

Optimal feedback policies in stochastic epidemic models

2024 IEEE 63rd Conference on Decision and Control, CDC 2024

Carratelli, Giovanni Pugliese; Cheng, Xiaodong; Parag, Kris V.; Lestas, Ioannis

<https://doi.org/10.1109/CDC56724.2024.10886584>

This publication is made publicly available in the institutional repository of Wageningen University and Research, under the terms of article 25fa of the Dutch Copyright Act, also known as the Amendment Taverne.

Article 25fa states that the author of a short scientific work funded either wholly or partially by Dutch public funds is entitled to make that work publicly available for no consideration following a reasonable period of time after the work was first published, provided that clear reference is made to the source of the first publication of the work.

This publication is distributed using the principles as determined in the Association of Universities in the Netherlands (VSNU) 'Article 25fa implementation' project. According to these principles research outputs of researchers employed by Dutch Universities that comply with the legal requirements of Article 25fa of the Dutch Copyright Act are distributed online and free of cost or other barriers in institutional repositories. Research outputs are distributed six months after their first online publication in the original published version and with proper attribution to the source of the original publication.

You are permitted to download and use the publication for personal purposes. All rights remain with the author(s) and / or copyright owner(s) of this work. Any use of the publication or parts of it other than authorised under article 25fa of the Dutch Copyright act is prohibited. Wageningen University & Research and the author(s) of this publication shall not be held responsible or liable for any damages resulting from your (re)use of this publication.

For questions regarding the public availability of this publication please contact openaccess.library@wur.nl

Optimal feedback policies in stochastic epidemic models

Giovanni Pugliese Carratelli, Xiaodong Cheng, Kris V. Parag, Ioannis Lestas

Abstract— We consider the problem of finding optimal policies that mitigate the effects of an epidemic. We develop computational tools for finding such policies for broad classes of stochastic epidemic models and investigate various features of such policies. In particular, we observe that optimal policies are predominantly constant for epidemics where the mitigation measures are associated with the infected population.

I. INTRODUCTION

Approximate deterministic models are frequently adopted in studies of epidemic control. Recent studies include the computation of the policies for initiating and exiting lock down interventions [1], [2], [3], [4], the analysis of the time sensitivity of their implementation [5], [6] and even cost-benefit analysis studies [7], [8]. Stochastic models, that take into account the randomness of the spread [9], are often employed to evaluate comprehensible intuition based policies [10] but are however seldomly used to devise policy design principles and limitations.

In [11], [12] time delays and nonlinearities in the infections are shown to be a limitation for negative feedback methods. The controllability of an emerging infectious disease is also quantified in [13]. The results associated with negative feedback policies for stochastic models of the spread [14], [15] are however more limited despite their practical significance. Studies on the optimal management of an epidemic via integer valued stochastic models are addressed via computational methods in [16], [17] and [18] while in [19] trade offs between socialising and risk of infection are studied on network Markov models.

In this study we try to investigate the effectiveness of real time interventions for stochastic epidemic models. In particular, and we find via computations of optimal policies for models with low levels of recoveries and mitigation measures associated with the infected population that constant interventions are often optimal in limiting the spread of a disease. The contribution of this paper is two fold:

- We develop computational tools for finding the optimal policies for epidemic models with arbitrary number of states modelled as a continuous time Markov Jump Process (MJP) with controlled rates.
- We consider the continuous time stochastic Susceptible Infected (SI) epidemic with a controlled transmission rate and we find that optimal policies are predominantly

G. Pugliese Carratelli and I. Lestas are with the Department of Engineering, University of Cambridge, United Kingdom; K.V. Parag is with the Department of Infectious Disease Epidemiology, Imperial College London, United Kingdom; X. Cheng is with the Department of Plant Science, Wageningen University, Netherlands. {gp459, icl20}@cam.ac.uk, k.parag@imperial.ac.uk, xiaodong.cheng@wur.nl.

constant for costs that depend linearly with the number of infected individuals but can otherwise depend non-linearly with the mitigation measures.

The manuscript is organised as follows. In Section II we introduce the notation and the epidemic model that will be considered. In Section III we define the problem we seek to address. Our main results are stated in Section IV.

II. SYSTEM MODEL AND NOTATION

A. Notation

$\mathbb{E}[\cdot]$	Expectation operator
$\mathbb{R}_{>}$	Set of positive real numbers $\{x \in \mathbb{R} : x > 0\}$
\mathbb{R}_{\geq}	Set of non-negative real numbers $\{x \in \mathbb{R} : x \geq 0\}$
$\mathbb{Z}_{>}$	Set of positive integers $\{1, 2, 3, \dots\}$
\mathbb{Z}_{\geq}	Set of non-negative integers $\{0, 1, 2, 3, \dots\}$

B. Mathematical preliminaries

We develop computational tools for finding the optimal policies that apply to broad classes of Markov Jump Process (MJP) models for epidemics. These are then used to investigate the properties of such policies for classes of optimal control problems formulated for stochastic SI models.

In particular, we consider the MJP version of the Susceptible Infected (SI) transmission model for an epidemic. The state of the system is the number of individuals in the susceptible and infected categories which evolves via infection, and recovery events occurring at certain rates. These events take place at random times that depend upon the state of the system and its parameters and are summarised in Table I.

This system is a birth-death like process that can be defined via the Kolmogorov Equation (also known as the Master equation). This is a partial difference equation for the probability at time t that the number of individuals in each category takes prescribed values. The Kolmogorov equation for the considered epidemic is stated below, but before that we introduce the variables describing the system and their physical interpretation.

More precisely, we consider a constant population of $N \in \mathbb{Z}_{>}$ individuals and the integer valued random variables $S(t)$ and $I(t)$, each defined on the support $\mathbb{S} = \{0, 1, \dots, N\}$. $S(t)$ and $I(t)$ denote respectively the number of Susceptible and Infected individuals at time t . We denote the infection prevalence with $I(t)/N$, and we denote the incidence of the disease with $S(t)/N$.

The state of the epidemic evolves via infection and recovery events occurring at the rates defined in rows 1a and 1b of Table I.

ID	Event	Transition	Rate
1a	SI Infection	$I \rightarrow I + 1$	$\mu\beta_f(I, u)$
1b	SI Recovery	$I \rightarrow I - 1$	γI

TABLE I: Events associated with the Susceptible Infected (SI) epidemic

The rate at which infected individuals recover, corresponding to event 1b in Table I, is $\gamma I(t)$ where $\gamma \in \mathbb{R}_{>}$. For each infection event a single individual transitions from the Susceptible to the Infected category. Infections occur at a rate of $\mu\beta_f(I, u)$ where $\mu \in \mathbb{R}_{>}$ is the transmission parameter and β_f is defined below in equation (2). We highlight that due to the proportional nature of the infection rate with respect to I the systems have an absorbing state at $I = 0$.

The variable $u \in \mathcal{U}$ modulates the transmission rate. The set \mathcal{U} is the set of transmission mitigation measures that may be adopted to tame the disease transmission. \mathcal{U} is a set of cardinality $N_u \in \mathbb{Z}_{>}$, *i.e.* $\mathcal{U} = \{u_1, u_2, \dots, u_{N_u}\}$. Note that in the following sections we address optimal control problems where we optimise over all possible values of \mathcal{U} .

The function $h(u) : \mathcal{U} \rightarrow [0, 1]$ reflects how a specific mitigation measure u reduces the transmission of the disease. We consider for notational purposes and to simplify the interpretation of the results, the following assumption

Assumption 1: The function $h(u) : \mathcal{U} \rightarrow [0, 1]$ is a non-increasing function.

Assumption 1 formalises that increasing values of u , that are associated with increasingly restrictive mitigation measures, reduce $h(u)$ and thus reduce the transmission of the disease.

We consider a constant population and hence we have $S(t) + I(t) = N$. Therefore $S(t) = N - I(t)$ and without any loss of generality we consider only the random variable $I(t)$. For each $i \in \mathbb{S}$ we denote by $\mathbb{P}(i, t)$ the probability that at time t we have $I(t) = i$.

The considered MJP satisfies the following Kolmogorov Equation

$$\frac{\partial \mathbb{P}(i, t)}{\partial t} = \mu [\beta_f(i-1, u)\mathbb{P}(i-1, t) - \beta_f(i, u)\mathbb{P}(i, t)] + \gamma [(i+1)\mathbb{P}(i+1, t) - i\mathbb{P}(i, t)] \quad (1)$$

where $\beta_f(i, u)$ is defined as

$$\beta_f(i, u) = \frac{i}{N}(N-i)h(u) \quad (2)$$

This is a well suited model for diseases characterised by weak levels of immunity after a recovery or at the initial stages of an epidemic when the number of infected is small with respect to the size of the population.

III. OPTIMAL CONTROL PROBLEM FORMULATION AND CONSIDERED COSTS

We consider the problem of finding an optimal feedback disease mitigation policy. We formulate a stochastic optimal control problem for the SI epidemic and we outline in Section III-A the considered costs.

We seek to minimise social costs due to the presence of i infected and due to the adoption of the transmission

mitigation measure u . In the considered problem we minimise the total expected costs in continuous time over an infinite horizon, where the stage cost per unit of time is $g_c(i, u) : \mathbb{S} \times \mathcal{U} \rightarrow \mathbb{R}_{\geq}$. The function g_c , discussed below, is the cost of applying the mitigation measure u with i infected individuals.

We consider deterministic feedback policies that are a function of the current state of the process. This is without loss of generality due to the Markov nature of the system, *i.e.* state feedback policies would be optimal even if the feedback policy was allowed to depend on the history of the process [20],[21]. We consider without any loss of generality only time invariant feedback laws since it can be shown that the policies for an infinite horizon problem are independent of time [22], [21]. A particular policy for the system in (1) is denoted by $u(i)$, $u : \mathbb{S} \rightarrow \mathcal{U}$. With \mathcal{F} we denote the set of time invariant state feedback policies, *i.e.* the set of functions u which for each value of the state give an element in \mathcal{U} .

Problem 1: Consider the system in (1) and the stage cost $g_c(i, u) : \mathbb{S} \times \mathcal{U} \rightarrow \mathbb{R}_{\geq}$. We consider the following problem

$$J^c(i_0) = \min_{u \in \mathcal{F}} J_u^c(i_0) \quad (3)$$

$J_u^c(i_0)$, defined in (4), is the cost associated with the evolution of the MJP in (1) from the initial condition i_0 under a time invariant state feedback policy u .

$$J_u^c(i_0) = \lim_{T_{N_t} \rightarrow \infty} \mathbb{E} \left[\int_0^{T_{N_t}} g_c(i(t), u) dt \middle| I(0) = i_0 \right] \quad (4)$$

As previously discussed the considered continuous time process has an absorbing state and terminates in finite time. It can be shown that the considered continuous time optimal control problem can be transformed to that of an appropriately constructed discrete time Markov Decision Process (MDP). This equivalent formulation is presented in Section IV-A.

The solution of Problem 1 corresponds to finding an optimal policy $u^*(i)$ that minimises a socio-economic cost $g_c(i, u)$ related to an epidemic. In particular, we associate for each infected individual a cost that society has to incur and a costs for interventions that reduce the transmission of the disease.

A. Considered costs

In this section we describe the class of cost functions that will be considered, and we show in the next section that for these costs the optimal policies do not vary with respect to the infection levels.

In particular, we consider epidemics where infected individuals may require medical treatment or may be subjected to isolation and quarantine thus incurring a cost to the economy/society. The transmission mitigation interventions u encompass a broad class of such measures for infected individuals including social distancing, medical isolation and quarantine. Such measures may require for example individuals infected not to work or they may be able to do so only in a limited manner.

Specifically, we consider cost functions g_c of the form

$$g_c(i, u) = c_1 + c_2 iz(u) + c_3 i \quad (5)$$

where $z(u) : \mathcal{U} \rightarrow \mathbb{R}_{\geq}$ is an arbitrary function with non-negative values. The term $c_1 \in \mathbb{R}_{\geq}$ is a fixed cost to be accounted for regardless of the number of infected or the adopted mitigation measures. The proportionality between i and $z(u)$ suggests that a cost of $c_2 z(u)$, $c_2 \in \mathbb{R}_{\geq}$ is incurred for each infected for the mitigation measure u . The term $c_3 i$, $c_3 \in \mathbb{R}_{\geq}$ relates to the sole presence of infected individuals and is linked to fatalities and hospitalisations when the latter are well below their capacity. The parameter c_3 depends, for example, upon the infected to fatality ratio or the infected to hospitalisation ratio. Note that without loss of generality the cost in (5) can be rescaled with respect to the sole hospitalisation cost parameter c_3 . We define the ratio

$$\zeta = c_2/c_3 \quad (6)$$

that is the ratio of marginal cost c_2 of using mitigation measures, with respect to the marginal costs for hospitalisations, diagnosis and fatalities c_3 .

Costs of the form in (5) apply for example to influenza like epidemics [16] or more broadly to regimes where the efforts to control the epidemics through isolation measures of infected individuals are applicable. It excludes non-selective policies where the entire population is subjected to restrictions such as lock down.

IV. RESULTS

In Section IV-A we develop computational tools for characterising the optimal policies that are applicable to a broad class of MJP epidemic models with an arbitrary number of states. In Section IV-B these tools are then used to characterise the optimal policies for the SI epidemic in Problem 1 and for the considered costs.

A. Formulation as an equivalent discrete time Markov Decision Process (MDP)

In order to characterise the optimal policies of Problem 1 we derive first an equivalent discrete time optimal control problem, stated below as Problem 2, via the method of *uniformisation* [22]. We consider a discrete time Markov Chain I_k that takes values in \mathbb{S} where k is the time step. The transition probabilities p_{ij}^u are given by

$$p_{ij}^u = \mathbb{P}(I_{k+1} = j | I_k = i) = \begin{cases} \frac{1}{\nu} \beta_f(i, u) & \text{if } j = i + 1 \\ \frac{1}{\nu} \gamma i & \text{if } j = i - 1 \\ 1 - \frac{\beta_f(i, u) + \gamma i}{\nu} & i = j \end{cases} \quad (7)$$

where $i, j \in \mathbb{S}$ and the transmission rate $\beta_f(i, u)$, is defined in (2). The constant $\nu \in \mathbb{R}_{>}$ is

$$\nu \geq \max_{i, u} [\mu \beta_f(i, u) + \gamma i] \quad (8)$$

An equivalent discrete time Problem to Problem 1 is defined below.

Problem 2: Consider the discrete time Markov Chain specified by the transition probabilities in (7). We seek to solve the following problem

$$J(i_0) = \min_{u \in \mathcal{F}} J_u(i_0) \quad (9)$$

where \mathcal{F} is the set of time invariant state feedback policies $u(i)$, $u : \mathbb{S} \rightarrow \mathcal{U}$. $J_u(i_0)$, defined in (10), is the cost associated with the evolution of the discrete time Markov Chain with transition probabilities as in (7) from the initial condition i_0 under the time invariant state feedback policy u .

$$J_u(i_0) = \lim_{N_d \rightarrow +\infty} \sum_{k=0}^{N_d-1} \mathbb{E}[g(I_k, u) | I_0 = i_0] \quad (10)$$

The control input $u(i)$ associated with Problem 1 only changes at the times in which the state changes. In order to solve Problem 1 it is thus sufficient to consider the values of the policies when state transitions take place. This allows to reduce Problem 1 to Problem 2, which is based on a discrete time formulation, as stated in Proposition 1 below.

Proposition 1: Consider Problem 1 and consider $\nu \in \mathbb{R}_{>}$ as in (8) then the optimal policy for Problem 2 with cost

$$g(i, u) = \frac{1}{\nu} g_c(i, u) \quad (11)$$

is also the optimal policy for Problem 1 where the cost is $g_c(i, u)$.

The proposition above follows from a procedure known as *uniformisation* described in e.g. [22, Vol. II, Ch. 5, p. 288] and [23, Ch. 9, p. 503].

Remark 1: It should be noted that analogous versions of Proposition 1 discussed above, hold more broadly for other MJP models for epidemics with any number of states whereby these can be converted to an equivalent discrete time optimal control problem.

Remark 2 (Computation of optimal policies): The transformation to a discrete time MDP with finite state and input values and an absorbing state allows to make use of appropriate iterations, such as the value iteration or policy iteration, to compute the optimal policy with arbitrarily high precision [22]. The former was used for the computation of the optimal policies presented in Section IV.

In order to compute the optimal policies we make use of the Bellman equation for optimality and the value iteration algorithm, which is an iteration that is guaranteed to converge to the minimum cost [24], [22].

The value iteration algorithm for Problem 2 is

$$J_{k+1}(i) = \frac{1}{\nu} \min_{u \in \mathcal{U}} [g_c(i, u) + \mu \beta_f(i, u) L_k(i) - \gamma i L_k(i-1) + \nu J_k(i)] \quad (12)$$

where we have introduced the quantity $L_k(i)$ defined¹ as

$$L_k(i) = J_k(i+1) - J_k(i) \quad (13)$$

¹We set $\gamma i L_k(i-1) = 0$ for $i = 0$ and $\beta_f(i, u) L_k(i) = 0$ for $i = N$.

and for simplicity in the notation we remove the subscript from i_0 in the rest of the manuscript.

We denote with $\mathcal{J}_k(i, u)$ the function to be minimised with respect to u at each time step in the right hand side of (12), *i.e.*

$$\mathcal{J}_k(i, u) = g_c(i, u) + \mu\beta_f(i, u)L_k(i) - \gamma iL_k(i-1) + \nu J_k(i) \quad (14)$$

and with $u_k^*(i)$ we denote the optimal policy for state i at step k of the value iteration algorithm.

It should be noted that Problem 2 is an infinite horizon optimal control problem with an absorbing state, positive costs, and with finite cardinality N_u of the mitigation measures \mathcal{U} . As stated in the following Lemma [22] (12) is known to converge to the minimum cost $J(i)$, which is the solution of the Bellman equation (15).

Lemma 1: Consider Problem 2. Then for any initial condition $J_0(i)$ the value iteration algorithm in (12) converges to the optimal cost $J(i)$ which is the unique solution to the Bellman equation

$$J(i) = \frac{1}{\nu} \min_{u \in \mathcal{U}} [g_c(i, u) + \mu\beta_f(i, u)L(i) - \gamma iL(i-1) + \nu J(i)] \quad (15)$$

where $L(i)$ denotes the quantity.

$$L(i) = J(i+1) - J(i) \quad (16)$$

Furthermore, a policy $u^*(i)$ is optimal if and only if it solves (15).

B. Properties of the optimal solution of Problem 1

We make use of the computational tools developed in Section IV-A and in Fig.1 we provide two numerical examples of the optimal policies for the SI epidemic in Problem 1 with costs as in (5). We observe that the optimal policy $u^*(i)$ is predominantly constant and non-increasing in the variable i (or equivalently with prevalence i/N). It follows that the controlled transmission rate $\beta_f(i, u^*(i))$ is non-decreasing with the number of infected. It should be noted that these properties appear to hold for a large range of system and cost parameters such as for different values of ζ , different functions $z(u) : \mathcal{U} \rightarrow \mathbb{R}_{\geq}$ and $h(u) : \mathcal{U} \rightarrow [0, 1]$ and for different values of the recovery rate $\gamma \in \mathbb{R}_{>}$. Furthermore, the set \mathcal{U} of control inputs and its cardinality N_u can also be chosen arbitrarily. Analytical results associated with the form of these policies, and the presence of these features in more broad classes of epidemic models and optimal control problems will be included in an upcoming more extended version of this work.

V. CONCLUSIONS

We have considered the problem of finding optimal mitigation policies in stochastic epidemic models, with a focus on epidemics where the mitigation measures are associated with the infected population. We have developed tools for computing such policies for a broad class of epidemic models and have investigated the optimal policies for the SI model. In particular, we have observed that when the mitigation

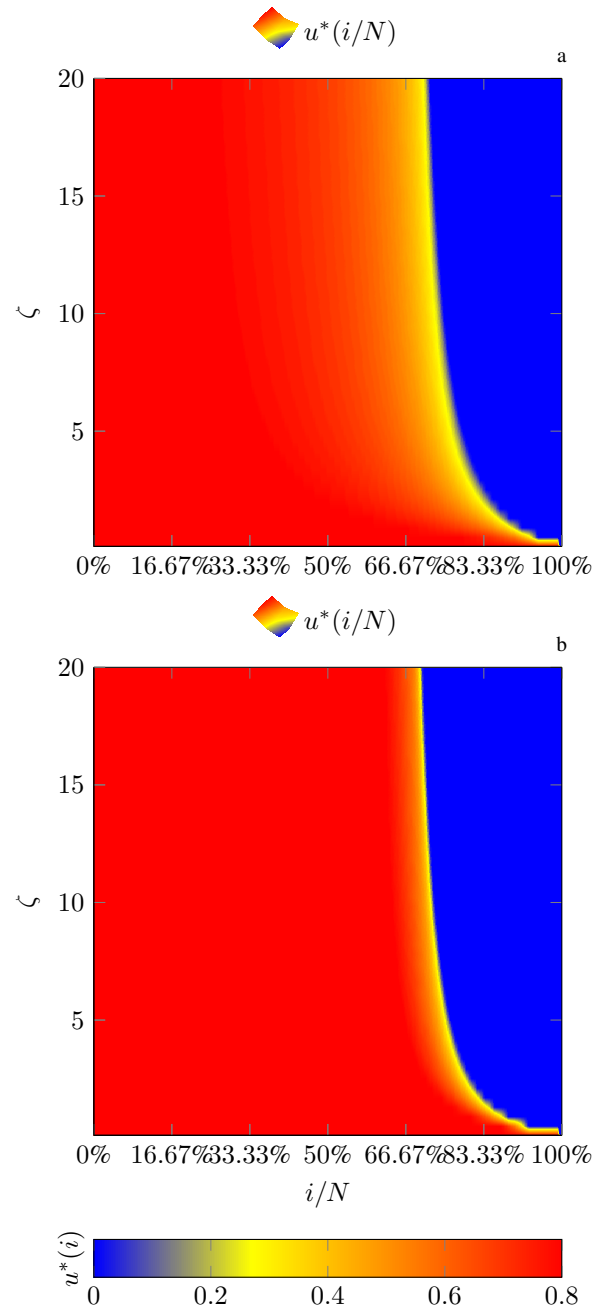


Fig. 1: Diagram 1a and 1b display the optimal policies $u^*(i)$. High values of u^* are associated with restrictive interventions and low values are associated with less restrictive interventions. The policies are non-increasing for increasing levels of prevalence and are predominantly constant. We select $h = (1 - q_{eff}u)^2$, which has been found to work well in practice [25]. Diagram 1a shows the policies for $q_{eff} = 1$ and diagram 1b shows the policies for $q_{eff} = 0.65$. The parameters are $N = 300$, $\gamma = 0.32$, $\mathcal{R}_0 = \mu/\gamma = 3.5$, $\mathcal{U} = \{0, 0.05, \dots, 0.8\}$

measures are associated with the infected individuals optimal policies are predominantly constant.

REFERENCES

- [1] M. Bin, P. Y. K. Cheung, E. Crisostomi, P. Ferraro, H. Lhachemi, R. Murray-Smith, C. Myant, T. Parisini, R. Shorten, S. Stein, and

- L. Stone, "Post-lockdown abatement of COVID-19 by fast periodic switching," *PLOS Computational Biology*, vol. 17, no. 1, jan 2021.
- [2] A. Kasis, S. Timotheou, N. Monshizadeh, and M. Polycarpou, "Optimal intervention strategies to mitigate the covid-19 pandemic effects," 2021.
- [3] A. Borri, P. Palumbo, F. Papa, and C. Possieri, "Optimal design of lock-down and reopening policies for early-stage epidemics through SIR-d models," *Annual Reviews in Control*, vol. 51, pp. 511–524, 2021.
- [4] C. Tsay, F. Lejarza, M. A. Stadtherr, and M. Baldea, "Modeling, state estimation, and optimal control for the US COVID-19 outbreak," *Scientific Reports*, vol. 10, no. 1, jul 2020.
- [5] D. H. Morris, F. W. Rossine, J. B. Plotkin, and S. A. Levin, "Optimal, near-optimal, and robust epidemic control," *Communications Physics*, vol. 4, no. 1, Apr. 2021.
- [6] F. Di Lauro, I. Z. Kiss, and J. C. Miller, "Optimal timing of one-shot interventions for epidemic control," *PLOS Computational Biology*, vol. 17, Mar. 2021.
- [7] R. Rowthorn and J. Maciejowski, "A cost–benefit analysis of the COVID-19 disease," *Oxford Review of Economic Policy*, vol. 36, no. Supplement_1, pp. S38–S55, 2020.
- [8] J. Maciejowski, R. Rowthorn, S. Sheffield, D. Vines, and A. Williamson, "Mitigation policy for the covid-19 pandemic: Intertemporal optimisation using an seir model," *SSRN Electronic Journal*, 2022.
- [9] N. A. Ruhi and B. Hassibi, "SIRS epidemics on complex networks: Concurrence of exact markov chain and approximated models," in *2015 54th IEEE Conference on Decision and Control (CDC)*. IEEE, dec 2015.
- [10] N. Ferguson, D. Laydon, G. Nedjati Gilani, N. Imai, K. Ainslie, M. Baguelin, S. Bhatia, A. Boonyasiri, Z. Cucunuba Perez, G. Cuomo-Dannenburg, A. Dighe, I. Dorigatti, H. Fu, K. Gaythorpe, W. Green, A. Hamlet, W. Hinsley, L. Okell, S. Van Elsland, H. Thompson, R. Verity, E. Volz, H. Wang, Y. Wang, P. Walker, P. Winskill, C. Whittaker, C. Donnelly, S. Riley, and A. Ghani, "Report 9: Impact of non-pharmaceutical interventions (npis) to reduce covid19 mortality and healthcare demand," 2020.
- [11] F. Casella, "Can the COVID-19 epidemic be controlled on the basis of daily test reports?" *IEEE Control Systems Letters*, vol. 5, no. 3, pp. 1079–1084, jul 2021.
- [12] R. Pates, A. Ferragut, E. Pivo, P. You, F. Paganini, and E. Mallada, "Respect the unstable: Delays and saturation in contact tracing for disease control," *SIAM Journal on Control and Optimization*, vol. 60, no. 2, pp. S196–S220, mar 2022.
- [13] K. V. Parag, "How to measure the controllability of an infectious disease?" Oct. 2023.
- [14] 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*, vol. 36, no. 1, pp. 26–46, feb 2016.
- [15] M. Mubarak, J. Berneburg, and C. Nowzari, "Stochastic vs. deterministic modeling for the spread of COVID-19 in small networks," in *2021 American Control Conference (ACC)*. IEEE, may 2021.
- [16] M. Ludkovski and J. Niemi, "Optimal dynamic policies for influenza management," *Statistical Communications in Infectious Diseases*, vol. 2, no. 1, jan 2010.
- [17] D. Merl, L. R. Johnson, R. B. Gramacy, and M. Mangel, "A statistical framework for the adaptive management of epidemiological interventions," *PLoS ONE*, vol. 4, jun 2009.
- [18] L. Ge, A. R. Kristensen, M. C. Mourits, and R. B. Huirme, "A new decision support framework for managing foot-and-mouth disease epidemics," *Annals of Operations Research*, vol. 219, no. 1, pp. 49–62, jul 2010.
- [19] M. Mubarak, J. Berneburg, and C. Nowzari, "Individual non-pharmaceutical intervention strategies for stochastic networked epidemics," in *2022 IEEE 61st Conference on Decision and Control (CDC)*. IEEE, dec 2022.
- [20] D. Blackwell, "Memoryless strategies in finite-stage dynamic programming," *The Annals of Mathematical Statistics*, vol. 35, no. 2, pp. 863–865, jun 1964.
- [21] M. L. Puterman, Ed., *Markov Decision Processes*. John Wiley & Sons, Inc., apr 1994.
- [22] D. Bertsekas, *Dynamic programming and optimal control*. Belmont, Mass: Athena Scientific, 2005.
- [23] S. L. Christos G. Cassandras, *Introduction to Discrete Event Systems*. Springer-Verlag GmbH, 2007.
- [24] R. Bellman, "On the theory of dynamic programming," *Proceedings of the National Academy of Sciences*, vol. 38, no. 8, pp. 716–719, aug 1952.
- [25] M. Barnett, G. Buchak, and C. Yannelis, "Epidemic responses under uncertainty," *Proceedings of the National Academy of Sciences*, vol. 120, no. 2, jan 2023.