t),I^{\mathbf{M}}(t),R^{\mathbf{M}}(t)\right)-\left(\tilde\mu_{\mathbf{M}},0,0,0\right)\right\|\\
&\leq\left\|S^{\mathbf{M}}(t)-\tilde\mu_{\mathbf{M}}\right\|+\left\|E^{\mathbf{M}}(t)\right\|+\left\|I^{\mathbf{M}}(t)\right\|+\left\|R^{\mathbf{M}}(t)\right\|\\
&\leq\left\|N_{\mathbf{M}}(t)-\tilde\mu_{\mathbf{M}}\right\|+2\left\|E^{\mathbf{M}}(t)\right\|+2\left\|I^{\mathbf{M}}(t)\right\|+2\left\|R^{\mathbf{M}}(t)\right\|\\
&\leq\quad\varepsilon+4\,\kappa\,\eta+2\left\|R^{\mathbf{M}}(t)\right\|
\end{aligned}$$

since, for all $t\geq 0$, $S^{\mathbf{M}}(t)+E^{\mathbf{M}}(t)+I^{\mathbf{M}}(t)+R^{\mathbf{M}}(t)=N_{\mathbf{M}}(t)$. Now, for all $t\geq 0$, $\left\|R^{\mathbf{M}}(t)\right\|\leq\sum_n R^{\mathbf{M}}_n(t)\leq\sum_n R^{\mathbf{M}}_n(\infty)$, as the individuals who arrive in some compartment R stay there indefinitely. Now, for all $t\geq 0$, we have $\sum_{n\in\mathcal{N}}\frac{dR_n(t)}{dt}=\sum_{n\in\mathcal{N}}\delta_n I_n(t)$, so that, for all $t\geq 0$, we have

$$\begin{aligned}
\sum_{n\in\mathcal{N}}R_n(\infty)-\sum_{n\in\mathcal{N}}R_n(0)&=\sum_{n\in\mathcal{N}}\delta_n\int_0^\infty I_n(t)dt\\
&\leq\kappa\,\eta\sum_{n\in\mathcal{N}}\delta_n\int_0^\infty\exp(-\lambda t)\,dt=\frac{\kappa\,\eta}{\lambda}\sum_{n\in\mathcal{N}}\delta_n.
\end{aligned}$$

As a result, for all $t\geq 0$, remembering $\sum_{n\in\mathcal{N}}R_n(0)\leq\eta N$, we have

$$\begin{aligned}
&\left\|\left(S^{\mathbf{M}}(t),E^{\mathbf{M}}(t),I^{\mathbf{M}}(t),R^{\mathbf{M}}(t)\right)-\left(\tilde\mu_{\mathbf{M}},0,0,0\right)\right\|\\
&\qquad\leq\varepsilon+4\,\kappa\,\eta+2\eta N+2\frac{\kappa\,\eta}{\lambda}\sum_{n\in\mathcal{N}}\delta_n.
\end{aligned}$$

Upon diminishing η, we have therefore proven our claim. □

We can now prove Corollary 6.

Proof. Let $\mathcal{B}_1$ be any bounded ball with radius $r>0$ around $\mathbf{M}_{\rm ref}$. Then, for any diffusion matrix $\mathbf{M}\in\mathcal{B}_1$, we have $c=\min_{\mathbf{M}\in\mathcal{B}_1}\|\tilde\mu_{\mathbf{M}}\|>0$, as the closure of $\mathcal{B}_1$ is compact, and $\mathbf{M}\mapsto\|\tilde\mu_{\mathbf{M}}\|$ is continuous. Let $\varepsilon>0$, $\varepsilon_1>0$ to be fixed later, and $\varepsilon_2=c\varepsilon_1$. Thanks to Lemma 11 applied to ε_2, we can find a ball $\mathcal{B}_2$ around $\mathbf{M}_{\rm ref}$, and $\eta>0$ such that, for every initial condition $\left(S(0),E(0),I(0),R(0)\right)$ satisfying $\left\|\left(S(0),E(0),I(0),R(0)\right)-\left(\tilde\mu_{\mathbf{M}_{\rm ref}},0,0,0\right)\right\|<\eta$, for any diffusion matrix $\mathbf{M}$ in $\mathcal{B}_2$, for all $t\geq 0$, we have $\left\|S^{\mathbf{M}}(t)-\tilde\mu_{\mathbf{M}}\right\|<\varepsilon_2$. Define $\mathcal{B}$ a ball around $\tilde\mu_{\mathbf{M}_{\rm ref}}$, of radius

$\min(r,\,\eta) > 0$. As a result, for every $\mathbf{M} \in \mathcal{B}$, for all $t > 0$ and initial condition satisfying the requirements above, we have

$$\left\| S^{\mathbf{M}}(t) - \tilde{\mu}_{\mathbf{M}} \right\| < \varepsilon_1 \min_{\mathbf{M}\in\mathcal{B}_1} \|\tilde{\mu}_{\mathbf{M}}\| \leq \varepsilon_1 \min_{\mathbf{M}\in\mathcal{B}} \|\tilde{\mu}_{\mathbf{M}}\| .$$

Invoking the equivalence of norms, we obtain $\left\| S^{\mathbf{M}}(t) - \tilde{\mu}_{\mathbf{M}} \right\|_\infty < \kappa\varepsilon \|\tilde{\mu}_{\mathbf{M}}\|_\infty$, for some $\kappa > 0$ which depends only on the norms, and is thus independent of η. Setting $\varepsilon_1 = \varepsilon/\kappa$, we obtain $\left\| S^{\mathbf{M}}(t) - \tilde{\mu}_{\mathbf{M}} \right\|_\infty < \varepsilon \|\tilde{\mu}_{\mathbf{M}}\|_\infty$. □

C. Proof of the Final Size Relation, Theorem 3

a) Comparison with "constant-matrices": As a consequence of Corollary 6, for all $t \geq 0$, we have, coordinate-wise,

$$\begin{aligned}
&\begin{pmatrix} \mathbf{M}-\boldsymbol{\gamma} & \boldsymbol{\beta}\mathrm{diag}\left(\tilde{\mu}_{\mathbf{M}}(1-\varepsilon)\right) \\ \boldsymbol{\gamma} & \mathbf{M}-\boldsymbol{\delta} \end{pmatrix} \\
&\leq \begin{pmatrix} \mathbf{M}-\boldsymbol{\gamma} & \boldsymbol{\beta}\mathrm{diag}\left(S^{\mathbf{M}}(t)\right) \\ \boldsymbol{\gamma} & \mathbf{M}-\boldsymbol{\delta} \end{pmatrix} \\
&\leq \begin{pmatrix} \mathbf{M}-\boldsymbol{\gamma} & \boldsymbol{\beta}\mathrm{diag}\left(\tilde{\mu}_{\mathbf{M}}(1+\varepsilon)\right) \\ \boldsymbol{\gamma} & \mathbf{M}-\boldsymbol{\delta} \end{pmatrix}.
\end{aligned}$$

Let us define

$$\mathbf{F}_{\mathbf{M}} = \begin{pmatrix} 0 & \boldsymbol{\beta}\,\mathrm{diag}\left(\tilde{\mu}_{\mathbf{M}}\right) \\ 0 & 0 \end{pmatrix}, \text{ and } \quad \mathbf{V}_{\mathbf{M}} = \begin{pmatrix} \mathbf{M}-\boldsymbol{\gamma} & 0 \\ \boldsymbol{\gamma} & \mathbf{M}-\boldsymbol{\delta} \end{pmatrix}.$$

To proceed by comparison, let us define the following two ODEs, which solutions will lower-bound, and upper-bound, the map $t \mapsto \left(E^{\mathbf{M}}(t),\, I^{\mathbf{M}}(t)\right)$:

$$\begin{cases} \left(\dfrac{dE_-^{\mathbf{M}}}{dt},\, \dfrac{dI_-^{\mathbf{M}}}{dt}\right) = \left((1-\varepsilon)\,\mathbf{F}_{\mathbf{M}} + \mathbf{V}_{\mathbf{M}}\right)\left(E_-^{\mathbf{M}}(t),\, I_-^{\mathbf{M}}(t)\right) \\ \left(\dfrac{dE_+^{\mathbf{M}}}{dt},\, \dfrac{dI_+^{\mathbf{M}}}{dt}\right) = \left((1+\varepsilon)\,\mathbf{F}_{\mathbf{M}} + \mathbf{V}_{\mathbf{M}}\right)\left(E_+^{\mathbf{M}}(t),\, I_+^{\mathbf{M}}(t)\right), \end{cases}$$

both starting at $\left(E^{\mathbf{M}}(0),\, I^{\mathbf{M}}(0)\right)$. In the following, we consider fixed a vector $v^{\mathbf{M}}$ such that $\left(E^{\mathbf{M}}(0),\, I^{\mathbf{M}}(0)\right)$ is directed by $v^{\mathbf{M}}$. Now, $(1-\varepsilon)\,\mathbf{F}_{\mathbf{M}} + \mathbf{V}_{\mathbf{M}}$ and $(1+\varepsilon)\,\mathbf{F}_{\mathbf{M}} + \mathbf{V}_{\mathbf{M}}$ are both Metzler matrices, since $\mathbf{V}_{\mathbf{M}}$ is non-negative off diagonal and $\mathbf{F}_{\mathbf{M}}$ has non-negative coefficients. As a result, we may use the comparison Theorem B.1 of Smith and Waltman [57], and we obtain that, for all $t \geq 0$, $\left(E_-^{\mathbf{M}}(t),\, I_-^{\mathbf{M}}(t)\right) \leq \left(E^{\mathbf{M}}(t),\, I^{\mathbf{M}}(t)\right) \leq \left(E_+^{\mathbf{M}}(t),\, I_+^{\mathbf{M}}(t)\right)$. Now, the matrices in the bounding systems are constant, so we can express their solutions with the matrix exponential and, as a result, for all $t \geq 0$, we have

$$\begin{aligned}
&\exp\left(((1-\varepsilon)\mathbf{F}_{\mathbf{M}} + \mathbf{V}_{\mathbf{M}})t\right)\left(E^{\mathbf{M}}(0),\, I^{\mathbf{M}}(0)\right) \\
&\qquad \leq \left(E^{\mathbf{M}}(t),\, I^{\mathbf{M}}(t)\right) \\
&\qquad \leq \exp\left(((1+\varepsilon)\mathbf{F}_{\mathbf{M}} + \mathbf{V}_{\mathbf{M}})t\right)\left(E^{\mathbf{M}}(0),\, I^{\mathbf{M}}(0)\right).
\end{aligned}$$

b) Integrating the comparison: Now, for every matrix A which eigenvalues all have a negative real part, we know that $\int_0^\infty \exp(As)\,ds = -A^{-1}$. Since $\mathcal{R}_0(\mathbf{M}_{\mathrm{ref}}) < 1$, all the eigenvalues of $F_{\mathbf{M}_{\mathrm{ref}}} + V_{\mathbf{M}_{\mathrm{ref}}}$ have real parts (strictly) less than 1. Now, the maps which associate to a matrix its eigenvalues are continuous. Upon diminishing the ball $\mathcal{B}$ around $\mathbf{M}_{\mathrm{ref}}$, we may therefore assume that, for all diffusion matrix $\mathbf{M}$ in $\mathcal{B}$, all the eigenvalues of $\mathbf{F}_{\mathbf{M}} + \mathbf{V}_{\mathbf{M}}$ have real parts - strictly - less than 1. Upon diminishing ε, we may finally assume that, for all diffusion matrix $\mathbf{M} \in \mathcal{B}$, the eigenvalues of $(1-\varepsilon)\,\mathbf{F}_{\mathbf{M}} + \mathbf{V}_{\mathbf{M}}$ and $(1+\varepsilon)\,\mathbf{F}_{\mathbf{M}} + \mathbf{V}_{\mathbf{M}}$ have - strictly - negative real parts. Thus, writing $\mathbf{I} = \int_0^\infty \left(E^{\mathbf{M}}(s),\, I^{\mathbf{M}}(s)\right) ds$, we have

$$\begin{aligned}
&-\left((1-\varepsilon)\,\mathbf{F}_{\mathbf{M}} + \mathbf{V}_{\mathbf{M}}\right)^{-1}\left(E^{\mathbf{M}}(0),\, I^{\mathbf{M}}(0)\right) \leq \mathbf{I} \\
&\qquad \leq -\left((1+\varepsilon)\,\mathbf{F}_{\mathbf{M}} + \mathbf{V}_{\mathbf{M}}\right)^{-1}\left(E^{\mathbf{M}}(0),\, I^{\mathbf{M}}(0)\right). \quad (4)
\end{aligned}$$

c) Using the next-generation matrix: Let us now multiply the inequality in Equation (4) by a square non-negative matrix $\mathbf{Z}$, to be fixed below. We therefore have

$$\begin{aligned}
&-\mathbf{Z}\left((1-\varepsilon)\,\mathbf{F}_{\mathbf{M}} + \mathbf{V}_{\mathbf{M}}\right)^{-1}\left(E^{\mathbf{M}}(0),\, I^{\mathbf{M}}(0)\right) \leq \mathbf{Z}\mathbf{I} \\
&\qquad \leq -\mathbf{Z}\left((1+\varepsilon)\,\mathbf{F}_{\mathbf{M}} + \mathbf{V}_{\mathbf{M}}\right)^{-1}\left(E^{\mathbf{M}}(0),\, I^{\mathbf{M}}(0)\right). \quad (5)
\end{aligned}$$

Using the invertibility of $\mathbf{V}_{\mathbf{M}}$, we know that $-\left((1-\varepsilon)\,\mathbf{F}_{\mathbf{M}} + \mathbf{V}_{\mathbf{M}}\right) = \left(-(1-\varepsilon)\,\mathbf{F}_{\mathbf{M}}\mathbf{V}_{\mathbf{M}}^{-1} - \mathrm{Id}_{|\mathcal{N}|}\right)\mathbf{V}_{\mathbf{M}}$. Since the next-generation matrix with large domain [29], $\mathbf{K}_{\mathbf{M}} = -\mathbf{F}_{\mathbf{M}}\mathbf{V}_{\mathbf{M}}^{-1}$ is non-negative, we can find a right eigenvector $v^{\mathbf{M}} \neq 0$ associated to the spectral radius $\mathcal{R}_0 = \mathcal{R}_0(\mathbf{M}) = \rho(\mathbf{K}_{\mathbf{M}}) = \rho(\mathbf{G}_{\mathbf{M}})$. In what follows, we fix one such $v = v^{\mathbf{M}}$, and drop the explicit dependencies on $\mathbf{M}$ for $v^{\mathbf{M}}$ and $\mathcal{R}_0(\mathbf{M})$ so as to simplify the notations. Hence, we obtain $-\left((1-\varepsilon)\,\mathbf{F}_{\mathbf{M}} + \mathbf{V}_{\mathbf{M}}\right)\mathbf{V}_{\mathbf{M}}^{-1} v = (1-\varepsilon)\,\mathcal{R}_0 v - v$.

Since $\mathcal{R}_0(\mathbf{M}_{\mathrm{ref}}) < 1$, and since the spectral radius of a matrix depends continuously on the matrix, upon reducing $\mathcal{B}$ further, we may assume that, for all diffusion matrix $\tilde{\mathbf{M}} \in \mathcal{B}$, we have $(1+\varepsilon)\,\mathcal{R}_0(\tilde{\mathbf{M}}) < 1$, so that both $1-(1+\varepsilon)\,\mathcal{R}_0(\tilde{\mathbf{M}})$ and, *a fortiori*, $1-(1-\varepsilon)\,\mathcal{R}_0(\tilde{\mathbf{M}})$ do not vanish.

By assumption, $\left(E^{\mathbf{M}}(0),\, I^{\mathbf{M}}(0)\right)$ is directed by v, so that it is also a right eigenvector of $\mathbf{K}_{\mathbf{M}}$ associated to $\mathcal{R}_0(\mathbf{M})$. Hence, we successively have

$$\begin{aligned}
&-\left((1-\varepsilon)\,\mathbf{F}_{\mathbf{M}} + \mathbf{V}_{\mathbf{M}}\right)\mathbf{V}_{\mathbf{M}}^{-1}\left(E^{\mathbf{M}}(0),\, I^{\mathbf{M}}(0)\right) \\
&\qquad = \left((1-\varepsilon)\,\mathcal{R}_0 - 1\right)\left(E^{\mathbf{M}}(0),\, I^{\mathbf{M}}(0)\right)
\end{aligned}$$

then, multiplying by $\left((1-\varepsilon)\,\mathbf{F}_{\mathbf{M}} + \mathbf{V}_{\mathbf{M}}\right)^{-1}$, we obtain

$$\begin{aligned}
&-\mathbf{V}_{\mathbf{M}}^{-1}\left(E^{\mathbf{M}}(0),\, I^{\mathbf{M}}(0)\right) \\
&= \left((1-\varepsilon)\,\mathcal{R}_0 - 1\right)\left((1-\varepsilon)\,\mathbf{F}_{\mathbf{M}} + \mathbf{V}_{\mathbf{M}}\right)^{-1}\left(E^{\mathbf{M}}(0),\, I^{\mathbf{M}}(0)\right).
\end{aligned}$$

Finally, dividing by $1-(1-\varepsilon)\,\mathcal{R}_0$ - which is not zero thanks to the above - and multiplying by the matrix $\mathbf{Z}$, we get

$$\begin{aligned}
&\left((1-\varepsilon)\,\mathcal{R}_0 - 1\right)^{-1}\mathbf{Z}\mathbf{V}_{\mathbf{M}}^{-1}\left(E^{\mathbf{M}}(0),\, I^{\mathbf{M}}(0)\right) \\
&\qquad = -\mathbf{Z}\left((1-\varepsilon)\,\mathbf{F}_{\mathbf{M}} + \mathbf{V}_{\mathbf{M}}\right)^{-1}\left(E^{\mathbf{M}}(0),\, I^{\mathbf{M}}(0)\right).
\end{aligned}$$

The same reasoning applies with the matrix $(1+\varepsilon)\,\mathbf{F}_{\mathbf{M}} + \mathbf{V}_{\mathbf{M}}$ in the upper-bounding system. Plugging it all back into Equation (5), we obtain

$$\begin{aligned}
&-\left(1-(1-\varepsilon)\,\mathcal{R}_0\right)^{-1}\mathbf{Z}\mathbf{V}_{\mathbf{M}}^{-1}\left(E^{\mathbf{M}}(0),\, I^{\mathbf{M}}(0)\right) \leq \mathbf{Z}\mathbf{I} \\
&\qquad \leq -\left(1-(1+\varepsilon)\,\mathcal{R}_0\right)^{-1}\mathbf{Z}\mathbf{V}_{\mathbf{M}}^{-1}\left(E^{\mathbf{M}}(0),\, I^{\mathbf{M}}(0)\right).
\end{aligned}$$

Now, set $\mathbf{Z} = \begin{pmatrix} 0 & 0 \\ 0 & \boldsymbol{\delta} \end{pmatrix}$ and, denoting $\mathbf{\Lambda}_{(i:j)}$ the submatrix of $\mathbf{\Lambda}$ obtained by extracting the i^{th} to j^{th} rows of $\mathbf{\Lambda}$, set $(\mathbf{B}_1,\, \mathbf{B}_2) = \left(\mathbf{V}_{\mathbf{M}}^{-1}\right)_{(|\mathcal{N}|+1:2|\mathcal{N}|)}$, a $|\mathcal{N}| \times 2|\mathcal{N}|$ matrix.

Restraining to the rows $(|\mathcal{N}|+1:2|\mathcal{N}|)$ corresponds to the infected coordinates - that of the vector $I^{\mathbf{M}}$. Thus, we obtain

$$\frac{-\left(\boldsymbol{\delta}\mathbf{B}_1,\,\boldsymbol{\delta}\mathbf{B}_2\right)}{1-(1-\varepsilon)\,\mathcal{R}_0}\left(E^{\mathbf{M}}(0),\,I^{\mathbf{M}}(0)\right) \leq \boldsymbol{\delta}\int_0^\infty I^{\mathbf{M}}(s)ds \leq \frac{-\left(\boldsymbol{\delta}\mathbf{B}_1,\,\boldsymbol{\delta}\mathbf{B}_2\right)}{1-(1+\varepsilon)\,\mathcal{R}_0}\left(E^{\mathbf{M}}(0),\,I^{\mathbf{M}}(0)\right). \quad (6)$$

d) Conclusion: Remember we want to bound $\Delta^{\mathbf{M}} = \sum_{n\in\mathcal{N}} R_n(\infty) = \sum_{n\in\mathcal{N}} \delta_n \int_0^\infty I_n^{\mathbf{M}}(s)ds = \int_0^\infty \boldsymbol{e}^T\boldsymbol{\delta} I^{\mathbf{M}}(s)ds$, where $\boldsymbol{e} = (1,\ldots,1)$ is the vector with all entries equal to unity. Multiplying by $\boldsymbol{e}$ every term of Equation (6), using an explicit expression of $\mathbf{V}_{\mathbf{M}}^{-1}$ - inversion of a 2×2 block matrix - and expanding the terms $\boldsymbol{e}^T\left(\boldsymbol{\delta}\mathbf{B}_1,\,\boldsymbol{\delta}\mathbf{B}_2\right)\left(E^{\mathbf{M}}(0),\,I^{\mathbf{M}}(0)\right)$, we obtain

$$\begin{aligned}&(1-(1+\varepsilon)\,\mathcal{R}_0)\,\Delta^{\mathbf{M}} \leq \\ &\boldsymbol{e}^T\left(\boldsymbol{\delta}\left(\mathbf{M}-\boldsymbol{\delta}\right)^{-1}\boldsymbol{\gamma}\left(\mathbf{M}-\boldsymbol{\gamma}\right)^{-1}E^{\mathbf{M}}(0) - \boldsymbol{\delta}\left(\mathbf{M}-\boldsymbol{\delta}\right)^{-1}I^{\mathbf{M}}(0)\right) \\ &\qquad\leq (1-(1-\varepsilon)\,\mathcal{R}_0)\,\Delta^{\mathbf{M}}.\end{aligned}$$

Using the fact $\boldsymbol{\delta}\left(\mathbf{M}-\boldsymbol{\delta}\right)^{-1} = -\,\mathrm{Id}_{|\mathcal{N}|} + \mathbf{M}\left(\mathbf{M}-\boldsymbol{\delta}\right)^{-1}$, and the analogous relation for $\boldsymbol{\gamma}\left(\mathbf{M}-\boldsymbol{\gamma}\right)^{-1}$, we may rewrite the term between parentheses as

$$\begin{aligned}&\left(\mathbf{M}\left(\mathbf{M}-\boldsymbol{\delta}\right)^{-1} - \mathrm{Id}_{|\mathcal{N}|}\right)\left(\mathbf{M}\left(\mathbf{M}-\boldsymbol{\gamma}\right)^{-1} - \mathrm{Id}_{|\mathcal{N}|}\right)E^{\mathbf{M}}(0) \\ &\qquad + \left(\mathrm{Id}_{|\mathcal{N}|} - \mathbf{M}\left(\mathbf{M}-\boldsymbol{\delta}\right)^{-1}\right)I^{\mathbf{M}}(0).\end{aligned}$$

Since $\mathbf{M}$ is a diffusion matrix, we know that the coordinates of the vectors $\mathbf{M}\left(\mathbf{M}-\boldsymbol{\delta}\right)^{-1}E^{\mathbf{M}}(0)$, $\mathbf{M}\left(\mathbf{M}-\boldsymbol{\gamma}\right)^{-1}E^{\mathbf{M}}(0)$, $\mathbf{M}\left(\mathbf{M}-\boldsymbol{\delta}\right)^{-1}\mathbf{M}\left(\mathbf{M}-\boldsymbol{\gamma}\right)^{-1}E^{\mathbf{M}}(0)$ and $\mathbf{M}\left(\mathbf{M}-\boldsymbol{\delta}\right)^{-1}I^{\mathbf{M}}(0)$ sum to zero. Hence, we finally obtain

$$\begin{aligned}&(1-(1-\varepsilon)\,\mathcal{R}_0)^{-1}\,\boldsymbol{e}^T\left(E^{\mathbf{M}}(0)+I^{\mathbf{M}}(0)\right) \leq \Delta^{\mathbf{M}} \\ &\qquad\leq (1-(1+\varepsilon)\,\mathcal{R}_0)^{-1}\,\boldsymbol{e}^T\left(E^{\mathbf{M}}(0)+I^{\mathbf{M}}(0)\right).\end{aligned}$$

Expressing the scalar products with the vector $\boldsymbol{e}$ as a sum, and remembering $\mathcal{R}_0 = \mathcal{R}_0(\mathbf{M})$, we obtain the expression given in the statement of Theorem 3.

□

References

[1] H. Youn, M. T. Gastner, and H. Jeong, "Price of anarchy in transportation networks: Efficiency and optimality control," *Physical Review Letters*, vol. 101, no. 12, p. 128701, 9 2008.

[2] R. Pastor-Satorras and A. Vespignani, *Internet: structure et évolution*. Belin, 2004.

[3] M. E. J. Newman, D. J. Watts, and S. H. Strogatz, "Random graphs of social networks," *Proc. Natl. Acad. Sci. USA*, vol. 99, no. 1, pp. 2566–2572, 2002.

[4] D. Perfido, M. Raciti, C. Zanotti, N. Chambers, L. Hannon, M. Keane, E. Clifford, and A. Costa, "Towards sustainable water networks: Automated fault detection and diagnosis," *The International Journal of Entrepreneurship and Sustainability numbers*, vol. 4, no. 3, pp. 339–350, 3 2017.

[5] P. Mieghem and J. Omic, "In-homogeneous virus spread in networks," *arXiv:1306.2588*, 06 2013.

[6] M. Garetto, W. Gong, and D. Towsley, "Modeling malware spreading dynamics," in *INFOCOM Joint Conference of the IEEE Computer and Communications*, vol. 3, 2003, pp. 1869–1879.

[7] C. Nowzari, V. M. Preciado, and G. J. Pappas, "Analysis and control of epidemics: A survey of spreading processes on complex networks," *IEEE Control Systems Magazine*, vol. 36, no. 1, pp. 26–46, 2016.

[8] W. O. Kermack and A. McKendrick, "A contribution to the mathematical theory of epidemics," *Journal of The Royal Society*, vol. 115, no. 772, pp. 700–721, 1927.

[9] O. Diekmann, H. Heesterbeek, and T. Britton, *Mathematical Tools for Understanding Infectious Disease Dynamics*, ser. Princeton series in theoretical and computational biology. Princeton University Press, 2012.

[10] R. Pastor-Satorras and A. Vespignani, "Epidemic spreading in scale-free networks," *Physical Review Letter*, vol. 86, no. 14, pp. 3200–3203, 2001.

[11] R. Pastor-Satorra, C. Castellano, P. Van Mieghem, and A. Vespignani, "Epidemic processes in complex networks," *Review of Modern Physics*, vol. 87, no. 3, pp. 925–979, 2015.

[12] J. Arino, "Diseases in metapopulations," in *Modeling and Dynamics of Infectious Diseases*, 2009, pp. 64–122.

[13] R. Levins, "Some demographic and genetic consequences of environmental heterogeneity for biological control," *Bulletin of the Entomological Society of America*, vol. 3, no. 15, pp. 237–240, 1969.

[14] P. Van den Driessche and J. Watmough, "Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission," *Mathematical Biosciences*, vol. 180, no. 1, pp. 29–48, 2002.

[15] L. J. S. Allen, B. M. Bolker, Y. You, and A. L. Nevai, "Asymptotic profiles of the steady states for an sis epidemic patch model," *SIAM Journal of Applied Mathematics*, vol. 67, no. 5, pp. 1283–1309, 2007.

[16] J. H. Tien, Z. Shuai, M. C. Eisenberg, and P. van den Driessche, "Disease invasion on community networks with environmental pathogen movement," *Journal of Mathematical Biology*, vol. 70, no. 5, pp. 1065–1092, 2015.

[17] J. Arino, "Spatio-temporal spread of infectious pathogens of humans," *Infectious Disease Modelling*, vol. 2, no. 2, pp. 218–228, 2017.

[18] D. Gao, "Travel frequency and infectious diseases," *SIAM Journal of Applied Mathematics*, vol. 79, no. 4, pp. 1581–1606, 2019.

[19] V. Colizza and A. Vespignani, "Invasion threshold in heterogeneous metapopulation networks," *Phys. Rev. Lett.*, vol. 99, no. 14, p. 148701, Oct 2007.

[20] ——, "Epidemic modeling in metapopulation systems with heterogeneous coupling pattern: Theory and simulations," *Journal of Theoretical Biology*, vol. 251, no. 3, pp. 450–467, 2008.

[21] E. Gourdin, J. Omic, and P. Van Mieghem, "Optimization of network protection agains virus spread," in *8th International Workshop on the Design of Reliable Communication Networks*, 2011, pp. 86–93.

[22] V. M. Preciado, M. Zargham, C. Enyioha, A. Jadbabaie, and G. J. Pappas, "Optimal vaccine allocation to control epidemic outbreaks in arbitrary networks," in *IEEE Conference on Decision and Control*, 2013, pp. 7486–7491.

[23] ——, "Optimal resource allocation for network protection: A geometric programming approach," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 1, pp. 99–108, 2014.

[24] C. Nowzari, V. M. Preciado, and G. J. Pappas, "Optimal resource allocation for control of networked epidemic models," *IEEE Transactions on Control of Network Systems*, vol. 4, no. 2, pp. 159–169, 6 2017.

[25] V. M. Preciado and M. Zargham, "Traffic optimization to control epidemic outbreaks in metapopulation models," in *2013 IEEE Global Conference on Signal and Information Processing*, 2013, pp. 847–850.

[26] M. Umar B. Niazi, C. Canudas de Wit, A. Kibangou, and P.-A. Bliman, "Optimal Control of Urban Human Mobility for Epidemic Mitigation," in *CDC 2021 – 60th IEEE Conference on Decision and Control*, 12 2021.

[27] G. Macdonald, "The analysis of equilibrium in malaria," *Trop. Dis. Bull.*, vol. 49, no. 9, pp. 813–829, sep 1952.

[28] O. Diekmann, J. Metz, and J. Heesterbeek, "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*, vol. 28, no. 4, pp. 365–382, 1990.

[29] O. Diekmann, J. A. P. Heesterbeek, and M. G. Roberts, "The construction of next-generation matrices for compartmental epidemic models," *Journal of The Royal Society*, vol. 7, no. 47, pp. 873–885, 2009.

[30] L. Zino and M. Cao, "Analysis, Prediction, and Control of Epidemics: A Survey from Scalar to Dynamic Network Models," *IEEE Circuits and Systems Magazine*, vol. 21, no. 4, pp. 4–23, 2021.

[31] V. M. Preciado and A. Jadbabaie, "Spectral analysis of virus spreading in random geometric networks," in *IEEE Conference on Decision and Control*, 2009, pp. 4802–4807.

[32] D. Knipl, "A new approach for designing disease intervention strategies in metapopulation models," *Journal of Biological Dynamics*, vol. 10, no. 1, pp. 71–94, 2016, pMID: 26561360.

[33] S. Ottaviano, F. De Pellegrini, S. Bonaccorsi, and P. Van Mieghem, "Optimal curing policy for epidemic spreading over a community network with heterogeneous population," *Journal of Complex Networks*, vol. 6, no. 5, pp. 800–829, 2018.

[34] P. E. Paré, C. L. Beck, and A. Nedić, "Epidemic Processes Over Time-Varying Networks," *IEEE Transactions on Control of Network Systems*, vol. 5, no. 3, pp. 1322–1334, 2018.

[35] C. Cenedese, L. Zino, M. Cucuzzella, and M. Cao, "Optimal policy design to mitigate epidemics on networks using an SIS model," *60th IEEE Conference on Decision and Control (CDC)*, pp. 4266–4271, 2021.

[36] S.-Y. Liu, A. Baronchelli, and N. Perra, "Contagion dynamics in time-varying metapopulation networks," *Phys. Rev. E*, vol. 87, no. 3, p. 032805, Mar 2013.

[37] C. Nowzari, M. Ogura, V. M. Preciado, and G. J. Pappas, "Optimal resource allocation for containing epidemics on time-varying networks," in *2015 49th Asilomar Conference on Signals, Systems and Computers*, 2015, pp. 1333–1337.

[38] M. Junling and J. E. David, "Generality of the final size formula for an epidemic of a newly invading infectious disease," *Bulletin of Mathematical Biology*, vol. 68, no. 3, pp. 679–702, 2006.

[39] J. Arino, F. Brauer, J. Watmough, and J. Wu, "A final size relation for epidemic models," *Mathematical biosciences and engineering*, vol. 4, no. 2, pp. 159–175, 2007.

[40] P. Magal, O. Seydi, and G. Webb, "Final size of an epidemic for a two-group SIR model," *SIAM Journal of Applied Mathematics*, vol. 76, no. 5, pp. 2042–2059, 2016.

[41] ——, "Final size of a multi-group sir epidemic model: Irreducible and non-irreducible modes of transmission," *Mathematical Biosciences*, vol. 301, pp. 59–67, 2018.

[42] V. Andreasen, "The final size of an epidemic and its relation to the basic reproduction number," *Bulletin of Mathematical Biology*, vol. 73, no. 10, pp. 2305–2321, 2011.

[43] F. Brauer, "Epidemic models with heterogeneous mixing and treatment," *Bulletin of Mathematical Biology*, vol. 70, no. 7, pp. 1869–1885, oct 2008.

[44] L. Almeida, P.-A. Bliman, G. Nadin, B. Perthame, and N. Vauchelet, "Final size and convergence rate for an epidemic in heterogeneous population," *Mathematical Models and Methods in Applied Sciences, World Scientific Publishing*, vol. 31, no. 5, pp. 1021–1051, 2021.

[45] D. Gao, "How does dispersal affect the infection size," *SIAM Journal of Applied Mathematics*, vol. 80, no. 5, pp. 2144–2169, 2020.

[46] J. Li, D. Blakeley, and R. J. Smith, "The failure of R_0," *Comput. Math. Methods Med.*, vol. 2011, p. 527610, aug 2011.

[47] J. Arino, J. R. Davis, D. Hartley, R. Jordan, J. M. Miller, and P. Van Den Driessche, "A multi-species epidemic model with spatial dynamics," *Mathematical Medicine and Biology*, vol. 22, no. 2, pp. 129–142, 2005.

[48] L.-B. Beaufort, P.-Y. Massé, A. Reboulet, and L. Oudre, "Network reconstruction problem for an epidemic reaction-diffusion," *arxiv preprint*, 2021.

[49] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*. Society for Industrial and Applied Mathematics Philadelphia, PA, USA, 2000.

[50] H. Caswell, *Sensitivity Analysis: Matrix Methods in Demography and Ecology*. Springer Open, 2019.

[51] G. H. Golub and C. D. Meyer, "Using the qr factorization and group inversion to compute, differentiate, and estimate the sensitivity of stationary probabilities for markov chains," *SIAM Journal on Algebraic Discrete Methods*, vol. 7, no. 2, pp. 273–281, 1986.

[52] B. T. Polyak, *Introduction to Optimization*, ser. Translations series in mathematics and engineering. Optimization Software - Inc., Publications Division, 1987.

[53] S. K. Zavriev and F. V. Kostyuk, "Heavy-ball method in nonconvex optimization problems," *Computational Mathematics and Modeling*, vol. 4, p. 336–341, 1993.

[54] P. Ochs, Y. Chen, T. Brox, and T. Pock, "ipiano: Inertial proximal algorithm for nonconvex optimization," *SIAM Journal on Imaging Sciences*, vol. 7, no. 2, pp. 1388–1419, 2014.

[55] T. P. Peixoto, "The graph-tool python library," *figshare*, 2014.

[56] G. Sallet, "Mathematical epidemiology," *Lecture Notes from Pretoria*, 2018.

[57] H. L. Smith and P. Waltman, *The Theory of the Chemostat – Dynamics of Microbial Competition*. Cambridge University Press, 1995.