Politecnico di Torino

ScuDo
Scuola di Dottorato ~ Doctoral School
WHAT YOU ARE, TAKES YOU FAR

Doctoral Dissertation

Doctoral Program in Electrical, Electronics and Communications Engineering

(36^{th}cycle)

Modeling Social Interactions in Complex Systems

A Journey from Opinion Formation to Epidemic Spread

By

Franco Galante

******

Supervisor(s):

Prof. Emilio Leonardi

Doctoral Examination Committee:

Prof. Sara Alouf, INRIA, *Referee*

Prof. Andrea Bianco, Politecnico di Torino

Prof. Paolo Giaccone, Politecnico di Torino

Prof. Anastasios Giovanidis, Ericsson Research, *Referee* & CNRS

Prof. Alberto Tarable, CNR

Politecnico di Torino

2024

Declaration

I hereby declare that, the contents and organization of this dissertation constitute my own original work and does not compromise in any way the rights of third parties, including those relating to the security of personal data.

Franco Galante
2024

* This dissertation is presented in partial fulfillment of the requirements for **Ph.D. degree** in the Graduate School of Politecnico di Torino (ScuDo).

To my beloved nonna,

Abstract

Our society is undergoing profound change, shaped by the rapid progression of technological advancements, which are not only redefining the dynamics of interaction, education, and labor but also influencing all other aspects of our lives. One of the driving forces behind the current social changes is the advent of online social networks (OSNs). They changed interpersonal interactions and are also redefining the way businesses, politicians, and media organizations engage with the wider public. A pressing challenge of our time is to understand how these changes affect our society and what potential dangers are associated with them. This thesis aims to provide a broad introduction to social modeling, highlight the connection between seemingly disparate domains, and introduce some of the fundamental tools of the *social modeler*. We will then focus on specific aspects of society, namely the mechanisms of opinion formation and virus spread in a population.

On online social platforms, a few individuals, commonly referred to as *influencers*, produce the majority of content consumed by users and hegemonize the landscape of the social debate. Traditional opinion models do not capture this asymmetry in communication. In Chapter 2, we develop an opinion model inspired by observations on social media platforms with two main objectives: first, describing this inherent communication asymmetry in OSNs, and second, modeling the effects of content personalization. We derive a Fokker-Planck equation for the temporal evolution of users' opinion distribution and analytically characterize the stationary system's behavior. Analytical results, confirmed by Monte-Carlo simulations, show how strict forms of content personalization tend to radicalize user opinion, leading to the emergence of *echo chambers* and favoring *structurally advantaged* influencers. As an example application, we apply our model to Facebook data during the Italian government crisis in 2019.

The new class of social media influencers plays a crucial role in shaping opinion. While we assumed in Chapter 2 that they have a fixed opinion on a topic, they actively compete for users' attention on social media. Through these targeted efforts, influencers seek to captivate users and build a loyal and engaged fan base, solidifying their position as an authoritative voice in the digital world and maximizing their impact on the population's opinion. The goal of Chapter 3 is twofold: first, we formalize the problem of maximizing social media impact and study the structure of the optimal solution. Then, taking inspiration from the optimal strategy, we develop a game with two opposing players trying to maximize their influence on users' opinions, for which we characterize the Nash equilibria in pure strategy. Although it may seem far-fetched, the dynamics of belief and epidemics are described in very similar terms.

In the last chapter of this thesis, we develop a SIR-like stratified model to better understand the options available to policymakers in the case of an outbreak of a new pandemic. We consider a scenario where efforts are made to control the infection and focus on two control strategies. The first aims to control the rate of new infections; the second directly controls hospitalizations and intensive care unit (ICU) occupancy. Using a first-order analysis, we show that controlling the transmission rate can be difficult due to a lack of accurate information. Furthermore, we observe that while hospitalizations and ICU occupancy are easily accessible and less noisy than the rate of new infections, a delay is introduced into the control loop that can endanger system stability. Our framework allows us to assess the joint impact of control strategies and vaccination campaigns on economic and social costs, taking into account: i) the heterogeneity of the population in terms of mortality rate and risk exposure, ii) the closed-loop control of the epidemiological curve, and iii) the progressive vaccination of individuals.

Contents

Chapter 1

Introduction and Motivation

The study of social phenomena has a long tradition, and its importance has increased significantly over the last decades. Today's society is characterized by a growing global interconnectedness facilitated by complex infrastructures that enable, among others, instant communication (e.g., the Internet) and fast and reliable means of transport (e.g., global airline networks). We have experienced drastic societal changes, partly due to technological advancements, such as unprecedented automation, a widespread resort to remote work, and sophisticated and accurate algorithms that permeate all facets of social dynamics. Furthermore, the recent introduction of groundbreaking technologies, such as generative Artificial Intelligence (AI) or Large Language Models (LLM), see ChatGPT and Bard, for example, has not only stunned the general public but also taken the scientific community by surprise due to their remarkable capabilities. These significant changes highlight the imperative for a collective effort in social modeling to anticipate future scenarios, uncover the structures and patterns of social behavior, and even quantify the risks associated with our "new reality".

1.1 The Importance of Modeling Social Phenomena

The broad field of social sciences includes but is not limited to sociology, psychology, and economics. It has the daunting task of exploring and understanding the complexity of societies and the interactions between individuals, cultures, and institutions. The ultimate goal is to find answers to humanity's most intimate

questions, to find regularity where chaos seems to reign, and to try to develop simple explanations for the most complex phenomena.

The social sciences are a melting pot for scholars, both because of the universal interest they attract and because of the great variety of tools required in this field of research. In a posthumous article by Ettore Majorana [1], the Italian physicist argued that the idea of abandoning determinism in the laws of nature makes "substantial the analogy between physics and social sciences" by recognizing the statistical nature of the underlying processes [1]. This article presents an interesting and informal discussion that starts from the merits of classical mechanics, which provides simple "exact laws" determined by the characterization of the system's state. In particular, the success of the classical approach to planetary motion has led many to believe that any complicated phenomenon should be reducible to a simple, similar mechanism. Majorana continues presenting the statistical view of mechanics, best illustrated by the statistical description of the ideal gas, which leads to the well-known equations of thermodynamics. He concludes by pointing out the statistical nature of the elementary principles themselves, as illustrated by the theories of quantum mechanics.

Over the last century, the efforts of scientists have intensified and many communities, notably also outside those traditionally dealing with social sciences, have contributed to the development of the field, including physicists (following in the footsteps of Majorana), mathematicians, and computer scientists. This trend has even intensified in recent years in response to a society undergoing unprecedented change in a short period, and also thanks to massive production and collection of behavioral data, extensively used in *data-driven* modeling approaches. The wealth of data and the resulting possibilities are far from fully explored, there are many issues associated with the use of data, first of all, its collection and availability, and then: data quality (i.e., incomplete or inaccurate data), data diversity, data protection, and the management of a large amount of data. In the context of social data, for example, many other questions arise in addition to the issue of trust, e.g.: "Who owns and controls data? [...] Are smart services inclusive?" [2]. In this thesis, data are used in the context of social interaction and epidemics to inform modeling decisions and strengthen the framework.

Data-driven and social modeling has experienced one of its biggest challenges in recent years, when the COVID-19 pandemic interested all national systems

simultaneously, forcing them to make predictions and informed policy decisions to control and contain the spread of the epidemic. The response to the pandemic was one of the few global collaborative social modeling efforts that temporarily set aside the ongoing friction between countries in favor of the collective benefit. Ideally, this would also be valuable in many other areas, such as healthcare, social media monitoring (e.g., limit the spread of fake news), and sustainable environmental practices.

1.2 The Toolkit for the Social Modeler

There are several tools used in social modeling. In this section, we will discuss some of them that will be the foundation for developing the contributions in this thesis. We will explore the theoretical concepts using examples of established models. First, we will discuss the Ising model, provide a *mean-field* solution for an arbitrary dimension, and then solve it exactly in one dimension. These two steps will allow us to introduce the mean-field approach and its inherent limitations (we will see that the qualitative behavior of the one-dimensional Ising model resulting from the mean-field does not match that of the exact solution). Second, we illustrate the link between the Ising model and opinion dynamics, the main topic of this thesis, by introducing the Sznajd model, one of the most popular models of opinion dynamics inspired by the Ising model. We will derive a partial differential equation describing the model's dynamics along the lines of the Fokker-Planck equation, which we will derive later in Chapter 2 for the Communication Asymmetry opinion model. Finally, we will discuss one of the most famous epidemic models: the Susceptible-Infected-Susceptible (SIS) model. We provide a stochastic formulation of the model, which naturally leads to an agent-based modeling framework. We consider its Discrete-Time Markov Chain (DTMC) formulation and derive a deterministic bound on the average number of infected individuals in the population. We will examine the link between this bound and the deterministic SIS model and discuss the differences between the two models. The interested reader may have a more formal and detailed introduction to the "tools for the social modeler" from the book "*Dynamical Processes on Complex Network*" [3].

In the upcoming sections, some concepts from physics will be needed, for which we will give an informal and intuitive definition since we are not interested in formally introducing them, but will use them as tools for our investigation. Any statistical mechanics textbook would provide further details to deepen the understanding of these concepts.

1.2.1 Mean-Field Solution of the Ising Model

The Ising model is a simple model for describing ferromagnetic materials, capturing the emergence of a *phase transition* in the material's behavior. In dimensions greater than 1, the Ising model has two distinct phases separated by a critical temperature T_c. The *paramagnetic* phase is a disordered phase in which thermal fluctuations dominate the spin orientation. Instead, the *ferromagnetic* phase corresponds to a spontaneous alignment of the spins in one direction.

The volume of the material is divided into elementary cells as in Fig. 1.1, which, in arbitrary dimension d, gives rise to a d-regular-lattice $\mathscr{L}$.

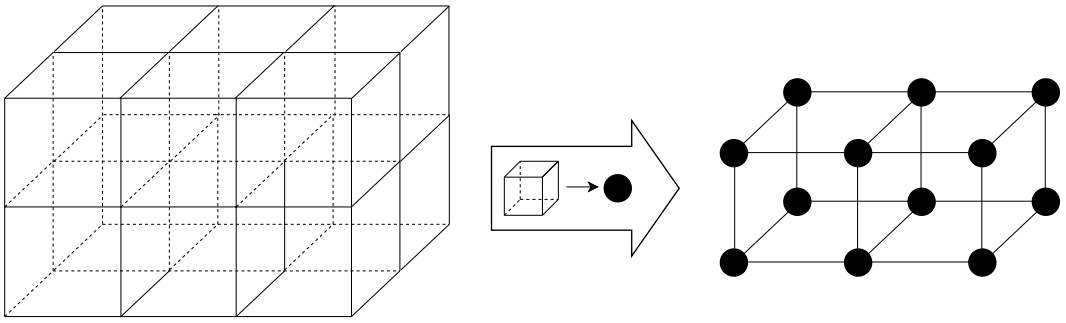


Fig. 1.1 Abstraction of the volume of the ferromagnetic material, each small volume corresponds to a node in the regular lattice.

Each site (node) in the lattice is characterized by a *spin* variable s_i, which describes the site's orientation of the *magnetic moment*. Hence, with respect to the external magnetic field h, s_i can assume one of two values:

$$s_i = \begin{cases} +1 \text{ parallel to } h \\ -1 \text{ antiparallel to } h \end{cases}$$

We must now define the Hamiltonian of the system. In physics, the Hamiltonian is a mathematical expression that represents the total energy of a system. In classical mechanics, for example, the Hamiltonian is defined as a function of the position and momentum of an object and is given by the sum of the kinetic energy and the potential energy of the object. In our case, considering N sites, the Hamiltonian for the Ising model can be written as follows:

$$\mathscr{H}(\boldsymbol{s}) = -J \sum_{\langle i,j \rangle} s_i s_j - h \sum_{i}^{N} s_i \tag{1.1}$$

where J is the ferromagnetic coupling, which we will consider constant for every pair of sites i, j, and with $\langle i, j \rangle$ we indicate all pairs of neighboring sites (considering each pair just once). Assuming the interactions have a strong short-range component and a rapidly decaying tail, the interaction happens only between the first neighbors. Lastly, h is the external magnetic field. We will follow the steps as in [4] and omit some details, the reader can refer back to [4] for details.

The *magnetization* m can be regarded as the empirical mean (over all sites) of the average orientation s_i. For the system, m serves as an *order parameter*, as it indicates how ordered the system is, since at $m = 0$ the spins are disordered, while whenever $m \neq 0$ the spins in the systems start to align. The magnetization is used to distinguish between the paramagnetic and the ferromagnetic phase. We can formally define it as follows:

$$m = \frac{1}{N} \sum_{i=1}^{N} \mathbb{E}[s_i] \tag{1.2}$$

where $\mathbb{E}[s_i]$ is the thermal expectation value of spin s_i, computed considering the *canonical ensemble*, i.e., the *statistical ensemble*[1] of a system that is in thermal equilibrium with a heat bath at a constant temperature T. When considering this system the probability $\mathbb{P}(\boldsymbol{s})$ of finding the system in a certain state $\boldsymbol{s} = (s_1, s_2, \ldots, s_N)^T$ (a possible configuration of the sites), corresponds to the Boltzmann distribution:

[1] A statistical ensemble is an abstraction of a system that considers a (large) "ensemble"of fictitious copies of a system, each of which represents a possible state the *real* system may be in.

$$\mathbb{P}(\boldsymbol{s}) = \frac{e^{-\beta \mathcal{H}(\boldsymbol{s})}}{Z} \tag{1.3}$$

where $\beta = \frac{1}{k_B T}$ is the inverse temperature, and Z is the *partition function*, which in short represents a normalization term in the probability for a given state $\boldsymbol{s}$. However, it has a deeper meaning, as it links the microscopic and macroscopic properties of the system. It includes the contribution of all states of the system, weighted by the Boltzmann factor. These concepts may seem quite abstract, but they are instrumental for our analysis. The interested reader can refer to any introductory book on statistical mechanics (for example [5]) for a deeper understanding of these concepts, which are intentionally introduced only intuitively here. Thus, the partition function is:

$$Z_N = \sum_{s_1=\pm 1} \sum_{s_2=\pm 1} \cdots \sum_{s_N=\pm 1} e^{-\beta \mathcal{H}(\boldsymbol{s})}. \tag{1.4}$$

Having defined the setting, the expectation of the spin s_i can be written as:

$$\mathbb{E}[s_i] = \frac{\sum_{s_1=\pm 1} \sum_{s_2=\pm 1} \cdots \sum_{s_N=\pm 1} s_i e^{-\beta \mathcal{H}(\boldsymbol{s})}}{Z_N} \tag{1.5}$$

In the following section, we will solve the Ising model in one dimension, and we note that it can also be solved in two dimensions. The solutions are not so simple and require some technical methods. However, if we make a *mean-field* assumption, the system simplifies drastically and can be solved in any dimension using just some algebra. The mean-field approximation consists of assuming that:

Assumption: *The system's thermal fluctuations are relatively small and can therefore be neglected to a certain extent.* [4]

Let us rewrite the spin s_i of each site highlighting the deviation δs_i from the mean value defined in Eq. (1.5):

$$s_i = \mathbb{E}[s_i] + \delta s_i \tag{1.6}$$

It is immediate to derive the interaction terms $s_i s_j$ in light of this rewriting:

$$s_i s_j = (\mathbb{E}[s_i] + \delta s_i)(\mathbb{E}[s_j] + \delta s_j) = \mathbb{E}[s_i]\mathbb{E}[s_j] + \mathbb{E}[s_j]\delta s_i + \mathbb{E}[s_i]\delta s_j + \delta s_i \delta s_j$$

As we assume that the fluctuations δ are small, we can neglect the second-order term, i.e., $\delta s_i \delta s_j = 0$. And by writing $\delta s_i = \mathbb{E}[s_i] - s_i$ we easily get:

$$s_i s_j \approx \mathbb{E}[s_j] s_i + \mathbb{E}[s_i] s_j - \mathbb{E}[s_i]\mathbb{E}[s_j] \tag{1.7}$$

Now, observe that each spin is equivalent to the others, making the system translational invariant. This implies that the empirical mean of the average magnetization m coincides with that of a single site: $m = \mathbb{E}[s_i]$. This allows us to rewrite Eq. (1.7) as a function of m itself:

$$s_i s_j \approx m[(s_i + s_j) - m] \tag{1.8}$$

In Eq. (1.8) it appears clear why the mean-field approach drastically simplifies the study of the system:

The initial system of interacting particles (note the product term $s_i s_j$ in the Hamiltonian) becomes a system of non-interacting particles, where each site only interacts with a "mean-field", capturing the average behavior of the neighboring sites [4].

With this simplification, we can rewrite the *mean-field* Hamiltonian as:

$$\mathscr{H}_{HF}(\boldsymbol{s}) = -Jm \sum_{\langle i,j \rangle} (s_i + s_j - m) - h \sum_{i=1}^{N} s_i$$

The summation over all pairs of nodes (taken once) can be viewed from the perspective of a node i with respect to its neighborhood $\mathscr{N}_i = \{j : (i,j) \text{ edge in } \mathscr{L}\}$. We can use the symmetry between site i and j, which implies $\sum_{\langle i,j \rangle} s_i + s_j = \sum_{\langle i,j \rangle} 2s_i$. Also note that the summands no longer depend on j and we consider a regular lattice where each node has $\sum_{j \in \mathscr{N}_i} 1 := q$ neighbors:

$$\mathscr{H}_{HF} = -Jm \frac{1}{2} \sum_{i=1}^{N} \sum_{j \in \mathscr{N}_i} (2s_i - m) - h \sum_{i=1}^{N} s_i = \frac{NqJm^2}{2} - (h + qJm) \sum_{i=1}^{N} s_i \tag{1.9}$$

This form of the Hamiltonian makes it clear that the sites in the mean-field approximation do not interact with each other but with an "effective magnetic

field" equal to $h + qJm$. After some manipulations (see [4] for details) we obtain the *self-consistency* equation:

$$m = \tanh\left[\beta(h + qJm)\right] \tag{1.10}$$

This is a transcendental equation that has no analytical solution but can be solved graphically. In the special case of $h = 0$, for example, it can be determined that the system undergoes a *phase transition* at a critical temperature $T_C = \frac{qJ}{K_B}$. When $T > T_C$ the only solution is $m = 0$, which indicates a disordered system (*paramagnetic phase*), while when $T_C < T$ there could be three solutions $\pm m_0$ and 0, but only $\pm m_0$ are stable, so the system is magnetized even in the absence of an external field ($h = 0$).

This provides a fairly immediate solution for the Ising model in any dimension. It accurately reproduces the qualitative behavior of the model in dimensions greater than one, correctly indicates a phase transition, and provides an approximate value for the critical threshold. However, as we will see in the next chapter, it incorrectly predicts a phase transition in the one-dimensional case. This example is very illustrative because it shows both the strengths and limitations of the mean-field approach.

1.2.2 The Ising Model in One Dimension

The mean-field approach from the previous section applies, as already mentioned, in any dimension. Let us now focus on the simplest case of $d = 1$, for which we consider *toroidal* boundary conditions (required to deal with the interactions between the spins at the extremes, i.e., we assume that $s_1 = s_N$). In the presence of an external field h, we can write the Hamiltonian as follows, which is simpler than Equation 1.1, since the spins can only interact with the left and right neighbors:

$$\mathscr{H} = -J\sum_{i=1}^{N} s_i s_{i+1} - h\sum_{i=1}^{N} s_i = -J\sum_{i=1}^{N} s_i s_{i+1} - \frac{h}{2}\sum_{i=1}^{N}(s_i + s_{i+1}) \tag{1.11}$$

Note that all spins become equivalent considering the above Hamiltonian and a ring structure due to the toroidal boundary conditions. To solve the Ising

model in one dimension, we apply the *transfer matrix* method, a powerful tool that allows us to rewrite the partition function.

Each element of the transfer matrix T is defined as:

$$T_{s_i,s_j} = e^{\beta\left[Js_is_j+\frac{h}{2}(s_i+s_j)\right]} \tag{1.12}$$

Recall that each spin s can only assume values in $\{\pm 1\}$, thus the transfer matrix T, considering all four possible combinations of the two spin values, becomes:

$$T = \begin{pmatrix} T_{++} & T_{+-} \\ T_{-+} & T_{--} \end{pmatrix} = \begin{pmatrix} e^{\beta(J+h)} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta(J-h)} \end{pmatrix} \tag{1.13}$$

We can write the partition function Z_N associated with the Hamiltonian in Eq. (1.11) and factor the exponential term. It becomes immediate that these factors can be written as elements of the transfer matrix T:

$$\begin{aligned} Z_N &= \sum_{s_1=\pm 1}\sum_{s_2=\pm 1}\cdots\sum_{s_N=\pm 1}\prod_{i=1}^{N} e^{\beta\left(Js_is_{i+1}+\frac{h}{2}(s_i+s_{i+1})\right)} \\ &= \sum_{s_1=\pm 1}\sum_{s_2=\pm 1}\cdots\sum_{s_N=\pm 1} T_{s_1s_2}T_{s_2s_3}\ldots T_{s_{N-1}s_N} \end{aligned}$$

It is useful to recall that the i,j element of the n-th power of a matrix A can be written as $(A^n)_{ij} = \sum_i\sum_k\sum_l\cdots\sum_z a_{ik}a_{kl}a_{l,m}\ldots a_{zj}$. Moreover, as we are considering toroidal boundary conditions ($s_1 = s_N$), we get:

$$Z_N = \sum_{s_1=\pm 1}\sum_{s_2=\pm 1}\cdots\sum_{s_N-1=\pm 1} T_{s_1s_2}T_{s_2s_3}\ldots T_{s_{N-1}s_1} = \sum_{s_1=\pm 1} T^N_{s_1s_1} = \mathrm{Tr}(T^N) \tag{1.14}$$

We have just established that the partition function can be written as the trace of the power N of the transfer matrix T. Recall that the trace of a matrix does not depend on its representation, so we can write the matrix T in diagonal form. T is square and symmetric, so it has real eigenvalues, which we denote as λ_+ and λ_-. This transformation makes the power N of the transfer matrix trivial, and the

partition function can be written as $Z_N = \lambda_+^N + \lambda_-^N$. By solving the characteristic polynomial associated with the matrix T, we obtain:

$$\lambda_\pm = e^{\beta J}\cosh\beta h \pm \sqrt{e^{-2\beta J} + e^{2\beta J}\sinh^2\beta h} \tag{1.15}$$

To derive the magnetization m, we will use the relationship between the magnetization m and the free energy density f, i.e., $m = -\frac{\partial f}{\partial h}$. Let us first derive the Helmholtz free energy as $F = -\frac{1}{\beta}\log Z_N$. We can compute the free energy per spin (the density), considering a large number of spins N, as follows:

$$f = \lim_{N\to\infty} -\frac{1}{N\beta}\log Z_N = \lim_{N\to\infty} -\frac{1}{N\beta}\log\left(\lambda_+^N + \lambda_-^N\right) \tag{1.16}$$

From Eq. (1.15) it is immediate to see that $\lambda_+ > \lambda_-$, thus $\lim_{N\to\infty}\left(\frac{\lambda_-}{\lambda_+}\right)^N = 0$. We can rewrite Eq. (1.16) and obtain an easy formula for f:

$$f = \lim_{N\to\infty} -\frac{1}{N\beta}\log\left[\lambda_+^N\left(1 + \frac{\lambda_-^N}{\lambda_+^N}\right)\right] \approx -\frac{\log\lambda_+}{\beta} \tag{1.17}$$

We can finally compute the magnetization m:

$$m = -\frac{\partial f}{\partial h} = \frac{\sinh\beta h}{\sqrt{\sinh^2\beta h + e^{-4\beta J}}} \tag{1.18}$$

This formula implies that there is no phase transition in one dimension. In fact, the system shows a paramagnetic behavior as $m = 0$ for $h = 0$ ($\sinh\beta h|_{h=0} = 0$) for any finite temperature. Only for $T \to 0$ the model becomes a ferromagnetic one as $\lim_{\beta\to\infty}\frac{\sinh\beta h}{\sqrt{\sinh^2\beta h + e^{-4\beta J}}} = \pm 1$. A similar derivation and some considerations on the two-dimensional Ising model can be found in [6].

1.2.3 The Sznajd Opinion Model

We introduced the mean-field approach by analyzing the Ising model. We found that it is generally easier to obtain results for the mean-field model (almost only algebraic calculations), while for the simplest case of the one-dimensional Ising model we needed a more technical derivation. However, the results of the mean-

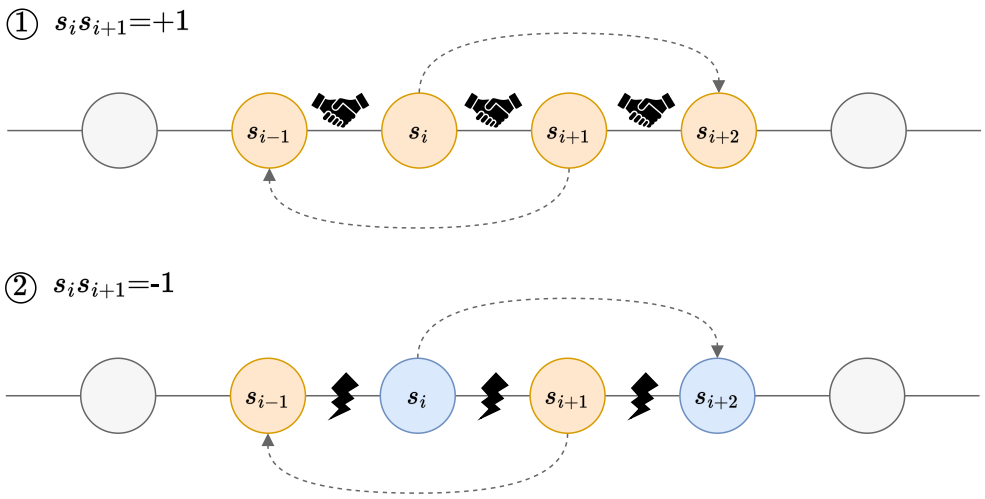


Fig. 1.2 Sketch of the spin updating process of the Sznajd model, for ① agreeing pair and ② disagreeing pairs.

field model are not always accurate. This can be checked by deriving an exact solution (if it is possible) or by extensive *Monte Carlo simulations*. But,

How is this related to opinion dynamics?

One of the most popular opinion models was proposed by physicists drawing inspiration from the Ising model: the so-called Sznajd model [7]. It was developed to describe the decision-making process of a closed community. Each site can be interpreted as an agent and spin value as an opinion or, better, a preference against an outcome. For example, $s_i = +1$ corresponds to the preference for candidate A in an election, while $s_i = -1$ stands for competitor B.

The setting is the same as in the one-dimensional Ising model, with N nodes (agents) over a line graph characterized by short-range interactions. The idea underlying the Ising model is that of "united we stand, divided we fall". If two agents s_i and s_{i+1} agree, i.e., $s_i s_{i+1} = +1$, they transfer their opinion to their neighbors. Therefore, s_{i-1} and s_{i+2} adopt the same opinion as the pair s_i, s_{+1}. On the other hand, if the two agents disagree, i.e., $s_i s_{i+1} = -1$, then the neighbors also tend to disagree with each agent in the pair, i.e., s_{i-1} takes the direction of s_{i+1} and s_{i+2} takes the direction of s_i. These rules describe the influence of a pair on the decisions of its neighbors [7]. Figure 1.2 gives a visual representation of this simple mechanism from the point of view of an agent pair.

It is easy to see that there are three equilibrium points for the dynamics: all spins point up (+1), all spins point down (-1), or all neighbors have opposite orientations. If we describe the model as a Markov chain (MC), these states would all be absorbing states for the MC. The authors study the model using Monte Carlo simulations and find that the three steady states mentioned above occur with probabilities of 0.25, 0.25, and 0.5, respectively. Thus, the Sznajd model gives rise to two possible final configurations: complete disagreement or consensus among the agents.

The Sznajd model has been extensively studied in the literature. Slanina and Lavička [8] consider two extensions of the model and redefine it over a general graph. They take a graph $\mathcal{G}(\mathcal{N},\mathcal{E})$, where $N = |\mathcal{N}|$ is the number of agents in the population and $\mathcal{E}$ represents the (undirected) connections between the agents that capture their relationships.

Algorithm 1: Sznajd Model on a Social Graph $\mathcal{G}$

Input: Social graph $\mathcal{G}(\mathcal{N},\mathcal{E})$, initial opinion/spin configuration $\boldsymbol{s}(0)$

while *true* **do**
  Randomly choose agent i;
  From $\mathcal{N}_i = \{j : (i,j) \in \mathcal{E}\}$ randomly choose agent j;
  if $s_i(t) = s_j(t)$ **then**
    Choose a common neighbor k of both i and j (in $\mathcal{N}_i \cap \mathcal{N}_j$) ;
    $s_k(t) = s_i(t)$;

Each agent in $\mathcal{G}$ has an opinion s_i (as in the original Sznajd model), which can take up to q different values. Here, we restrict ourselves to $q = 2$ for simplicity. The state of the system is represented by the opinion value of each agent: $\Sigma(t) = [s_1(t), s_2(t), \ldots, s_N(t)]$. The system is thus a discrete Markov process that evolves as follows (see Algorithm 1): at each discrete time t, an agent $i \in \mathcal{N}$ is randomly selected. Then, a neighbor $j \in \mathcal{N}_i := h \in \mathcal{N} : (i,j) \in \mathcal{E}$ of agent i is again selected at random. If the two agents disagree, nothing happens, while if $s_i(t) = s_j(t)$ a common neighbor k ($k \in \mathcal{N}_i \cap \mathcal{N}_j \setminus \{i,j\}$) takes the opinion of the pair: $s_k(t) = s_i(t) = s_j(t)$. This extension of the model is particularly interesting because it allows us to derive a partial differential equation (PDE) for the probability distribution of the magnetization $m = \frac{1}{N}\sum_{i=1}^{N} s_i$ when $\mathcal{G}$ is a complete graph.

If we consider $\mathcal{G}$ as a fully connected graph, the dynamics are somewhat simplified, since after choosing the agent i, neighbor j is any node other than i and the common neighbor k is again any node in $\mathcal{G}$ other than i and j. Moreover, each agent becomes equivalent to every other and the state of the system can be fully characterized by the empirical magnetization, which we can define as $m = \frac{N_+ - N_-}{N}$, where $N_+ := \sum_{i=1}^{N} \delta_{s_i,+1}$ and $N_- := \sum_{i=1}^{N} \delta_{s_i,-1}$ are the number of agents whose opinion/spin points up (+1) or down (-1), respectively. This setup is described by [8] as a *mean-field approximation*[2], in the sense that "the Ising model on the complete graph can be considered as an approximation for the Ising model on hypercubic lattice of high dimensionality" [8].

Given a state $\Sigma(t)$, the DTMC describing the model can undergo one of three transitions: a self-loop to the state $\Sigma(t)$ itself, when the selected agents have disagreeing opinions (Case I), a transition to a state with higher magnetization, whenever the two selected agents have agreeing positive opinions (Case II), and a symmetric transition to a lower magnetization, as the agents have agreeing negative opinions (Case III). In the two latter situations the magnetization changes by $\frac{2}{N}$ as one agent flips its opinion from one state to the other.

Let us first characterize the probability of *Case II*. To increase the magnetization it is needed to take two agreeing neighbors i and j, and that the common neighbor k disagrees with them:

$$\mathbb{P}\,(\text{Case I}) = \mathbb{P}\left(m \to m + \frac{2}{N}\right) = \mathbb{P}\left(s_i(t) = +1, s_j(t) = +1, s_k(t) = -1\right) \tag{1.19}$$

where s_i, s_j, s_k are the opinions of the agents picked sequentially as for Algorithm 1. To evaluate Eq. 1.19, let us recall the definition of the magnetization $m = \frac{N_+ - N_-}{N}$ and the trivial equality $N_+ + N_- = N$. From these equations, it is easy to obtain the fraction of spins in a certain state as a function of the magnetization:

$$\frac{N_+}{N} = \frac{m+1}{2} \quad \text{and} \quad \frac{N_-}{N} = \frac{1-m}{2}$$

[2] This notion is slightly different from the one we studied earlier in this chapter. However, the Curie-Weiss model (another mean-field approximation of the Ising model) assumes that any spin can interact with any other (all-to-all communication), but the interactions are very weak ($\frac{s_i s_j}{N}$). The same mean-field results in Section 1.2.1 can be obtained.

From which it is immediate:

$$\begin{aligned}\mathbb{P}\left(m \to m+\frac{2}{N}\right) &= \frac{N_+}{N}\frac{N_+-1}{N-1}\frac{N_-}{N-2} \\ &= \frac{N}{N-1}\frac{N}{N-2}\frac{1-m^2}{8}\left[m+1-\frac{2}{N}\right]\end{aligned} \tag{1.20}$$

And similarly for *Case III* we obtain:

$$\mathbb{P}\left(m \to m-\frac{2}{N}\right) = \frac{N}{N-1}\frac{N}{N-2}\frac{1-m^2}{8}\left[1-m-\frac{2}{N}\right] \tag{1.21}$$

Note that the two transition probabilities are not symmetric, which is linked to the presence of a phase transition [8]. Finally, by subtraction, we get:

$$\mathbb{P}(m \to m) = 1-\frac{N}{N-1}\frac{N}{N-2}\frac{1-m^2}{4}\left[1-\frac{2}{N}\right] \tag{1.22}$$

At this point, we would like to derive a PDE for the evolution of the probability density of the magnetization $\mathbb{P}(m,t)$ in an infinitesimal step τ:

$$\begin{aligned}\mathbb{P}(m,t+\tau) =& \mathbb{P}(m \to m)\,\mathbb{P}(m,t) + \left(m \to m+\frac{2}{N}\right)\mathbb{P}\left(m+\frac{2}{N},t\right) \\ &+ \left(m \to m-\frac{2}{N}\right)\mathbb{P}\left(m-\frac{2}{N},t\right)\end{aligned}$$

We can consider the *thermodynamic limit,* as the size of the system becomes large ($N \to \infty$) and considering an appropriate *scaling factor* for the time (specified later), we obtain:

$$\begin{aligned}\mathbb{P}(m,t+\tau) =& \left(1-\frac{1-m^2}{4}\right)\mathbb{P}(m,t) + \frac{1-m^2}{8}(m+1)\mathbb{P}\left(m+\frac{2}{N},t\right) \\ &+ \frac{1-m^2}{8}(1-m)\mathbb{P}\left(m-\frac{2}{N},t\right)\end{aligned}$$

$$\mathbb{P}(m,t+\tau)-\mathbb{P}(m,t)=\frac{1-m^2}{8}\left[\left(\frac{2}{N}\right)^2\frac{-2\mathbb{P}(m,t)+\mathbb{P}\left(m+\frac{2}{N},t\right)+\mathbb{P}\left(m-\frac{2}{N},t\right)}{(\frac{2}{N}^2)}\right.$$
$$\left.-m\left(2\frac{2}{N}\right)\frac{\mathbb{P}\left(m+\frac{2}{N}\right)-\mathbb{P}\left(m-\frac{2}{N}\right)}{2\frac{2}{N}}\right] \tag{1.23}$$

If we divide both sides of Equation (1.23) by τ, and scale appropriately the time as $t=\frac{2\tau}{N}$, we can take the limit and recognize a time partial derivative on the left-hand side and two magnetization partial derivatives on the right-hand side, for which the higher order terms $\frac{1}{N^2}$ can be neglected, so that we get:

$$\frac{\partial}{\partial t}\mathbb{P}(m,t)=-\frac{\partial}{\partial m}\left[m(1-m^2)\,\mathbb{P}(m,t)\right], \tag{1.24}$$

which is studied in [8], to which we refer the interested reader. We note that the logic behind the derivation of Eq. (1.24) is similar to that which we will use in Chapter 2 later in this thesis.

1.2.4 The Susceptible-Infected-Susceptible (SIS) Model

We now discuss one of the classical epidemic models, the Susceptible-Infected-Susceptible. First, we present its stochastic formulation, which can be modeled as a discrete-time Markov chain, and then its deterministic formulation. We will find that there is a close relationship between the two models and the latter is a kind of mean-field of the stochastic formulation. This section will show that epidemics and opinion dynamics, although seemingly distant from each other, are closely related as much of the underlying structure and analysis is identical. Furthermore, SIS and SIR (Susceptible-Infected-Removed) models have been used in the context of opinion dynamics to describe the diffusion of beliefs and innovations. For example, Cinelli et al. [9] used a SIR model in the context of opinion dynamics to model the spread of information and to investigate the emergence of *echo chambers* in online social networks. In the last Chapter of this thesis, we will introduce the SIR model, which is another classical model and can be considered as a complement to the SIS model, as they can be used to

model viruses that exhibit different biological behaviors (reinfection vs. acquired immunity).

Stochastic SIS Model

In epidemic models, individuals are classified according to their disease status. *Susceptible* individuals are those who, upon contact with an *Infected* individual, i.e., one who carries the disease, may develop the disease themselves. The Susceptible-Infected-Susceptible model describes exactly this situation, where individuals do not acquire immunity to the disease and become susceptible again after recovery. This is one of the reasons why the SIS model was used for modeling sexually transmitted diseases. Note that deaths are not explicitly modeled (while they are included in the SIR model).

Let us now introduce the random variables[3] $\mathscr{S}(t)$ and $\mathscr{I}(t)$, which respectively denote the number of susceptible and infected individuals in the given population at a given time t. We assume that time is discrete, a multiple of a generic interval Δt, therefore we have $t \in \{0, \Delta t, 2\Delta t, \dots\}$. We consider a fixed population of size N, thus $\mathscr{S}(t) = N - \mathscr{I}(t)$, which results in only one independent random variable: $\mathscr{I}(t) \in 0, 1, \dots, N$. We want to describe the stochastic process[4] $\{\mathscr{I}(t)\}_{t=0}^{\infty}$, for which we define the *state* probability $s_i(t)$:

$$s_i(t) = \mathbb{P}(\mathscr{I}(t) = i) \tag{1.25}$$

Following [10], we can formalize this model as a discrete-time Markov chain. For a given time t, the process is characterized by the probability vector $\boldsymbol{p}(t) = \{p_0(t), p_1(t), \dots, p_N(t)\}$ and satisfies the Markov property, i.e., the process at the subsequent time depends only on the process at the current time:

$$\mathbb{P}(\mathscr{I}(t+\Delta t) \,|\, \mathscr{I}(0), \dots \mathscr{I}(t)) = \mathbb{P}(\mathscr{I}(t+\Delta t) \,|\, \mathscr{I}(t))$$

The time interval Δt must be sufficiently small so that only one event (or none) occurs during Δt. The event can consist of an infected individual infecting a

[3] We use calligraphic letters to distinguish the *stochastic* susceptible and infected individuals from the *deterministic* quantities that we will use later in this section and in Chapter 4.

[4] A stochastic process is a collection of random variables, here indexed by time t.

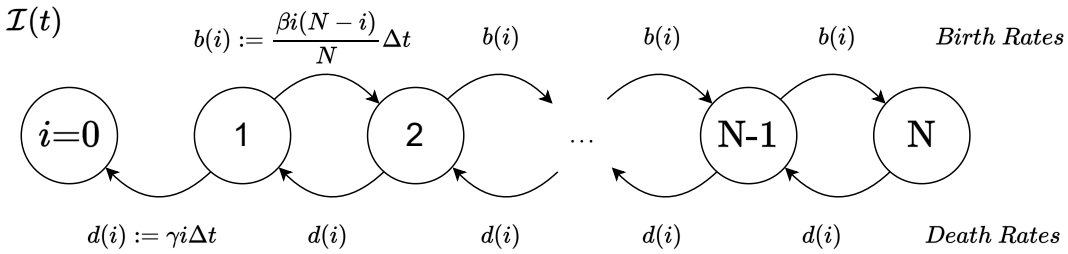


Fig. 1.3 Markov chain representation of the SIS stochastic model.

susceptible individual, which increases the number of infected $\mathscr{I}(t) = i$ by one. The probability that such an event occurs is proportional to a disease-specific parameter β (the higher β, the more contagious the disease), the chance of a susceptible individual encountering an infected individual: $\mathscr{S}(t)\mathscr{I}(t) = si$ and the time interval of observation Δt:

$$p_{i,i+1} := p_{i\to i+1} = \frac{\beta i(N-i)}{N}\Delta t, \tag{1.26}$$

where we have taken advantage of the fact that the population is fixed, i.e., $s = N - i$. The event during the interval Δt can be that an infected individual recovers at a rate equal to γ, i.e., with probability $\gamma\Delta t$ and proportional to $\mathscr{I}(t) = i$:

$$p_{i,i-1} := p_{i\to i-1} = \gamma i \Delta t \tag{1.27}$$

We are left to consider the case in which neither of the above events occurs, and the state of the system does not change. The probability of such an event is:

$$p_{i,i} := p_{i\to i-1} = 1 - \left(\frac{\beta i(N-i)}{N}\Delta t + \gamma i \Delta t\right) \tag{1.28}$$

In a DTMC, these probabilities characterize the *transitions* between states and are usually called *transition probabilities* $p_{s_i\to s_j}$. They are the last ingredient we need to specify for our Markov chain. We have implicitly assumed that the probabilities are independent of the specific time t, therefore we consider a *homogeneous* process.

Interestingly, the process is a *birth and death* process (see Figure 1.3). Note that in this Markov chain, the state $\{0\}$ is an *absorbing state*, while all other states are *transient states*, therefore the state probability in the limit is $\lim_{t\to\infty} \boldsymbol{p}(t) =$

$(1,0,\ldots,0)^T$. This means that the population is asymptotically approaching the disease-free equilibrium [10]. We can rewrite the transition probability as a birth probability $b(i)$ and a death probability $d(i)$:

$$b(i) = \frac{\beta i(N-i)}{N}\Delta t \quad \text{and} \quad d(i) = \gamma i \Delta t \tag{1.29}$$

In general, the probability of the process to be in state i at time $t+\Delta t$ can be easily determined by considering the probability of the three possible events that lead to it: being in state $i-1$ and having a new infection, being in state $i+1$ and having a recovery, or finally just staying in state i:

$$\begin{aligned} s_i(t+\Delta t) &= s_{i-1}p_{i-1\to i} + s_{i+1}p_{i+1\to i} + s_i(t)p_{i\to i} \\ &= s_{i-1}(t)b(i-1) + s_{i+1}(t)d(i+1) + s_i(t)[1-b(i)-d(i)] \end{aligned} \tag{1.30}$$

The authors of [10] derive a relation for the expected value of the number of infected, i.e., $\mathbb{E}[\mathscr{I}(t)] = \sum_{i=0}^{N} i s_i(t+\Delta t)$, by making use of Eq. (1.30):

$$\begin{aligned} \mathbb{E}[\mathscr{I}(t+\Delta t)] = &\sum_{i=1}^{N} i s_{i-1} b(i-1) + \sum_{i=0}^{N-1} i s_{i+1} d(i+1) \\ &+ \sum_{i=0}^{N} i s_i(t) - \sum_{i=0}^{N} i s_i(t) b(i) - \sum_{i=0}^{N} i s_i(t) d(i) \end{aligned}$$

If we make a change of variables in the first two summations, so that we realign all addends, we can readily obtain:

$$\begin{aligned} \mathbb{E}[\mathscr{I}(t+\Delta t)] &= \mathbb{E}[\mathscr{I}(t)] + \sum_{i=0}^{N} [(i+1)-i]\, s_i(t) b(i) + \sum_{i=1}^{N} [(i-1)+i]\, s_i(t) b(i) \\ &= \mathbb{E}[\mathscr{I}(t)] + \sum_{i=0}^{N} s_i(t) \frac{\beta i(N-i)}{N}\Delta t - \sum_{i=1}^{N} s_i(t)\gamma i \Delta t \\ &= \mathbb{E}[\mathscr{I}(t)] + [\beta-\gamma]\Delta t \sum_{i=0}^{N} s_i(t) i - \frac{\beta}{N}\Delta t \sum_{i=0}^{N} s_i(t) i^2 \\ &= \mathbb{E}[\mathscr{I}(t)] + [\beta-\gamma]\Delta t \mathbb{E}[\mathscr{I}(t)] - \frac{\beta}{N}\Delta t \mathbb{E}[\mathscr{I}^2(t)] \end{aligned} \tag{1.31}$$

Lastly, we derive a straightforward bound observing that $\mathbb{E}\left[x^2\right] - \mathbb{E}^2[x] \geq 0$ hence $\mathbb{E}\left[x^2\right] \geq \mathbb{E}^2[x]$ we can bound from above the right-hand-side:

$$\mathbb{E}[\mathscr{I}(t+\Delta t)] \leq \mathbb{E}[\mathscr{I}(t)] + [\beta-\gamma]\Delta t\,\mathbb{E}[\mathscr{I}(t)] - \frac{\beta}{N}\Delta t\,\mathbb{E}^2[\mathscr{I}(t)] \tag{1.32}$$

We can now rearrange the equation, divide both sides by Δt and take the limit for $\Delta t \to 0$. We obtain the differential equation:

$$\begin{aligned}\frac{d\mathbb{E}[\mathscr{I}(t)]}{dt} &\leq \left(\beta-\gamma\right)\mathbb{E}[\mathscr{I}(t)] - \frac{\beta}{N}\mathbb{E}^2[\mathscr{I}(t)] \\ &= \frac{\beta}{N}\left(N - \mathbb{E}[\mathscr{I}(t)]\right)\mathbb{E}(\mathscr{I}(t)) - \gamma\mathbb{E}[\mathscr{I}(t)]\end{aligned} \tag{1.33}$$

Equation (1.33) is a deterministic differential equation that contains the expectation of the number of infected individuals. It implies that the average value of the infected for the stochastic model is less than or equal to the solution of the deterministic equation. Fascinatingly, the deterministic differential equation we obtained corresponds exactly to the deterministic SIS model. So from a stochastic instance of our model, we have again obtained a sort of *mean-field* approximation of our system, which is governed by deterministic equations.

As a final remark, we would like to emphasize that the stochastic version of the SIS model can be easily simulated by agent-based modeling, as can the Sznajd opinion model in both formulations. Agent-based modeling is a computational modeling technique in which each agent is equipped with a set of predefined rules for interacting with the rest of the population that can be easily studied and implemented through computer simulations. The approach, which is also widely used in this thesis, is to study the emergent behavior of the stochastic system through simulations and, if possible, derive a deterministic mean-field relation to evaluate the mean behavior of the system.

Deterministic SIS Model

We present here the deterministic formulation of the classical Susceptible-Infected-Susceptible model, which is probably the simplest of the so-called *compartmental* models. The population is divided into different "compartments" according to dis-

ease status (as above): individuals can be *Susceptible* (S) to the virus or *Infected* (I). The system of differential equations describing the system's dynamics is:

$$\begin{aligned} \frac{dS}{dt} &= -\frac{\beta}{N} SI + \gamma I \\ \frac{dI}{dt} &= \frac{\beta}{N} SI - \gamma I \end{aligned} \tag{1.34}$$

As we hinted in the previous section, this deterministic model is closely related to its stochastic counterpart. Namely, if we replace the expected value of the number of infected $\mathbb{E}[\mathscr{I}(t)]$ in Eq. (1.33) by I and remember that $S = N - I$ due to mass conservation, we get exactly the above deterministic bound:

$$\frac{dI}{dt} = \frac{\beta}{N}(N-I)I - \gamma I \tag{1.35}$$

This differential equation can be solved and from the solution, it is possible to divide the behavior of the system into two different phases depending on the *basic reproduction number* $\mathscr{R}_0 := \frac{\beta}{\gamma}$, which is a disease-specific value[5]. Let us first rewrite the right-hand side of Eq. (1.35) and separate the variables:

$$\frac{dI}{dt} = -\frac{\beta}{N} I\left(I - \frac{N(\beta-\gamma)}{\beta}\right) \implies \frac{dI}{I(I-b)} = a \cdot dt$$

where for ease of presentation we introduced $a := -\frac{\beta}{N}$ and $b := \frac{N(\beta-\gamma)}{\beta}$. To solve the equation, we can use the partial decomposition method and then integrate both sides of the equality. We can easily determine that $\frac{1}{I(I-a)} = -\frac{1}{b}\frac{1}{I} + \frac{1}{b}\frac{1}{I-b}$:

$$\int_{I(0)}^{I(t)} -\frac{1}{b}\frac{1}{I} + \frac{1}{b}\frac{1}{I-b} dI = \int_0^t a \cdot dt$$

$$\frac{1}{b}\left[\ln\frac{b}{I}\right]_{I(0)}^{I(t)} = at \implies \frac{1-\frac{b}{I(t)}}{1-\frac{b}{I(0)}} = e^{bat}$$

Finally, after some manipulations, we obtain the exact solution of the SIS:

[5] Note that the factor $\frac{1}{N}$ is only a normalization factor. Indeed, the SIS system can be rewritten in terms of the fractions of individuals in each compartment by dividing both sides of the equation by $\frac{1}{N}$ rendering (first equation only): $\frac{ds}{dt} = -\beta si + \gamma i$

$$I(t) = \frac{N(\beta-\gamma)I(0)e^{(\beta-\gamma)t}}{\beta I(0)\left[e^{(\beta-\gamma)t}\right] + N(\beta-\gamma)} \tag{1.36}$$

From the exact solution, we can gain interesting insights into the asymptotic behavior of the system. Namely, if $\beta < \gamma$ (i.e., $\mathscr{R}_0 = \frac{\beta}{\gamma} < 1$), then $\lim_{t\to\infty} I(t) = 0$ hence the epidemic dies out (*disease-free equilibrium*). While if $\beta > \gamma$, i.e. $\mathscr{R}_0 > 1$, we obtain $\lim_{t\to\infty} I(t) = N\left(1 - \frac{1}{\mathscr{R}_0}\right)$, which denotes a point at which the virus becomes endemic in the population, i.e. part of the population will be infected at any given time (*endemic equilibrium*). The reproduction number therefore plays a very important role in an epidemic, as it determines whether the disease dies out or becomes *endemic* in the population. In Chapter 4 we will see that the reproduction number also has a similar meaning in the Susceptible-Infected-Removed (SIR) model.

Let us make one final remark. We have seen the connection between the stochastic model and the deterministic model, but the behavior of the deterministic model is somewhat different from what we have observed for the stochastic model, where the system always tends to the disease-free equilibrium (regardless of the value of the basic reproduction number $\mathscr{R}_0$). This is related to the fact that the absorption result for the Markov chain is asymptotic and the rate of convergence to this equilibrium can be very slow [10]. The interested reader can refer to "Mathematical Epidemiology"by Fred Brauer et al. [10] for a comprehensive introduction to epidemic modeling.

1.3 Main Contributions and Organization

The above examples have been chosen to provide both an overview of the fields of opinion dynamics and epidemiology and to show some of the key tools that the "social modeler" might need in its investigation. The presentation also makes it clear that the two fields, although appearing very different, are usually explained in quite similar terms. We will now briefly discuss the contribution of this work. It is a modeling journey that begins with opinion formation in online social networks, continues through the competition of online influencers, and ends with the the control of the COVID-19 pandemic.

The second chapter of this thesis is dedicated to online interactions. Although the advent of social media platforms has drastically changed the way individuals interact with each other, little attention has been paid to the development of opinion models that go beyond the classical models (most of which are appropriate for face-to-face or small group interactions) and capture the specific characteristics of *online* interactions. In Chapter 2, we move in this direction by proposing a multidimensional opinion model that explicitly divides users into two classes: regular users and influencers. The latter are popular and influential people on the platform who produce the majority of the content consumed by users on modern social networks. The amount of content produced on such social media platforms is immense. Therefore, to keep users engaged within the social network, the platform performs filtering to select the posts to offer. This filtering mechanism can reinforce the natural tendency of users to interact with like-minded individuals, one of the main drivers for building social networks known as *homophily*[11]. This, in turn, can lead to the formation of *echo chambers* [9], where people who share a similar point of view interact with each other while being isolated from the rest of the users. We propose the Communication Asymmetry opinion model, the goal of which is to model:

- The asymmetry of OSNs: a relatively small portion of well-known users, i.e., the influencers, can reach a vast number of far less known individuals.
- The closed-loop between influencers' and regular users' dynamics, triggered by the content personalization mechanism applied by the social media platform in one direction and user feedback (e.g., likes) in the other.

In addition, the range of topics discussed in OSNs is very wide and influencers in particular tend to post about a specific topic, which can be their area of expertise (e.g., tennis or more broadly sport for tennis players) or the reason why they are known over a platform (e.g., cooking for someone presenting recipes). We have empirically observed this tendency in OSNs (Facebook and Instagram) and proposed the concept of *reference direction*, over which content personalization occurs.

For the case of a single topic, we derive a Fokker-Planck equation as a second-order approximation for the opinion formation process and prove the existence and uniqueness of the stationary solution. Furthermore, since this approach is

not constructive, we develop a less accurate first-order approximation, i.e., the fluid limit, which provides a closed-form expression for the stationary solution. We then use a Monte Carlo approach to analyze the impact of influencers' structural parameters and the interplay with content personalization in a multi-topic environment. The results highlight some of the dangers associated with content filtering. We show that it favors *structurally advantaged* influencers and can radicalize users' opinions, leading to the formation of *echo chambers*.

In Chapter 2, we have always considered influencers as stubborn, conveying to their user base a fixed opinion that corresponds to their belief about a particular matter. However, we argue that it is interesting to consider the case where influencers are strategic and choose a posting pattern to attract the population towards a certain final opinion (e.g., due to economic interests) rather than stubbornly holding the same fixed opinion. Therefore, in Chapter 3 we use a simplified version of the Communication Asymmetry opinion model from Chapter 2, which allows influencers to strategically change their *expressed* opinion.

Our goal is to model the competition between different influencers over an OSN and to investigate the best possible strategies they can employ to maximize their influence on the population. In the first step of our investigation, we address the fundamental question of how an influencer, considered in isolation, should best proceed to increase their influence on a specific group of users. An influencer's "strategy" is a sequence of posts that convey an opinion and that they publish on their social media profile. The "best" strategy is the one that brings the greatest benefit, however it is defined (by a utility function). Although a person's opinion generally does not fluctuate frequently, an influencer's stance can be affected by both external and personal factors. For example, it is common for influencers to retract certain viewpoints due to the pressure of public opinion. Another example: influencers may change their collaborative partnerships and promote other brands' products. Similarly, politicians often adjust their positions to reflect their electorate's views on certain issues. Therefore, assuming that an influencer's opinion can span the entire opinion space, we show that the *greedy* strategy which maximizes influence at each post emission is not always *optimal*. Our experiments suggest that it is optimal to group the user base, which has different initial views on the topic, around a common viewpoint and then move the group towards the desired opinion.

The final Chapter of this thesis is devoted to the development and investigation of a novel stratified epidemic model which naturally accounts for the correlations between risk exposure to the virus and mortality. Furthermore, the model assumes a tightly controlled environment in which the rate of new infections is controlled by both pharmaceutical (e.g., vaccine) and non-pharmaceutical interventions (e.g., lockdown measures). Especially during the recent COVID-19 pandemic, governments worldwide faced the challenge of developing effective strategies to contain the virus while minimizing the economic and societal impact. It was crucial to address the question of whether a suppression or a mitigation strategy would be more beneficial. In Chapter 4, we present our extended SIR-like model in which we define a control problem that balances economic loss and social loss (as measured by the number of deaths). We also investigate the impact of control lags in the feedback loop and how they endanger system stability.

We include vaccinations in our framework, as several vaccines against the virus became available about a year after the COVID-19 pandemic. The availability of these vaccines was initially limited and their effectiveness in preventing infection was not fully understood. Given these uncertainties, developing strategies to prioritize vaccine distribution became a challenge. The interested reader can refer to [12] for insights and analysis on the complexity of developing effective vaccine prioritization strategies in uncertain times. We have defined simple vaccine prioritization strategies (that can be easily implemented by governments) and assessed the interplay between non-pharmaceutical interventions and vaccination campaigns. The joint effect of these two aspects has been largely neglected in the literature and we aim to fill this gap. We aim to explore in depth the different strategies, measures, and approaches that policymakers can consider and implement to respond effectively to the challenges of a pandemic in order to protect public health, minimize the impact on society and ensure the well-being of individuals. Lastly, even if we consider the case of COVID-19, we remark that the framework is flexible enough to model any other virus.

The thesis is complemented by several appendices which deal with interesting but less central aspects in connection with the main part. The work ends with some concluding remarks in which the findings and contributions of this work are discussed.

Chapter 2

A Model for Online Interactions

Part of the work presented in this chapter has already been published in [13]:

- Galante, F., Vassio, L., Garetto, M., & Leonardi, E. (2023). Modeling communication asymmetry and content personalization in online social networks. In Online Social Networks and Media (Vols. 37–38, p. 100269). Elsevier BV. https://doi.org/10.1016/j.osnem.2023.100269

In the opinion dynamics landscape, numerous approaches have been proposed to capture the intricacies of human interactions however, a comprehensive approach tailored to online social networks (OSNs) is still missing. Online interactions have some unique characteristics that strongly distinguish them from other, more traditional forms of interaction.

The Internet has revolutionized global communication, enabling worldwide, cost-effective communication like no other technology in human history. Instant messaging apps, video conferencing tools, and online social networks have harnessed the potential of the Internet and other modern information technologies, bringing together an ever-expanding user base that, to this day, encompasses the vast majority of the world's population. Most people use social media daily, with seven in ten Facebook users saying they visit the site at least once a day [14]. The study also shows that other platforms, such as YouTube and Instagram, are widely used (e.g., 81 % of respondents have been using YouTube). It is important to note that the study only surveyed people aged 18 and over and did not include the population group that uses social media the most, namely young people, which vastly underestimates social media usage.

One of the most striking features of OSNs is the pronounced communication asymmetry between users, where only a few influential individuals produce a significant amount of content on social media platforms and are far more popular than the other users. In addition, the social media platform performs personalization of the content that is suggested to users. This happens because the number of posts (even if it is just your own "friends" on Facebook) is so large that usually only the posts with the highest relevance are seen. Therefore, social media platforms try to suggest the most relevant content to maximize user engagement.

This chapter presents and examines a model that accounts for these characteristic features and provides a flexible and fine-grained description of the influencers that can be used to describe a whole range of social phenomena. The model is examined both analytically and experimentally using measures from popular social media platforms: Facebook and Instagram. The chapter first briefly reviews the relevant literature on opinion formation, presents contributions from various research communities, and highlights recent advances in the measurement and modeling of OSN. Section 2.2 defines the general notation used in this and the next chapter, along with a table summarizing all the symbols introduced in this chapter. Sections 2.3 and 2.4 describe the Communication Asymmetry opinion model and motivate some of the modeling decisions by analyzing social media data. Section 2.5 is dedicated to the analysis. First, we derive a Fokker-Planck equation for the opinion distribution and establish conditions for the existence and uniqueness of the solution. Then, we provide a first-order approximation to determine the stationary solutions of the dynamical system. In Section 2.6, we consider a more general bidimensional instance of the model trough a Monte Carlo approach and examine the model as a function of the characteristics of the opinion leaders and the population. Finally, in Section 2.7, we relate the posting frequency of an influencer to its popularity growth using data from Instagram. We apply our model to the Italian government crisis in August 2019. The chapter ends with a brief discussion on the implications of the work.

2.1 Related Work and Context

The first steps in the study of opinion dynamics were taken in the late 1950s by a number of social psychologists [15][16][17]. Ash introduced the concept of *social*

pressure [15], a conformist tendency in individuals, while French used directed graphs to model interpersonal relationships [16]. A landmark in the field is Festinger's work on *social comparison*[17]. Individuals tend to evaluate their position by comparing it to others, and it is inversely proportional to the distance between viewpoints. Opinion models are continuous or discrete, according to the description of the opinion variable. As for most models [18], the seminal work in the field is continuous. For example, the DeGroot model [19] considers a networked social system in which individuals interact with their neighbors. Individuals average their current opinion with the opinion of their neighbors. Subsequently, Friedkin and Johnsen [20] developed a linear model which encompasses both the processes of *social conformity* and *social conflict* leading to behavior that goes beyond simple consensus. Dandekar et al. [21] further examined averaging dynamics in light of polarization and proposed a model that included *biased assimilation*. In the early 2000s, Hegselmann and Krause [22] and Deffuant and Weisbuch [23] proposed two similar models, introducing the idea of bounded confidence: individuals interact only with peers whose beliefs are not too different. The proposed models are nonlinear and challenging to study [24]. A great deal of attention has also been paid to discrete models. A prominent example is the voter model [25][26] and its extensions accounting for evolving networks [27][28], individuals with multiple opinions [29] or spontaneous changes of opinion [30]. A consistent bulk of research on opinion dynamics comes from the physics literature, among which early contributions are Ben-Naim [31] and Toscani [32]. Ben-Naim and Toscani consider two mechanisms of opinion formation: compromise, and introspection (in other models, e.g., [30], modeled as noise), which the authors believe represents the impact of external sources of information (e.g., media). The Sznajd model [7] is a generalization of the Ising model, which implements *social validation* and for which [8] derives analytical results. For a comprehensive review of classic opinion models, we refer to the survey by Castellano et al. [33].

Most of the seminal literature on opinion dynamics is suited to describe the decision-making process in small groups of individuals, e.g., a board of directors, or to capture relatively regular patterns determined by the daily personal interactions of individuals. Models such as the voter model have been studied extensively on regular lattices [34] [35]. The structure of interactions, especially those online, is far from homogeneous. As mentioned earlier, an inherent asymmetry in communication exists in OSNs where a limited number of individuals (influencers)

monopolize the discussion. The voter model has been studied over heterogeneous networks (e.g., [36] [37]) to account for this diversity. On such networks, there can exist *hubs* (strongly connected nodes) playing a role similar to influencers in our framework. However, the authors did not explicitly make such a distinction. Other works have divided the population into classes, e.g., [38] introduced *stubborn* agents, and if such individuals have opposing opinions, they hinder the possibility of the population converging to consensus. Recent work further draws attention to online platforms by adapting classical frameworks to the specificity of online interactions. [39–41] have developed an opinion model that embodies *algorithmic personalization*. [39] also compares model predictions and dynamics observed on Facebook and Twitter. Our work differs from [39–41] as we consider distinct classes of users, precisely characterizing influencers and closing the interaction loop between users and the platform by a *feedback* function.

Attempts to validate opinion models are scarce due to several reasons: i) the mapping of opinions into values, ii) an adequate definition of links between agents, and iii) the change of opinion after an interaction is hardly measurable. A recent survey [42] examines the latest research concerning the use of data in opinion dynamics. Usually the approach is either through observational data [43] [9] [39] or controlled sociological experiments [44] [45]. The first type allows us to scale to a large number of users and is more related to our work. The political environment has classically been a florid field for opinion dynamics due to the possibility of attributing a person's opinion to the political orientation of the chosen candidate [43]. Also, a noisy voter model was used to fit data from US elections [46]. The authors of [47] estimate the political ideology of users on Twitter on a single axis (left-right) using the ground truth reconstructed by roll-call votes of members of parliament and their network of followees-followers. In [48], opinions are estimated through a sentiment analysis tool applied to the text of posts on Twitter. Posts are first classified into topics according to their keywords (as we do, see Section 2.4), and then a continuous sentiment score is given between negative and positive. The authors of [49] consider users' opinions in a multidimensional ideological space. Through doc2vec clustering, they identify the axes of 4 dimensions and then map users to these 4-topics opinions. Validation is made through well-known positions of famous Reddit groups. Other recent approaches [50] have used shared *news* on Facebook to assess the extent to which individuals are exposed to opposing views through their (online) friendship relationships, using

users' self-reported ideological affiliations to infer opinions. Recent work [9] [39] has directly used data from online social networks, such as Gab, Facebook, Reddit, and Twitter, to observe the emergence of echo chambers [9] and to validate a model encompassing algorithmic personalization in the process of opinion formation [39].

2.2 Notation

Here we establish the notation that will be used in this and the next chapter of the thesis. We adopt the following vectorial notation. We denote vectors by bold characters, whereas we denote their components with normal-font characters whose subscript is the index in the vector, e.g., $\boldsymbol{a} = \{a_k\}_k$. Lowercase letters indicate parameters and dynamical variables associated with an individual. In general, index i runs over the set of influencers while the index u runs over that of regular users. For those parameters/variables that can be associated with individuals of both classes (either influencers or regular users), the above indices are indicated between superscript parentheses, e.g., $a^{(i)}, a^{(u)}$, to identify the class to which the individual belongs immediately. If necessary, the dependence of variables on other system parameters is made explicit by specifying the independent variables between parentheses, e.g., $\alpha(\cdot,\cdot)$. Italic capital letters denote sets, e.g., $\mathscr{I}$ is the set of all influencers in the population, while $|\mathscr{I}|$ is its cardinality. Capital letters represent outcomes of stochastic experiments whose characteristic parameters are lowercase letters: e.g., $\Omega(\omega(\cdot,\cdot))$. The operator $\mathbb{E}[\cdot]$ represents an expected value, and a bar over a variable, e.g., $\bar{A}$, indicates the average value of the corresponding random variable. Whenever we need to express the probability of an event, we use the notation $\Pr[\cdot]$. We use $\mathbb{1}_{\{\cdot\}}$ for the indicator function. Lastly, time is denoted by t if considered continuous and by n if discrete.

Table 2.1 is intended to provide a comprehensive and self-contained overview of all the symbols presented in this chapter, together with a concise description of what they represent.

Table 2.1 Notation Summary Table

Symbol	Description
$\mathscr{I}$	Set of influencers within the OSN population
$\mathscr{U}$	Set of regular users within the OSN population
$\mathscr{X}$	Set of opinions, in general $\mathscr{X} \subset \mathbb{R}^d$ but typically $[0,1]$
$\mathscr{N}_r$	$\mathscr{N}_r = \{0,1,...,d-1\} \setminus \{r^{(i)}\}$ set of non reference direction
n	Discrete time instant
λ	Aggregate posting rate
$\boldsymbol{x}^{(u)}(n)$	Opinion vector of a regular user at time n
$\boldsymbol{z}^{(u)}(n)$	*Prejudice* vector of a regular user at time n
$\boldsymbol{x}^{(i)}(n)$	Opinion vector of an influencer at time n
$r^{(i)}$	*Reference direction* of influencer i
$s_j^{(i)}$	J-th *secondary direction* for influencer i in $\mathscr{N}_r$
$d_r(n)$	Distance over the reference direction $d_r(n) := \left\| x_r^{(u)}(n) - x_r^{(i)}(n) \right\|$
$d_j(n)$	Opinion distance $d_j(n) := \left\| x_j^{(u)} - x_j^{(i)} \right\|$
$c^{(i)}(n)$	*Consistency* of influencer i
$f^{(i)}$	Posting probability of influencer i at any time n
$\boldsymbol{p}(n) = \{p_i\}_{i \in \mathscr{I}}$	Vector of all influencer's popularities p_i
$\boldsymbol{\pi}(n)$	Vector of normalized popularities $\pi_i(n) := \frac{p_i(n)}{\sum_j p_j(n)}$
$\Omega(\omega)$	Bernoulli random variable, if $\Omega = 1$ the post reaches the user
$\omega(d_r(n), \boldsymbol{\pi}(n))$	*Visibility function* modulating the extent to which posts are seen
ρ	Parameter which controls the level of personalization in ω
$\Theta(\theta)$	Bernoulli random variable, $\Theta = 1$ if the user *likes* the post
$\theta(d_j(n))$	Feedback parameter as a function of $d_j(n)$
Θ_T	Total feedback of the population reached by a post $\mathscr{U}^{post}$
α	First opinion update weight
β	Second opinion update weight (or *inertia*)
$\gamma := 1 - \alpha - \beta$	Third opinion update weight
$\delta := \frac{\alpha}{\alpha+\gamma}$	*Degree of stubbornness*
$F(x,z,t)$	Cumulative regular users distribution, $F(x,z,t) := \Pr[X(t) < x, Z < z]$
$v_x(x,z)$	Velocity of the users/particles in the Fokker-Planck Equation
$\sigma_x^2(x,z)$	Variance of the velocity in the Fokker-Planck Equation

2.3 The Communication Asymmetry Opinion Model

In this section, we present the Communication Asymmetry model in its most general formulation, explain its novel elements, and discuss the limitations due to the inherent simplifications of the modeling process. The proposed model explicitly divides the influential and regular individuals into two classes, that of *influencers* and that of *regular users*. Moreover, it implements the closed-loop between the regular users and the influencers mediated by the platform, which

uses user feedback to update the popularity of influencers, which in turn regulates the way posts are suggested to social media users.

2.3.1 Description of the Model

We propose a continuous opinion model with two interacting classes of agents. Specifically, the population consists of $|\mathscr{U}|$ *regular users* and $|\mathscr{I}|$ *influencers*. This division mimics what happens in real social networks, where a small portion of the population, the influencers, has a much larger number of people following their posts on the online social network. The opinion space is $\mathscr{X} \subset \mathbb{R}^d$, where each dimension represents an uncorrelated topic on which users have a belief. Hence, an opinion is a d-dimensional vector $\boldsymbol{x}^{(u)}(n) \in \mathbb{R}^d$, which evolves as a result of the interaction between a regular user and the influencers on every possible topic. Moreover, we model the tendency toward an a priori opinion $\boldsymbol{z}^{(u)}$, and we refer to it as the *prejudice* of a user. Unless otherwise specified, we will assume that the user's initial opinion corresponds to the prejudice: $\boldsymbol{x}^{(u)}(0) = \boldsymbol{z}^{(u)}$. We assume that the generation of new posts, i.e., messages in the OSN, is a Poisson Point Process (PPP) with intensity λ, where each event of the PPP corresponds to the creation of a new post from an influencer $i \in \mathscr{I}$. The corresponding embedded discrete time will be denoted by the integer $n \in \mathbb{N}_+$, where n is the n-th post. Figure 2.1 illustrates the social media platform's role in *filtering* the posts sent[1] to regular users and receiving *feedback* from those users. These two aspects implement, respectively, a *selective exposure* effect: namely the tendency of both the platform and the users to suggest/access similar content, and a *confirmation bias*, namely the tendency of users to value content that is close to one's point of view (see [9] and resources therein).

Influencers are considered stubborn, i.e., $\boldsymbol{x}^{(i)}(n) = \boldsymbol{x}^{(i)}(0) = \boldsymbol{x}^{(i)} = \boldsymbol{z}^{(i)} \in \mathbb{R}^d$, $\forall n > 0$ and $i \in \mathscr{I}$. As we will show in Section 2.4, each influencer has a main topic of interest on which it publishes most of its posts and typically coincides with the topic it is mainly known for on the OSN. We call it the *reference direction* $r^{(i)} \in \{0, .., d-1\}$ of the influencer. Another parameter characterizing influencer i is its *consistency* $c^{(i)}(n)$, which denotes the (possibly time-varying) probability of

[1] We use the terms *send*, *suggest* and *reach* interchangeably, referring to a post shown to a user by the platform. We will refer to *regular users* simply as *users*. Moreover, the terms *agent* or *individual* indicate a social network user of either class.

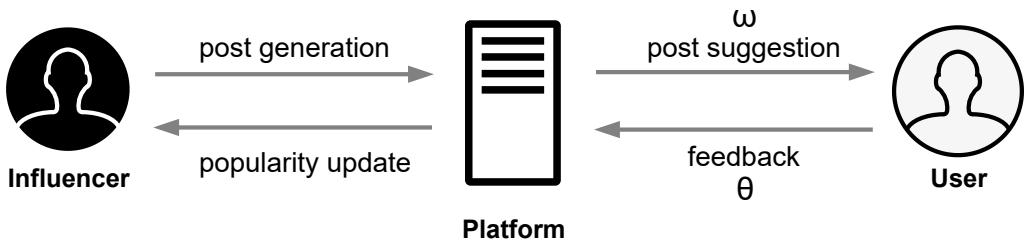


Fig. 2.1 Illustration of the interaction between users, influencers, and social media platforms highlighting the closed loop between user feedback θ and content filtering ω, which will be further characterized below.

posting on the reference direction. High-consistency individuals preferably post on their reference topic. We denote by $f^{(i)}$ the probability that a post is generated by influencer i at any time instant n, with $\sum_{i\in\mathscr{I}} f^{(i)} = 1$. At last, we introduce the *popularity* vector $\boldsymbol{p}(n) \triangleq \{p_i(n)\}_{i\in\mathscr{I}}$, containing the current popularity of all influencers at time n, before the emission of the post at time n. We also introduce the normalized version of this vector $\boldsymbol{\pi}(n) = \{\pi_i(n)\}_{i\in\mathscr{I}}$ where the components are the normalized popularities $\pi_i = \frac{p_i}{\sum_{j\in\mathscr{I}} p_j}$.

The dynamic variables of users, i.e., their opinion $\boldsymbol{x}^{(u)}$, and influencers, i.e., their popularity p_i, are updated upon every post generation according to Algorithm 2. Figure 2.2 gives a schematic representation of the model.

At any time instant n, an influencer i, selected according to distribution $f^{(i)}$, generates a post. The influencer i posts on its reference direction $r^{(i)} = j$ with a probability equal to its consistency $c^{(i)}$. Otherwise, it posts on one of the *secondary* directions $j \in \{0,1,...,d-1\} \setminus \{r^{(i)}\} = \mathscr{N}_r$ according to a given distribution, $\Pr[\ j = k\ ]$ for k in $\mathscr{N}_r$. In the following, we assume this distribution to be uniform over the set of non-reference directions. The emitted post carries the j-th component of the influencer's opinion vector $\boldsymbol{x}^{(i)}$. We suppose posts accurately reflect the influencer's true belief and no noise affects user perception.

To decide whether the post reaches a given user (independently from other users), we extract a Bernoulli random variable Ω with parameter ω. The user receives the message when $\Omega(\omega) = 1$. The parameter ω can be interpreted as a *visibility* function from the influencer's perspective, as it affects the subset of users reached by its posts. In principle, a post can reach any user as the interaction network is a dynamic complete bipartite network between the set of nodes

Algorithm 2: The Communication Asymmetry Opinion Model

Input: $|\mathscr{I}|$ influencers, $|\mathscr{U}|$ users, filtering function ω, feedback function θ
Output: opinion of each regular user $\boldsymbol{x}^{(u)}(n)$, $\forall u$
popularity of each influencer $p_i(n)$, $\forall i$

while *true* **do**
  Select influencer i according to $f^{(i)}$;
  Select a posting direction $j = r^{(i)}$ with probability $c^{(i)}$, otherwise j is selected uniformly on $j \in \{0,..,d-1\} \setminus \{r^{(i)}\}$;
  $p_i(n+1) = p_i(n)$;
  for *each regular user u in the population* **do**
    $x_j^{(u)}(n+1) = x_j^{(u)}(n)$;
    if $\Omega\left(\omega(|x_{r^i}^{(i)} - x_{r^i}^{(u)}|, \pi_i(n))\right) = 1$ **then**
      Get feedback $\Theta\left(\theta(|x_j^{(u)} - x_j^{(i)}|)\right)$;
      if $\Theta = 1$ **then**
        $x_j^{(u)}(n+1) = \alpha z_j^{(u)} + \beta x_j^{(u)}(n) + (1-\alpha-\beta)x_j^{(i)}$;
        Update popularity of i: $p_i(n+1)$ += $1/|\mathscr{U}|$;

$\mathscr{U}$ and $\mathscr{I}$ whose links are defined by $\Omega(\cdot)$ (see Figure 2.3). The *visibility* function ω is a decreasing function of the opinion distance on the reference direction $d_r(n) = |x_r^{(u)}(n) - x_r^{(i)}(n)|$. This dependence embeds the concept of homophily, one of the main drivers of interaction on social networks[11]: individuals with strongly divergent opinions interact less frequently than like-minded individuals. Moreover, ω is increasing in the popularity ratio π_i. The higher the relative popularity of an influencer, the more users it can reach. This posts-users matching process constitutes the *content personalization* we consider in this work (see Remark 2). Note that the *filtering* process for selecting the subset of users who receive the post is based on the opinion distance along the reference direction between each user and the influencer who made the post. This because we expect that influencers mainly attract users whose opinions are similar to their main topic. For instance, politicians primarily attract users interested in the political landscape and with similar orientations. Adopting this distance to perform the user selection couples the dynamics in different directions, which would otherwise evolve independently of each other (see Remark 1).

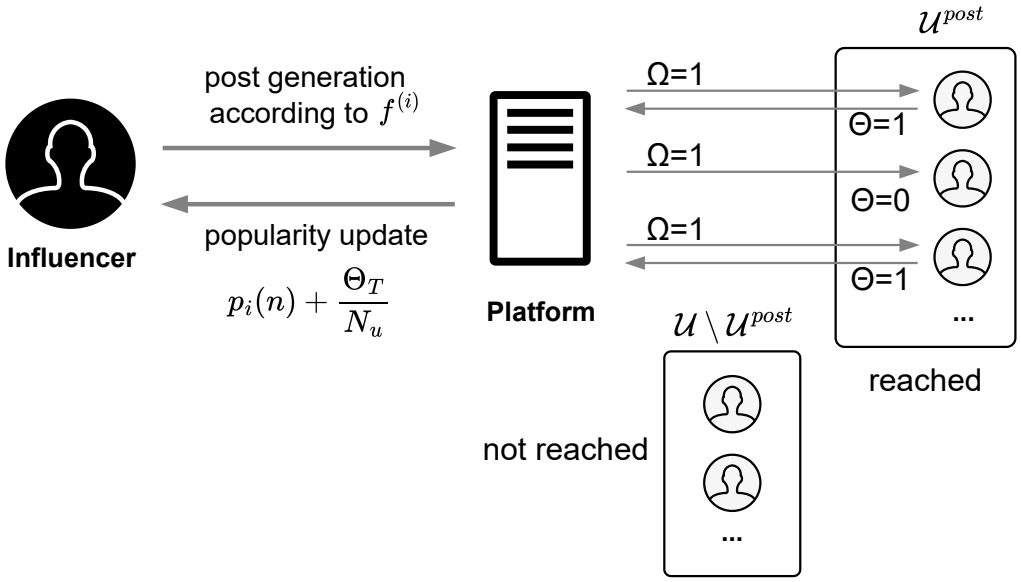


Fig. 2.2 Schematic representation of the model dynamics. The figure highlights the proportions of users who view a particular post $\mathscr{U}^{post}$, i.e., those for which the random variable Ω equals 1. They react with their feedback Θ (e.g., likes), which depends on the opinion distance between them and the influencer i. Then the platform updates the influencer's i popularity.

Users express their *feedback* to a post on the platform through a Bernoulli random variable $\Theta\left(\theta(|x_j^{(u)} - x_j^{(i)}|)\right) \in \{0,1\}$ whose parameter θ depends on the difference in opinion on the *actual* direction j of the contribution. Only posts that receive positive feedback, i.e., $\Theta = 1$, can influence the user's opinion, reflecting the tendency to ignore unappreciated content. The social media platform collects feedback from all reached users to update the popularity p_i of the posting influencer.

The update rule for the popularity of the posting influencer i reads as follows:

$$p_i(n+1) = p_i(n) + \frac{\Theta_T(\theta, \mathscr{U}^{post})}{|\mathscr{U}|} \tag{2.1}$$

$$\Theta_T(\theta, \mathscr{U}^{post}) = \sum_{u \in \mathscr{U}^{post}} \Theta\left(\theta(|x_j^{(u)}(n) - x_j^{(i)}(n)|)\right) \tag{2.2}$$

where $\mathscr{U}^{post}$ is the subset of users who were made aware of the post by the platform, i.e., those for whom $\Omega(\omega)$ takes the value one. The summation in the

formula gives the aggregate feedback of all users who saw the post, which is normalized by the size of the population of regular users $|\mathscr{U}|$ to update the popularity. This normalization is introduced only to avoid excessive popularity growth of the influencers when the number of users becomes large. It does not affect the system dynamics, which depends only on the normalized popularity π_i, which is not affected by the scaling factor $1/|\mathscr{U}|$.

The core of the dynamics is the opinion update rule, which prescribes how the user's opinion changes on the direction j of the post:

$$x_j^{(u)}(n+1) = \begin{cases} \alpha z_j^{(u)} + \beta x_j^{(u)}(n) + \gamma x_j^{(i)} & \text{if } \Omega\left(\omega(d_r, \pi_i)\right) = 1, \Theta\left(\theta(d_j)\right) = 1 \\ x_j^{(u)}(n) & \text{otherwise} \end{cases} \tag{2.3}$$

When the post reaches the user ($\Omega = 1$) who likes it ($\Theta = 1$), then the updated opinion is a convex combination, i.e., $\alpha + \beta + \gamma = 1$, of the current opinion $x_j^{(u)}(n)$, the prejudice $z_j^{(u)}$, and the opinion $x_j^{(i)}$ conveyed by the influencer through the post. The opinion is not updated if the user does not receive the post (Ω=0) or does not like it even if it reaches them ($\Theta = 0$). While it is common in the literature to express the opinion update as the convex combination in Eq. (2.3), we present an equivalent formulation that sheds more light on the meaning of the update parameters, which are practically two:

$$x_j^{(u)}(n+1) = (1-\beta)\left[\delta z_j^{(u)} + (1-\delta) x_j^{(i)}\right] + \beta x_j^{(u)}(n), \quad \delta, \beta \in [0,1] \tag{2.4}$$

(2.4) can be easily derived from (2.3) by noting that $\alpha + \gamma = 1 - \beta$, and defining $\delta \triangleq \frac{\alpha}{\alpha+\gamma}$. Thus, β represents the *inertia* of users, i.e., how slowly they change, and δ their *degree of stubbornness*, i.e., the relative impact of external opinions w.r.t their prejudice.

Remark 1. *The distance on the reference direction drives filtering action because we assume the platform is unaware of the specific topic associated with the post just created. At the same time, homophilic connections between users and influencers primarily depend on opinion similarity on the main topic of discussion. Note the joint effect in the model of the distance between the user's opinion and the influencer's opinion on the reference direction and the distance along the direction defined by the post's topic. Both contribute to determining the likelihood for the user to provide positive feedback to the message.*

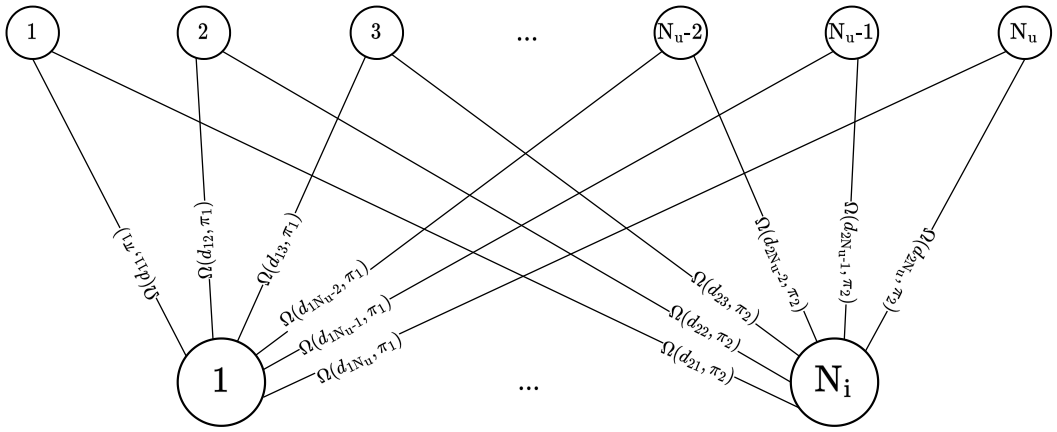


Fig. 2.3 The bipartite dynamic structure of the underlying graph between regular users (top) and influencers (bottom) defined by the visibility function $\Omega\left(d_{iu}(n) = |x_r^{(u)}(n) - x_r^{(i)}(n)|, \pi_i(n) = \frac{p_i(n)}{\sum_i p_i(n)}\right)$ which depends on both the user-influencer opinion distance and the relative popularity of the influencer at a given time instant n.

Remark 2. *Most OSNs have explicit subscriptions to influencers, i.e., the follow mechanism. Our approach does not explicitly represent such long-term relationships. However, by applying the function* $\Omega(\omega(\cdot))$*, we dynamically determine the set of users reached by each influencer. Essentially, followers are regenerated at each post-emission. The resulting network is a dynamic bipartite graph, see Figure 2.3, whose structure reflects a given degree of homophily of users' connections. Typically homophily is one of the elements that mainly influences users' choices when they select individuals to connect to [11]. In particular, for some domains (e.g., product adoption, which is also a good fit for our competing scenario in Section 2.6) homophily has emerged as the key driver governing the structure of the network [51].*

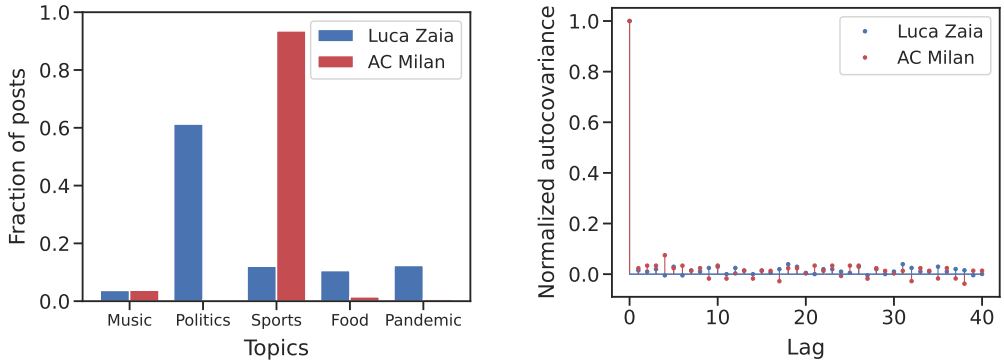
At last, observe that, nowadays, most social media platforms (e.g., Facebook, Instagram, Twitter) do not only offer their users content they explicitly subscribe to but also what they may like. The selection of such users is based on their previous activity on the platform. This mechanism reinforces the homophilic structure of the network and resembles what we are modeling.

Remark 3. *In our framework, regular users are passive, as they merely consume content produced by influencers: this constitutes a rather simplistic assumption. First, users can share the posts they receive, which increases their reach. Secondly,*

users themselves write posts that reflect their opinions, influencing other users. The impact of active users is briefly discussed in Section 2.8.

2.4 Observations from Online Social Networks

This section motivates some of our modeling choices by analyzing real-world social networks. We monitored on Facebook and Instagram the posts of 649 influencers for over 5 years. For a detailed description of the dataset used, see Appendix A. One of the most important features introduced in this work is the concept of *reference direction*, i.e., the main topic an influencer is interested in and on which they publish most of their posts, which is validated here. Moreover, we examine the post-generation process to justify the choice of a Poisson Point Process to describe it.



(a) (b)

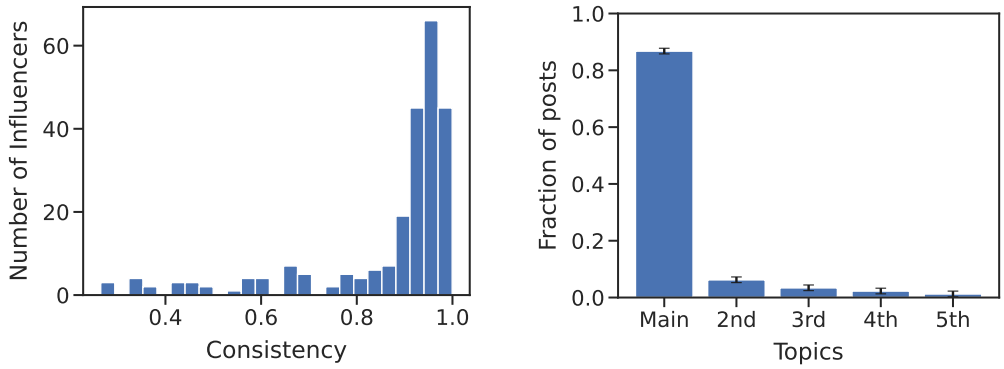
Fig. 2.4 (2.4a) Percentage of labeled posts on each of the considered topics for Luca Zaia, an Italian politician, and AC Milan, an Italian football club. (2.4b)Normalized autocovariance function for both influencers on a *secondary* topic, i.e., music.

2.4.1 The Reference Direction

In this section, we show that influencers prefer to post about a specific topic. We have developed a post classifier that flags posts based on their topic. See Appendix A.1 for details on the classification and filtering process on the data. We should point out that classifying posts on OSNs into topics is not straightforward, and interpreting the results should be done cautiously. First, the range

of possible subjects discussed in a social network is practically countless. For practical reasons, we will only focus on a subset of five topics: Sports, Politics, Food and Cooking, Music, and Pandemics. These can be considered popular and general enough to cover a substantial fraction of the influencers' posts taken into consideration in our dataset.

After classification, we examined the distribution of posts on the topics for each influencer. In Figure 2.4a, we show two example influencers. In these two cases, the influencers have one topic on which they write most of their posts. Luca Zaia, an Italian politician, posts mainly about politics, and AC Milan, a soccer club, discuss sports predominantly. This behavior supports the existence of a reference direction for influencers. Figure 2.5a shows the distribution of the proportion of posts dealing with the main topic of each influencer. Recall that this proportion was called *consistency* in the jargon of our model.



(a) (b)

Fig. 2.5 (2.5a) Distribution of the fraction of posts published on the main topic of interest by the subset of influencers considered in this experiment, i.e., their consistency. (2.5b) The average percentage of labeled posts on each topic, in decreasing order for all the influencers considered. The 95% confidence interval for each average value is reported in the figure.

Most influencers have a clear *reference topic* on which they write more than half of their posts, i.e., with high consistency. Figure 2.5b shows the average per-topic percentage of all influencers in the dataset in descending order, regardless of the specific topic. On average, almost 90% of the posts are in the reference direction. We discovered that influencers with low consistency values are affected by the presence of news outlets in the considered profiles, for which the lack of a sharp main topic is sensible.

2.4.2 Post Generation Independence on Secondary Directions

Users interact in an OSN by posting content (i.e., text, images, videos) and receiving suggestions about what other users of the OSN posted, according to the filtering process set up by the social media platform. We examine the normalized autocovariance[2] between posts on each topic by looking at the chronological sequence of the messages of the individual influencers. We perform this analysis only on *secondary* topics, i.e., those that differ from the influencer's *reference*. We do it since influencers post less frequently on these topics, and it would be easier to detect a bursty behavior pattern (which would not be well captured by the Poisson process). Regarding the main topic, since the consistency of the influencers is generally relatively high, we expect the covariance to be rather small, (see Figure 2.5a). Indeed, the covariance on the main topic tends to zero by construction as $c^{(i)}$ approaches one.

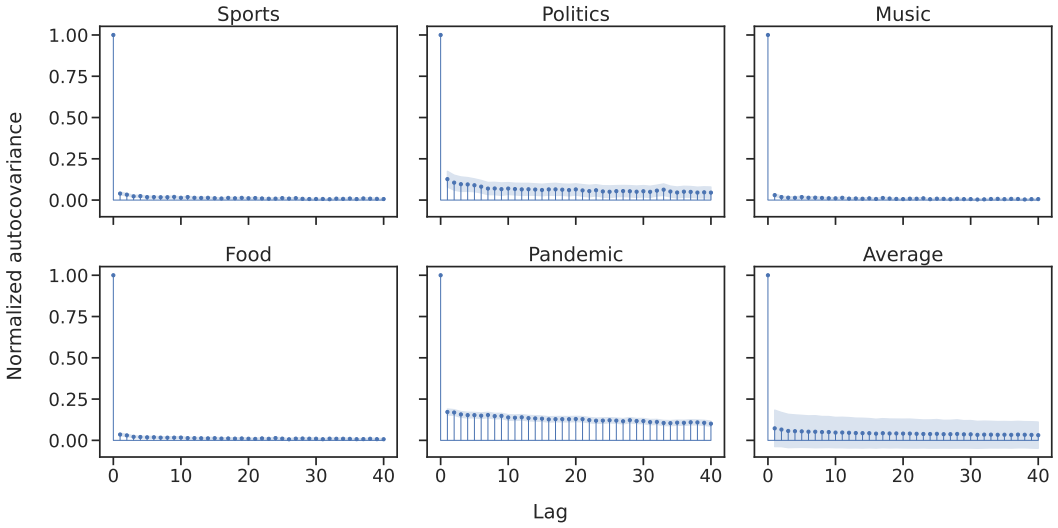


Fig. 2.6 Mean normalized autocovariance values of the post-generation process for each secondary topic of all the influencers. The last plot represents the average value over all topics. The 95% confidence interval is shown in each plot.

In the previous section, we were able to assign a reference direction $r^{(i)}$ to each influencer. Here we look at the time series of the Influencers' labeled posts. For each secondary direction $s_j^{(i)}$, we define an indicator function $\mathbb{1}_{\{post_{label}=s_j^{(i)}\}}$ that takes the value 1 if the post was labeled as $s_j^{(i)}$ and 0 otherwise. For each

[2] Given a wide-sense stationary process $\{X_n\}_n \in \mathbb{Z}$ with average μ and variance σ^2, the normalized autocovariance is given by: $\rho_{XX}(i) = \frac{\mathbb{E}[(X_n-\mu)(X_{n+i}-\mu)]}{\sigma^2}$

influencer, we thus obtain four sequences (recall that we consider five topics in total) of Bernoulli random variables indicating whether a post belongs to that particular direction. We calculated the normalized autocovariance function $a(t)$ for these sequences. Figure 2.4b shows two examples of such functions, limited to 40-time lags, for the profiles of Luca Zaia and AC Milan. The time is discretized, i.e., the actual time between postings is not considered: only the posting events matter. An autocovariance that equals zero everywhere except at $\tau = 0$ would represent uncorrelated samples. In our case, the autocovariance takes moderate values in most cases ($\ll 1$). Therefore, it is reasonable to assume that the post-generation is independent, and a Poisson Point Process is an appropriate choice. Lastly, note that the autocovariance function for the *pandemic* topic takes larger values than for the other topics (see Figure 2.6), suggesting that the samples are weakly correlated. This fact is due to the exceptional public interest in the topic and because the outbreak of the epidemic only interested the last part of the considered time horizon.

2.5 Asymptotic Analysis of the Model

This section is devoted to the analytical study of the model. In particular, results are derived through a mean-field approach obtained by letting the number of users $|\mathscr{U}| \to \infty$. In this situation, we will show that, under mild assumptions, the system converges to a unique steady state, independently from the initial condition. Moreover, in some cases, it is possible to analytically characterize the equilibrium value for the influencers' mean-popularity ratios $\bar{\pi}_i$ as well as users' mean opinion value $\bar{x}(z)$.

Furthermore, transient analysis of the system can be carried out by describing the dynamics of the users through a Fokker-Plank equation. For simplicity, we restrict our investigation to the situation where the opinion space is one-dimensional. However, we remark that it is possible to extend the analysis to the more general case by following the same approach.

2.5.1 Mean-Field Approach

When the number of users grows large, it is convenient to characterize the system state by the users' opinion *distribution* over the space. Moreover, from now on, we will refer to system dynamics over continuous time t.

Let $(X(t), Z(t)) = (X(t), Z)$ be the current position (opinion) and prejudice of a randomly selected user. We introduce the cumulative distribution function $F(x,z,t) = \Pr[X(t) < x, Z < z]$. The corresponding probability density function is $f(x,z,t) = \frac{\partial^2}{\partial x \partial z} F(x,z,t)$. Note that, by hypothesis, there are no dynamics along the z-axes, thus $h(z) = \int_x f(x,z,t)\mathrm{d}x$ does not depend on t and corresponds to the initial distribution of users' prejudice. In Section 2.5.2, we will derive a Fokker-Plank equation for the evolution of the opinion distribution over time and space.

For what concerns the evolution of the popularity of a generic influencer i, recall that we distinguish between its absolute popularity value $p_i(t)$ and the normalized value $\pi_i = \frac{p_i(t)}{\sum_j p_j(t)}$. Influencer's popularities concentrate around their average $\bar{p}_i(t)$ as $|\mathscr{U}|$ grows large, as it can be easily shown. We can write down the equation for the evolution of the mean popularity $\bar{p}_i(t)$:

$$\frac{\mathrm{d}\bar{p}_i(t)}{\mathrm{d}t} = \frac{1}{|\mathscr{U}|} \lambda f^{(i)} \int_x \int_z f(x,z,t)\, \theta\left(|x - x^{(i)}|\right) \omega\left(\bar{\pi}_i, |x - x^{(i)}|\right) \mathrm{d}z\, \mathrm{d}x \tag{2.5}$$

Indeed, the rate at which the popularity of influencer i grows is proportional to its posting rate (term $\lambda f^{(i)}$) times the probability that a generic user at (x,z) provides positive feedback to the post generated at time t (integral term). Moreover, recall from Algorithm 2 that each positive feedback increases the absolute popularity of the influencer by $1/|\mathscr{U}|$.

2.5.2 Fokker-Planck Equation for the Opinion Distribution

The Fokker-Planck (FP) equation [52] is a standard tool that describes the evolution of an asymptotically large population of particles moving over a given domain according to a Brownian motion, which is locally characterized by the instantaneous average velocity and the relative variance. In general, both average and variance may depend on the point and time at which they are evaluated.

Furthermore, the FP approach has been successfully applied in the literature to approximate the evolution of a large population of particles moving according to more general laws. The stochastic process describing the movement of particles is approximated by a Brownian motion which fits the first two moments (average and variance of the velocity). For these reasons, FP description is referred to in the literature as a second-order approximation. Here, we essentially identify each user opinion with a particle moving over the space $\mathscr{X}^2$ (i.e., the set (x,z) with $x,y \in [a,b]$), and characterize by an average velocity $v_x(x,z,t)$ along the x-axes, and a variance $\sigma_x^2(x,z,t)$, in Section 2.5.3. Therefore, the Fokker-Plank equation for the probability density function $f(x,z,t)$ (where $x,z \in [a,b]$) is given by:

$$\frac{\partial f(x,z,t)}{\partial t} = -\frac{\partial v_x(x,z,t) f(x,z,t)}{\partial x} + \frac{1}{2}\frac{\partial^2 \sigma_x^2(x,z,t) f(x,z,t)}{\partial x^2} \tag{2.6}$$

2.5.3 Identification of the Parameters $v_x(x,z,t)$ and $\sigma_x^2(x,z,t)$

To compute $v_x(x,z,t)$, in the continuous-time FP approximation, we assume that for the effect of a post, users/particles reach their new position by moving at a constant speed during the interval ΔT equal to the average time $1/\lambda$ that elapses between the generation of two successive posts. Therefore, assuming that at time t a post is generated by user i, the following equation describes how the opinion of a user with prejudice z evolves from t to $t+\Delta T$:

$$x(t+\Delta T) = \alpha z + \beta x(t) + \gamma x^{(i)}(t)$$

Thus, the increment is:

$$\Delta x(i) = x(t+\Delta T) - x(t) = \alpha(z - x^{(i)}(t)) + (1-\beta)(x^{(i)}(t) - x(t)) \tag{2.7}$$

where $\Delta x(i)$ is the change in position of a user in position x, providing positive feedback to a post of influencer i. We can compute its average velocity as:

$$\begin{aligned}
\mathbb{E}[v_x(x,z,t) \,|X(t)=x, Z=z] &= \mathbb{E}\left[\frac{[X(t+\Delta T) - X(t) \mid X(t)=x, Z=z]}{\Delta T}\right] \\
&= \sum_i \lambda f^{(i)} \Delta T \, \theta\left(|x - x^{(i)}|\right) \omega\left(\bar{\pi}_i(t), |x-x^{(i)}|\right) \frac{\Delta x(i)}{\Delta T} \\
&= \sum_i \lambda f^{(i)} \theta\left(|x - x^{(i)}|\right) \omega\left(\bar{\pi}_i(t), |x^{(i)} - x|\right) \Delta x(i)
\end{aligned} \tag{2.8}$$

where $\theta\left(|x-x^{(i)}|\right)$ is the probability of providing positive feedback (users move only in this case), while $\omega\left(\bar{\pi}_i(t),|x-x^{(i)}|\right)$ is the probability with which a user in x is exposed to a post created by influencer i at time t. Indeed, users only move if they are exposed to the post and provide positive feedback. Note that, to avoid a cumbersome notation, we have omitted the dependency on the time of the distance term $|x-x^{(i)}|$.

The variance of the velocity is given by the relation:

$$\begin{aligned}\sigma_x^2(x,z,t) &= \sum_i \lambda f^{(i)} \Delta T \theta\left(|x-x^{(i)}|\right)\omega\left(\bar{\pi}_i(t),|x-x^{(i)}|\right)\frac{(\Delta x(i)-\mathbb{E}[v_x(x,z,t)]\Delta T)^2}{\Delta T^2} \\ &= \frac{1}{(\Delta T)^2}\sum_i f^{(i)}\theta\left(|x-x^{(i)}|\right)\omega\left(\bar{\pi}_i(t),|x-x^{(i)}|\right)(\Delta x(i)-\mathbb{E}[v_x(x,z,t)]\Delta T)^2\end{aligned}$$

2.5.4 Steady State Analysis

Now we direct our attention to the existence of stationary solutions for the system. Stationary solutions of (2.6) necessarily satisfy:

$$\frac{\partial}{\partial x}\left(-v_x(x,z)f(x,z)+\frac{1}{2}\frac{\partial\sigma_x^2(x,z)f(x,z)}{\partial x}\right)=0$$

where $v_x(x,z)$ and $\sigma_x^2(x,z)$ must be constant over time. This requires the normalized popularities to be static (i,e. $\omega(\cdot)$ to be constant over time). From the previous equation, integrating both sides with respect to x, we get:

$$\left(-v_x(x,z)f(x,z)+\frac{1}{2}\frac{\partial\sigma_x^2(x,z)f(x,z)}{\partial x}\right)=c_0(z) \tag{2.9}$$

where $c_0(z)$ is a uni-dimensional arbitrary in z. Now, observe that, for every z, previous equation is a first-order linear ODE in x, and therefore an explicitly solution for $f(x,z)$ can be obtained:

$$\begin{aligned}f(x,z) = \Big(&c_1(z)\exp(A(x,z)-A(a,z))+ \\ &c_0(z)\exp(-A(x,z))\int_a^x \exp(A(\theta,z))\mathrm{d}\theta\Big)h(z)\end{aligned} \tag{2.10}$$

where

$$A(x,z)=\int_a^x \eta(u,z)\mathrm{d}u \qquad \eta(x,z)=-2\frac{v_x(x,z)-\frac{1}{2}\frac{\partial\sigma_x^2(x,z)}{\partial x}}{\sigma_x^2(x,z)}.$$

Function $c_0(z)$ can be obtained by imposing boundary conditions:

$$\left(-v_x(x,z)f(x,z)+\frac{1}{2}\frac{\partial}{\partial x}\sigma_x^2(x,z)f(x,z)\right)\Big|_{x=a,b}=0. \qquad \forall z$$

which leads to $c_0(z)=0$, while function $c_1(z)$ is determined by imposing the normalization condition:

$$\int f(x,z)\mathrm{d}x=h(z).$$

Observe that when $\sigma_x^2(x,z)\to 0$ and $\frac{\partial\sigma_x^2(x,z)}{\partial x}\to 0$, from (2.9), with $x_0(z)=0$, we obtain that necessarily the mass concentrates around the points for which $v_x(x,z)=0$. Such points, improperly referred to in the following as *equilibrium points*, will be characterized analytically later on.

Turning our attention to popularity dynamics, recall that stationary conditions necessarily imply normalized popularities to be constant over time:

$$\bar{\pi}_i(t)=\bar{\pi}_i \qquad \forall i$$

On the other hand, absolute popularities naturally grow over time. However, the ratio between any two of them (say i,j) must converge to a constant value c_{ij} equal to the ratio of their corresponding normalized popularities:

$$\frac{\bar{p}_i(t)}{\bar{p}_j(t)}=c_{ij}=\frac{\bar{\pi}_i}{\bar{\pi}_j} \qquad \forall i,j\in\mathscr{I}, i\neq j \tag{2.11}$$

Now observe that in stationary conditions the right-hand side. of (2.5) does not depend on time. Therefore (2.5) admits the following trivial solution:

$$\bar{p}_i(t)=\left(\lambda f^{(i)}\int_x\int_z \theta(|x-x^{(i)}|)\omega(\bar{\pi}_i,|x-x^{(i)}|)\mathrm{d}F(x,z)\right)\frac{t}{|\mathscr{U}|}+\bar{p}_i(0) \tag{2.12}$$

Therefore, we meet conditions (2.11) for any $t \geq 0$, iff normalized popularities of influencers $\{\bar{\pi}_i\}_i$ satisfy the following system of equations:

$$\lambda f^{(i)} \int_x \int_z \theta(|x - x^{(i)}|)\omega(\bar{\pi}_i, |x - x^{(i)}|)\mathrm{d}F(x,z) = c\bar{\pi}_i \qquad \forall i, \text{ for some } c \in \mathbb{R}^+$$
$$\text{s.t. } \bar{\pi}_i \geq 0 \text{ and } \sum_i \bar{\pi}_i = 1. \tag{2.13}$$

and the initial condition $\{p_i(0)\}_i$ satisfies (2.11), i.e., $p_i(0) = k\bar{\pi}_i$ for some $k > 0$.

Let

$$k_i(\bar{\pi}_i) \triangleq \lambda f^{(i)} \int_x \int_z \theta(|x - x^{(i)}|)\omega(\bar{\pi}_i, |x - x^{(i)}|)\mathrm{d}F(x,z) \quad \bar{\pi}_i \in [0,1] \tag{2.14}$$

We can show that:

Theorem 1. *Solutions of* (2.13) *always exist whenever* $k_i(\cdot) \in C_1[0,1]$, $k_i(\cdot)$ *is increasing, continuous and strictly concave.*

The proof is reported in Appendix B.1.

We remark that when $k_i(0) > 0 \; \forall i$, the solution is always unique with $\bar{\pi}_i \in (0,1)$. Instead when $k_i(0) = 0$ for some i, the solution is not guaranteed to be unique.

Now, the problem is how to jointly solve for stationary solutions of $\{\bar{\pi}_i\}_i$ and $F(x,z)$. In a schematic way, on the one hand, we have shown that given $\bar{\boldsymbol{\pi}} = \{\bar{\pi}_i\}_i$, and $h(z)$, we can uniquely determine a $F_{\bar{\boldsymbol{\pi}}}(x,z) = \mathscr{H}(\bar{\boldsymbol{\pi}})$, where $F_{\bar{\boldsymbol{\pi}}}(x,z) = \int_{-\infty}^{x}\int_{-\infty}^{z} f_{\bar{\boldsymbol{\pi}}}(y,w)\,\mathrm{d}y\,\mathrm{d}w$ is the opinion distribution of users resulting from fixed influencers' popularities $\bar{\boldsymbol{\pi}}$ (by (2.10)).

On the other hand, under the conditions: $k_i(\cdot) \in C_1[0,1]$, $k_i(\cdot)$ is increasing and strictly concave, $k_i(0) > 0 \; \forall i$, given $F(x,z)$, we can obtain a $\bar{\boldsymbol{\pi}}_F = \mathscr{G}(F(x,z))$ that uniquely corresponds to i (Theorem 1). The existence of a unique fixed point for the joint system of (stationary) users' opinions and influencers' popularities is guaranteed under the condition that the operator $\mathscr{H} \circ \mathscr{G}(\cdot)$ is a contraction over a complete space.

Theorem 2. *Under the assumption that both* $\omega(\cdot,\cdot)$ *and* $\theta(\cdot)$ *exhibit a sufficiently weak dependence on their variables, the operator* $\mathscr{H} \circ \mathscr{G}(\cdot)$ *is a contraction over a complete space, and therefore a unique stationary solution exists.*

The proof is reported in Appendix B.2.

2.5.5 Asymptotic Analysis of the Fluid Limit

Previous theoretical analysis is, unfortunately, non-constructive, meaning that it does not allow for direct computation of stationary solutions of our dynamical system. To complement the previous analysis, we propose a methodology to numerically compute stationary solutions (even in multi-dimensional scenarios) under stricter assumptions. In particular, if the FP approach is a second-order approximation (matching the first to moments of the instantaneous velocity), the approach proposed in this section, and referred to as *fluid limit*, is a first-order approach (matching only the first moment of the velocity and assuming the variance, as well as its spatial derivative, to be negligible). Therefore we expect that the *fluid limit* provides reasonably good predictions when $|\mathscr{U}| \to \infty$, $\lambda \to \infty$, and $1-\beta \to 0$ (keeping $\lambda(1-\beta)$ constant). Previous assumptions, indeed, imply that $\sigma_x^2(x,z) \to 0$ and $\frac{\partial \sigma_x^2(x,z)}{\partial x} \to 0$.

Mean Opinion Assuming Normalized Popularities Convergence

As already observed in Section 2.5.4, recall that, given $\bar{\boldsymbol{\pi}} = \{\bar{\pi}_i\}_i$, the distribution of users with a given prejudice z concentrates around the *equilibrium points*, i.e., points $\bar{x}(z)$ at which $v(x,z)$, as given in Eq. (2.8), is null (i.e. $v(\bar{x}(z),z) = 0$). Therefore, points $\bar{x}(z)$ must satisfy:

$$0 = \sum_i f^{(i)} \omega\left(\bar{\pi}_i, |\bar{x} - x^{(i)}|\right) \theta\left(|\bar{x} - x^{(i)}|\right) \left(\alpha(z - x^{(i)}) + (1-\beta)(x^{(i)} - \bar{x})\right) \tag{2.15}$$

Defining for compactness $d^{i,\bar{x}} = \left|\bar{x} - x^{(i)}\right|$ and recalling $\gamma = 1-\alpha-\beta$, from Eq. (2.15) we get:

$$\bar{x}(z) = \frac{\alpha}{1-\beta} z + \frac{\gamma}{1-\beta} \frac{\sum_{i \in \mathscr{I}} f^{(i)} \omega\left(\bar{\pi}_i, d^{i,\bar{x}}\right) \theta\left(d^{i,\bar{x}}\right) x^{(i)}}{\sum_{i \in \mathscr{I}} f^{(i)} \omega\left(\bar{\pi}_i, d^{i,\bar{x}}\right) \theta\left(d^{i,\bar{x}}\right)} \tag{2.16}$$

We can rewrite this relation in terms of the degree of stubbornness of Eq. (2.4) as $\bar{x} = \delta z + (1-\delta) \frac{\sum_{i \in \mathscr{I}} f^{(i)} \omega(\bar{\pi}_i, d^{i,\bar{x}}) \theta(d^{i,\bar{x}}) x^{(i)}}{\sum_{i \in \mathscr{I}} f^{(i)} \omega(\bar{\pi}_i, d^{i,\bar{x}}) \theta(d^{i,\bar{x}})}$, $\delta \in [0,1]$. By doing so, it should be clear that the assumption $\beta \to 1$ is well-founded, reinforcing the idea that β is associated with the *inertia* of the system (see the end of Section 2.3.1) but does not affect the equilibrium points. This assumption is required to avoid too large oscillations of

users' opinions in response to a single post generated by an influencer, which may reduce the accuracy of our mean-field approximation.

The hypothesis is not restrictive: since β represents the weight individuals give to their current opinion, we can reasonably assume that users do not dramatically change their opinion in response to single post events.

Normalized Popularities Assuming Opinion Convergence

Here we assume that users with prejudice z are concentrated in the opinion point $\bar{x}(z)$, and we look for the stationary popularity ratios $\bar{\pi}_i$. To simplify the expressions, we introduce the quantity $F_i(\bar{\pi}_i) \triangleq \int_z f^{(i)} \omega(\bar{\pi}_i, d^{i,\bar{x}(z)}) \theta(d^{i,\bar{x}(z)}) h(z) \mathrm{d}z$.

Observe that solutions of (2.13) are necessarily of the form:

$$\bar{\pi}_i = \frac{F_i(\bar{\pi}_i)}{\sum_{j \in \mathscr{I}} F_j(\bar{\pi}_j)} \tag{2.17}$$

where c appearing in (2.13) is given by $c = \frac{1}{\sum_{j \in \mathscr{I}} F_j(\bar{\pi}_j)}$. Under the assumption that $\omega(\cdot,\cdot)$ is concave in its first argument (for any choice of the second), Theorem 1 guarantees the existence of such solutions for every choice of function $\bar{x}(z)$. Moreover, even in the more general case, i.e., when $\omega(\cdot,\cdot)$ is non-concave in its first argument, solutions of (2.17) can be found numerically in many cases, through a fixed point iteration method.

To conclude, observe that a pair $(\bar{x}(z), \{\bar{\pi}_i\}_i)$ represents a stationary solution if it jointly satisfies (2.16) and (2.17). The existence of such a solution can be, again, only verified *numerically* through a fixed point approach.

At last, note that, in the special case in which all users have the same prejudice z we can rewrite (2.16) as:

$$\bar{x} = \frac{\alpha}{1-\beta} z + \frac{\gamma}{1-\beta} \frac{\sum_{i \in \mathscr{I}} F_i(\bar{\pi}_i, \bar{x}) x^{(i)}}{\sum_{i \in \mathscr{I}} F_i(\bar{\pi}_i, \bar{x})} = \frac{\alpha}{1-\beta} z + \frac{\gamma}{1-\beta} \sum_{i \in \mathscr{I}} \bar{\pi}_i x^{(i)} \tag{2.18}$$

which provides a direct formula for the mean opinion $\bar{x}$ in terms of the normalized popularities $\bar{\pi}_i$ and the influencers' opinions $x^{(i)}$.

2.6 Monte-Carlo Approach for Model Behavior

This section presents a selection of results obtained through a Monte-Carlo approach. We first vary model parameters to explore the impact of content personalization and influencers' characteristics in a bidimensional opinion scenario. Then, going back to the one-dimensional case, we validate the *fluid limit* approximation we introduced in Section 2.5.5. To what concerns the model's behavior, we focus on the two main dynamic variables of the system: the average opinion $\bar{\boldsymbol{x}}$ of regular users and the normalized popularities $\{\bar{\pi}_i\}_i$ for influencers. Note that the quantities shown in this section, i.e., the pair $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\pi}})$, are empirical averages over multiple runs, and over all the regular users, as far as $\bar{\boldsymbol{x}}$ is concerned. Hence, they can be regarded as empirical, finite system approximations of the quantities defined in the previous section, which refer to the limiting case of an infinite population of users with the same prejudice z, and where $\lambda \to \infty$ and β approaches 1. Moreover, in some cases, we omit the results on average user opinion to save space as it is tightly coupled with the normalized popularities.

Lastly, to facilitate the interpretation of results, we restrict ourselves to the case of two *competing* influencers. We are interested in determining the conditions under which an individual attains higher $\bar{\pi}_i$ than the other. We say that influencer i *wins* over the opponent, which is not anymore visible over the platform, when $\bar{\pi}_i \to 1 \implies \omega(\cdot, \bar{\pi}_j) \to 0 \; \forall j \in \mathscr{I}$. Note that this scenario is by no means trivial and is relevant in various applications, for example, in marketing (e.g., two brands promoting the same product) or in election campaigns (e.g., two candidates of different parties). We provide further details on the scenario in Section 2.6.1. In Section 2.6.2 we show the final opinion distributions of the regular users in a few paradigmatic cases. Then, in Section 2.6.3, we present the behavior as function of publication frequency $f^{(i)}$, and in Section 2.6.4 as function of consistency $c^{(i)}$. To conclude, we validate the first-order fluid approximation comparing it to Monte Carlo results of the system in Section 2.6.6.

2.6.1 Description of the Scenario

The default parameters of our reference scenario are reported in Table 2.2, unless otherwise explicitly stated. As mentioned earlier, we consider the case of two *competing* influencers, i.e., $|\mathscr{I}| = 2$ and a bi-dimensional opinion space. We

assume that $x_j^{(0)} = 0$ and $x_j^{(1)} = 1 \ \forall j \in \{1,2\}$ hence the influencers are placed on two antipodal vertices of the square $[0,1]^2$. We consider influencers with different reference directions $r^{(0)} \neq r^{(1)}$. Furthermore, most regular users initially take a *moderate* position on both topics. More precisely, initial opinions, which coincide with prejudices, are distributed according to a Beta distribution, independently on each axis, with shape parameters $a = b = 10$, see Figure 2.7.

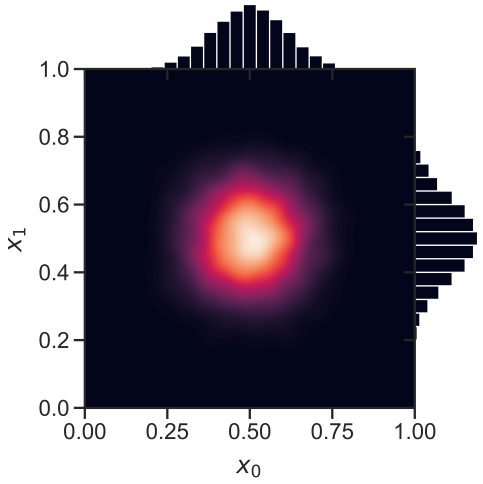


Fig. 2.7 Initial opinion (prejudice) distribution of the regular users, recall $\boldsymbol{x}^{(u)}(0) = \boldsymbol{z}^{(u)}$.

We take as $\omega(\cdot)$ a Gaussian function similar to the *trust* function in [53], but modulated by $\bar{\pi}_i$:

$$\omega(d_r^{i,u}, \bar{\pi}_i) = e^{-\rho \frac{\left(x_r^{(u)} - x_r^{(i)}\right)^2}{\bar{\pi}_i}} \tag{2.19}$$

Here, the coefficient ρ is a parameter that controls the extent to which the social media platform filters content, i.e., which expresses the homophilic degree over the network in a synthetic way. Small values of ρ correspond to *smooth* personalization, i.e., influencers can reach users whose opinion strongly differs from theirs. Conversely, high values of ρ correspond to *sharp* personalization: only close users (in the opinion space) are reachable with non-negligible probability. The function $\theta(\cdot)$ is assumed to be a decreasing, linear function of the opinion difference:

$$\theta(d_j^{i,u}) = 1 - \left| x_j^{(i)} - x_j^{(u)} \right| \tag{2.20}$$

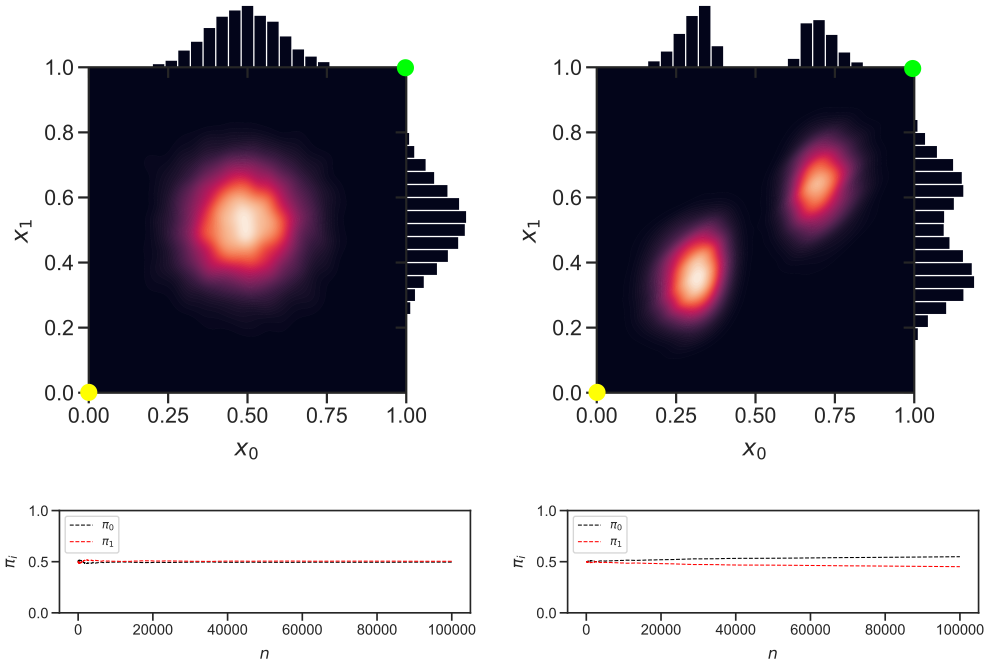
Table 2.2 Parameters and functions shared across experiments

Symbol	Value - Form	Description
$\lvert\mathscr{I}\rvert$	2	Number of influencers
$x_j^{(0)}$	0	Opinion of influencer 0 on direction j
$x_j^{(1)}$	1	Opinion of influencer 1 on direction j
$r^{(0)}$	0	Reference direction of influencer 0
$r^{(1)}$	1	Reference direction of influencer 1
$p_{0,1}(0)$	100	Initial absolute popularity of both influencers
$\lvert\mathscr{U}\rvert$	10000	Number of regular users
N_{iter}	100000	Number of iterations for each simulation
α	0.05	First weight in the updating rule in Eq. 2.3
β	0.93	Second weight in the updating rule in Eq. 2.3
$\theta(\cdot)$	$1-\left\lvert x_j^{(i)}-x_j^{(u)}\right\rvert$	Functional form of the *feedback* function
$\omega(\cdot)$	$e^{-\rho\frac{\left(x_r^{(u)}-x_r^{(i)}\right)^2}{\bar{\pi}_i}}$	Functional form of the *visibility* function
a	10	First parameter of the initial Beta distribution
b	10	Second parameter of the initial Beta distribution
$z_j^{(u)}$	$x_j^{(u)}(0)$	Prejudice coincides with initial opinion

2.6.2 Opinion Configurations for Different Combinations of Reference Directions

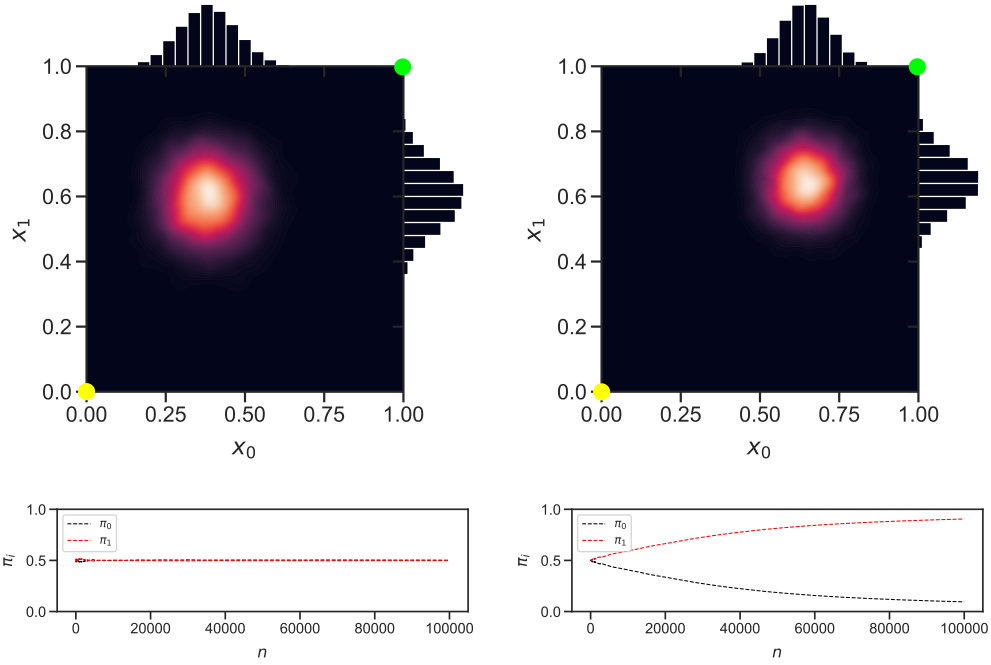
In the following sections, we will focus primarily on the influencer perspective by observing $\bar{\pi}_i$ as a function of their parameters. Here, we present possible final opinion configurations of the users' population, in a symmetric scenario in terms of influencers' characteristics, i.e., frequency of publication $f^{(0)} = f^{(1)}$ and consistency $c^{(0)} = c^{(1)}$. As before, they hold opinions $\boldsymbol{x}^{(0)} = (0,0)$ and $\boldsymbol{x}^{(1)} = (1,1)$. We consider both the case of same (Fig. 2.8a and 2.8b) and different (Fig. 2.8c and 2.8d) reference directions, assessing the impact of *smooth* (Fig. 2.8a and Fig. 2.8c) and strict content personalization (Fig. 2.8b and Fig. 2.8d).

In Figure 2.8a, we observe only a negligible perturbation with respect to the initial distribution (see Figure 2.7). In this case, the platform practically does not filter the content, so every post reaches all users. From a regular user's perspective, individuals are exposed to nearly identical forces, i.e., opposite stimuli from the two influencers, which almost entirely cancel each other out. In Figure 2.8b, the impact of *sharp* personalization is clear: the filtering effect introduced by the platform leads to the emergence of two echo chambers, whose membership is determined mainly by the user's prejudice. Each user reaches an equilibrium



(a) Same reference dir. $r^{(0)} = r^{(1)} = 0$
Smooth personalization $\rho = 5.0$

(b) Same reference dir. $r^{(0)} = r^{(1)} = 0$
Strict personalization $\rho = 5.0$



(c) Different reference dir. $r^{(0)} = 0$, $r^{(1)} = 1$
Smooth personalization $\rho = 5.0$

(d) Different reference dir. $r^{(0)} = 0$, $r^{(1)} = 1$
Strict personalization $\rho = 5.0$

Fig. 2.8 Users' opinion distribution and normalized popularities. The two influencers have same consistency $c^{(0)} = c^{(1)} = 0.8$ and same frequency $f^{(0)} = f^{(1)} = 0.5$. The distributions are obtained as the time average of the opinion distribution in one realization of the process.

point at which the resultant attraction induced by the two influencers is balanced by the attraction exerted by its prejudice. Interestingly, users also tend to cluster in the non-reference direction (x_1 in Fig. 2.8b) and align their opinion with the influencer associated with the echo chamber they end up in. We remark that this is a metastable condition, i.e., the influencers have not yet reached a stable equilibrium, as the gap in the π_i in Fig. 2.8b hints. One of the two influencers will eventually *win* (similar to Fig. 2.8d) but in a much longer time horizon, which may be unreasonable.

Figures 2.8c and 2.8d refer to the case of different reference directions: the two influencers do not primarily compete on the same topic. In Figure 2.8c, it is clear that there is no competition on their reference directions as the two influencers are able to attract users to their *reference opinion*, i.e., $x_{r^{(0)}} = 0$ for $i = 0$ and $x_{r^{(1)}} = 1$ for $i = 1$. This is a particularly relevant case, whose occurrence is linked indissolubly to the newly introduced concept of reference direction. In the last scenario, Figure 2.8d, the influencer $i = 1$ *wins*, i.e., $\bar{\pi}_1 \to 1$, which brings public opinion closer to its belief $\boldsymbol{x}^{(1)}$ on both issues. The final users' opinion does not coincide with $\boldsymbol{x}^{(1)}$ because users are anchored by their prejudice. Note that here *sharp* personalization leads to a situation where only one individual monopolizes the public scene. To better understand the dynamics, we simulated this scenario with 10 different simulator seed selections: 5 times influencer $i = 1$ won, 3 times influencer $i = 0$ won, and in 2 cases, the system did not reach full convergence after $N_i = 100000$ iterations. The nature of the equilibrium point $\left(\pi_0 = \frac{1}{2}, \pi_1 = \frac{1}{2}\right)$ appears to be *unstable*. Stochastic fluctuations of the system state bring it to one of the two *asymptotically stable* configurations: $(\pi_0 = 1, \pi_1 = 0)$, where $i = 0$ wins, or $(\pi_0 = 0, \pi_1 = 1)$, where $i = 1$ wins.

2.6.3 Behavior as a Function of the Frequency of Publication

The frequency of publication $f^{(i)}$ is one of the basic parameters that characterize influencers. The higher $f^{(i)}$, the higher the *structural advantage* of the influencer because it more frequently reaches users through posts, attracting them to its own opinion. In this section, we examine the value of mean normalized popularity $\bar{\pi}_0$ as a function of $f^{(0)}$. Note that in the case of two influencers, $f^{(1)} = 1 - f^{(0)}$. We performed this experiment by fixing the consistency of the two influencers:

$c^{(0)} = c^{(1)} = 0.8$, which is approximately the average consistency observed on real-world data (Figure 2.5a).

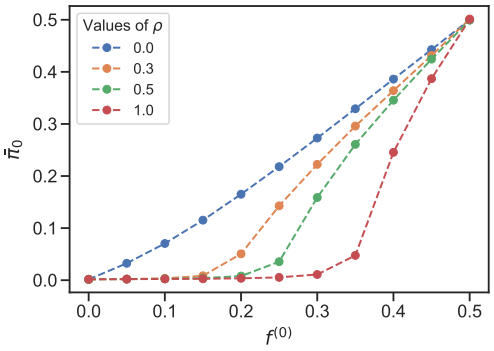
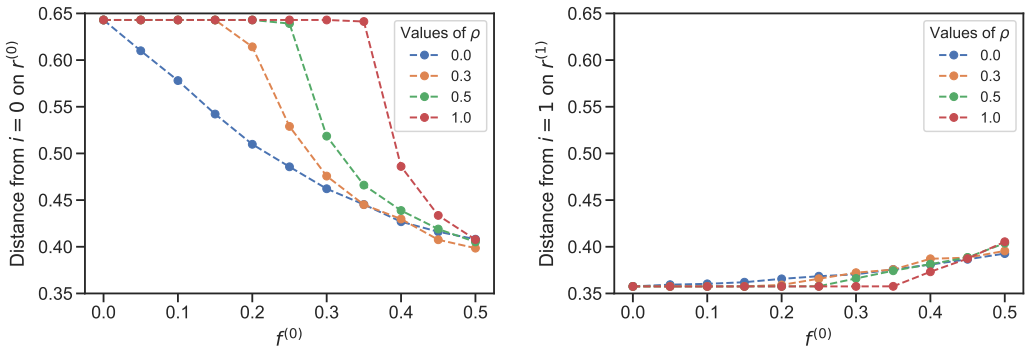


Fig. 2.9 Popularity ratio $\bar{\pi}_0$ of influencer 0 as function of the publication rate $f^{(0)}$. Each point is obtained by averaging over 100 time samples and 10 different process realizations. We consider different levels of personalization by varying the parameter ρ. The two influencers have the same consistency $c^{(0)} = c^{(1)} = 0.8$. Recall that $f^{(0)} = 1 - f^{(1)}$ and the identity of the influencer is arbitrary. This symmetry allows us to readily infer the behavior of the normalized popularity $\bar{\pi}_0$ for $f^{(0)}$ in the range of $[0.5, 1.0]$.



(a) (b)

Fig. 2.10 Opinion distance of the influencers' opinion on their reference direction $x^{(i)}_{r^{(i)}}$ and the average opinion of the regular users' population on the same direction $\bar{\boldsymbol{x}}_{r^{(i)}}$. Various degrees of personalization are considered, tuning the parameter ρ. The setting is the same as that of Figure 2.9.

In Figure 2.9, we consider different levels of personalization by varying the parameter ρ in the exponent of the visibility function ω. We see that the higher the degree of personalization (i.e., the higher the value of ρ), the lower the normalized popularity of influencer $i = 0$, for any given $f^{(0)}$. This result suggests that

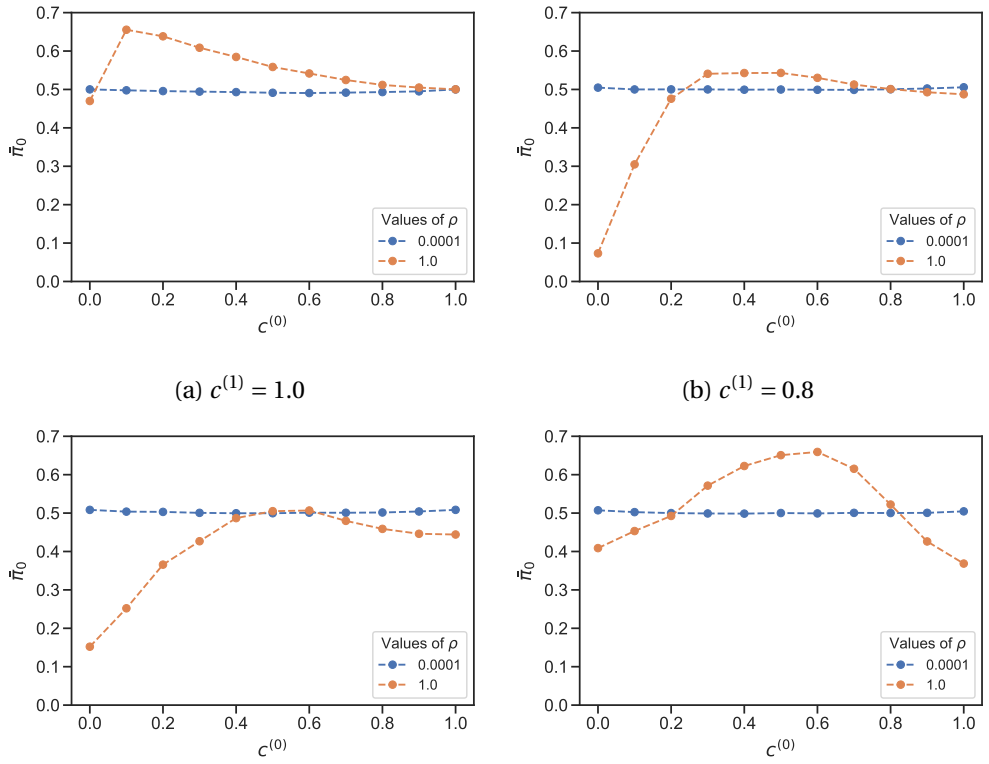
algorithmic personalization favors the *structurally advantaged* individual, i.e., the one with higher $f^{(i)}$. This mechanism, in turn, leads to more radical positions in the population of regular users, as the platform preferentially exposes them to the belief of the advantaged influencer. Figure 2.10 clearly shows this behavior. Note that for high values of ρ, the average user opinion exhibits a significant bias toward the structurally advantaged influencer. Such bias persists up to a critical value of posting frequency. For example, when the personalization parameter is $\rho = 1.0$, the critical posting frequency value is roughly 0.35; when the personalization parameter is $\rho = 0.5$, the critical posting frequency value is approximately 0.25. Below this critical threshold the advantaged influencers *wins*, i.e., its normalized popularity $\bar{\pi}_i$ approaches 1, completely shadowing the opposing influencer. In Figure 2.10b, the opinion variation is limited since $f^{(1)} = 1 - f^{(0)} > 0.5$ implies that $i = 1$ exerts a strong influence over $r^{(1)}$ and suffers little competition from $i = 0$ (as $c^{(0)} = c^{(1)} = 0.8$ and $r^{(0)} \neq r^{(1)}$). For instance, for $f^{(0)} = 0$ the opinion values are those admitted by users' prejudice and a single *winning* influencer (similar to what we discuss 2.6.6). Fig. 2.10a and 2.10b provide complementary information. Moreover due to the aforementioned symmetry, the reader can easily understand how the system would evolve for $f_0 \in [0.5, 1]$.

We argue that content filtering in OSN potentially threatens opinion diversity. This premise is inextricably linked to the goal of usage maximization [41] pursued by the social media platform. Indeed, many platforms indeed prefer to suggest just *similar* content rather than exposing individuals to radically different opinions, hence often avoiding the so-called *serendipity*.

2.6.4 Behavior as a Function of the Consistency

In Section 2.4.1, we showed the existence of a reference direction for real influencers. Here, we investigate the impact on dynamics of the extent to which an influencer publishes on its reference direction, i.e., its consistency $c^{(i)}$. In this experiment, we consider two influencers with the same posting frequency $f^{(0)} = f^{(1)} = 0.5$, different reference directions $r^{(0)} = 0$, $r^{(1)} = 1$, and we let $c^{(0)}$ vary while keeping $c^{(1)}$ fixed. We report plots for a few choices of $c^{(1)}$ since the popularity pattern as a function of $c^{(0)}$ depends on the characterization of the competing influencer $i = 1$. Figure 2.11 shows that consistency does not significantly affect the normalized popularities when personalization is *smooth* ($\rho = 0.0001$). In con-

trast, it becomes relevant when the platform applies *sharp* personalization to the content ($\rho = 1$). We consider two cases. First, that of high $c^{(1)}$ (Figs 2.11a and 2.11b), which is in line with the empirical evidence of Section 2.4.1. Second, the low-consistency scenario (Figs 2.11c and 2.11d) which has an interesting interpretation that goes beyond the scope of this chapter and is briefly discussed in Section 2.8.2.



(a) $c^{(1)} = 1.0$ (b) $c^{(1)} = 0.8$

(c) $c^{(1)} = 0.5$ (d) $c^{(1)} = 0.2$

Fig. 2.11 Popularity ratio $\bar{\pi}_0$ of i_0 as function of its consistency $c^{(0)}$, while considering $f^{(0)} = f^{(1)} = 0.5$ and keeping fixed the consistency of the second influencer at (2.11a) $c^{(1)} = 1.0$, (2.11b) $c^{(1)} = 0.8$, (2.11c) $c^{(1)} = 0.5$, and (2.11d) $c^{(1)} = 0.2$. The two colors represent two different levels of personalization (i.e., *smooth* and *sharp*). Each point is obtained by averaging over 100 time samples and 10 different realizations of the process.

In Figure 2.11a, $c^{(1)} = 1$ means that influencer $i = 1$ posts exclusively in the reference direction $r^{(1)}$. The corner cases, in which both influencers post all their posts in one direction, are i) $c^{(0)} = 1 - c^{(1)}$, where both post on $r^{(1)}$ but $i = 1$ has a slight advantage as it is posting on its reference direction where filtering occurs, and ii) $c^{(0)} = c^{(1)}$, which is a symmetric scenario. This dichotomy is also found

in Fig. 2.11b, 2.11c and 2.11d. Whenever the orange curve approaches $\bar{\pi}_i = 0.5$ and $c^{(0)} = c^{(1)}$ or $c^{(0)} = 1 - c^{(1)}$, the influencer with the largest share of posts in the reference direction has a slight advantage. Again in Fig. 2.11a, if $0 < c^{(0)} < 1$, the influencer $i = 0$ posts in both directions and does not face competition over $r^{(0)}$. Therefore, $i = 0$ attracts users towards its reference opinion $x_0^{(0)}$, which in turn increases the chance of reaching users while competing with the other influencer in the non-reference direction (content filtering is performed with respect to the distance on the reference direction). In this rather extreme case, the lower the consistency $c^{(0)}$, the higher the proportion of posts on $r^{(1)}$, and the higher the final value of $\bar{\pi}_0$ as more posts compete for users' attention with $i = 1$. The other scenarios are not as easy to interpret. However, all 2.11a2.11b 2.11c 2.11d are consistent in pointing out that influencer $i = 0$ has a *structural advantage* roughly when its consistency is $\min\left(c^{(1)}, 1 - c^{(1)}\right) < c^{(0)} < \max\left(c^{(1)}, 1 - c^{(1)}\right)$.

Figure 2.11 suggests that a value of consistency around 0.5 is nearly optimal for any value of the opponent's $c^{(1)}$. This observation reflects the natural tendency of people to seek varied content. We evaluated $\arg\max_{c^{(0)}} \bar{\pi}_0(c^{(1)})$, and this is indeed true for $c^{(1)} < 0.8$. However, for high values of $c^{(1)}$, $i = 0$ is better off reducing its consistency, i.e., post less on its reference direction $r^{(0)}$ and more on the reference topic $r^{(1)}$ of the opponent (recall $r^{(0)} \neq r^{(1)}$). This behavior points towards another potential hazard of content personalization. If influencer $i = 0$ is well-known on a platform as it deals with topic $r^{(0)}$ and starts posting massively on the other topic $r^{(1)}$ ($c^{(0)}$ drops low), it will gain an advantage over the opponent. This is because the platform filters posts considering the reference direction $r^{(0)}$ where $i = 0$ faces little competition from $i = 1$ (since $c^{(1)}$ is high) and therefore can attract users more easily over that topic. This results in an increase in popularity, allowing it to reach more users. This is an indication of how an influencer can leverage its importance in the reference direction to attract users in a non-reference topic favored by content personalization.

2.6.5 Behavior as Function of the Updating Weights

The behavior of the system depends not only on the characteristics of the influencers and the composition of public opinion but also on the parameters controlling the opinion update rule in equation (2.3). The update is a convex combination of the prejudice, the current opinion, and the opinion conveyed

by the post. We chose to hold fixed the weight β (inertia) and vary the *degree of stubbornness*, $\frac{\alpha}{\alpha+\gamma}$.

We considered an unbalanced scenario in which influencer $i = 1$ has a structural advantage, i.e., $f^{(1)} = 0.7 > f^{(0)}$. Figure 2.12 again shows that personalization favors the structurally advantaged individual (consistent with Section 2.6.3). Note that the x-scale is logarithmic to highlight the sudden drop of $\bar{\pi}_0$ for $\frac{\alpha}{\alpha+\gamma} \approx 10^{-3}$ (corresponding to small values of α) when *sharp* personalization is applied. The shape of the two curves is quite similar, only the decrease is observed at different values of $\frac{\alpha}{\alpha+\gamma}$. *Smooth* personalization allows the coexistence of influencers on the whole domain, while with *sharp* personalization, for a wide range of parameters, influencer $i = 1$ *wins*.

In both cases, there is an initial phase (for low values of $\frac{\alpha}{\alpha+\gamma}$) in which the two influencers coexist, and this is followed by a drop in the normalized popularity of the disadvantaged influencer. This can be explained by the fact that small values of $\frac{\alpha}{\alpha+\gamma}$ imply that a negligible weight is given to the prejudice, and therefore regular users concentrate around the two influencers' opinions on their reference direction. This can be easily confirmed by looking at the final opinion configuration of users, who concentrate in the upper corners of the opinion space (around $[0,1]$ and $[1,1]$). This is because the influencer $i = 1$, whose opinion is $x^{(1)} = [1,1]$ is stronger than the other in terms of popularity and is able to pull users along its non-reference direction as well. We remark that when users are very close in opinion to a particular influencer, it is difficult for the other to persuade them, as the probability of this happening is proportional to the product $\omega \cdot \theta$, both of which are a function of opinion distance. In these scenarios, the distance from the *further* influencer is $d_j \approx 1$, which drastically reduces the probability of reaching the users. Thus, as long as $\frac{\alpha}{\alpha+\gamma}$ is small enough, both influencers can build their user base. These situations represent rather degenerate cases where the population almost disregards their prejudice in favor of the opinion conveyed by the post. It might be interesting to consider users with varying degrees of *volatility* who are able to pull along the opinion of their neighborhood.

As the degree of stubbornness increases, users are more entrenched in their prejudice and therefore no longer concentrate on a small neighborhood of the influencer's opinions. This favors the structurally advantaged influencer, as the other (i.e., $i = 0$) is unable to build its user base because users do not get close

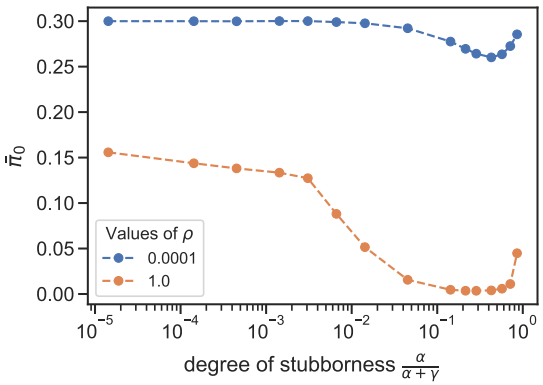


Fig. 2.12 Normalized popularity $\bar{\pi}_0$ as a function of the degree of stubbornness $\frac{\alpha}{\alpha+\gamma}$, the points are obtained considering 50 realizations of the process and averaging over 100 discrete time instants. Again, smooth ($\rho = 0.0001$) and strict ($\rho = 1$) personalization are considered.

enough to it (see Figure 2.12 for $i = 1$ and $\rho = 0.0001$ we have $\bar{\pi}_0 \to 0$). The subsequent rise in $\bar{\pi}_0$ depends on the fact that when the degree of stubbornness approaches one, users give importance only to their prejudice, and therefore they do not deviate too much from their initial position. As a consequence, it can not trigger the positive feedback between users' opinions and influencers' popularity that leads to the complete victory of one influencer.

2.6.6 Validation of the Fluid Limit Approximation

In this last section, we compare predictions of the simplified *fluid limit* with Monte-Carlo simulations of the stochastic model described in Algorithm 2. We consider the one-dimensional setting, as in Section 2.5, and only two two *competing* influencers. A similar analysis could be performed in scenarios with any number of influencers located anywhere in the opinion space. However, this would be computationally more challenging as there may be multiple stationary points, each with its own basin of attraction.

Let us preliminarily specialize the equations presented in Section 2.5.5 for the mean opinion $\bar{x}(z)$ (Eq. (2.18)) and the normalized popularities $\bar{\pi}_i$ (Eq. (2.17)). Note that for $|\mathscr{I}| = 2$, $\bar{\pi}_0 = 1 - \bar{\pi}_1$, so it is sufficient to study $\bar{\pi}_1$.

As for the mean user opinion $\bar{x}(z)$, Eq. (2.18) allows us to write the asymptotic mean as a function of $\bar{\pi}_1$ and the opinions of the two influencers $x^{(0)}, x^{(1)}$:

$$\bar{x}(z) = \frac{\alpha}{1-\beta} z + \frac{\gamma}{1-\beta} \left[(1-\bar{\pi}_1)\, x^{(0)} + \bar{\pi}_1 x^{(1)} \right] \tag{2.21}$$

Substituting the functional forms of the *visibility* ω and *feedback* θ as given in Eq. 2.19 and 2.20, into equation (2.17), we obtain the following expression for the normalized popularity $\bar{\pi}_1$:

$$\bar{\pi}_1 = \frac{f^{(1)} e^{-\rho \frac{\left(x^{(1)} - \bar{x}\right)^2}{\bar{\pi}_1}} \left(1 - |x^{(1)} - \bar{x}|\right)}{\sum_{i \in \{0,1\}} f^{(i)} e^{-\rho \frac{\left(x^{i} - \bar{x}\right)^2}{\bar{\pi}_1}} \left(1 - |x^{(i)} - \bar{x}|\right)} = f(\bar{\pi}_1, \bar{x}) \tag{2.22}$$

Moreover, if we combine the above expression with equation (2.21) for $\bar{x}$, we get $\bar{\pi}_1 = f(\bar{\pi}_1)$, which can be solved numerically through a fixed-point approximation (FPA) (a graphical representation is shown on Fig. 2.14). The outcome of this FPA and the corresponding simulation results are compared in Figure 2.13.

The combination of equations (2.21) and (2.22) cannot be solved in closed form in the general case. However, there are at least two scenarios in which this is possible, separately considered in the following subsections.

When an Influencer "Wins"

We consider an influencer a "winner" if its normalized popularity $\bar{\pi}_i$ approaches 1. Suppose that the influencer whose opinion is $x^{(1)} = 1$ *wins*, then $\bar{\pi}_1 \to 1$. This implies $\bar{\pi}_0 \to 0$ and thus $\omega \to 0^+$: the influencer with $x^{(0)} = 0$ is seen by a negligible fraction of users and in practice, only influencer $i = 1$ remains visible. Note that in the extreme case in which influencer 1 wins, users see only $x^{(1)}$, and asymptotically all users move towards it. In this case, the final opinion $\bar{x}(z)$ can be easily calculated with a recursion of the update rule (2.3):

$$x^{(u)}(n) = \sum_{i=0}^{n} \beta^i \left(\alpha z + \gamma x^{(1)} \right) + \beta^n x(0)$$

For $n \to \infty$ and considering $\beta < 1$ (the case $\beta = 1$ coincides with the trivial case where users remain fixed at their initial opinion) we get:

$$x^{(w)} = \frac{\alpha}{1-\beta} z + \frac{\gamma}{1-\beta} x^{(1)}, \tag{2.23}$$

which is in agreement with (2.21) if one sets $\bar{\pi}_1 = 1$. This corresponds to one of the extreme cases that we will use later to examine the model behavior as a function of the personalization parameter ρ. It should be noted that this construction relies on the knowledge of the *winning* influencer, which is unknown in advance. However in the fluid limit, we expect that the winning influencer, if any, is the one that has a structural advantage over the others at the beginning (e.g., a higher posting rate $f^{(i)}$, see Figure 2.9).

Constant Personalization Function

The other extreme case we consider is the one in which $\rho = 0$. In this case, the personalization function ω no longer depends on $\bar{\pi}_i$, and it is easy to see from Table 2.2 that it returns $\omega \equiv 1$. Moreover, we consider $x^{(1)} = 1, x^{(0)} = 0$, which further simplifies (2.21). The above formulas (Eq. 2.22 and Eq. (2.21) can then be solved in closed form. In particular, equation (2.22) for the normalized popularity $\bar{\pi}_1$ becomes:

$$\bar{\pi}_1 = \frac{f^{(1)}\left(q + m\bar{\pi}_1\right)}{f^{(0)}\left(1 - (q + m\bar{\pi}_1)\right) + f^{(1)}\left(q + m\bar{\pi}_1\right)}$$

where $m \triangleq \frac{\gamma}{1-\beta}$ and $q \triangleq \frac{\alpha}{1-\beta} z$ for compactness. This leads to a second-order equation which can be easily solved for $\bar{\pi}_1$:

$$\bar{\pi}_1^2 \, m(f^{(1)} - f^{(0)}) + \bar{\pi}_1 \left[f^{(0)}(1-q) + f^{(1)}(q-m) \right] - f^{(1)} q = 0 \tag{2.24}$$

Simulation, Fluid Limit and Fixed-point Approximation

Finally, we can compare the analytical results derived in Section 2.5 with Monte-Carlo simulations of the model. We provide numerical and graphical solutions of equation (2.22), shedding light on the impact of the algorithmic personalization performed by the platform. The scenario is a simplified version of that described in Section 2.6.1. The main difference is that here we consider a one-

dimensional opinion space $[0,1]$, with all users having the same prejudice, i.e., $z^{(u)} = z = 0.4, \forall u \in \mathcal{U}$ matching their initial opinion $x^{(u)}(0)$. The two *competing* influencers have opinions at the extremes of the domain ($x^{(0)} = 0$, $x^{(1)} = 1$), and their posting frequencies are $f^{(1)} = 0.7$ and $f^{(0)} = 0.3$, i.e., influencer $i = 1$ has a *structural advantage* over influencer $i = 0$. Note that in a one-dimensional space, the reference direction $r^{(i)}$, and hence the consistency $c^{(i)}$, lose their significance. To avoid obtaining trivial results in which influencer 1 obviously wins, regular users are initially placed closer to the disadvantaged influencer $i = 0$.

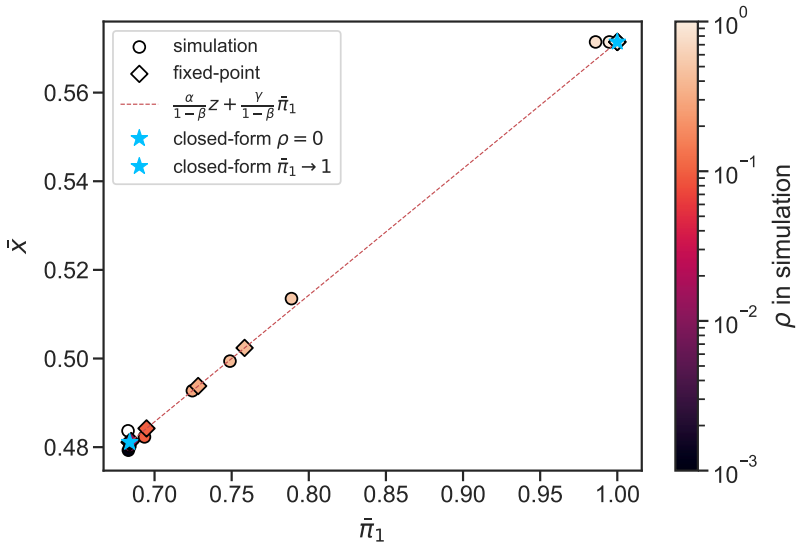


Fig. 2.13 Comparison between analytical results, including the exact extreme points calculated in 2.6.6 and 2.6.6, and the linear relationship between $\bar{x}$ and $\bar{\pi}_1$ according to Equation (2.21). Diamonds represent the fixed-point approximation for the solution of equation (2.22). Simulation results of the stochastic dynamics, represented by circles, were obtained by averaging 100 realizations of the process as described by Algorithm 2. We consider a scenario in which $\alpha = 0.05, \beta = 0.93$, with two influencers at the extremes of the domain, with $f^{(0)} = 0.3, f^{(1)} = 0.7$ and the same initial absolute popularity $p_0 = p_1 = 100$. Numerical values from simulation and fixed-point approximation are reported in the table alongside the plot in Fig. 2.14.

Comprehensive validation and comparison of the approaches used to obtain the system equilibria are shown in Figure 2.13. First, the stochastic model described by Algorithm 2 is "simulated" by obtaining 100 different sample whose length is 500000 elementary steps. The variables of interest $\bar{x}(z)$ and $\bar{\pi}_1$ are obtained by averaging the process over both discrete times steps n and sample paths and are represented by circle marks. Second, equation (2.21), which is a special-

ization of (2.18) obtained from the fluid limit, indicates that the state of the system lies on a line in the plane $\bar{\pi}_1, \bar{x}$ (dashed line in Figure 2.13). Third, the extreme cases of the model analyzed above, for which we derived a closed-form solution, are represented by star-like marks. Lastly, diamonds are solutions of Eq. (2.22) employing the fixed-point approximation (FPA).

Simulation and FPA

ρ	$\bar{\pi}_1$ *SIM*	$\bar{\pi}_1$ *FPA*
0.0	0.682	0.684
0.001	0.683	0.684
0.01	0.684	0.685
0.1	0.693	0.695
0.3	0.725	0.728
0.4	0.749	0.758
0.5	0.789	0.999
0.8	0.986	0.999
1.0	0.995	1.0

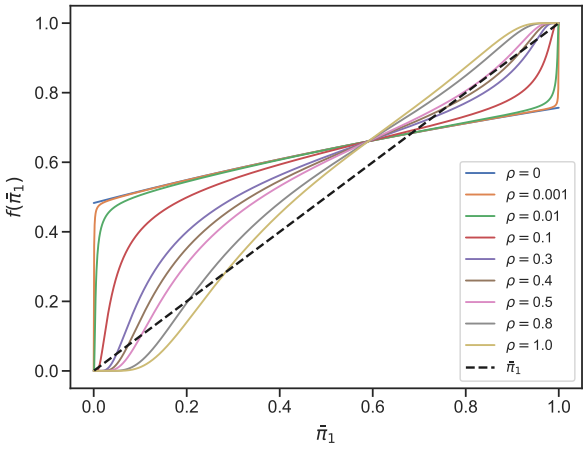


Fig. 2.14 Graphical solution of $\bar{\pi}_i = f(\bar{\pi}_i)$, (2.22). Stable solutions correspond to intercepts between $f(\bar{\pi}_i)$ and the bisector, such that $f'(\bar{\pi}_i) < 1$. We observe that non-trivial solutions (i.e., solutions in which both influencers survive) exist, roughly in the interval $[0.7, 0.8]$, provided that ρ is not too large (i.e., $\rho < 0.5$). For $\rho > 0.5$, the only stable solution is $\bar{\pi} = 1$. This explains the results in Figure 2.13. Simulation results reported in the Table, confirm the validity of the analytical predictions (FPA).

We observe that, for given ρ, simulation marks match well with analytical marks. The only exception is for $\rho = 0.5$, for which simulations provide $\bar{\pi}_1 \approx 0.79$, whereas the analysis provides $\bar{\pi}_1 \approx 1$ (see also the table on Fig. 2.14). This mismatch is due to the fact that $\rho = 0.5$ is close to a *phase transition*, at which the system switches from a regime in which two stable solutions exist (in particular, one in which both influencers survive) to a regime in which influencer $i = 1$ wins. In such a situation, the population is exposed to the opinions of a single individual, hindering diversity on the social platform. This behavior is better illustrated in Fig. 2.14, where the curve corresponding to $\rho = 0.5$ is almost tangent to the bisector. It should be noted that the empty region in Figure 2.13 is directly related to this behavior since no stable solutions can exist for those values of $\bar{\pi}_i$. In fact, there is no *stable* intersection with the bisector in Figure 2.14 in the corresponding interval.

2.7 Online Social Network Data

This section examines data collected from Facebook and Instagram social networks (see Appendix A) and compares the observed behavior with some of the findings of our Communication Asymmetry model.

2.7.1 Correlation Between Posting Frequency and Popularity

In previous sections, especially in Section 2.6, we discussed *structural advantage* from the influencer's point of view. A key advantage parameters is the publication frequency $f^{(i)}$: the higher $f^{(i)}$, the greater the advantage (see Figure 2.9). In this section, we attempt to validate this finding by correlating the frequency of publication of influencers with their popularity growth, using the total number of *followers*, i.e., the number of people subscribed to the *profile*, as a proxy for popularity. We consider temporal sequences from Instagram on a sample set of 110 influencers.

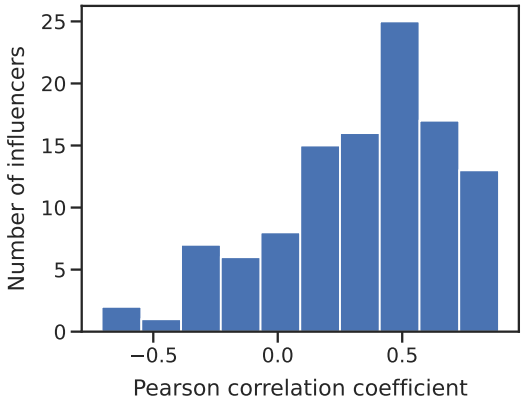


Fig. 2.15 Distribution of the Pearson correlation coefficient between the monthly number of posts and popularity growth (in terms of number of *followers*).

For each influencer, we considered a temporal granularity of one month, determined the number of posts during this period, and calculated the relative change in the number of *followers* considering the values at the beginning and end of the interval. Then for each user, we calculated the Pearson correlation coefficient between the number of posts and the relative variation of followers in the month. In Figure 2.15, we show the distribution of these correlation coefficients. Results suggest that there exists, in general, a positive correlation between the two quantities,

i.e., influencers with aggressive posting habits tend (but not always) to get more followers, which likely favors them when in competition with other influencers on social media platforms. This is consistent with the model predictions shown in Section 2.6.3.

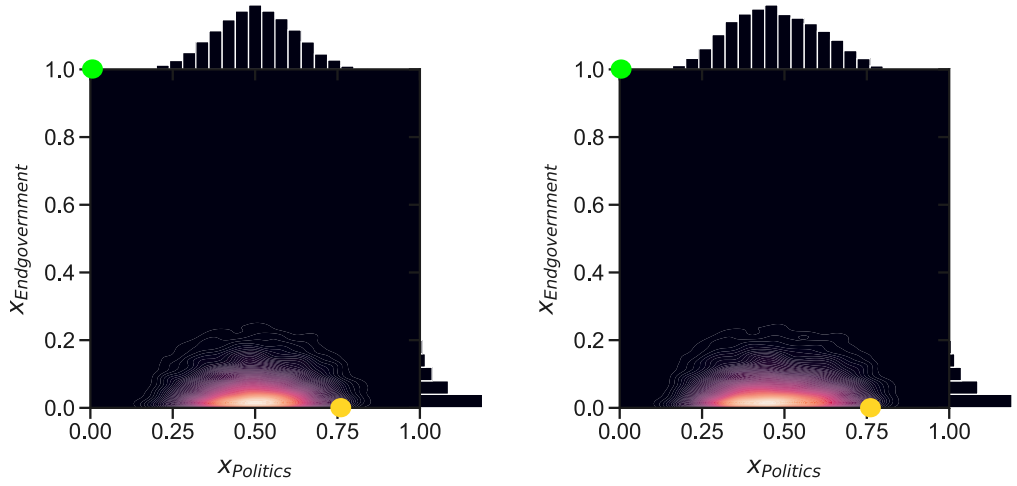
2.7.2 Case Study: Italian Government Crisis in August 2019

In June 2018, a few months after the general elections, Giuseppe Conte was appointed Italian Prime Minister. Two parties formed his supporting coalition: Movimento 5 Stelle (his party, holding the relative majority of the Italian Parliament) and Lega, whose leader was Matteo Salvini. In August 2019, Salvini decided to withdraw Lega's support to the government, starting a crisis aimed at driving Italians to new elections and gaining more votes. However, Movimento 5 Stelle reached an agreement with other parties to form a new government, and on September 5, 2019, Giuseppe Conte became Prime Minister for the second time, excluding Lega from the new administration.

In this section, we apply the proposed model to reproduce the sudden rise of Giuseppe Conte's popularity in social networks during this government crisis. We exploit the multidimensional capability of the model considering two directions: *Politics*, reference topic for Salvini and Conte, and attitude toward government fall, *End government* (see Figure 2.16).

In the opinion space, we assume Salvini has a more radical political viewpoint ($x_{Politics} = 0$), while Conte has an opposing and more moderate position ($x_{Politics} = 0.76$). We set these values in a somewhat discretionary manner. However, we provide a sensitivity analysis in C, proving that results are robust. Conversely, it is safe to assume that the two politicians take opinions at the extreme of the spectrum on the attitude toward government fall, i.e., Salvini has $x_{Endgovernment} = 1$ and Conte has $x_{Endgovernment} = 0$. Moreover, we consider a population with a *moderate* initial opinion on *Politics* (centered at $x_{Politics} = 0.5$, see Figure 2.16a). On *End government*, we sought a distribution that could explain the sudden popularity leap of Conte. We found that the user population must be strongly biased towards Conte's opinion (Figure 2.16).

Some further simplifying assumptions are necessary to apply the model. We assume that the two politicians have a consistency $c^{(i)}$ of exactly one (real values



(a) Initial distribution

(b) After the transient

Fig. 2.16 (2.16a) Initial population's distribution along the *Politics* and *End government* direction and (2.16b) after transient, at the start of the observed period. The opinion position of the two leaders in the space is depicted with a green (Salvini) and a yellow (Conte) point.

are often close to this value, see Figure 2.5a). Moreover, Giuseppe Conte and Matteo Salvini are the only influencers. Although this hypothesis is restrictive, in the scenario studied, the two influencers were the main (active and popular) protagonists during the government crisis. Moreover, we consider the simplest scenario in which personalization is not employed: $\rho = 0$ and thus $\omega \equiv 1$. We consider a feedback function of the form $\theta = e^{-8.25(x^{(u)}-x^{(i)})^2}$ for both opinion directions. For an exhaustive list of the parameters, we refer to Table 2.3.

A period of eleven weeks is considered, from July 7 to September 22, during which data was collected weekly from Facebook. A total of 1162 posts were published, of which 125 were by Conte. The frequency of publication $f^{(i)}$ is calculated as the number of posts by an influencer relative to the total number of posts ($f^{(Conte)} = 0.108$, $f^{(Salvini)} = 0.892$).

Figure 2.17 shows the timeline of the experiment. The two influencers start with the same initial popularity. We consider a transient of $N_t = 10000$ discrete time-units, after which the stationary normalized popularities π_i approximately correspond to the empirical normalized popularities obtained by dividing the number of followers of each influencer by the total number of the two. After the

Table 2.3 Parameters and functions for the Case Study

Symbol	Value - Form	Description
N_i	2	Number of influencers
$x_0^{(Conte)}$	0.76	Opinion of Giuseppe Conte on direction j
$x_0^{(Salvini)}$	0.0	Opinion of influencer 1 on direction j
$f^{(Conte)}$	0.108	Opinion of Giuseppe Conte on direction j
$f^{(Salvini)}$	0.892	Opinion of influencer 1 on direction j
$r^{(Conte),(Salvini)}$	0	Reference direction of both influencers
$p_{Conte,Salvini}(0)$	20	Initial absolute popularity of both influencers
N_u	10000	Number of regular users
N_{iter}	15000	Number of iterations for each simulation
N_t	10000	Duration of the transient phase
w	550	Length of the government crisis
α	0.3	First weight in the updating rule in Eq. 2.3
β	0.65	Second weight in the updating rule in Eq. 2.3
$\theta(\cdot)$	$e^{-8.25(x^{(u)}-x^{(i)})^2}$	Functional form of the *feedback* function
$\omega(\cdot)$	$\rho = 0 \implies \omega \equiv 1$	Functional form of the *visibility* function

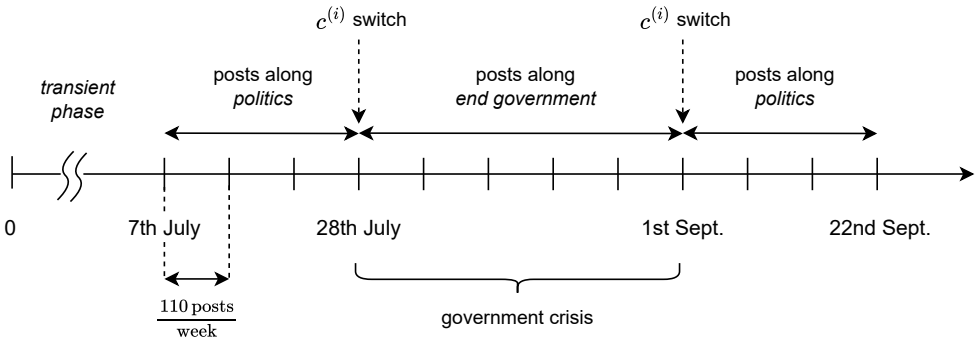


Fig. 2.17 Timeline of modeled scenario from July 7, to September 22. From July 28 to September 1 we have a consistency switch, with posts along *End government* direction.

transient, we can see in Figure 2.16b that the distribution of public opinion is skewed towards Salvini, who has a higher popularity ratio due to his much higher publication frequency. After the transient, the crisis starts, and both influencers post in the *End government* direction (i.e., we observe a consistency switch for both influencers) during a time window of five weeks that approximates the duration of the government crisis, after which the two politicians switch back to posting on the *Politics* direction.

Even with these simplifications, it is possible to reproduce the observed social behavior as a whole: it corresponds to a situation where an influencer is in stark

contrast to the opinions of its user base and loses ground with respect to the other influencer. In the model, only a very unbalanced population distribution towards Conte's opinion (against the government fall) can explain the sudden increase in Conte's popularity, despite the remarkable differences in popularity ratios in favor of Salvini. Figure 2.18 compares the simulation results of the described setting and Facebook's measurements.

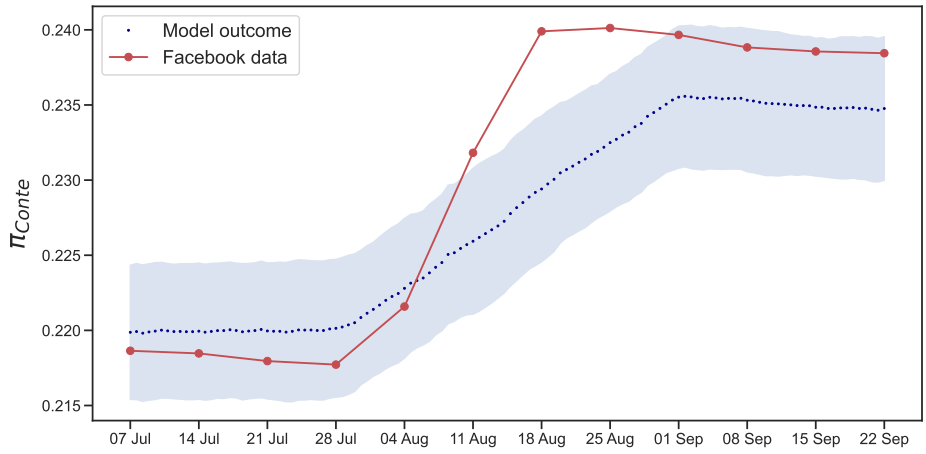


Fig. 2.18 Popularity ratio π_{Conte} for Conte, from Facebook data (red) and from the model (blue) along with its 95% confidence interval, computed over 10 realizations of the process. One can see how the model follows the increase in popularity during August 2019.

Clearly, our model does not precisely fit empirical observations but provides qualitative insights into the possible causes of the rather sudden popularity shift that was observed. Many of the model's parameters are unknown, such as the opinion distribution, the weights of the updating rule, or the feedback function. However, by measuring some parameters, e.g., $f^{(0)}$, the phases' duration, π_{Conte}(7 July), and making reasonable assumptions about the others, i.e., α, β, θ, $\boldsymbol{x}^{(i)}$, the emerging behavior is consistent for several choices of the parameters as demonstrated in C. We conclude that the observed popularity trends can be explained mainly by considering the fear of political instability in the user base.

2.8 Possible Model Extensions

In this final section, we discuss some possible direct extensions of the model, which on the one hand make it more realistic, but on the other hand further

complicate its analysis. First, we consider the interactions between *regular* users and then imagine the possibility of strategic influencers, which we further develop and explore in Chapter 3.

2.8.1 Modeling Interactions Between Users

In our model, we have deliberately ignored the interactions between *regular* users. This decision can be justified by the need to keep the model simple enough to be analyzed with the tools presented so far and by the hypothesis that, opinion dynamics are primarily driven by the interaction between influencers and regular users.

Nevertheless, our model could be extended to include interactions between regular users, and here we outline a possible way forward. First, to preserve the scalability of the model, we do not represent individual interactions between regular users as done by other authors, but we represent the average effect of such interactions on the generic regular user u through a mean-field approach. Second, we distinguish between two classes of additional interactions that arise from the activity of regular users: *direct* (pairwise) interactions between regular users and *influencer-triggered* user interactions.

For what concerns the first class, we introduce a monotonic non-increasing function $\zeta(d)$, representing the probability that two *regular* users whose current opinion distance is d influence each other. Note that $\zeta(d)$ measures the degree of homophily of regular users as a function of their separation in the opinions space: more skewed $\zeta(d)$ shapes correspond to larger degrees of homophily. Moreover, we again assume that user u provides positive feedback to a message generated by user u' at time t on topic j with probability $\theta(|x_j^{(u)(t)} - x_j^{(u')(t)}|)$. It follows that, to account for the overall effect on u of pairwise interactions, we need to compute the *center of mass* of neighbors receiving positive feedback from u along direction j:

$$\begin{aligned}\widehat{x}_j^{(u)}(n) &= \sum_{u' \neq u} \boldsymbol{x}^{(u')}(n)\theta(x_j^{(u)}(n) - x_j^{(u')}(n)|)\zeta(||\boldsymbol{x}^{(u)}(n) - \boldsymbol{x}^{(u')}(n)||) \\ &\to \int f(\boldsymbol{x}, t)\theta(|x_j^{(u)} - x_j|)\zeta(||\boldsymbol{x}^{(u)} - \boldsymbol{x}||)\mathrm{d}\boldsymbol{x}\end{aligned}$$

Then we extend the user opinion update rule (2.3) as follows:

$$x_j^{(u)}(n+1) = \begin{cases} \alpha z_j^{(u)} + \beta x_j^{(u)}(n) + \gamma x_j^{(i)} + \delta \widehat{x}_j^{(u)} & \text{if } \Omega\left(\omega(d_r, \pi_i)\right) = 1, \Theta\left(\theta(d_j)\right) = 1 \\ \frac{\alpha}{1-\gamma} z_j^{(u)} + \frac{\beta}{1-\gamma} x_j^{(u)}(n) + \frac{\delta}{1-\gamma} \widehat{x}_j^{(u)} & \text{otherwise} \end{cases} \tag{2.25}$$

with $\alpha + \beta + \gamma + \delta = 1$. Different ratios δ/γ correspond to different relative strengths between *direct* interactions between regular users and *influencer-triggered* users interactions.

Influencer-triggered user interactions represent a cascading effect of reactions by a user on a post generated by an influencer, appearing on the timeline of its friends. This is another form of (indirect) interaction among regular users which might play a significant role in the overall opinion dynamics. It is reasonable to assume that all follow-up reactions to posts of influencer i mainly circulate among the *followers* of i with the resulting effect of amplifying the strength of the original post on the users providing positive feedback to it. As a consequence, one can take into account the impact the *influencer-triggered* users interactions by playing on weights α, β, γ and δ, i.e., by increasing the relative weight of γ.

2.8.2 Strategic Influencers

The influence exerted on non-reference directions depends heavily on the size of the user base that the influencer's post reaches, which in turn depends on how popular an influencer is in their main area of expertise. For example, famous public figures (e.g., athletes, and models) may express their opinions on potentially sensitive topics and resonate more than experts due to their popularity in their own field. This mechanism can be deliberately exploited by an influencer to gain popularity and/or strategically exert influence on a topic other than their own reference topic. In addition, influencers can also develop a posting pattern that ensures they maximize their influence on the user population.

For example, we have seen in Section 2.6.4 how *low-consistency* values allow an influencer to gain a structural advantage when the consistency of the opposing influencer (who has a different reference topic) approaches one. We must note at this point that a consistency $c^{(i)} < 0.5$ in a two-dimensional environment is slightly in contrast with the definition of the reference direction itself if the parameters

are static. However, if we consider a time-varying consistency, an influencer that is well-known on a certain field j, i.e., $r^{(i)}(n_1) = j$ may modify its consistency at time n_2 if it learns to have little competition on its reference direction.

Our model can be rather easily adapted to study such scenarios by representing influencers as *strategic agents* whose selected topics and expressed opinions in their posts can be dynamically adjusted to maximize a given utility function. A first step in this direction is presented in Chapter 3, where we consider only a one-dimensional opinion space and thus we focus on the sequence of *expressed* opinions which maximizes a certain utility function.

2.9 Concluding Remarks

In this chapter, a novel multidimensional opinion model tailored to online interactions was presented. Thanks to Monte Carlo experiments, we observed how the *personalization* of content by the social media platform plays a central role in online interactions. The emergence of echo-chambers can occur in the case of sharp personalization, as long as the influencers compete over the same topic (i.e., have the same *reference direction*). In the same setting, we observed also the rapid rise to the top of one of the influencers, in the case of different reference directions. It was observed (both in the model and in the empirical analysis on Instagram) that the frequency of publication correlates positively with the increase in popularity. The behavior as a function of the consistency value was the most difficult to interpret and needs further investigation. Our results suggest that since influencers usually have a high consistency score (see Fig 2.5a), a competitor can gain a popularity advantage by posting massively in their secondary direction for a certain period of time and returning to their main topic after a certain period of time. Finally, in a limited but evocative example, we have seen how the model can be used in real-life situations by applying it to the 2019 Italian government crisis.

Chapter 3

Modeling Influencers Competition

Part of the work presented in this chapter has been presented at the workshop of the 35th International Teletraffic Congress (ITC) and published in [54]:

- Galante, F., Garetto, M., & Leonardi, E. (2024). Competition of Influencers: A Model for Maximizing Online Social Impact. In Proceedings of the 16th ACM Web Science Conference (WEBSCI '24). Association for Computing Machinery, New York, NY, USA, 343–353. https://doi.org/10.1145/3614419.3644031

One of the simplifications of the Communication Asymmetry model introduced in Chapter 2 is that users are considered to be stubborn and therefore hold a fixed opinion over the entire time horizon under consideration. This hypothesis can be reasonable in certain scenarios where the dynamics of opinions are particularly slow and views are unlikely to be subject to drastic and erratic changes, such as in long-running brand competitions (e.g., Nike vs. Adidas) or even when contrasting religious and non-religious beliefs. In other situations, online social networks (OSNs) appear to be very fluid environments in which individuals can express very diverse opinions that can change significantly over time. The influencer class in particular has an incentive to exhibit such changeable behavior for several reasons. It should be borne in mind that many of these personalities have a strong interest in the performance of their social media profiles, either because they are used directly to generate income (e.g. through advertising) or to convey a message (e.g. a political one). Influencers are therefore interested in maximizing their influence on the social network and the messages they convey may be subject to frequent change, for example, once they change their brand

partnership or also due to pressure from public opinion on sensitive issues on which they have previously held radical views.

In the previous chapter, we discussed the possibility of choosing the posting pattern strategically. Either to gain popularity on the platform or to facilitate the dissemination of radical or counterfactual ideas, perhaps by capitalizing on the visibility of another unrelated topic. Here, we take a first step in this direction and try to explore how an influencer can maximize its influence on an OSN in the case of a single topic (one-dimensional situation), which has proven to be challenging enough in itself. We leave the more interesting multidimensional case aside for future work, noting that its highly nonlinear nature and the correlation between popularity on the *reference* direction and the visibility granted by the platform makes the analytical study of the model prohibitive. We focus on a simplified case in which we have duopolistic competition over a single topic and a large number of users who follow the posts of the two influencers, whose posts are filtered by the social media platform. We use the framework of game theory to study the competition and develop a trellis-like approach to solve the impact maximization problem, which we formalize in Section 3.3, in the case of an influencer in isolation.

The chapter begins with a discussion of the research context and introduces some relevant related literature. Then the notation, which largely overlaps with that of the previous chapter, is briefly introduced, together with a summary table of all the symbols used in this chapter. Section 3.3 formalizes the online influence maximization problem, also specializing the opinion model to the one-dimensional case. Section 3.4 introduces the trellis-like approach to finding the optimal solution for this problem. Subsequently, the optimal strategy is compared with a greedy strategy. Sections 3.5 and 3.6 present the game for influencer competition and its extension, which also considers the popularity of influencers. Section 3.7 concludes the chapter.

3.1 Related Work and Context

Our work can be considered at the intersection between the literature on opinion dynamics [55] and the more classic economic literature on competition [56]. Widely used models of opinion dynamics include linear models proposed by DeG-

root [57] and Friedkin–Johnsen [58], and non-linear models by Deffuant–Weisbuch [59] and Hegselman-Krause [60]. We refer the reader to Section 2.1 of this thesis or to [61] for a recent review on different models of opinion propagation, and to [62] for results specific to non-linear (bounded confidence) dynamics.

Influence maximization is a different problem, slightly related to ours [63]. It aims at finding a (small) set of k agents so that to maximize the adoption of a certain product. This problem differs largely from ours in that it considers an explicit network structure and uses a simpler underlying (cascade) opinion model whose states are binary. Many extensions appeared, [64] considers a *negativity* bias, and [65] non only competitive behavior. In the following, we limit ourselves to mentioning works more similar in spirit to ours, i.e., those dealing with non-linear (e.g., bounded confidence) dynamics, and focused on opinion manipulation by a restricted set of strategic agents aiming at maximizing their impact on a population of users.

A bounded confidence model of opinion dynamics on a fixed network structure comprising both influencers and followers is proposed and analyzed by simulation in [66]. In [67] the impact of charismatic leaders is taken into account in bounded confidence dynamics as a constant exogenous signal. Interestingly, they discovered that higher signals may have less effect in attracting other agents. Opinion manipulation through (possibly time-varying) exogenous inputs is analyzed in [68] for an Eulerian (i.e., by considering a probability distribution of agents) bounded-confidence system. In [69], the authors consider a continuous-time bounded confidence model with a single leader, showing that it is possible to control the leader velocity to ensure final consensus at a prescribed opinion value.

We mention that optimal control approaches have also been proposed to optimize opinion manipulation. For example, [70] takes this direction considering the case of a single influencer. In [71], a set of (coordinated) controllers is optimally placed to minimize the convergence time of the system to a final stable state in the influence graph.

To the best of our knowledge, there are only a few game-theoretical attempts to study the problem of opinion formation/manipulation. Papers [72–75] consider a scenario in which two competing marketers play a resource allocation game, whose goal is to establish how many resources to allocate to each potential customer in the network. Customer dynamics obey a voter model in [72],

to a consensus model over a graph in [75, 74], and to a Friedkin-Johnsen model in [73]. A Nash control formulation is proposed in [76] for the case of two opposing influencers. Opinions spread on an underlying graph in an epidemic manner.

Differently from previous papers, we consider non-linear opinion dynamics where the interaction between influencers and users is mediated by a social media platform. Influencers operate on a fixed budget and have to decide what opinion to express within each of N posts.

3.2 Notation

In this chapter we try to stick to the notation introduced in the previous chapter, bearing in mind that we are not dealing with vectorial quantities here. Lowercase letters are used for parameters and dynamical variables associated with an individual. The index i runs over the set of influencers, while the index u runs over the set of regular users. They are specified between superscript brackets, e.g. $a^{(i)}, a^{(u)}$, to specify which user class the variable a refers to. Italic capital letters denote sets, e.g., $\mathscr{I}$ is the set of all influencers in the population, while $|\mathscr{I}|$ is its cardinality. Capital letters stand for results of stochastic experiments whose characteristic parameters are lowercase letters: e.g. $\Psi\left(\psi(\cdot,\cdot)\right)$. Unlike in the previous chapter, we specify the discrete time n as the subscript value of a certain variable, e.g., $p_n^{(i)}$ is the popularity of influencer i at time n.

3.3 Formulation of the Problem

First, we present a slightly modified instance of the Communication Asymmetry (CA) opinion model from Chapter 2 and specialize it to the case of a one-dimensional opinion space. We will first use a simplified version of the model that describes the interaction between popular individuals (i.e., influencers) and the pool of regular users without considering the reinforcing mechanism implied by popularity updating. In section 3.6.1 we will use a more complete version of the CA model that takes into account the close-loop between influencers and regular users. The main focus of this section is to introduce the *social impact maximization* problem for an influencer considered in isolation. This answers the question:

Table 3.1 Notation Summary Table

Symbol	Description
$\mathscr{I}$	Set of influencers within the OSN population
$\mathscr{U}$	Set of regular users within the OSN population
$\mathscr{X}$	Set of opinions, in general $\mathscr{X} \subset \mathbb{R}^d$ but typically $[0,1]$
n	Discrete time instant
N	Total time horizon
$x_n^{(u)}$	Opinion of a regular user at time n
$z^{(u)}$	*Prejudice* of a regular user
$x_n^{(i)}$	Opinion of an influencer at time n
d	User-influencer opinion distance $d := \lvert x_n^{(u)} - x_n^{(i)} \rvert$
$\mu_n(x)$	Regular users opinion distribution
$f(\cdot)$	Influencer's benefit function for the impact maximization problem
$\Psi(\psi)$	Bernoulli random variable capturing the platform filtering, if $\Psi = 1$ the opinion delivered through the post influences the user
w	Width of the *rectangular* (0-1) ψ function
α	First opinion update weight
β	Second opinion update weight (or *inertia*)
$\delta := \frac{\alpha}{\alpha+\gamma}$	*Degree of stubbornness*
$\mathscr{B}$	Set of discrete opinion values
$\mathscr{M}$	Set of regular users' groups with the same prejudice
$x^{(i;T)}$	Target opinion for influencer i (x^T if $\lvert\mathscr{I} = 1\rvert$)
$x^{(i;E)}$	Exploratory opinion for influencer i
$\delta^{(i)}$	Difference between the target and exploratory opinion
$\mathscr{A}$	Set of actions for the game
$a^{(i)}$	Action of influencer i
$\mathscr{BR}(a^{(i)}, a^{(-i)})$	Best response of influencer i to the other players
P	Payoff matrix of the game
$p_n^{(i)}$	Influencer's i popularity
$T_n^{(i)}$	Total feedback on a influencer's post

"What would be the posting pattern of an influencer who wants to maximize her influence in an online social network where the online social platform filters the content?" .

3.3.1 The One-Dimensional Communication Asymmetry Model

The element at the core of our approach is clearly the CA opinion model from the previous chapter. To ensure the coherence of this chapter, we briefly summarize the model here, furthermore, we specialize it to the case of a one-dimensional opinion space $\mathscr{X} = [0,1]$, and consider a simple characterization of the influencers compared to Chapter 2. Moreover, we add a mechanism of "self-thinking" for

the users, i.e. whenever a user is not reached by an influencer's post, he updates his opinion and drifts to his original view on the topic. So we recall that the audience on social media can be divided into two macro-categories: the popular and influential individuals: the *influencers*, and the "other" users: the *regular* users, who represent the vast majority of OSN users.

Let us denote by $\mathscr{I}$ (indexed by i) the set of influencers and by $\mathscr{U}$ (indexed by u) the set of regular users. These social network users at a certain time instant $n \in \mathbb{N}_0^+$ hold opinions $x_n^{(i)} \in \mathscr{X}$ and $x_n^{(u)} \in \mathscr{X}$, respectively[1]. We assume that at each discrete time instant n influencer $i \in \mathscr{I}$ publishes a post over the social media network conveying her *expressed* opinion. In this chapter, as opposed to [13], we consider influencers with time-varying opinions, i.e., the influencers generate a sequence of posts $\{x_n^{(i)}\}$, $n \in [0,..,N]$, where N represents a considered (finite) time horizon. An influencer will adapt her expressed point of view strategically in order to maximize her[2] impact on the population (see Section 3.3.2).

A given post is not guaranteed to reach a regular user $u \in \mathscr{U}$, due to the *filtering* effect of the social media platform. This represents the *content personalization* performed by most platforms (e.g., the ranking of posts to be shown on the user's timeline). We assume a post effectively reaches a regular user with a probability $\psi\left(|x^{(i)} - x^{(u)}|\right)$ which depends on the distance between the influencer's *expressed* opinion and the user's opinion, modeling a homophilic behavior, whereby individuals are more likely to interact with others who share similar beliefs.

A user holding opinion $x_n^{(u)}$ at instant n updates his opinion according to the following rule:

$$x_{n+1}^{(u)} = \begin{cases} \alpha z^{(u)} + \beta x_n^{(u)} + \gamma x_n^{(i)} & \text{if } \Psi\left(\psi(|x_n^{(u)} - x_n^{(i)}|)\right) = 1 \\ \frac{\alpha}{\alpha+\beta} z^{(u)} + \frac{\beta}{\alpha+\beta} x_n^{(u)} & \text{otherwise} \end{cases} \tag{3.1}$$

where α, β, γ are fixed parameters in $[0,1]$ such that $\alpha + \beta + \gamma = 1$. $\Psi(\psi)$ is a Bernoulli random variable with parameter ψ and determines whether a post will be visible to a given user or not. Note that when a user receives a new post from the influencer, i.e, when $\Psi = 1$ (first row of Eq. (3.1)), he moves to a new position

[1] In this chapter, for convenience, we adopt a slightly different notation with respect to Chapter 2, representing the time as a subscript rather than between parentheses.

[2] In this chapter, to avoid the confusion which would arise using the gender-neutral "they", we will, arbitrarily, use she/her for an *influencer*, and he/his for a *regular* user.

in the opinion space, which is a convex combination of three contributions: i) his prejudice $z^{(u)}$, i.e., the preconceived opinion about a certain matter[3]; ii) his current opinion $x_n^{(u)}$; and iii) the influencer's opinion $x_n^{(i)}$ expressed in the post.

Otherwise (second row of Eq. (3.1)), the influencer's contribution is not present, either because the post has not been proposed to the user or he has not been influenced (e.g., not liked it), so a renormalization of the weights is required. This case in Eq. (3.1) models a process of *self-thinking*, namely that users who are not reached by an influencer's post gradually return to their prejudice $z^{(u)}$.

The above updating rule is simple but does not allow a direct interpretation of the parameters. Noting that the free parameters in the first line of Eq. (3.1) are only two, it is possible to rewrite the update rule:

$$x_{n+1}^{(u)} = \begin{cases} (1-\beta)\left[\delta z^{(u)} + (1-\delta)x_n^{(i)}\right] + \beta x_n^{(u)} & \text{if } \Psi\left(\psi\right) = 1 \\ \frac{\delta}{\delta+\frac{\beta}{1-\beta}} z^{(u)} + \frac{\frac{\beta}{1-\beta}}{\delta+\frac{\beta}{1-\beta}} x_n^{(u)} & \text{otherwise} \end{cases} \tag{3.2}$$

where $\delta, \beta \in [0,1]$ have a direct interpretation as the *inertia* (β) of the user, i.e., the weight the users give to their current opinion, and the *degree of stubbornness* (δ), i.e., the weight on the user's preconceived opinion.

Remark 4. *(Large population) In the large population limit (i.e., when $|\mathcal{U}| \to \infty$) fluctuations of aggregate random variables around their average smooth out. Therefore macroscopic dynamics tend to become deterministic.*

Thanks to the *large-population* assumption, we do not have to track the microscopic interactions described in Eq. (3.2) but we can consider a distribution of regular users characterized by the probability density function $\mu_n(x)$, whose evolution is driven by influencers' posts emission. In particular, every time influencer i generates a new post, a fraction $\psi(|x - x^{(i)}|)$ of the population placed at x will be hit by the influencer's messages while the remaining fraction of users will not be reached by it. This assumption greatly simplifies the analysis, in particular in Section 3.6.2 where we consider a more complicated function ψ, while it is not strictly necessary for the rest of the work (see Remark 5). In our framework, the

[3] This quantity could also be interpreted as the *field effect* of the total population, which would represent a process of peer interaction not otherwise included in the model.

assumption is not restrictive, as our focus lies in the mean-field effects observed across a large population of individuals.

3.3.2 Social Impact Maximization for an Influencer in Isolation

We are interested in determining the influencer posting pattern that maximizes her online social impact on a population of regular users. Even considering the case where a single influencer seeks to maximize her impact over a finite time horizon, assuming there are no other influencers, is insightful and complicated enough to be worth exploring. We will then use the observations gathered in this simplified setting to develop our game of online competition.

Our novel *social impact maximization* problem, for the case of a single influencer, can be formulated as follows. Recall that we consider a fixed time horizon N, where the influencer has to choose the temporal sequence of opinions $\{x_n^{(i)}\}_{n=1}^N$ to convey through her posts to attract regular users towards a desired *target* opinion x^T in the opinion space. This value can represent, for example, the *true* opinion of the influencer regarding a certain topic or a certain consumption behavior to be instilled in the population. Regular users obey the dynamics in Eq. (3.1).

We assume that the influencer knows how users would react to her posts, i.e., the parameters of Eq. (3.1), and in particular the shape of $\psi(\cdot)$ as a function of the opinion distance $d \triangleq |x_n^{(u)} - x_n^{(i)}|$, which dictates whether her posts are received by users in the first place (platform filtering). Moreover, we assume that the influencers know the initial distribution of users $\mu_0(x)$ (e.g., through polls, surveys, reviews and other forms of users' feedback).[4]

The benefit an influencer obtains from a particular distribution of users' opinions may vary. This variability can be captured by an arbitrary function $f(\cdot)$ that provides the influencer's benefit from a generic user at a given distance from the target opinion. Thus, in its greatest generality, the problem can be formulated as

[4] It would be possible to incorporate a *noisy* version of such distribution into the model, along the lines of what happens in politics, where polls provide a noisy measure of the true distribution of public opinion. We leave this possibility for future work.

follows:

$$\max_{\{x_n^{(i)}\}_{n=1}^N} \mathbb{E}_x \left[f(|x_N^{(u)} - x^T|) \right] = \int f(|x - x^T|) \mathrm{d}\mu_N(x) \tag{3.3}$$

s.t. dynamics in (3.2)

where $\mu_N(x)$ is the *final* distribution reached by users over the opinion space at time N. Note however that the maximization in Eq. (3.3) is over the entire sequence of N posts generated by the influencer. The influencer *benefit function* $f(\cdot)$ is reasonably a non-increasing function of the distance from the target opinion.

The formulation in Eq. (3.3) leads to a complex optimization problem, given the generality of the initial user distribution, the probabilistic movement of users (in the *finite population* case), the arbitrary choice of the influencer's opinion at each step, and the arbitrariness of function $f(\cdot)$. Therefore, we will now make a series of simplifications that eventually lead to a problem that is solvable in polynomial time with N.

First, we assume that $f(\cdot)$ is a linearly decreasing function of the distance from the target point x^T, i.e., $f = -|x_N^{(u)} - x^T|$. While this simplification does not effectively reduce the complexity, it allows us to get a reasonable case study that does not require us to specify details of the shape function $f(\cdot)$. Therefore, by linearity of the mean and as $\max -g(\cdot) = \min g(\cdot)$, the problem becomes:

$$\min_{\{x_n^{(i)}\}_{n=1}^N} \mathbb{E}_x \left[|x_N^{(u)} - x^T| \right] = \int \left| x - x^T \right| \mathrm{d}\mu_N(x) \tag{3.4}$$

s.t. dynamics in (3.2)

This reformulation also corresponds to a simpler interpretation: the influencer aims to bring the overall opinion of the population of regular users as close as possible to her target opinion.

Our main simplification assumes a binary (0-1) behavior for the event related to whether a post reaches a user at a certain distance d from the opinion expressed in the post. This is achieved, for example, by the following natural choice for the function ψ:

$$\psi\left(d = \left| x_n^{(u)} - x_n^{(i)} \right|; w\right) = \begin{cases} 0 & \text{if } d > w \\ 1 & \text{if } d \leq w \end{cases} \tag{3.5}$$

where w is a fixed width parameter. This means that a post conveying the opinion of the influencer $x_n^{(i)}$ is deterministically read at time n by users who have an opinion that is at most w away from it. This formulation of the model essentially implements the well known concept of *bounded confidence* (see e.g. [60] and [59]), where here we consider a single entity (the influencer) publishing posts to attract regular users. We emphasize that bounded confidence dynamics have proven quite difficult to analyze, so most of the known results have been obtained through Monte Carlo simulations.

Remark 5. *Due to the assumption of a 0-1 $\psi(\cdot)$ function in Eq. (3.5), the model becomes deterministic even in the case of $|\mathcal{U}| < +\infty$ (finite population). Indeed, a user either moves or does not move, depending solely on his distance from the influencer's expressed opinion.*

Even with this simplifying assumption, the optimization problem remains significantly challenging. Indeed, it can be formulated as a Markov Decision Process (MDP), in which the number of states of the underlying Markov Chains (MCs), i.e., the MCs obtained by fixing the influencer sequence of posts, is combinatorially exponential with the number of users.

To gain initial insights into the problem and as a useful benchmark, we consider the *greedy* solution that, at each time step, selects the influencer's opinion $x_n^{(i)}$ that produces the best instantaneous improvement of the objective function. More formally, let $\Delta x_n^{(u)} \triangleq x_{n+1}^{(u)} - x_n^{(u)}$ be the *opinion shift* of a user u holding opinion $x_n^{(u)}$ at time n:

$$\Delta x_n^{(u)} = \begin{cases} (1-\beta)\left[\delta z^{(u)} + (1-\delta)x_n^{(i)} - x_n^{(u)}\right] & \text{if } \Psi(\psi) = 1 \\ \frac{\delta}{\delta + \frac{\beta}{1-\beta}}\left(z^{(u)} - x_n^{(u)}\right) & \text{if } \Psi(\psi) = 0 \end{cases}$$

For simplicity of exposition, but without loss of generality, consider the case in which all users are initially to the left of target point x^T, so that positive values of $\Delta x_n^{(u)}$ translate into equivalent improvements of the objective function, whereas negative values translate into equivalent utility losses.

Then, given the users' distribution $\mu_n(x)$, the *greedy* algorithm selects, at each step n, the influencer opinion $x_n^{(i)}$ maximizing the overall users' opinion shifts:

$$x_n^{(i)} = \arg\max_{x_n^{(i)}} \mathbb{E}\left[\Delta x_n^{(u)}\right] = \arg\max_{x_n^{(i)}} \int \Delta x_n^{(u)} \, \mathrm{d}\mu_n(x) \tag{3.6}$$

Claim 1. *The greedy strategy is not always optimal.*

As might have been expected, the above *greedy* strategy is, in general, suboptimal. We will demonstrate this in the next section in a simple but representative scenario. The reason for the suboptimality lies in the fact that the greedy algorithm does not "look into the future," ruling out solutions that initially reduce the overall utility but, in the long run, lead to a better final configuration of users in the opinion space. Understanding when the *greedy* strategy may be suboptimal and by how much is of great interest both theoretically and practically. In cases where the *greedy* strategy does not lead to an optimal outcome, it is essential to apply strategies that sacrifice short-term gains in favor of long-term benefits.

3.4 The Optimal Strategy

The optimal strategy can be computed (numerically) under the simplifying assumptions introduced in the previous section by resorting to a discretization of the opinion space and the user distribution. In particular, let us assume that both the users' opinion and the influencer's opinion expressed in each post can only take values in a discrete set of cardinality B. In practice, we divide the opinion space $\mathscr{X} = [0,1]$ into bins of constant width, and assume that only the mid-point of each bin is a feasible opinion value for the influencer's expressed opinion. For the sake of simplicity, we take as target points $x^{(T)} \in \{0,1\}$ for the influencers operating in the system: these points can be interpreted as two opposing political views or as two different brands offering the same product to customers. In the case of a single influencer, we assume $x^T = 1$ (the case $x^T = 0$ is completely symmetrical).

Let $\mathscr{B}$ be the set of feasible opinion values, and j be the index running on it. We will also discretize the prejudice of users, assuming that it belongs to a finite set $\mathscr{Z} \subseteq \mathscr{B}$ of prejudice values, of cardinality Z. Note that the population is indeed described by two distributions, the time-varying opinion and the static prejudice.

At last, we assume that the distribution $\mu_n(x)$ of users over the opinion space can be well approximated by considering the users belonging to a finite set $\mathscr{M}$ of "groups" of cardinality M, indexed by m. Users of a given group m share the same (time-varying) opinion $x_n^{(m)} \in \mathscr{B}$, and the same (static) prejudice value $z^{(m)} \in \mathscr{Z}$. Groups represent discrete "masses" of users moving together as a single unit, that cannot split into smaller sub-units over time. This is guaranteed by the assumption stated in Eq. (3.5), which leads to deterministic opinion movements. Note that without the 0-1 assumption for ψ, the position of a group could not be used to identify the state of the system and that as long as $0 < \psi < 1$, there would be an exponential growth of group subdivisions over time.

It follows that the system state at time n can be fully specified by the vector $\{x_n^{(m)}\}_m$, and that there are B^M possible system states. At each time step, influencer i has to choose an opinion $x^{(i)} \in \mathscr{B}$ to convey in her n-th post, and we can separately evaluate the effect of this post on each group. This can be efficiently done by exploiting the Trellis-like structure of system dynamics described next.

3.4.1 Trellis-like Structure for the Optimal Strategy

Under the simplifying assumptions introduced before, the optimal solution can be computed in polynomial time for any n by exploiting the trellis-like structure sketched in Figure 3.1 for the toy case of $M = 2$ groups, $m \in \{0,1\}$, and $B = 2$ opinion values, $x_n^{(m)} \in \{a,b\}$, leading to 4 possible system states. Let, for short, $S_n^{(s)} \in \mathscr{S}$ denote the possible system states at time n. From each state $S_n^{(s)} \in \mathscr{S}$, it is possible to evaluate all reachable states and the *action* $x^{(i)}$ that leads to the transition. This allows us to define a transition matrix T, shown on the left in Figure 3.1. Note that there may be multiple $x^{(i)}$ leading to the same target state, and for our purposes these transitions are equivalent. Indeed, we are interested only in the final *best* state, and not in finding all particular sequences of traversed states leading to it.

With the transition matrix in hand, it is possible to unfold the process over time, starting from a given initial state S_0^1 (see right part of Figure 3.1). The resulting trellis-like structure allows us to account for all paths starting from the initial state and efficiently compute the one to the final *best* state.

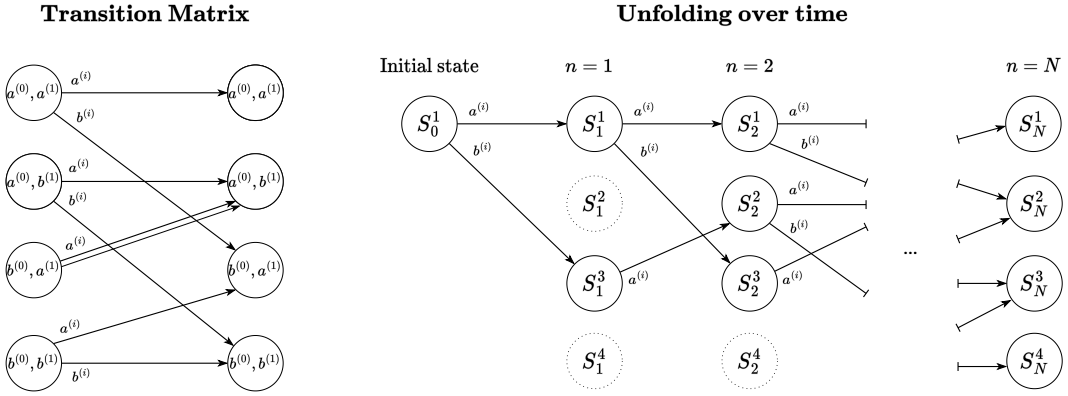


Fig. 3.1 A simple illustration with only four states ($B = 2$) to demonstrate the method used to determine the optimal solution.

Without loss of generality, we can consider the case where $x^T = 1$ and the objective of our social impact maximization problem becomes $\max_{\{x^{(i)}\}_1^N} \mathbb{E}[x_N]$. Therefore, the *best* state is the one that leads to the highest average opinion in the population for $n = N$. In principle, the path leading to this state is not unique, and thus we consider a path as *optimal* if it leads to the *best* state in the shortest possible time.

3.4.2 The Case of Two User Groups

Now that we have a method for deriving the *optimal* solution to the online impact maximization problem, we can prove the correctness of Claim 1. For this, we have considered a very simple system that is computationally tractable and contains enough features to be of interest. We consider only two *regular* users groups whose prejudice and initial opinion are $z^{(m)}, x_0^{(m)}$, $m \in \{0,1\}$ respectively, and a large number B of possible opinion values, leading to B^2 possible system states.

The discrepancy between greedy and optimal strategies can be evaluated for a variety of parameters. However, for the sake of compactness, we limit ourselves to two representative scenarios whose parameters are summarized in Table 3.2.

There are two ways in which the *optimal* solution can be better. It is either faster, i.e., it reaches the *best* state with a smaller number of posts (first scenario, Fig. 3.2a), or it leads to a higher value of $\mathbb{E}[x_N]$, i.e., a better *best* final state (second scenario Fig. 3.2d).

As a general rule of thumb, we have derived the following empirical rule from the optimal numerical solution of the system:

Remark 6. *The optimal strategy for the online impact maximization problem is to first bring the (two) user groups close together, and then gradually persuade them towards the target opinion* x^T.

Table 3.2 Scenarios of the greedy vs optimal comparison

#	$z^{(0)}$	$x_0^{(0)}$	$z^{(1)}$	$x_0^{(1)}$	α	β	N	w
1	0.5	0.005	0.5	0.995	0.05	0.3	50	0.1
2	0.005	0.005	0.995	0.995	0.05	0.3	50	0.1

Note that as an effect of our choice of the platform filtering function ψ, group u is affected by the influencer's expressed opinion only when $|x^{(u)} - x^{(i)}| < w$. This means that a group whose opinion $x^{(u)}$ is too far from the influencer's expressed opinion $x^{(i)}$ is not affected by the post. Recall, however, that a group can still change its current opinion in the absence of the post's influence, as it is also attracted to its prejudice. In the next Section, we will go into more detail about what happens in the two-group scenario and explain the underlying system dynamics. We will also briefly discuss the effects of the parameters.

3.4.3 Numerical Experiments: Impact of the Parameters

The first scenario considers two user groups with equal prejudice at 0.5 (representing a moderate position toward a certain topic or indifference between the two extreme choices) and extreme initial opinions around 0 and 1, respectively. The second scenario instead assumes that the prejudices of the two groups coincide with their initial opinions, which are again set at the extremes of the opinion domain. It is interesting to note that the two strategies lead to different results in both cases, albeit in different ways. This becomes clear by looking at Figures 3.2b - 3.2c (for the first scenario) and figures 3.2e-3.2f (for the second scenario), showing detailed locations of the two user groups, as well as the influencer opinion, at each time step n, separately for the greedy and optimal strategy.

In the first scenario, the greedy strategy (plot 3.2b) leads the influencer to focus on the first, closest group $x^{(u=1)}$ and ignore the second, distant group $x^{(u=0)}$,

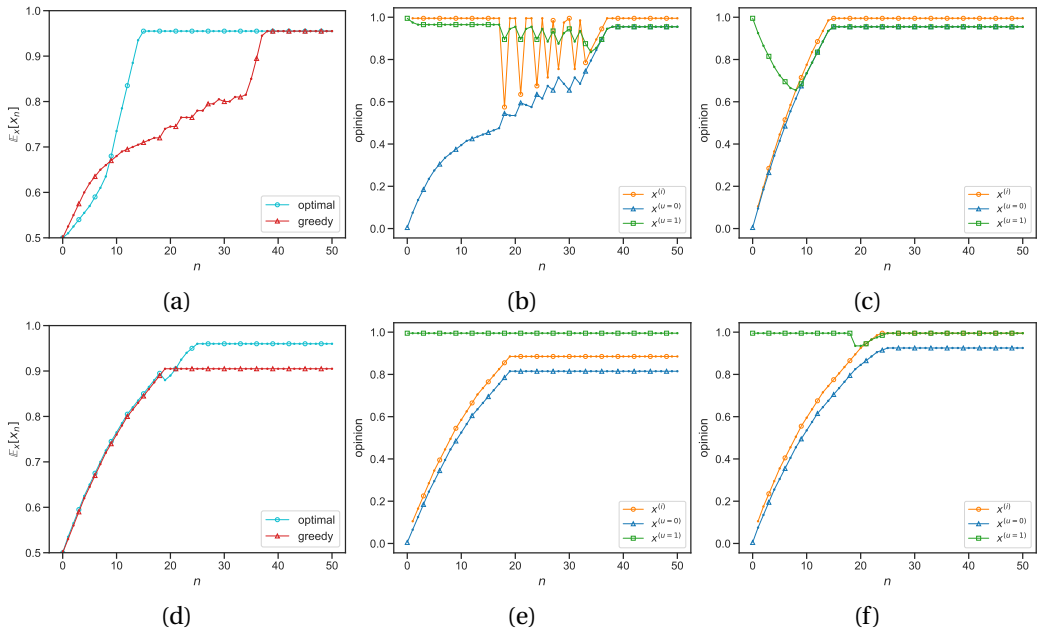


Fig. 3.2 (3.2a) Scenario # 1 in Table 3.2, (3.2b) greedy, (3.2c) optimal. (3.2d) Scenario # 2 in Table 3.2. (3.2e) greedy (3.2f) optimal.

which, in the absence of influencer stimulus, starts to gradually shift towards its prejudice $z^{(0)} = 0.5$.

Only when the second group is close enough, the greedy influencer finds it temporarily convenient to jump close to the second group, and immediately after go back to $x^{(i)} = 1$ to bring the first group close to the extreme of the opinion domain. This results in the erratic behavior of the orange trajectory in plot 3.2b.

These "hectic" moves make the greedy strategy inefficient, and in fact, the greedy strategy is largely outperformed by the optimal strategy (plot 3.2c) in terms of the number of steps (posts) to reach the *best* state. Note that, on the contrary, the optimal strategy focuses on the second distant group, which at some point (around $n = 10$) merges with the first group. The coalesced groups are then efficiently moved together towards the best state (see plot 3.2c).

In the second scenario, the greedy strategy (plot 3.2e) never allows the first group $x^{(u=1)}$ to step away from its (already taken) radicalized opinion at $x^{(u=1)=0.995}$. Note that since $w = 0.1$ the first group would move away from its initial position whenever the influencer conveys an opinion $0.895 \leq x^{(i)} < 0.995$

On the contrary, the optimal solution (plot 3.2f) accepts to temporarily worsen the opinion of the closer group, and by so doing it is able to pull the second group up to a better (closer to the target point) final position. This requires, at some point, a non-greedy step (see the non-monotonocity[5] in plot 3.2d).

The difference between the two strategies increases when w takes small values and gradually diminishes as w increases. For $w > 0.5$, the greedy and optimal strategies coincide in virtually all cases, as larger w values lead to a more significant influence over the population. The extreme scenario is the one in which the entire population is always reached by a post. The optimal strategy in this case is to publish the target opinion exclusively, as there is no advantage in taking other viewpoints. Finally, it is worth mentioning that the weight coefficients α and β in the opinion update determine both the inertia of the system and the maximum achievable opinion value of each user group. Slow dynamics (high values of $\alpha + \beta$) are more challenging to study numerically because of the smaller opinion shifts Δx produced, which require to use a denser discretization of the opinion space (larger number of bins B).

Results (in terms of greedy vs optimal performance) are consistent despite the choice of these parameters, considering that the final opinion value depends very weakly on β (which primarily impacts the convergence time), while α essentially determines the best target opinion that *any* strategy can eventually achieve in the long run.

3.5 The Game: Competing for Influence

The above framework is interesting in that it shows how to optimally target a population of users, and it provides evidence that it is worth accepting short-term losses to achieve long-term gains. However, it completely disregards competition from other influencers on the social media platform. Inspired by the fact that in the *online impact maximization* problem an influencer has an incentive to first take a more moderate position in order to then exert influence over a larger user base, we will develop a game that embodies this idea. Indeed, Remark 6 suggests that it is better for the influencer to first group users, i.e. to express an

[5] Sometimes we observe non-monotonic behavior on the greedy trajectory, which can occur when *any* greedy step produces a negative increment of the objective function.

opinion that allows it to reach "distant" users even, potentially, at the expense of de-radicalizing users close to its target opinion, and then gradually draw them towards its target. This consideration allows us to propose a fixed structure for an influencer's strategy that greatly simplifies the problem and makes it tractable.

We define a game for the duopolistic competition of influencers aiming at maximizing their online social impact. Indeed, we consider a set $\mathscr{I} = \{0,1\}$ of two *players* (influencers), each of which has a *target* opinion $x^{(i;T)}$: the opinion around which the influencer wants to attract users, recall Eq. (3.3). We will assume, without loss of generality, to have players having diametrically opposing views: $x^{(i=0;T)} = 0$ and $x^{(i=1;T)} = 1$, representing, for example, two opposing political parties.

The players/influencers are characterized by their willingness to deviate from their target opinion $x^{(i;T)}$ when trying to reach new users. This aspect is modeled by the parameter $\delta^{(i)}$, which allows us to define $x^{(i;E)} = x^{(i;T)} - \text{sgn}\left(x^{(i;T)} - 0.5\right) \cdot \delta^{(i)}$, which we will refer to as the *exploratory*[6] opinion. We thus assume that the influencers will only assume one of the two opinions $x^{(i;T)}, x^{(i;E)}$. Moreover the "exploration" phase will always precede the "targeting" phase, the situation is sketched in Figure 3.3.

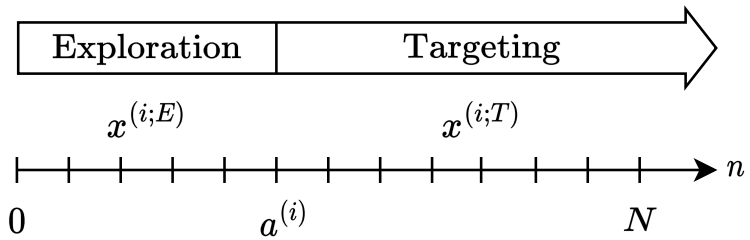


Fig. 3.3 Two-phase strategy $\{x^{(i)}\}_{i=1}^{N}$ structure.

Our assumptions may appear rather restrictive, however, they are motivated in part by the Remark 6 and by the necessity of reducing the space of possible *actions* for each player for tractability purposes. As a matter of fact, the set of *actions* becomes $\mathscr{A}^{(i)} = \{0,1,...,N\}$, and the action $a^{(i)}$ determines how long the "exploration" phase lasts for influencer i, defining a particular $\{x^{(i)}\}$ sequence. Recall that knowing (the system's parameters) and the influencer post sequence,

[6] This variable represents the opinion that the influencer expresses while trying to approach a group of users who are far apart in their opinions and whom it would otherwise not reach due to the filter function ψ.

the movement of the regular users is described by Eq. (3.2). Note that, once we fix the sequences of posts emitted by the two influencers (and knowing the system's parameters), the movement of the regular users can be deterministically predicted by opinion dynamics in Eq. (3.2), by superposition of the effects.

The last element to be specified to characterize our game is the *payoff* (or *utility*) function, which in our setting is a function $u_i : \mathscr{A} \times \mathscr{A} \to \mathbb{R}$ and specifies the preferences of the players over the outcomes of the game, given the strategy of the other player(s). In our case, the payoff function corresponds to the objective function in Eq. (3.3), where the opinion configuration depends on the combined *actions* of the two influencers: $x_N^{(u)}\left(a^{(0)}, a^{(1)}\right)$.

The game is a *simultaneous* game, i.e. both players choose their strategies at the same time and then stick to their choice for the entire duration of the game. It is also a game with *complete information*, i.e. both players know perfectly the rules of the game, i.e., they know the set of actions playable by the other player and the effect that such actions exert on the population of users.

3.5.1 The Two-Groups Scenario

We first consider a simplified but illustrative scenario, similar to the one in Section 3.4.2, for which it is possible to provide an exact procedure to determine the Nash Equilibria (NE) under the assumption that only pure strategies may be adopted.

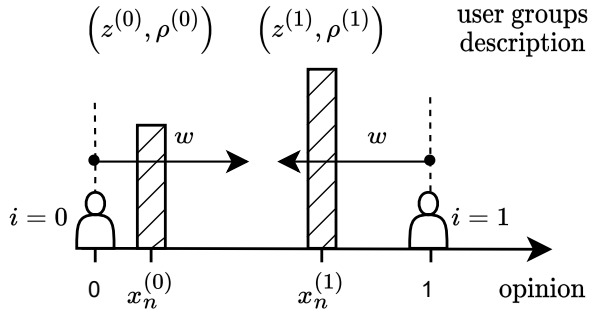


Fig. 3.4 Schematic representation of the simplified scenario.

The restriction to two user groups is not strictly necessary, and we could consider a population with M groups, each with a (static) prejudice $z^{(m)} \in \mathscr{Z}$. Indeed, the complexity of the procedure to determine the NE would only scale linearly

with M, as will become clearer later. However, we decided to limit the number of groups to two for continuity with Section 3.4.2 and ease of interpretation. We (first) consider a 0-1 ψ function to ensure identical reactions from the bulk of users belonging to the same group. In so doing, we avoid the exponential growth of the user groups over time generated by their splitting. In Section 3.6 the case of a general user distribution and ψ function is discussed. Each group is characterized by $z^{(m)} \in \mathcal{Z}$ (assuming $x_0^{(m)} = z^{(m)}$) and its "proportion" defined as $\rho^{(m)} \triangleq \frac{\int \mathrm{d}\mu_N(x|z^{(m)})}{\int \mathrm{d}\mu_N(x)}$ with respect to the overall population. Lastly, we consider influencers whose target opinions lay at the extreme of the opinion domain: $x^{(i;T)} \in \{0,1\}$. Fig. 3.4 sketches the setting. We first look into some straightforward solutions to the problem:

Proposition 1. *If* $|x^{(i;T)} - z^{(m)}| \leq w$, $\forall m \in \mathcal{M}, \forall i \in \mathcal{I}$ *then* $(0,0)$ *is a Nash equilibrium (NE) of the game.*

Proof. Without loss of generality, we can consider i' such that $x^{(i';T)} = 1$ (recall $x^{(i;T)} \in \{0,1\}$). Given that $x_0^{(m)} = z^{(m)}$, $\forall m \in \mathcal{M}$ and $|1 - x_0^{(m)}| \leq w$, $\forall m \in \mathcal{M}$, then $a^{(i')} = 0$ is a *dominant* strategy for i', i.e., $\arg\max_{a^{(i')}} \mathbb{E}_x[x_N^{(u)}(\{a^{(i)}\}_i)] = 0$, $\forall a^{(i)}, i \in \mathcal{I} \setminus i'$. Indeed, the total opinion increment $\Delta x(\{x_n^{(i)}\}_n) \triangleq \Delta x^{(0)} + \Delta x^{(1)}$ for a general $x_n^{(i)}$ sequence is weakly smaller than $\Delta x(\{x^{(i;T)}\}_n)$, for any strategy of the other players. Every player would play this dominant strategy, so $(0,0)$ is a NE. □

It is clear that such situations are not of interest. Note also that in the exploration phase, a player favors the other player in some way (i.e. the closer user group moves further away from the player). A player would only do this if there is a possibility of reaching a larger user base, which is otherwise hindered by filtering.

We make a structural assumption that is in no way restrictive but has the dual goal of avoiding trivial solutions (see proposition 1) and ensures that the influencer does not deviate too much from her target opinion. This captures the fact that the influencer needs to still keep in contact with (and somehow maintain under control) the more radical individuals (those closer to $x^{(i;T)}$).

Assumption 1. *Only the closest group in opinion m is reachable (*$\psi(|x^{(i;T)} - x^{(m)}|) = 1 \neq 0$*) in the targeting phase while b) both (m, M) are reachable in the exploring phase.*

3.5.2 Exact Solution for the Two-Groups Scenario

Since the set of possible actions $\mathscr{A}$ is finite and as $|\mathscr{I}| = 2$, it is possible to find the Nash equilibria for the two-player game or to determine when no Nash equilibrium exists.

Theorem 3. *The procedure in Algorithm 3 identifies all the Nash equilibria in pure strategy of the game, if any exist.*

Proof. The rationale for the algorithm is that in a Nash equilibrium, each player plays a *best response* $\mathscr{BR}_i(a^{(i)}, a^{(-i)})$ to the other player(s). In a two-player game, it is possible to define a matrix $\mathbf{P}$ in which we have as rows the strategies of player $i = 0$ and as columns those of player $i = 1$. Each element p_{ij} is defined as (u_0, u_1). This matrix is well-defined because the action space $\mathscr{A}$ has finite cardinality, $\mathbf{P}$ is $N+1 \times N+1$. For each column k (strategy of player $i = 1$), we compute $\arg\max_{a^{(0)}} u_0(a^{(0)}, a^{(1)} = k)$, which is the *best response* of player $i = 0$ to the action k of player $i = 1$. We do the same over the rows and then consider the elements (i, j), if any, for which the procedure identified an argmax over both the rows and the columns. These are the Nash equilibria in pure strategy, as both players play their *best response* to each other's strategy. □

3.5.3 Characterization of the Nash Equilibria as a Function of the Population Characteristics

Figure 3.5 shows the Nash equilibria computed with the method in Algorithm 3, as a function of the user *inertia* β and the degree of stubbornness δ and under the Assumption 1. The experimental setting includes two equal user groups with $z^{(0)} = 0.25, z^{(1)} = 0.75$, the influencers are identical ($\delta^{(0)} = \delta^{(1)}$) and the considered time horizon $N = 5$, with a rectangular ψ function with width $w = 0.7$. We see that the less the population can be influenced, the more time the influencers spend on exploration. This can be explained by the fact that the closer an influencer is to a group of users, the more influence she can exert. This is also related to the fact that $(0,0)$ is not a NE for the game. Since influencers in the exploration phase have a higher influence and can reach more users, any influencer who knows that the *competing* influencer can switch to a *milder* opinion would also carry

Algorithm 3: Determine Nash equilibria

Data: players $\mathscr{I}$, groups of users $\mathscr{M}$, ψ function, actions $\mathscr{A}$
// Compute the payoff matrix P
foreach $\left(a^{(0)}, a^{(1)}\right) \in \mathscr{A} \times \mathscr{A}$ **do**
  $v \leftarrow \arg\min_i a^{(i)}$
  $V \leftarrow \mathscr{I} \setminus m$
  for $n \in \{1, ..., N\}$ **do**
    if $n < m$ **then**
      $x_{n+1}^{(m)} = UpdateBelief(x_n^{(v;E)}, x_n^{(V;E)}, \forall m \in \mathscr{M}$
    else if $m < n < M$ **then**
      $x_{n+1}^{(m)} = UpdateBelief(x^{(v;T)}, x^{(V;E)}), \forall m \in \mathscr{M}$
    else
      $x_{n+1}^{(m)} = UpdateBelief(x_n^{(v;T)}, x_n^{(V;T)}, \forall m \in \mathscr{M}$
  $P\left[a^{(0)}, a^{(1)}\right] \leftarrow \left(1 - \mathbb{E}_x\left[x_N^{(u)}\right], \mathbb{E}_x\left[x_N^{(u)}\right]\right)$
// Compute Best Responses and Nash equilibria
$\mathscr{BR}_0(a^{(1)}) \leftarrow \arg\max_{a^{(0)}} \mathbf{P}\left[\cdot, a^{(1)}\right]$
$\mathscr{BR}_1(a^{(0)}) \leftarrow \arg\max_{a^{(1)}} \mathbf{P}\left[a^{(0)}, \cdot\right]$
$NE = \left\{(a^{(0)}, a^{(1)}) : a^{(0)} \in \mathscr{BR}_0(a^{(1)}),\ a^{(1)} \in \mathscr{BR}_0(a^{(0)})\right\}$

out an exploration phase. In such a symmetric setting it is reasonable to expect symmetric equilibria (as in Fig. 3.5), and this is summarized in Proposition 2:

Proposition 2. *In the two-group scenario, in the symmetric case ($\delta^{(i)} = \delta\ \forall i$, $z^{(0)} = 1 - z^{(1)}$, $\rho^{(0)} = \rho^1$) the actions of the two players in the Nash equilibrium are the same if the NE is unique.*

Proof. It is rather straightforward to see that, in this situation, the best response of player 0 is $BR_0 = \arg\max_i u\left(\mu_n(x) \middle| a^{(1)}\right)$, and it is equal to the best response of player 1 due to symmetry. Thus, if the NE is unique, the two players play the same strategy. □

Under Assumption 1, the equilibria of the game are non-trivial, i.e., different from $(0,0)$, and there is for the players a strategic incentive to compete.

So far we discussed a fully symmetric situation but, it is interesting to consider also an asymmetric distribution, i.e., $z^{(0)} = 0.15$, $z^{(1)} = 0.75$, in Figure 3.6 we depict the Nash equilibria. This setup implies that player 0 has an initial advantage in

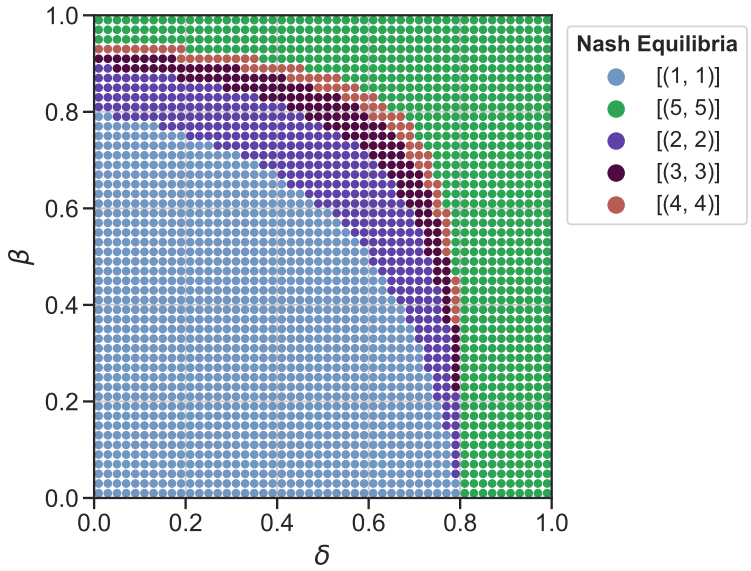


Fig. 3.5 Nash equilibria in a symmetric scenario as a function of the population characteristics β, δ. The Nash Equilibria are a list (of one element) of tuples of the form $\left(a^{(0)}, a^{(1)}\right)$.

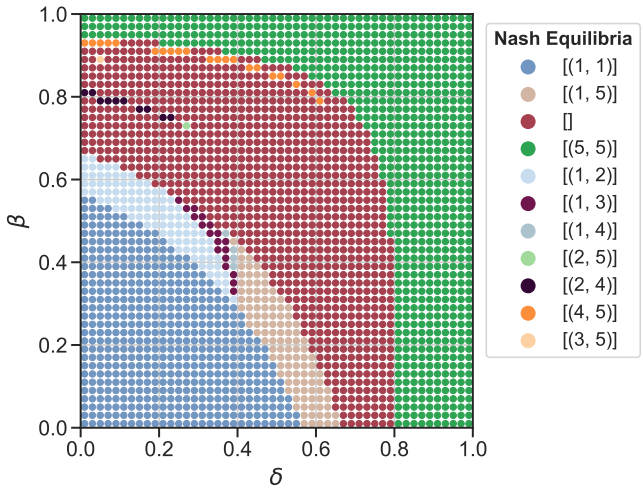
terms of user distribution. The structure of the above Nash equilibria suggests that the one in a disadvantageous position (here player 1) has a strategic incentive to prolong the exploration phases in order to approach the average collective opinion value of the population and thus also reach the most distant group and try to persuade it.

3.6 Towards a More Realistic Scenario

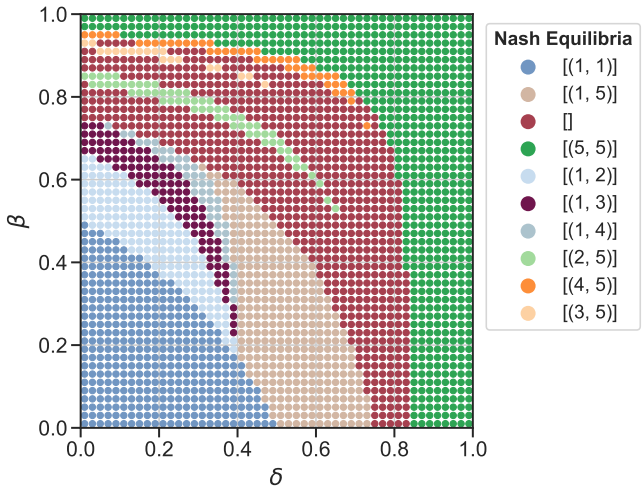
Online social networks are characterized by pronounced asymmetries in the interaction between entities. Furthermore, in [13], the authors model the closed feedback between regular users and influencers. User feedback determines the *popularity* of influencers, which in turn is tied to their ability to reach users, i.e., their *visibility* over the platform. In this section, we use the full specification of the model in [13] in our framework.

3.6.1 Closed-loop Opinion Model

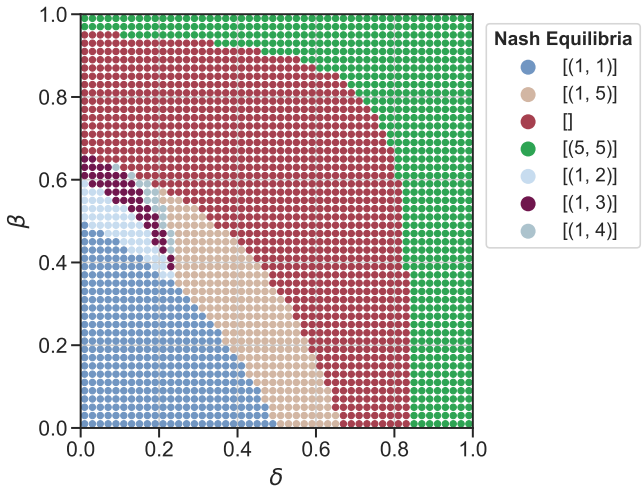
The users are subject to the dynamics described in Section 3.3.1 with the addition of the closed loop between regular users and influencers. For simplicity, we already



(a)



(b)



(c)

Fig. 3.6 Nash equilibria as a function of β and δ for different shifts $\delta^{(i)}$ from the target opinion $x^{(i;T)}$, we organize the shifts for the two influencers as $(\delta^{(0)}, \delta^{(1)})$ and we consider (0.2,0.2) in (3.6a), (0.1,0.2) in (3.6b), and (0.2,0.3) in (3.6c).

present the model in its deterministic form, see Remark 4. The user *feedback* can be directly derived by the value of ψ, as it describes when a user is reached by a certain post and the probability of him moving in the direction of the influencer's opinion because he likes it (positive feedback). Therefore we can define the total feedback provided by the users to a post of an influencer as $T_n^{(i)} = \int \psi_n^{(i)} \mathrm{d}\mu_N(x)$. The feedback allows us to define the popularity $p_n^{(i)}$ update:

$$p_{n+1}^{(i)} = p_n^{(i)}(t) + T_n^{(i)} \tag{3.7}$$

To close the loop, it is necessary to make ψ a function of popularity, i.e. the more popular an influence is in the OSN, the more users it can reach. To this end, consider the factorization of the function ψ into two contributions ($\psi = \omega \cdot \theta$), where the first factor models the *homophily* of interactions across OSNs together with the filtering of the platform, and the second describes the degree to which a user likes a post[7]. Considering θ a decreasing function of opinion distance also models the fact that users with very different opinions are less likely to be convinced. So even if an influencer is very popular, she will find it difficult to convince users who are distant in opinion.

3.6.2 Approximate Solution of the Game for Arbitrary ψ and User Distribution

The closed-loop scenario is more realistic but clearly more complicated. The procedure in Algorithm 3 cannot be applied directly because a group of users m cannot be tracked perfectly. At each time instant n, a certain group "splits"into a subgroup that is influenced and updates its opinion (with probability $\psi(d, p_n^{(i)})$), and the complementary group, which does not move. The number of subgroups grows exponentially as $2^N M$. To avoid the exponential explosion, one can discretize the opinion space into $\mathscr{B}$ bins and keep the proportion of users from the groups in $\mathscr{M}$ in each of the bins. With this simplification, it is possible to apply Algorithm 3 and consider any ψ function and also any (discrete) distribution of

[7] Only the product $\omega\theta$ is relevant for the total feedback since a user must be reached ($\Omega = 1$) to give positive feedback ($\Theta = 1$) in the general stochastic model.

users. The ψ function in the product form we used in our experimental setting is:

$$\psi_n^{(i)}(d, p_n^{(i)}) = \omega(d, p_n^{(i)})\,\theta(d) = e^{\nu \cdot \frac{d^2}{p_n^{(i)} / \sum_i p_n^{(i)}}}\,(1-d) \tag{3.8}$$

only the first factor depends on the relative popularity of the given influencer, and the parameter ν controls to what extent the visibility of an influencer decays with the distance.

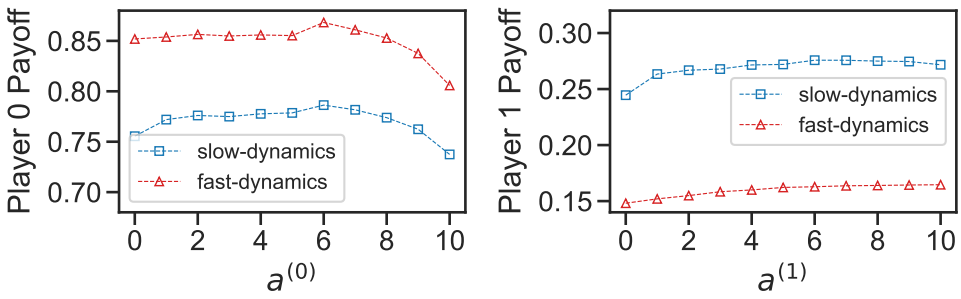
From now on, we consider two possible scenarios, the "slow-dynamics", in which popularity evolves considerably slower compared to opinions, and the "fast-dynamics", in which popularity evolves quicker and has a greater impact on the visibility over the OSN. It is immediate to do so by appropriately normalizing the total feedback update $T_n^{(i)}$ in Eq. (3.7), choosing a large enough constant.

3.6.3 The Effects of Strategic Behavior and Popularity Evolution

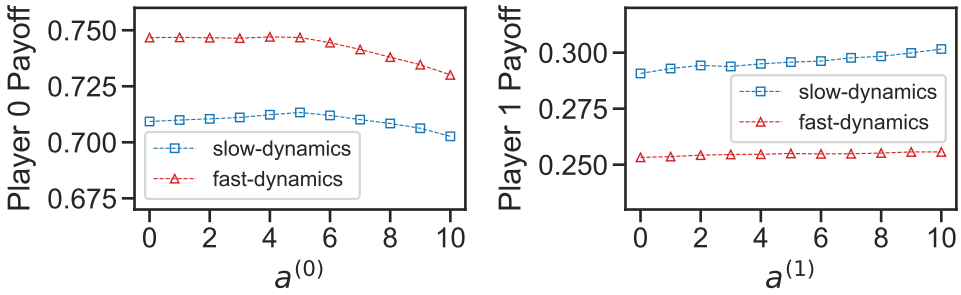
We developed a game for which the solution is limited to a very simplified structure. We will briefly discuss whether this simplified strategic behavior leads to a sizeable advantage for the players who behave strategically. To do this, we consider that one of the players is *stubborn*, i.e., she only posts her target opinion $x^{(i;T)}$. We can do this by looking at the first column and the first row of the payoff matrix P, which correspond respectively to a stubborn player 1 and a stubborn player 2.

We consider a Beta-shaped prejudice distribution (see gray lines in Fig 3.8a - 3.8b), skewed towards $x = 0$ and whose characterizing parameters $a = 2$ and $b = 4$. Similarly to previous sections, we assume that the initial opinion coincides with the users' prejudice. The two players have $x^{(i;T)}$ of 0 and 1 respectively and both have $\delta^{(i)} = 0.1$. We consider two possible choices for the underlying opinion update weights, i.e., $\alpha = 0.1, \beta = 0.5$ (Scenario 1) and $\alpha = 0.1, \beta = 0.8$ (Scenario 2). In this setting, we slightly modified the dynamical behavior so that the popularity can evolve macroscopically. For each action, we consider $N = 10$ posts are emitted carrying either the target opinion $x^{(i;T)}$ or the exploratory one $x^{(i;E)}$.

The results are shown in Figure 3.7 and reveal two aspects: first, the player who has a *structural advantage*, i.e., is favored by the initial distribution, benefits from the fast-dynamic settings. Second, the two players would exhibit somewhat opposite behavior, in that the favored player is harmed by long exploration



(a)



(b)

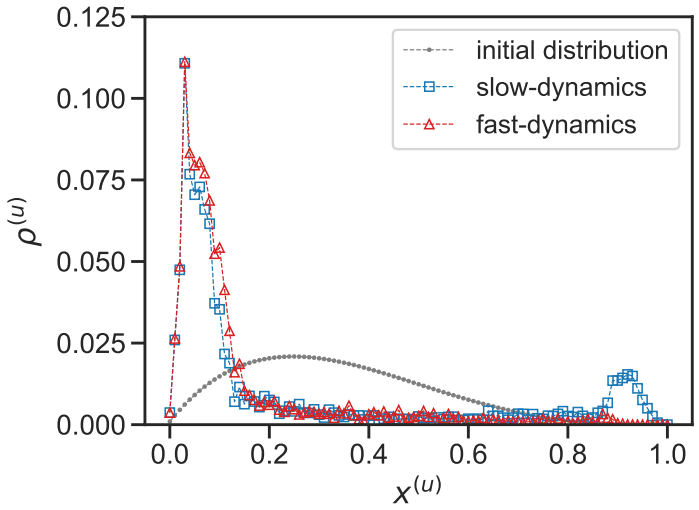
Fig. 3.7 Player's payoffs considering the other influencer as stubborn with $x_n^{(i;T)}$, $\forall n$ for Scenario 1 (top) and Scenario 2 (bottom).

phases, while the other receives better payoffs with long exploration phases. This supports the claim that the one in a disadvantaged position would benefit from being more "aggressive" and compromising her target opinion to get closer to the majority opinion in the population. This is also supported by the fact that in our experiments when looking at the Nash equilibria, we have that: $a_{NE}^{(0)} \leq a_{NE}^{(1)}$.

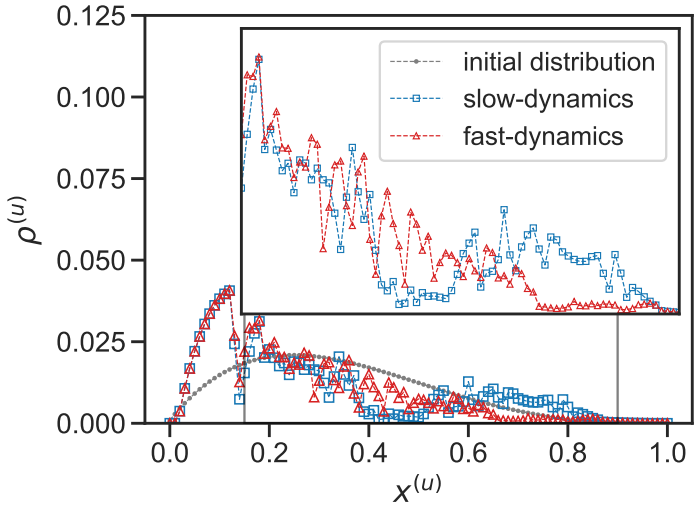
Finally, from Figure 3.8, it is clear that the rapid evolution of popularity erases competition, as the *advantaged* influencer is able to become more visible on the platform (see Eq. (3.8)) than the other, and eventually attracts all users. Note the bimodal nature of the final distribution in the slow-dynamics setting (blue curve) and the flat nature of the distribution in the fast-dynamics setting (red curve).

3.7 Concluding Remarks

In this chapter, we have developed a game-theoretic framework to model influencer competition. We have seen that even the case of an influencer in isolation is



(a)



(b)

Fig. 3.8 Final opinion distribution at the Nash equilibria for (b) Scenario 1 (c) Scenario 2, with different popularity growth speeds, in gray the prejudice (and initial opinion) distribution of the population is reported.

particularly complex and we were able to develop a method to derive the optimal solution only in a simplified setting. Thanks to our trellis-like approach, we have found that the best strategy to influence a population of users is to first group them around a common opinion and then attract all users towards the target opinion.

Furthermore, we studied the game between two players (influencers) and provided a tool to determine the Nash equilibria in pure strategy. We found

that in unbalanced situations, the structurally disadvantaged influencer has to compromise more on their opinion and adopt a more “aggressive” behavior in order not to succumb to the other influencer. In a more complex scenario, where each player’s popularity is updated according to feedback from regular users, the filtering effect combined with the rapid evolution of popularity tends to erase competition in favor of the structurally advantaged player.

Chapter 4

A SIR-like Stratified Epidemic Model

Part of the work presented in this chapter has already been published in [77]:

- Galante, F., Ravazzi, C., Garetto, M., & Leonardi, E. (2023). Planning Interventions in A Controlled Pandemic: The COVID-19 Case. In IEEE Transactions on Network Science and Engineering. doi: 10.1109/TNSE.2023.3343807

In this chapter, we present a novel SIR-like model that allows us to assess the impact of delays in disease control measures and the impact of vaccination. We begin the chapter with an overview of classical epidemic models. The proposed model improves on the literature by incorporating: i) variability in mortality rates and risk exposure between different population segments, ii) closed-loop control mechanisms to regulate the epidemiological curve, and iii) progressive vaccination campaigns. More precisely, differently from previous works on the subject (see [78], [79], [80], [81], [82], [83], [84]), our modeling framework explicitly represents the heterogeneity of risk exposure across population segments, through the definition of a population-specific distribution $f_{r,p}$ (see Appendix D).

We compare simple vaccination strategies based on assigned population-segment priorities and intervals between first and second-dose administration. The combined aspects of rate/ICU occupancy control in feedback and vaccination prioritization are not explicitly addressed in the literature. Indeed, the two approaches are usually studied separately. The lack of research addressing the integration of non-pharmaceutical control in feedback and vaccination prioritization points to a significant gap in the literature. Especially in the context of the

COVID-19 pandemic, where vaccination campaigns were implemented alongside control measures. This chapter aims to fill, at least in part, this gap.

This chapter opens by presenting a detailed overview of epidemic models, from the celebrated SIR to more complex compartmental models. Section 4.3 introduces the proposed epidemic model that describes the population in terms of their risk exposure and susceptibility to the virus. Section 4.4 presents the control measures for the epidemics considered in this chapter, then Section 4.5 examines the control exercised over new infections, and Section 4.6 examines the control exercised over hospital and ICU occupancy. Finally, we consider a comprehensive scenario spanning three years and inspired by the COVID-19 pandemic in Section 4.7, to compare different strategies using our model. We conclude the chapter with Section 4.8.

4.1 Related Work and Context

In this introductory section, we provide a comprehensive overview of the epidemic models from the literature, greatly extending the aspects discussed in the Introduction for the Susceptible-Infected-Susceptible model.

4.1.1 Modeling Epidemic Spread

The so-called SIR model [85] is paradigmatic in epidemiology and has been widely adopted to model infectious diseases for which recovered individuals acquire lasting or at least sufficiently durable immunity. Its complementary model, the SIS, which we learned about in the Introduction, is used when exposure to the virus does not confer immunity. In particular, the SIR model and its extensions have proven useful in modeling the dynamics of epidemic diseases such as seasonal influenza [86] and swine flue [87], and marked a significant milestone in utilizing mathematical models to understand disease dynamics and forecast the spread of epidemics. Since the outbreak of COVID-19 in late 2019, it has been an effective tool for studying the spread of the novel coronavirus. This model is the prototype of a broader class of models that partition the population according to disease status, called *compartmental* models [88]. One of the keys to the success of the SIR model is its simplicity, considering only three compartments: Susceptible $S(t)$,

Infected $I(t)$ and Removed $R(t)$, whose dynamics are described by the following system of ordinary differential equations:

$$\begin{aligned} \frac{dS(t)}{dt} &= -\beta \cdot \frac{S(t) \cdot I(t)}{N} \\ \frac{dI(t)}{dt} &= \beta \cdot \frac{S(t) \cdot I(t)}{N} - \gamma \cdot I(t) \\ \frac{dR(t)}{dt} &= \gamma \cdot I(t) \end{aligned} \tag{4.1}$$

N in the system of equations represents the total number of individuals in the population. β is a fundamental parameter and indicates the average number of contacts per person per time. This factor multiplies the term $\frac{S(t)I(t)}{N}$, which is linked with the probability of a virus transmission event (i.e., an infectious individual infects a susceptible one), assuming homogeneous mixing of the population. The parameter γ indicates the rate at which an individual exits the infectious state, either by recovering from the disease or dying. These compartmental models are better represented by block diagrams which highlight the transitions among states. For example, Figure 4.1 depicts the SIR model (A) and some of its extensions (B-C) which we discuss in the following.

However, the model's simplicity comes at the expense of oversimplifying the complexities of the disease processes [89]. For example, it assumes (similarly to the SIS) a homogeneous mixing of individuals. Therefore, it does not consider the correlation between daily contacts and other specific characteristics of the population (e.g., age). In [90], age-specific contact patterns are derived for different countries, which allows for a more accurate description of the population's interactions. The population is no longer considered uniform but grouped according to the individual's age. Moreover, the population is assumed to be closed, with no in-migration or out-migration. Accounting for this would not be complicated. It is sufficient to have information regarding a particular population's birth and death rates, together with immigration (individuals coming to the population from another population) and emigration information. In most cases, these effects almost balance out, leaving the overall number of individuals approximately the same. Indeed, the closed population assumption is reasonable as long as the time horizon is not too long (a few years). In addition, the SIR model does not account for the period during which an individual has been exposed to the virus but does not yet have sufficient infectious levels for transmission to others.

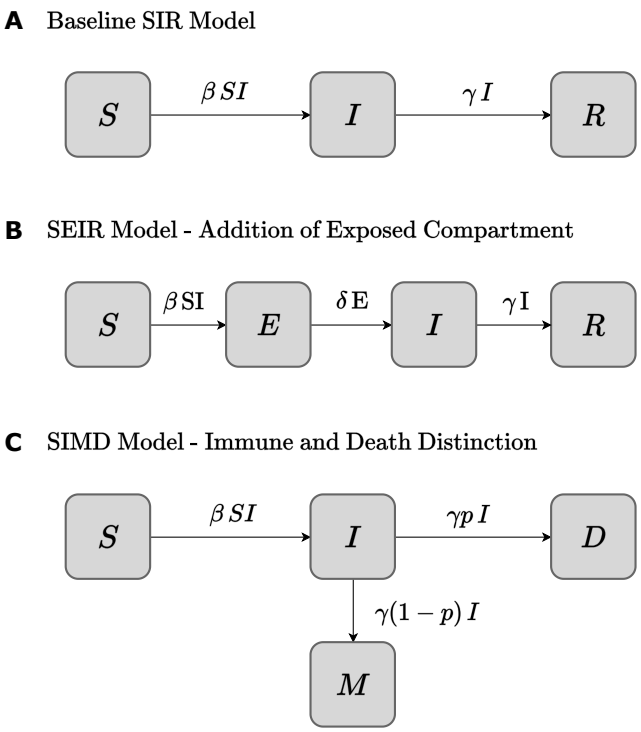


Fig. 4.1 Diagrams highlighting the transitions among compartments for the SIR, SEIR, and SIRD models introduced in this section.

A straightforward extension is to incorporate an additional compartment, usually denoted with E (standing for Exposed), which defines what is known as the SEIR model [91], Figure 4.1 (B). This only represents a delay in the dynamics for those individuals who have been in contact with the virus resulting in an infection and will, after $\frac{1}{\delta}$ units of time on average, become infectious for others. Another straightforward extension of the basic SIR model is is to distinguish in the *Removed* compartment between those who die as a result of the virus (Deaths) and those who acquire immunity against the virus (iMmune). Thus the model is referred to as SIMD, Figure 4.1 (B). It is sometimes called SIRD in the literature, depending on the convention used to indicate immune individuals, i.e., M from iMmune or R from Removed. The model extension is relatively simple and only relies on knowing the probability p with which an infected individual dies after infection. As it is clear from the block diagram (panel C in Fig. 4.1), this is achieved by simply multiplying the rates by the *probability of dying* p. Models that account for both aspects have long been used in the literature. Much of the most recent works on COVID-19 [92] [93] use SIR-like models and even add additional compartments,

such as one for asymptomatic individuals, i.e., infected individuals who do not manifest any COVID-19-specific symptoms, and are potentially more dangerous than symptomatic individuals who may reduce their contacts rate due to the insurgence of symptoms.

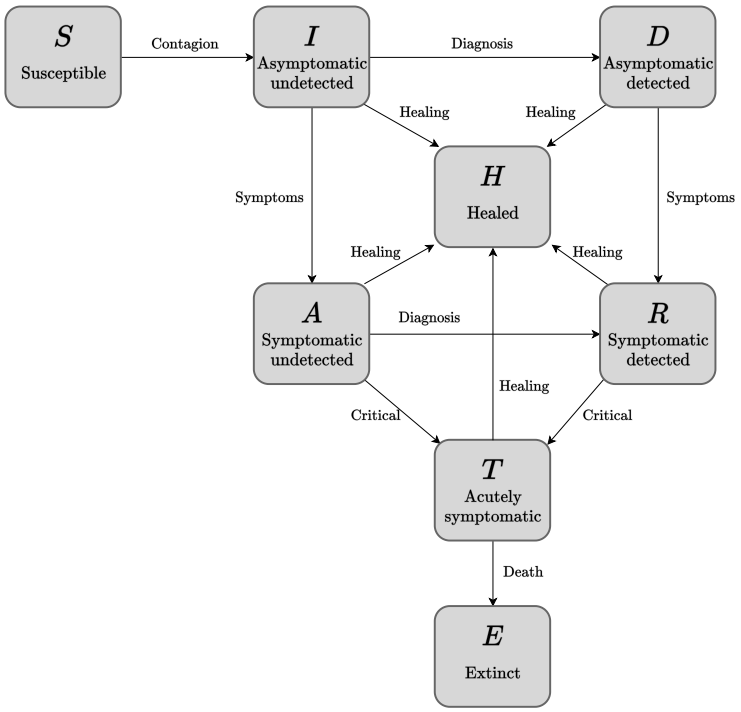


Fig. 4.2 Schematic representation of the SIDARTHE model, similar to that presented in [94]. We did not report the transition rates as we do not show the model's equations, the word description over the arrows provides more intuition.

One of the most comprehensive models recently appeared in the literature to model COVID-19 [94] considers eight states that distinguish between detected and undetected infectious cases, varying severity of illness (symptomatic and asymptomatic cases), non-life-threatening cases and potentially life-threatening cases requiring ICU admission. It is called the SIDARTHE model, and we present its block diagram in Figure 4.2. The states are indeed susceptible (S), infected (I), diagnosed (D, which represents detected asymptomatic cases), ailing (A, more severe, i.e., symptomatic, cases which have not been detected), recognized (R, symptomatic cases detected), threatened (T, acute symptomatic detected), healed (H) and extinct (E, death as a result of severe infection). This model has been purposefully developed to capture the peculiarities of COVID-19 (e.g., the distinction made between detected and undetected cases) and fits well data related to COVID-19

spread in Italy. With respect to the model we present in the next section, this approach does not consider heterogeneity in the population and the epidemiological states are specific to COVID-19, our approach aims to be more general.

Another fundamental aspect that has not been considered in the models above is the effect of vaccinations on the dynamics. In the case of COVID-19, vaccines started to become available towards the end of 2020, and great debate sparked around how vaccines should be prioritized. Our model also aims to answer this question and provide valuable insights into the possible trade-offs the decision-maker could face. Clearly, in the literature, there already exist models that consider vaccinations. For example [95] distinguishes between vaccinated who are protected by the vaccine, vaccinated without protection against COVID-19, and unvaccinated because of a positive serotest (note that upon recovery, infected individuals acquire immunity towards the virus) or refusal to vaccinate (i.e., no-vax). We consider these aspects in our model in one of its extensions presented in the next section and depicted in Figure 4.6.

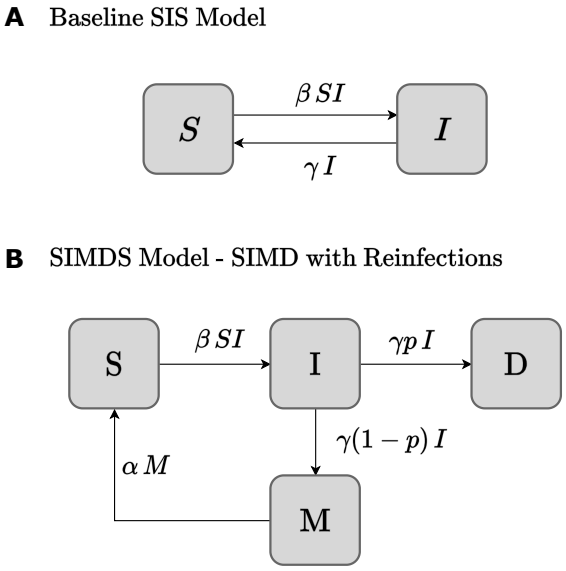


Fig. 4.3 Schematic representation of models with reinfections.

One final observation: the SIR (and SIR-like) underlying assumption of long-lasting immunity does not apply to all viruses. In the case of SARS-CoV-2, it is not yet clear how long immunity to the virus will last. Townsend et al. [96] claim that "reinfection by SARS-CoV-2 under endemic conditions would likely occur between 3 months and 5 years after peak antibody response". The susceptible-infected-

susceptible (SIS, see block diagram in Figure 4.3 (B)) model is the most simple model when reinfection comes into play which we have encountered and studied in the Introduction. The SIS model assumes that no protection against the virus is acquired upon infection and, at the end of the acute phase, an individual could be *immediately* be reinfected again. The SIMDS model, which is a refined version of the SIS model and extends the SIMD model, encompasses the possibility of a *temporary* immunity against the virus. This is captured by the M state (iMmune), where individuals do not contract the virus due to the obtained immunity. After an exponential time with parameter α, an immune individual loses its protection and returns to the susceptible state. This framework seems to be more appropriate in the context of COVID-19 since it appears that the acquired immunity is only temporary.

4.1.2 Control via Non-Pharmaceutical Interventions

For the mode proposed later in this chapter, we assume a central planner perspective, where the government can impose control measures on the population for the overall benefit of public health. Several papers published in the 1970s, such as [97–101], focused on studying optimal control problems in the context of the classical SIR model. Building upon these foundations, subsequent works like [102] and [103] expanded on the research and extended the models. These studies specifically addressed the challenge of minimizing the size of an epidemic outbreak and the cost of interventions. The control mechanisms explored in these models included regulating social distancing levels and implementing measures like isolation. These control strategies were also subject to rate constraints, ensuring the rate remained below a specified threshold. Papers on optimal control problems [104, 105] consider the minimization of a composite function taking into account the epidemic cost, related to the size of the outbreak or number of deaths, and the economic cost. The control perspective has been widely embraced in recent literature regarding the COVID-19 pandemic. Consequently, there has been considerable discussion about the effects of lockdown measures on healthcare, society, and the economy. The problem of minimizing the cost of a lockdown under the only constraint of maintaining the infection below a certain threshold to cope with ICU congestion problems is also considered in [106].

There is a growing recognition of the need to proactively design *simple* and effective control measures to combat and mitigate the spread of infectious diseases in the possibility of future pandemics. We refer the readers to [107], which offers an overview of various mathematical models for epidemic processes, encompassing traditional group models with no assumed graph structures, network-based models, stability analysis, parameter estimation methods, and simulation models.

4.1.3 Control via Vaccination

The paper [108] focuses on the optimal control of vaccination dynamics during an influenza epidemic. It provides insights into the design of vaccination strategies to effectively control the spread of the disease, considering factors such as limited vaccine supply and variations in transmission and severity across different groups. In [109], optimal vaccination and treatment strategies are studied in a multi-group epidemic model. The analysis explores the trade-offs between vaccination coverage and treatment allocation to maximize overall disease control, considering the interactions between different population groups. Finally, [110] explores the optimal timing and allocation of vaccinations based on age groups to maximize the effectiveness of the vaccination campaign and minimize the spread of infectious diseases. The research conducted in [111] and [112] contributes to the field by addressing the challenge of resource allocation for vaccination efforts in the context of epidemic control. By utilizing optimization techniques, these studies provide insights into the most effective strategies for targeting specific nodes in a contact network or groups within a population, considering both the budgetary constraints and the dynamics of disease transmission.

4.2 Notation

In this chapter, we use a subscript to indicate the socio-demographic class to which a particular variable refers. The superscript is used to convey additional information, such as a specific characteristic of the symbol or to indicate that it refers to the minimum/maximum/optimal value. The time is continuous and is denoted with t between brackets. Since the model we present consists of several compartments (and depends on many parameters), Table 4.1 is not intended

to be an exhaustive summary of all symbols, but rather a collection of the most frequently used ones.

Table 4.1 Notation Summary Table

Symbol	Description
r	Risk exposure to the virus
p	Probability of death due to the virus
$f_{r,p}$	Joint probability distribution describing socio-demographic groups
N	Size of the population, number of individuals
$S_{r,p}(t)$	Susceptible individuals of class (r,p)
$I_{r,p}(t)$	Infected individuals of class (r,p)
$M_{r,p}(t)$	Immune individuals of class (r,p)
$H_{r,p}(t)$	Hospitalized individuals of class (r,p)
$T_{r,p}(t)$	Under intensive treatment individuals of class (r,p)
$D_{r,p}(t)$	Death individuals of class (r,p)
$\sigma((t)$	Exogenous factors in the infection strength
$\lambda_U(t)$	Uncontrolled rate of infection
μ	Reinfection rate
γ	Removal (recovery/death) rate
ϕ	Hospital discharge rate
τ	ICU discharge rate
$p_{r,p}^{C_1,C_2}$	Probability of moving from compartment C_1 to C_2
$\widehat{H}$	Number of available hospital beds
$\widehat{T}$	Number of available ICUs
$\mathscr{R}_0$	Basic reproduction number
Δ	Time between two vaccines administration
ζ	Vaccine administration rate
VE^x	Vaccine efficacy of the x-th dose
$V_{r,p}^{xm}$	Vaccinated individuals in (r,p) who gained protection after the x-th dose
$V_{r,p}^{xs}$	Vaccinated individuals in (r,p) who are still susceptible after the x-th dose
$C_{r,p}^{v}$	Individuals in compartment C who got infected after vaccination
$\rho(t)$	Control parameter
$\lambda(t)$	(Controlled) intensity of new infections
$\mathfrak{C}(\rho)$	Economic cost
t_{max}	Total time horizon

4.3 Model Description

The epidemic model we propose aims to provide a flexible framework that can be applied to a wide range of viruses. We use COVID-19 as a use case, also due to the large data availability (see Appendix D). The proposed model considers different disease severity levels, differentiating between *hospitalized* individuals (H) and

those who need to be under intensive care (T). It considers the possibility of losing immunity, with a rate of μ, and, in its most general version (see Section 4.3.2), explicitly models the vaccination process, distinguishing between individuals who have undergone *partial* and *complete* vaccination.

Another distinguishing feature of our model lies in the distribution $f_{r,p}$, which characterizes population heterogeneity relating the *risk exposure* r and the *mortality rate* p. Furthermore, our modeling framework describes a pandemic in a tightly controlled scenario in which non-pharmaceutical (e.g., social distancing, lockdown) and pharmaceutical (i.e., vaccination) interventions are in place. This better describes an ongoing pandemic scenario.

4.3.1 Base Model

We start by introducing the base version of our compartmental model for the spread of a disease in a non-homogeneous population of size N without any intervention (either pharmaceutical or non-pharmaceutical). Socio-demographic groups are described by the joint distribution $f_{r,p}$ of the risk exposure r, corresponding to the contact rate of individuals, and their probability of death p.

We consider six epidemiological states. Let $S_{r,p}(t)$, $I_{r,p}(t)$, $M_{r,p}(t)$, $H_{r,p}(t)$, $T_{r,p}(t)$, and $D_{r,p}(t)$ denote the number of individuals characterized by (r,p) who at time t are susceptible, infected, immune, hospitalized, under intensive treatment and dead, respectively. The system presented below can be derived from stochastic processes, as in [113]. Thus, there is an underlying agent-based model in which all states have a statistical interpretation. In particular, the time an individual spends in the infected, hospitalized, intensive care unit and immune compartments is exponentially distributed with mean values $1/\gamma$, $1/\phi$, $1/\tau$, and $1/\mu$, respectively. Figure 4.4 visually represents the transitions between different compartments, which can be described as follows:

- Susceptible to Infected: Susceptible individuals (S) become Infected (I) when they come in contact with infected individuals. The corresponding transition rate depends on the risk exposure r (contact rate) and the number of infected individuals in the population.

- Infected to Immune: Infected individuals (I) can recover from the disease and acquire immunity, transitioning to the iMmune state (M). The *recovery* rate γ governs the transition and depends on the average infection duration before recovery.

- Infected to Hospitalized: Some Infected individuals (I) may develop severe symptoms and require Hospitalization (H). Various factors influence the transition rate, such as the capacity of the health care system, the proportion of infected individuals requiring hospitalization, and also disease severity in relation to the *fragility* of individuals.

- Hospitalized to Under Intensive Treatment: Hospitalized individuals (H) who require intensive care Treatment may be transferred to the intensive therapy state (T). The corresponding rate depends on factors such as the availability of intensive care units and the duration of hospitalization before the transfer to ICU.

- Under Intensive Treatment to Deceased: Unfortunately, some infected individuals under intensive Treatment (T) may succumb to the disease and move to the Deceased state (D).

In Appendix F we consider a straightforward extension of the model, showing how direct transitions $I \to D$ and $H \to D$ can be added to the model.

The following set of ordinary differential equations describes the dynamics:

$$
\begin{aligned}
\dot{S}_{r,p}(t) &= -\sigma(t)\left(\sum_{r',p'} r' I_{r',p'}(t)\right)\frac{r S_{r,p}(t)}{\sum_{r',p'} r' N f_{r',p'}} + \mu M(t) \\
\dot{I}_{r,p}(t) &= \sigma(t)\left(\sum_{r',p'} r' I_{r',p'}(t)\right)\frac{r S_{r,p}(t)}{\sum_{r',p'} r' N f_{r',p'}} - \gamma I_{r,p}(t) \\
\dot{H}_{r,p}(t) &= \gamma p_{r,p}^{IH} I_{r,p}(t) - \phi H_{r,p}(t) \\
\dot{T}_{r,p}(t) &= \phi p_{r,p}^{HT} H_{r,p}(t) - \tau T_{r,p}(t) \\
\dot{D}_{r,p}(t) &= \tau p_{r,p}^{TD}(t) T_{r,p}(t) \\
\dot{M}_{r,p}(t) &= \gamma(1 - p_{r,p}^{IH}) I_{r,p}(t) + \phi(1 - p_{r,p}^{HT}) H_{r,p}(t) + \tau(1 - p_{r,p}^{TD}(t)) T_{r,p}(t) - \mu M(t)
\end{aligned}
\tag{4.2}
$$

where $\sigma(t) \geq 0$ represents all exogenous (uncontrolled) factors changing the infection strength (e.g., seasonal effects). In this chapter, we will assume for simplicity that $\sigma(t) = \sigma$ is constant.

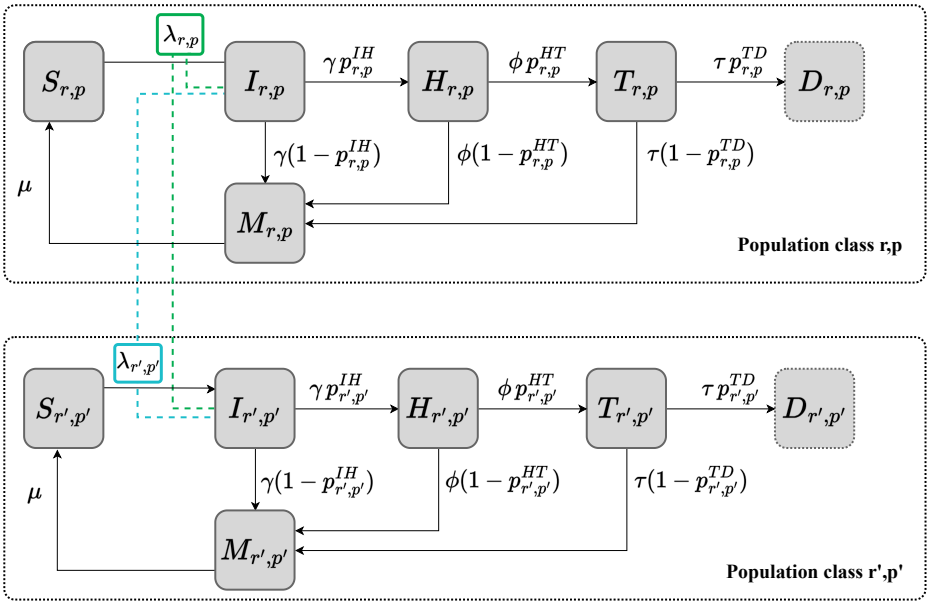


Fig. 4.4 Schematic representation of the proposed model.

The total (uncontrolled) rate of new infections is equal to:

$$\lambda_U(t) = \sigma(t)\left(\sum_{r,p} r I_{r,p}(t)\right)\frac{\sum_{r,p} r S_{r,p}(t)}{\sum_{r,p} r N f_{r,p}}.$$

The total number of susceptible people is $S(t) = \sum_{r,p} S_{r,p}(t)$. Similarly, we introduce the total number of people in the other compartments: $I(t)$, $H(t)$, $T(t)$, $M(t)$, $D(t)$. Probabilities $p^{IH}_{r,p}$, $p^{HT}_{r,p}$ and $p^{TD}_{r,p}(t)$ denote the probability that an individual of type (r,p) moves between the two compartments indicated in the superscript. We make probability $p^{TD}_{r,p}(t)$ depend on $T(t)$, i.e., on the instantaneous total number of people in ICUs since the death probability dramatically increases when ICUs are saturated. Denoted with $\widehat{T}$ the number of available ICUs, when $T(t) \leq \widehat{T}$, the overall death probability of an infected person is assumed to be equal to p:

$$p^{IH}_{r,p} \cdot p^{HT}_{r,p} \cdot \hat{p}^{TD}_{r,p} = p \qquad \text{if } T(t) \leq \widehat{T}, \tag{4.3}$$

where $\hat{p}^{TD}_{r,p}$ is the probability to transit from state T to state D in 'normal' conditions, i.e., when $T(t) \leq \widehat{T}$. Therefore, $p^{TD}_{r,p}(t) = \hat{p}^{TD}_{r,p}$ as long as $T(t) \leq \widehat{T}$.

When $T(t) > \widehat{T}$, we assume that the death probability of people who cannot receive treatment is increased by a factor θ, hence $p^{TD}_{r,p}(t)$ is dynamically adjusted as follows:

$$p^{TD}_{r,p}(t) = \hat{p}^{TD}_{r,p}\frac{\widehat{T}}{T(t)} + \min\{1, \theta \cdot \hat{p}^{TD}_{r,p}\}\frac{T(t)-\widehat{T}}{T(t)}. \tag{4.4}$$

We consider the case in which individuals might lose immunity with rate μ, thus becoming susceptible again. It should be noticed that the mass preservation $\dot{S}_{r,p}(t) + \dot{I}_{r,p}(t) + \dot{M}_{r,p}(t) + \dot{H}_{r,p}(t) + \dot{T}_{r,p}(t) + \dot{D}_{r,p}(t) = 0$ holds for all $t \geq 0$.

The presented model is an extension of the standard SIR model. This can become more evident by looking at the block diagram in Figure 4.4. To make it even more straightforward, it is possible to rewrite the first two equations in the system of equations (F.1) defining $\beta_{r\leftrightarrow r'} := \sigma(t)\frac{r'r}{\sum_{r'',p''} r''Nf_{r'',p''}}$:

$$\begin{aligned}
\dot{S}_{r,p}(t) &= -\sum_{r',p'} \beta_{r\leftrightarrow r'} I_{r',p'}(t) S_{r,p}(t) + \mu M(t) \\
\dot{I}_{r,p}(t) &= \sum_{r',p'} \beta_{r\leftrightarrow r'} I_{r',p'}(t) S_{r,p}(t) - \gamma I_{r,p}(t)
\end{aligned} \tag{4.5}$$

where the newly introduced parameter $\beta_{r\leftrightarrow r'}$ represents the pattern of interaction between population segments with risk exposure r and r'.

References as [114, 115, 113, 116, 117] explore SIR-like models with various extensions, including population heterogeneity, additional compartments, and considerations of specific epidemics like COVID-19. It is worth noting that while it can be expected that some individuals will naturally reduce their interactions out of fear of illness, our model does not explicitly incorporate this behavior as explored in [118] and [119]. Additionally, we do not consider the concept of "cost of anarchy" as explored in [120]. Unlike [121], our model does not distinguish between infected individuals who remain undetected and those who are detected, nor does it consider this distinction for those who recover. Nonetheless, our model introduces several innovative features, summarized in the following three remarks.

Remark 7 (Heterogeneity of population in terms of fatality rate and risk exposure)**.** *At the country level, there are substantial differences in population characteristics, including age distribution, general state of health, and frequency of interpersonal interactions. These differences in population contact patterns can potentially in-*

fluence disease transmission and accelerate the outbreak. Unique distributions, denoted as $f_{r,p}$, for different countries were derived from data on contact patterns and case fatality rates. For more information on how these distributions were determined and to see several examples (e.g. Italy, China), see Appendix D.

The continuous model used in this study can be interpreted as a mean field approximation of an epidemic model that operates over a dynamic network [122–124]. According to this interpretation, the risk exposure parameter represents the average number of contacts (per time unit) an individual experiences with others over a fixed time window. Therefore it can be seen as the degree of the corresponding node within a network, in which nodes represent individuals, and the edges represent the contacts between them. Note that pairs of individuals establishing contacts are randomly selected, as for the configuration model. This approach allows us to understand the dynamics of epidemics in terms of the interactions between individuals in a network setting.

Remark 8 (Quadratic dependence on the risk exposure r)**.** *Note that individuals with large r, i.e., pronounced social attitudes, represent at the same time the component of the population with the highest risk of infection and the highest chance of transmitting the disease. Therefore, the "impact" of every individual to the spread of the infection depends quadratically on r.*

Remark 9 (Edge-perspective analysis)**.** *Defining $\widetilde{I}(t) = \sum_{r,p} r\, I_{r,p}(t)$ as the number of infected contacts, multiplying the first and second equation in (4.2) by r and summing over r and p, we obtain $\dot{\widetilde{I}}(t) = \gamma\,(\mathscr{R}(t) - 1)\,\widetilde{I}(t)$ where $\mathscr{R}(t) = \frac{\sigma \sum_{r,p} r^2 S_{r,p}(t)}{\gamma \sum_{r,p} r N f_{r,p}}$.*

At early stages of the epidemic, we can approximate $S_{r,p}(t) \approx N f_{r,p}$ (see Remark 10), obtaining: $\dot{\widetilde{I}}(t) = \gamma\,(\mathscr{R}_0 - 1)\,\widetilde{I}(t)$ where we define the related basic reproduction number $\mathscr{R}_0 = \frac{\sigma}{\gamma}\mathbb{E}[r^2]/\mathbb{E}[r]$. As it is clear from the system of equations describing the evolution of the state variables, an edge-perspective analysis provides a fundamental tool to study the dynamics as a natural generalization of the SIR model.

The system of equations (4.2) can be greatly simplified when $S_{r,p}(t) \approx N f_{r,p}$, and this is instrumental for our subsequent analysis. Indeed:

Remark 10. *Observe that whenever we can find a time interval $[0, T]$ in which we have $\frac{S_{r,p}(t)}{N f_{r,p}} \approx 1\ \forall r, p, t \in [0, T]$, then (4.2) can be dramatically simplified (i.e.,*

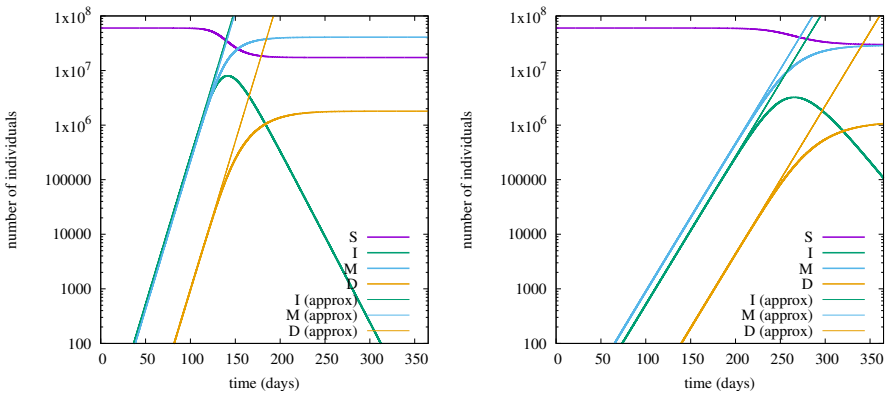
linearized) by replacing $S_{r,p}(t)$ with $Nf_{r,p}$ on the r.h.s. In particular, in such a case we can set: $\lambda_U(t) \approx \overline{\lambda}_U(t) := \sigma(t)\left(\sum_{r,p} r I_{r,p}(t)\right)$. Now, since by construction we have $S_{r,p}(t) \leq Nf_{r,p}\ \forall t, r, p \lambda_U(t) \leq \overline{\lambda}_U(t)\ \forall t$ and $S_{r,p}(0) = Nf_{r,p}\ \forall r, p$, denoted with $\mathbb{E}[r] = \sum_{r,p} r f_{r,p}$ the average risk exposure, from (4.2) *we get*

$$\begin{aligned} S_{r,p}(t) =& Nf_{r,p} - \int_0^t \lambda_U(\tau)\frac{rS_{r,p}(\tau)}{N\mathbb{E}[r]}\mathrm{d}\tau + \int_0^t \mu M(\tau)d\tau \\ \geq& Nf_{r,p}\left[1 - \int_0^t \overline{\lambda}_U(\tau)\frac{r}{N\mathbb{E}[r]}d\tau\right] + \int_0^t \mu M(\tau)d\tau \end{aligned}$$

Therefore, as long as it holds:

$$\max_r \int_0^t \overline{\lambda_U}(\tau)\frac{r}{N\mathbb{E}[r]}\mathrm{d}\tau \ll 1$$

we can approximate $S_{r,p}$ with $Nf_{r,p}$ for every (r,p). Lastly, observe that $\int_0^t \overline{\lambda_U}(\tau)\frac{r}{N\mathbb{E}[r]}d\tau$ is an upper bound to the fraction of individuals in class (r,p) who got infected in $[0,t)$.



$\mathscr{R}_0 = 2$ $\mathscr{R}_0 = 1.5$

Fig. 4.5 Evolution of $S(t)$, $I(t)$, $M(t)$, $D(t)$ in an uncontrolled scenario with $\mathscr{R}_0 = 2$ (left plot), or $\mathscr{R}_0 = 1.5$ (right plot). Comparison between exact solution (thick lines) according to (4.2), and approximate solution with $S_{r,p}(t) = Nf_{r,p}$ (thin lines).

To check numerically the limits of the validity of approximation $\frac{S_{r,p}(t)}{Nf_{r,p}} \approx 1$, we consider our reference scenario (see Table 4.3) related to the Italian population with $N = 60$ million, letting the epidemic evolve uncontrolled. In Fig. 4.5 we compare the exact solution of (4.2) (thick line) with the approximate solution in which $S_{r,p}(t)$ remains fixed and equal to $Nf_{r,p}$ (thin lines). As expected, $I(t)$, $M(t)$, $D(t)$ grow exponentially in the approximate solution (note the log vertical scale), matching the exact solution over an initial time window in which the total

number of infected, roughly equal to the total number of individuals $M(t)$ who have recovered, is comparatively small with respect to the total population N (say smaller than 5%). Indeed, when $M(t)$ becomes large, the number $S(t)$ of susceptible starts to drop (roughly after 100 days, with $\mathscr{R}_0 = 2$, or after 200 days, with $\mathscr{R}_0 = 1.5$), and the two solutions deviate from each other. Indeed, in the exact system, the epidemic eventually dies out due to herd immunity. Allowing the uncontrolled spread of the virus is not sustainable and, in the case of COVID-19, has not been considered a viable option by any developed country. Therefore, this regime is not of interest, the rate of new infections should be controlled and this makes our assumption reasonable.

4.3.2 Model with Vaccinations

In addition to non-pharmaceutical interventions, vaccination campaigns are a critical measure to contain the virus. Vaccines are assumed to guarantee partial protection. According to classification in [125], we consider two efficacy descriptors: reduction in the probability of becoming infected (vaccine efficacy on susceptibility) and reduction in the pathogenicity (vaccine efficacy to prevent or diminish symptoms). For simplicity, we neglect the vaccine response transient, and we consider a single type of vaccine administered in two doses separated by a fixed interval of Δ days. We assume that the administration rate of either dose is fixed, equal to ξ, so the entire population can be potentially vaccinated (with two doses) after $\mathscr{T}_v$ days. Hence we set $\xi = N/(\mathscr{T}_v - \Delta)$. Let $\mathsf{VE}^1, \mathsf{VE}^2$ be the vaccine efficacy on susceptibility after one or two doses, respectively. Moreover, we assume that mortality is reduced by a factor $q_{\texttt{post}}$ after a single dose of vaccine.

We assume that N^{novax} individuals refuse vaccination, uniformly distributed over the population. Their state evolution is still described by equations (4.2). Let $S^{\text{novax}}_{r,p}(t)$ be the number of no-vax people in class (r,p) still susceptible at time t.

We assume the individuals do not return to the susceptible state after infection or vaccination. This extension is easy, we omit it for brevity.

Vaccinations require the addition of a few more compartments: Let $V^{\text{1m}}_{r,p}(t)$ be the number of people in class (r,p) who have received just the first dose, which is already effective against the virus, i.e., they can no longer be infected. Let $V^{\text{1s}}_{r,p}(t)$

be the number of people in class (r,p) still susceptible after receiving just the first dose. Let $V^{2\mathrm{m}}_{r,p}(t)$ be the number of people in class (r,p) who have received both doses and are immune. At last, let $V^{2\mathrm{s}}_{r,p}(t)$ be the number of people in class (r,p) who have received both doses but are still susceptible. Due to strict prioritization among classes, a given class (r,p) receives the first dose at full rate ξ only within a specific time window: $[\mathscr{T}^{\min}_{r,p}, \mathscr{T}^{\max}_{r,p}]$ (to be specified later):

$$\xi^{(1)}_{r,p}(t) = \begin{cases} 0 & t < \mathscr{T}^{\min}_{r,p} \\ \xi & \mathscr{T}^{\min}_{r,p} \le t < \mathscr{T}^{\max}_{r,p} \\ 0 & t \ge \mathscr{T}^{\max}_{r,p} \end{cases}$$

Let $V^{1}_{r,p}(t) = \int_{t-\Delta}^{t} \xi^{(1)}_{r,p}(t)\mathrm{d}t$ be the number of people in class (r,p) who have received just the first dose of vaccine at time t. The second dose of vaccine is administered at a rate:

$$\xi^{(2)}_{r,p}(t) = \frac{V^{1s}_{r,p}(t) + V^{1m}_{r,p}(t)}{V^{1}_{r,p}(t)} \xi^{(1)}_{r,p}(t-\Delta)$$

only to individuals who have received the first dose and have not been infected in the meantime. At last, let

$$\widehat{S}(t) = \textstyle\sum_{r,p} r\,(S_{r,p}(t) + V^{1\mathrm{s}}_{r,p}(t) + V^{2\mathrm{s}}_{r,p}(t) + S^{\mathrm{novax}}_{r,p}(t))$$

be the total number of susceptible edges at time t.

Note that $\lambda(t) = \frac{\sigma}{\rho(t)} \left(\sum_{r,p} r I_{r,p}(t)\right) \frac{\widehat{S}(t)}{\mathbb{E}[r]N}$. Since people who receive at least one dose are less likely to die, we need to keep track of them, hence vaccinated people who get infected traverse a separate chain of compartments $I^{v}_{r,p}(t), H^{v}_{r,p}(t), T^{v}_{r,p}(t)$ with respect to those who do not receive any dose (see Figure 4.6). Dynamics governing the evolution of $H^{v}_{r,p}(t), T^{v}_{r,p}(t)$ are analogous to those in (4.2) with the only difference that $p^{TD}_{r,p}(t)$ is replaced by $p^{TD}_{r,p}(t)/q_{\texttt{post}}$. The vaccination window for each class is computed based on the class priority: $\mathscr{T}^{(1),\max}_{r,p} = \inf\{t : S_{r,p}(t) = 0\}$; $\mathscr{T}^{(1),\min}_{r,p} = \max_{(r',p')\in HP(r,p)}\{\mathscr{T}^{\max}_{r',p'}\}$, where $HP(r,p)$ is the set of classes with higher priority than (r,p).

The new vaccination dynamics is an extension of (4.2), described by the following system of equations:

$$
\begin{aligned}
\dot{S}_{r,p}(t) &= -\lambda(t)\frac{rS_{r,p}(t)}{\widehat{S}(t)} - \xi^{(1)}_{r,p}(t)\\
\dot{I}_{r,p}(t) &= \lambda(t)\frac{rS_{r,p}(t)}{\widehat{S}(t)} - \gamma I_{r,p}(t)\\
\dot{H}_{r,p}(t) &= \gamma p^{IH}_{r,p} I_{r,p}(t) - \phi H_{r,p}(t)\\
\dot{T}_{r,p}(t) &= \phi p^{HT}_{r,p} H_{r,p}(t) - \tau T_{r,p}(t)\\
\dot{I}^{v}_{r,p}(t) &= \lambda(t)\frac{r(V^{1\mathrm{s}}_{r,p} + V^{2\mathrm{s}}_{r,p})}{\widehat{S}(t)} - \gamma I^{v}_{r,p}(t)\\
\dot{H}^{v}_{r,p}(t) &= \gamma p^{IH}_{r,p} I^{v}_{r,p}(t) - \phi H^{v}_{r,p}(t)\\
\dot{T}^{v}_{r,p}(t) &= \phi p^{HT}_{r,p} H^{v}_{r,p}(t) - \tau T^{v}_{r,p}(t)\\
\dot{M}_{r,p}(t) &= \gamma(1-p^{IH}_{r,p})(I_{r,p}(t) + I^{v}_{r,p}(t))\\
&\quad +\phi(1-p^{HT}_{r,p})(H_{r,p}(t) + H^{v}_{r,p}(t))\\
&\quad +\tau\left((1-p^{TD}_{r,p}(t))T_{r,p}(t) + (1-p^{TD}_{r,p}(t)/q_{post})T^{v}_{r,p}(t)\right)\\
\dot{V}^{1\mathrm{m}}_{r,p}(t) &= \xi^{(1)}_{r,p}(t)\mathsf{VE}^{1} - \xi^{(2)}_{r,p}(t)\frac{V^{1\mathrm{m}}_{r,p}(t)}{V^{1s}_{r,p}(t) + V^{1m}_{r,p}(t)}\\
\dot{V}^{1\mathrm{s}}_{r,p}(t) &= (1-\mathsf{VE}^{1})\xi^{(1)}_{r,p} - \lambda(t)\frac{rV^{1\mathrm{s}}_{r,p}(t)}{\widehat{S}(t)} - \frac{\xi^{(2)}_{r,p}(t)V^{1\mathrm{s}}_{r,p}(t)}{V^{1s}_{r,p}(t) + V^{1m}_{r,p}(t)}\\
\dot{V}^{2\mathrm{m}}_{r,p}(t) &= \xi^{(2)}_{r,p}\frac{V^{1\mathrm{m}}_{r,p}(t) + \frac{\mathsf{VE}^{2}-\mathsf{VE}^{1}}{1-\mathsf{VE}^{1}}V^{1\mathrm{s}}_{r,p}(t)}{V^{1s}_{r,p}(t) + V^{1m}_{r,p}(t)}\\
\dot{V}^{2\mathrm{s}}_{r,p}(t) &= \xi^{(2)}_{r,p}\frac{1-\mathsf{VE}^{2}}{1-\mathsf{VE}^{1}}\frac{V^{1\mathrm{s}}_{r,p}(t)}{V^{1s}_{r,p}(t) + V^{1m}_{r,p}(t)} - \lambda(t)\frac{rV^{2\mathrm{s}}_{r,p}(t)}{\widehat{S}(t)}\\
\dot{D}_{r,p}(t) &= \tau p^{TD}_{r,p}(t)\left(T_{r,p}(t) + \frac{T^{v}_{r,p}(t)}{q_{post}} + T^{\text{no-vax}}_{r,p}(t)\right)
\end{aligned}
\tag{4.6}
$$

As a final consideration, we emphasize that the rate ξ is typically limited by the capacity of both the vaccine production and distribution infrastructure. However, to increase the effectiveness of the vaccination campaign, the vaccination rate must always be as high as possible under the above conditions.

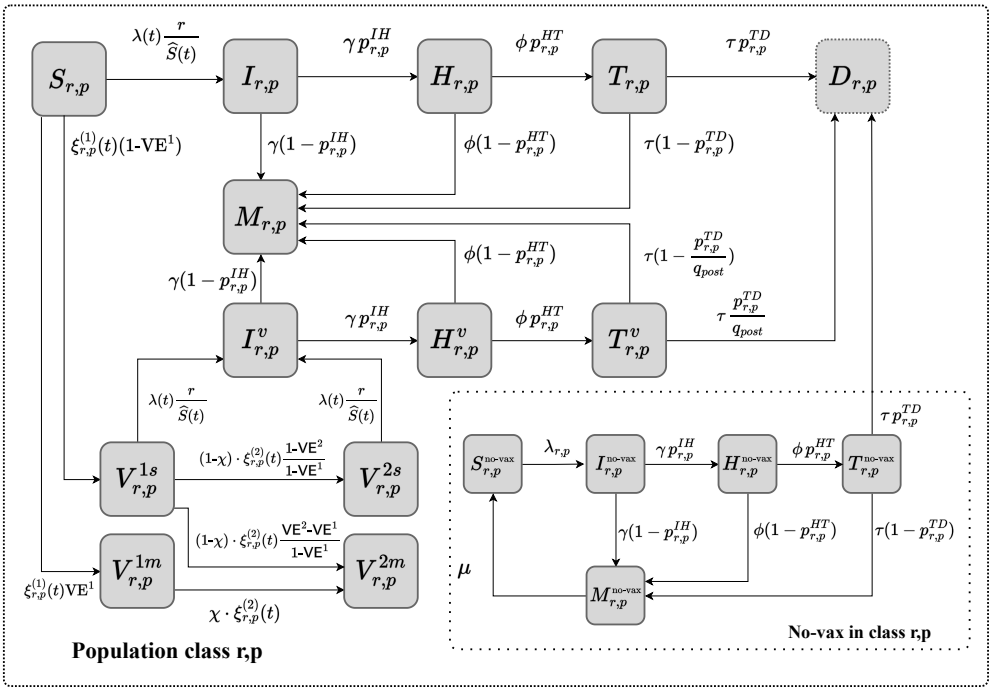


Fig. 4.6 Schematic representation of the proposed model with vaccinations, Eq. (4.6). For ease of representation we show the dynamics of only one class (r,p) we remark that the dynamics of different classes are intertwined both through the infection rate, and the process of vaccinations prioritization. We defined $\chi := \frac{V^{1\mathrm{m}}_{r,p}(t)}{V^{1s}_{r,p}(t)+V^{1m}_{r,p}(t)}$.

4.4 Epidemic Control

To mitigate the epidemic, several interventions are possible: (a) investments in the public health system, e.g., increasing the available number of ICUs $\widehat{T}$ and hospitalization facilities $\widehat{H}$, (b) non-pharmaceutical interventions, i.e., public health measures preventing and/or controlling virus transmission (Section 4.4.1); (c) vaccination that aims to reduce both the transmission and clinical severity of the disease (Section 4.4.2).

Our analysis will focus on quantifying the cost and the impact of different control strategies that jointly exploit non-pharmaceutical interventions and vaccination. In this section, first, we formalize the control problems in their most general formulation and then we introduce our *simple* strategies. Indeed, our goal is not to develop a mathematical theory of optimal control for epidemics but to provide a practical framework that informs public policy in controlling the spread of epidemics. We intend to offer decision-makers a means to compare and

evaluate a set of feasible controls, allowing them to make informed choices based on the outcomes and trade-offs associated with different control strategies. We will discuss *optimal control* approaches in Appendix E

4.4.1 Control via Non-Pharmaceutical Interventions

Non-pharmaceutical interventions are measures aimed at controlling the virus by managing certain behaviors in the population, such as using PPE (e.g. masks), implementing lockdowns, and promoting telework, to name a few. In our framework, we do not model the effects of social distancing and other countermeasures at a microscopic (class-specific) level. Instead, we summarize their effects by a single control parameter $\rho(t)$ that scales down the overall rate of potential (uncontrolled) new infections.

Specifically, we include the control in the model described by (4.2) by setting the actual intensity of new infections $\lambda(t)$ equal to $\frac{\lambda_U(t)}{\rho(t)}$, leading to an effective reproduction number:

$$\mathscr{R}^{\rho}(t) = \frac{\sigma}{\rho(t)\gamma} \frac{\sum_{r,p} r^2 S_{r,p}(t)}{\sum_{r,p} N r f_{r,p}}.$$

In this scenario, we will distinguish two main contributions to the cost: the social and the economic cost. It is crucial to note that the distinction between social and economic costs is not always clear-cut. Lockdown measures, while aimed at minimizing the social cost of the pandemic in terms of reducing deaths, have economic repercussions. Similarly, the economic cost of the pandemic, such as job losses and reduced economic activity, has social implications. Moreover, for technical reasons in some cases, we add a third component related to healthcare stress to the cost. Accordingly, we define:

(a) the *social cost*, evaluated in terms of the cumulative number of deaths [126];

(b) the *stress on the healthcare system* induced by the disease's severity;

(c) the *economic cost* $\mathfrak{C} = \mathfrak{C}(\rho)$, since widespread lockdowns cause a massive negative impact on the economy.

In Figure 4.7 we show some examples of economic costs as a function of the control parameter ρ. The economic costs are assumed monotone increasing with $\mathfrak{C}(1) = 0$. In the optimal control formulation (see [127] and reference therein) a

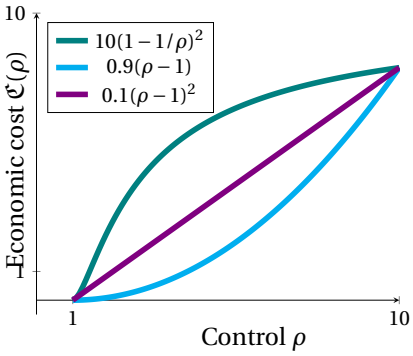


Fig. 4.7 Examples of economic costs as a function of control parameter

terminal cost is generally defined by taking a linear combination of the above costs and the policymaker aims at solving the following optimization problem:

$$\rho^\star(t_{\max}) = \underset{\rho:[0,t_{\max}]\to[1,\infty)}{\arg\min} \kappa_1 \frac{D(t_{\max})}{N} + \int_0^{t_{\max}} \left[\kappa_2 \left(\frac{T(t)}{N} \right)^\zeta + \kappa_3 \mathfrak{C}(\rho(t)) \right] \mathrm{d}t$$
$$\text{s.t. dynamics in (4.2)} \tag{4.7}$$

where the exponent ζ is typically assumed greater than 1, while $\kappa_1, \kappa_2, \kappa_3 \geq 0$ are the parameters that weigh the social, the healthcare stress and the economic costs in the objective function, according to how much one values one over the others.

Similarly to [128], we consider two *simple* control strategies:

Rate Control (Control on New Infections)**.** *The rate of new infections is tightly controlled and kept at a certain desired level λ_C. The main goal is to avoid congestion in the sanitary system by controlling the circulation of the virus.*

HT Control (Control on Hospitalizations and intensive Therapy occupancy)**.** *It directly uses the current level of hospitalization/intensive therapy occupancy as a control signal. Such a signal is readily available and less noisy than the rate of new infections. However, it may introduce a delay in the control loop, which may endanger system stability.*

4.4.2 Control via Vaccination Prioritization

The vaccination policy involves assigning a priority to each (r,p) class therefore determining which classes should be given priority.

Remark 11. *Any possible prioritization (permutation) $\pi_{(r,p)}$ of classes (r,p) corresponds to a different vaccination policy, The optimal control problem defined in* (4.7) *can be easily extended to account for vaccinations as follows:*

$$(\rho^\star(t_{\max}), \pi^*_{(r,p)}) = \underset{\substack{\rho:[0,t_{\max}]\to[1,\infty)\\ \pi_{(r,p)}}}{\operatorname{argmin}} \kappa_1 \frac{D(t_{\max})}{N} + \int_0^{t_{\max}} \left[\kappa_2 \left(\frac{T(t)}{N} \right)^\zeta + \kappa_3 \mathfrak{C}(\rho(t)) \right] \mathrm{d}t$$

s.t. dynamics in (4.6)

Again, we consider two *simple* vaccination policies the Most Vulnerable First (MVF) and the Most Social First (MSF):

MVF Policy (Most Vulnerable First)**.** *The MVF policy aims to protect the most clinically vulnerable people, with the goal of minimizing the number of deaths. It prioritizes classes with a higher value of p. For the same p, classes with higher r are vaccinated first.*

MSF Policy (Most Social First)**.** *The MSF policy prioritizes people with a high contact rate, aiming to minimize the force of infection. Classes with a higher value of r are prioritized. For the same r, classes with higher p are vaccinated first.*

The MSF policy is similar in spirit to the degree-based vaccination policy in contact networks [129], which targets the high-degree nodes first before moving on to lower-degree nodes. The interval Δ is another design parameter: prolonging the interval between doses, say from 3 to 12 weeks, might be a sensible choice under limited vaccine supplies, de facto minimizing hospitalization and deaths, especially when the efficacy of the first dose is sufficiently high.

4.5 Control on New Infections

In this section, we show that if function $\mathfrak{C}(\cdot)$ is convex, we can devise a simple strategy to minimize the overall economic cost. As already observed, a key role in

the epidemic dynamics is played by $\widetilde{I}(t)$, which, roughly speaking, represents the number of potentially infected contacts (see Remark 9). Thus, a sensible strategy is to control such a quantity. In our derivations, we assume that $S_{r,p}(t) \approx N f_{r,p}$, see Remark 10.

This assumption allows for simplifications in the mathematical modeling and analysis. Indeed, given the definition of $\widetilde{I}(t)$, multiplying the second equation in (4.2) by r and summing over r and p, we get:

$$\dot{\widetilde{I}}(t) = \gamma \left(\frac{\sigma}{\rho(t)\gamma} \frac{\sum_{r,p} r^2 S_{r,p}(t)}{\sum_{r,p} N r f_{r,p}} - 1 \right) \widetilde{I}(t).$$

under the assumption $S_{r,p}(t) \approx N f_{r,p}$, and defining $\mathscr{R}^{\rho}(t) = \frac{\mathscr{R}_0}{\rho(t)}$, we obtain the equation:

$$\dot{\widetilde{I}}(t) = \gamma \left(\frac{\mathscr{R}_0}{\rho(t)} - 1 \right) \widetilde{I}(t) = \gamma \left(\mathscr{R}^{\rho}(t) - 1 \right) \widetilde{I}(t). \tag{4.8}$$

4.5.1 Minimizing the Economic Cost in a Fixed Window

Fixing a target value $\widetilde{I}^{\star}$ for $\widetilde{I}(t)$, to be met within a prefixed a time horizon $t_{\max}$, Proposition 3 establishes optimality conditions.

Proposition 3. *Let $\mathfrak{C}(\rho)$ be a monotone increasing and convex function in $\rho \in [1, +\infty]$ and assume $S_{r,p}(t) \approx N f_{r,p}$. Among all trajectories, such that $\widetilde{I}(t_{\max}) = \sum_{r,p} r I_{r,p}(t_{\max}) = \widetilde{I}^{\star}$, the one that minimizes the overall economic cost in $[0, t_{\max}]$, is the one corresponding to:*

$$\mathscr{R}^{\rho}(t) = 1 + \frac{1}{\gamma} \log\left(\widetilde{I}^{\star} / \widetilde{I}(0) \right) \quad \forall\, t \in [0, t_{\max}],$$

and,

$$\rho(t) = \frac{\sigma}{\gamma} \frac{\mathbb{E}[r^2]}{\mathbb{E}[r]} \left[1 + \frac{1}{\gamma t_{\max}} \log\left(\widetilde{I}^{\star} / \widetilde{I}(0) \right) \right]^{-1} \quad \forall\, t \in [0, t_{\max}].$$

Proof. Consider Eq. (4.8) and note that the unique solution of the associated Cauchy problem with initial condition $\widetilde{I}(0)$ is given by:

$$\widetilde{I}(t) = \widetilde{I}(0) \exp\left(\gamma \int_0^t \left(\mathscr{R}^{\rho}(\tau) - 1 \right) \mathrm{d}\tau \right).$$

Imposing the constraint $\widetilde{I}(t_{\max}) = \widetilde{I}^\star$ leads to:

$$\frac{1}{t_{\max}} \int_0^{t_{\max}} \mathscr{R}^\rho(\tau) \mathrm{d}\tau = 1 + \frac{1}{\gamma\, t_{\max}} \log\left(\widetilde{I}^\star / \widetilde{I}(0)\right). \tag{4.9}$$

Now, focusing on a generic trajectory satisfying (4.9), we have: $\frac{1}{t_{\max}} \int_0^{t_{\max}} \mathfrak{C}(\rho(\tau)) \mathrm{d}\tau = \frac{1}{t_{\max}} \int_0^{t_{\max}} \chi(\mathscr{R}^\rho(\tau)) \mathrm{d}\tau$ with $\chi = \mathfrak{C} \circ \rho$, and $\rho(\mathscr{R}^\rho) = \frac{\sigma \mathbb{E}[r^2]}{\gamma \mathbb{E}[r] R^\rho(t)}$. Since $\mathfrak{C}$ is a monotonic increasing and convex function in $\rho \in [1, +\infty]$ then χ is a convex function over its domain, and by Jensen inequality, we conclude $\frac{1}{t_{\max}} \int_0^{t_{\max}} \chi(\mathscr{R}^\rho(\tau)) \mathrm{d}\tau \geq \chi\left(\frac{1}{t_{\max}} \int_0^{t_{\max}} \mathscr{R}^\rho(\tau) \mathrm{d}\tau\right)$ Therefore, from (4.9) the choice given by:

$$\rho(t) = \frac{\sigma \mathbb{E}[r^2]}{\gamma \mathbb{E}[r]} \left[1 + \frac{1}{\gamma\, t_{\max}} \log\left(\widetilde{I}^\star / \widetilde{I}(0)\right)\right]^{-1}, \forall\, t \in [0, t_{\max}]$$

minimizes the cost. □

Observe that the economic cost of previously defined optimal policy monotonically decreases while increasing the target $\widetilde{I}^\star$.

Corollary 1. *Under the assumptions that $\mathfrak{C}(\rho)$ is a monotone increasing and convex function and $S_{r,p}(t) \approx N f_{r,p}$, among all control strategies that maintain the number of infected less than or equal the initial value $\widetilde{I}(0)$, the overall economic cost is minimized when $\mathscr{R}^\rho(t)$ is kept equal to 1.*

Proof. From Proposition 3 we have that among all strategies guaranteeing $\widetilde{I}(t_{\max}) = \widetilde{I}(0)$, the one forcing $\mathscr{R}^\rho(t) = 1$ is cost-optimal. The proof is completed by observing that such a strategy guarantees $\widetilde{I}(t) \leq \widetilde{I}(0)$ for every $t \in [0, t_{\max}]$. □

Remark 12. *$\mathscr{R}^\rho(t) = 1$ can be achieved by controlling the rate of new infections and maintaining it equal to the target $\lambda_C = \gamma \widetilde{I}(0) \mathbb{E}[r] / \mathbb{E}[r^2]$. The resulting control function is $\rho(t) = \lambda_U(t-\varepsilon)/\lambda_C = \lambda(t-\epsilon)\rho(t-\varepsilon)/\lambda_C$, where ε is an arbitrarily small positive constant.*

In conclusion, given an initial condition $\widetilde{I}(0)$, a maximum allowable number of infected contacts $\widetilde{I}^\star$ and a time horizon $t_{\max}$, if the goal is to keep $\widetilde{I}(t) \leq \widetilde{I}^\star \quad \forall\, t \in [t^\star, t_{\max})$, with $t^\star$ as small as possible, the following strategy appears to be the natural answer: if $\widetilde{I}^\star > \widetilde{I}(0)$, set $\mathscr{R}^\rho(t) = 1 + \frac{1}{\gamma\, t_{\max}} \log\left(\widetilde{I}^\star / \widetilde{I}(0)\right), \forall\, t \in [0, t_{\max}]$. This strategy, indeed, minimizes the economic cost in $[0, t_{\max}]$, among all strategies

that guarantee $\widetilde{I}(t) \leq \widetilde{I}^\star, \forall\, t \in [0, t_{\max}]$, (i.e., $t^\star = 0$). If, instead, $\widetilde{I}^\star < \widetilde{I}(0)$, we can not guarantee $t^\star = 0$, and therefore to minimize $t^\star$ it is necessary to minimize $\mathscr{R}^\rho(t)$ in $[0, t^\star)$ and then to set $\mathscr{R}^\rho(t) = 1, \forall\, t \in [t^\star, t_{\max}]$. Indeed, this is the strategy that minimizes the economic cost in $[0, t_{\max}]$, among all strategies minimizing $t^\star$. Previous arguments can be formalized in the following proposition.

Proposition 4. *Given $\widetilde{I}(0)$, $\widetilde{I}^\star$ and $t_{\max}$, whenever our goal is to keep $\widetilde{I}(t) \leq \widetilde{I}^\star, \forall\, t \in [t^\star, t_{\max})$, with $t^\star$ as small as possible, the strategy described above is cost-optimal.*

It is worth remarking that the optimality criteria depend on various factors, including the dynamics of the system, the objective function, and the constraints. In the context of our study, the choice of time horizon $t_{\max}$ plays a crucial role. If $t_{\max}$ is set too low, it might limit the effectiveness of the control measures. We emphasize that the selection of $t_{\max}$ should be carefully considered based on the specific context and dynamics of the epidemic under investigation. The parameter $\tilde{I}^\star$, instead, represents the maximum admissible number of infected over the considered time horizon $t_{\max}$ and is indissolubly related to the transmission rate λ. It represents the number of infected it is possible to sustain (in a country, for example). This again showcases the tradeoffs between social cost (in terms of deaths) and economic cost (the entity of restriction measures), small values of $\tilde{I}^\star$ would result in a small number of deaths but also in a high economic cost. Our analysis will provide insights for a fixed value of $\tilde{I}^\star$, chosen by the decision maker according to the tradeoff between social and economic cost and the acceptable number of infected by the overall healthcare infrastructure. We acknowledge that this choice could potentially be optimized to minimize the cost, and while we do not explicitly optimize it in our study, we will discuss the implications of different target numbers and their impact on the control measures (see Section 4.5.3).

4.5.2 Rate Control with Feedback Delay

Policymakers cannot instantaneously react to changes in the rate of new infections due to several reasons: i) new infections are discovered by tests performed several days after infection, and high-risk individuals are more likely to undergo testing [130], ii) new regulations take time to be introduced and become effective, iii) decisions are based on trends obtained by averaging epidemiological curves, iv) the actual process of new infections in unknown (think of asymptomatic but

infectious people). Consequently, the measured process is a delayed, noisy subsample of the actual process. Therefore, we consider the case in which the actual, instantaneous effectiveness of mobility restrictions, modeled by $\rho(t)$, is given by: $\rho(t) = \max\left\{1, \frac{\int \mathfrak{f}_d(\tau)\lambda_U(t-\tau)\,\mathrm{d}\tau}{\lambda_C}\right\}$ where $\mathfrak{f}_d(\cdot)$ is a feedback delay distribution.

One of our main results is that the system becomes unstable if the feedback delay is too large with respect to $1/\gamma$ (the average time in the infectious state). To simplify the analytical derivations, we start with the case of deterministic feedback delay of constant duration d (days). Then we extend the result to a delay distribution $\mathfrak{f}_d$.

Theorem 4 (Stability analysis with constant delay)**.** *Assume*

$$\rho(t) = \max\left\{1, \frac{\lambda_U(t-d)}{\lambda_C}\right\} = \max\left\{1, \frac{\lambda(t-d)\rho(t-d)}{\lambda_C}\right\}$$

and $S_{r,p}(t) \approx N f_{r,p}$. If the delay $d < \frac{\pi}{2}\gamma$ then the system is locally stable, otherwise the system is unstable.

Proof. Since under the assumption $S_{r,p}(t) \approx N f_{r,p}$, the equation governing the evolution of the number of infected edges under delayed rate control becomes:

$$\dot{\widetilde{I}}(t) = \frac{\widetilde{I}(t)}{\widetilde{I}(t-d)}\lambda_C\frac{\mathbb{E}[r^2]}{\mathbb{E}[r]} - \gamma\widetilde{I}(t) \tag{4.10}$$

System stability can be analyzed by considering small perturbations around the equilibrium point $\widetilde{I}^* = \frac{\lambda_C}{\gamma}\frac{\mathbb{E}[r^2]}{\mathbb{E}[r]}$, i.e., $\widetilde{I}(t) = \widetilde{I}^* + \eta(t)$, with $\eta(t) \ll \widetilde{I}^*$. Exploiting the approximation $\frac{1}{1+x} \sim 1-x$, when $x \approx 0$, from (4.10) we obtain:

$$\dot{\widetilde{I}}(t) = \gamma\widetilde{I}^*\frac{1+\frac{\eta(t)}{\widetilde{I}^*}}{1+\frac{\eta(t-d)}{\widetilde{I}^*}} - \gamma(\widetilde{I}^* + \eta(t)) \approx -\gamma\eta(t-d)$$

where we have discarded the second-order term $\eta(t)\eta(t-d)$. We end up with the simple differential equation with delay:

$$\dot{\eta}(t) = -\gamma\,\eta(t-d) \tag{4.11}$$

Taking the Laplace transform $\mathscr{L}\{\eta(t)\}$ we obtain $\mathscr{L}\{\eta(t)\} = \frac{\eta(0)}{s+\gamma e^{-sd}}$. Equation (4.11) admits solutions of the form $\eta(t) = Ae^{bt}\cos(\omega t + \theta)$ under the conditions:

$$\begin{cases} b = -\gamma e^{-bd}\cos(\omega d) \\ \omega = \gamma e^{-bd}\sin(\omega d) \end{cases} \tag{4.12}$$

While A and θ can take any value, i.e., can be used to match desired values of $\eta(0)$ and $\eta'(0)$, b and ω are uniquely determined by the feedback delay d. Besides the trivial solution $b = \omega = 0$, there exists a stationary solution $b = 0$, $\omega = \gamma$ for the special case $d = \frac{\pi}{2\gamma}$. If $d < \frac{\pi}{2\gamma}$, from the first constraint we have that $b < 0$, corresponding to dumped oscillations. For $\frac{\pi}{2\gamma} < d < \frac{3\pi}{2\gamma}$, we have instead amplifying oscillations ($b > 0$). Therefore, $d = \frac{\pi}{2\gamma}$ is the critical value for stability. □

The analysis can be extended to a delay distribution $\mathfrak{f}_d$.

Theorem 5 (Stability analysis with delay distribution)**.** *Let us assume that* $\rho(t) = \max\{1, \int \mathfrak{f}_d(\tau)\lambda_U(t-\tau)\,\mathrm{d}\tau/\lambda_C\}$ *and* $S_{r,p}(t) \approx N f_{r,p}$. *Let* $\mathscr{Z} = \{z \in \mathbb{C} : z + \gamma\mathfrak{F}_d(z) = 0\}$, *where* $\mathfrak{F}_d(z)$ *is the Laplace transform of the delay distribution. Then, if* $\mathrm{Re}(z) < 0$ $\forall z \in \mathscr{Z}$, *the system is locally stable.*

Proof. Repeating the same approximations as before for small variations around the equilibrium $\widetilde{I}^*$, we obtain the differential equation with delay distribution:

$$\dot{\eta}(t) = -\gamma \int \mathfrak{f}_d(\tau)\eta(t-\tau)\,\mathrm{d}\tau \tag{4.13}$$

Taking the Laplace transform, we get $H(s) = \eta(0)/(s + \gamma\mathfrak{F}_d(s))$. Note that when $\mathfrak{f}_d(\tau) = \delta(\tau - d)$, we obtain the case with constant delay. We evince that we need the set of zeros $\mathscr{Z} = \{z \in \mathbb{C} : z + \gamma\mathfrak{F}_d(z) = 0\}$ to lie in the left half-plane to ensure stability. □

In the following corollaries, whose proof is Appendix F we explore two interesting cases of feedback delay distributions.

Corollary 2 (Exponential delay distribution)**.** *If* $\mathfrak{f}_d(\tau) = u(\tau)\delta e^{-\delta(\tau)}$, *then the system is always (locally) stable.*

Corollary 3 (Shifted exponential delay distribution)**.** *Let* $\mathfrak{f}_d(\tau) = u(\tau - d)\delta e^{-\delta(\tau-d)}$. *For any given* $\delta > 0$, *there exists a critical delay* $d^* = \frac{1}{\gamma}f(\delta)$, *such that the system is*

(locally) stable if $d < d^*$, otherwise the system is unstable. As δ grows from 0 to ∞, d^* grows from $1/\gamma$ to $\pi/(2\gamma)$.

The shifted exponential distribution can represent a system where: i) an exponentially weighted moving average (with parameter δ) is used to estimate the current trend of the epidemiological curve, ii) some fixed delay d is introduced before the control becomes effective. Our results suggest that system stability is crucially tied (by a factor between 1 and $\pi/2$ that depends on δ) to the mean sojourn time $1/\gamma$ in the infectious state. If d is too large with respect to $1/\gamma$, the control based on the force of infection is prone to instability.

In a finite population system, as time goes on, we can no longer assume that $S_{r,p}(t) \approx N f_{r,p}$, since the number of initially susceptible individuals is progressively reduced by the number of people who get infected (see (4.2)). Moreover, $S_{r,p}(t)$ can vary because of vaccinations and the finite duration of immunity. Nevertheless, we can still apply the above results by resorting to a *time-scale separation approach*, i.e., by assuming that $S_{r,p}(t)$, though not equal to $N f_{r,p}$, are almost constant at the time scale over which we analyze stability.

Remark 13. *The assumption of a constant number of susceptible individuals is fairly accurate when dealing with the dynamics of a large population over a relatively small window of time. In such a case, the rate of infection spread may have a minimal impact on the overall number of susceptible individuals in relative terms (mathematical considerations as in Remark 10 apply). We emphasize that the rate of newly infected people should be sufficiently small to ensure that the total number of infected/recovered people in the considered window is negligible with respect to the number of susceptible individuals.*

Indeed, recall from Remark 9 that the evolution of the total number of infected edges can be written as:

$$\dot{\tilde{I}}(t) = \gamma \left(\frac{\mathscr{R}(t)}{\rho(t)} - 1 \right) \tilde{I}(t) \tag{4.14}$$

where $\mathscr{R}(t) = \frac{\sigma \sum_{r,p} r^2 S_{r,p}(t)}{\gamma \sum_{r,p} r N f_{r,p}}$ is the basic reproduction in the general case. This equation is formally identical to (4.8) upon substituting $\mathscr{R}_0$ with $\mathscr{R}(t)$. Since our stability results do not depend on $\mathscr{R}_0$, they apply also to a system in which $\mathscr{R}(t)$ can be considered approximately constant at the time scale at which we analyze the system stability (i.e., time scale of $1/\gamma$).

4.5.3 Sensitivity Analysis - The impact of λ

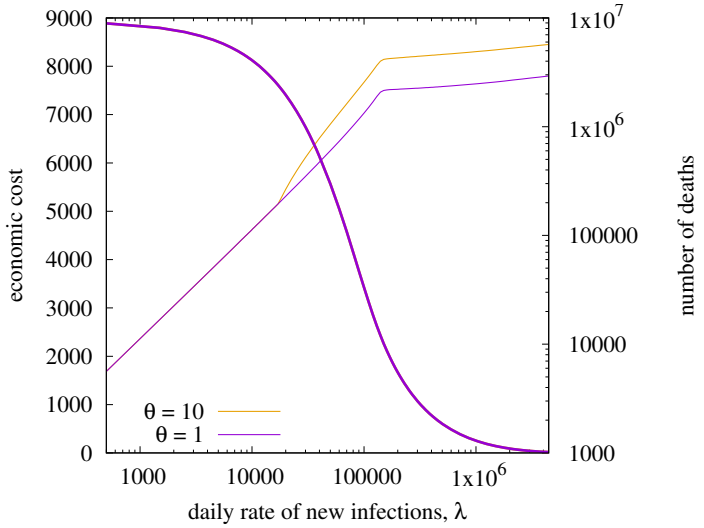


Fig. 4.8 Parametric curves of economic cost vs social cost, as we vary λ, with $\theta \in \{1, 10\}$.

Fig. 4.8 reports both the economic cost and number of deaths observed over a time horizon $t_{\max} = 1$ year as a function of the control parameter λ. We refer the reader to Table 4.3 for a detailed description of all the infection parameters. When controlling the infection rate, the suppression strategy, i.e., minimizing infection rate λ, appears to be the most reasonable choice since it minimizes the number of deaths incurring an almost constant economic cost for, e.g., all values of $\lambda < 10000$. Indeed note that, once the system is stabilized around a fixed infection rate[1] λ^*, the economic cost is the same for any λ^*, as long as $S(t) \approx N$.

In the case of COVID-19, some countries, e.g., China, have adopted the suppression strategy, which is particularly effective when restrictions can be geographically localized to small areas with limited impact on the national economy. Of course, this cannot be a solution in the long term unless the virus is totally eradicated or conditions change, e.g., herd immunity is reached through vaccinations. Indeed, note that all results discussed so far refer to a fixed time horizon $t_{\max} = 1$ year.

To understand how the optimal strategy might change as we increase the time horizon $t_{\max}$, it is convenient to look at the plot in Fig. 4.9, showing parametric

[1] Further, note that with proper control, the cost incurred during the transient phase necessary to bring the system to operate at a given λ is negligible with respect to the long-term accumulated cost.

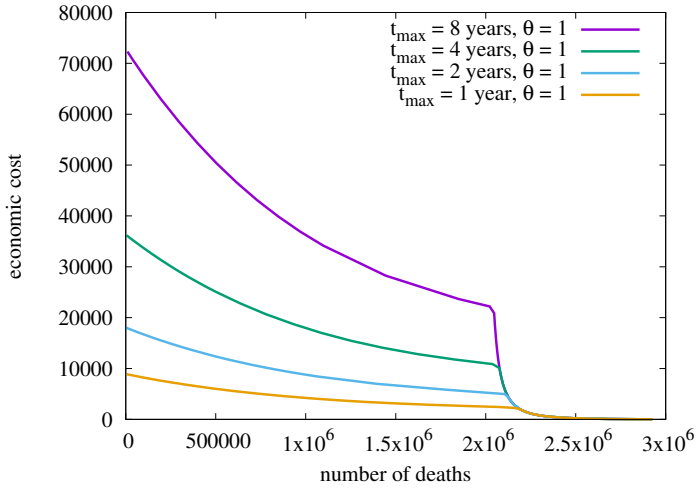


Fig. 4.9 Parametric curves of economic cost vs social cost, as we vary λ, for different time horizons $t_{\max}$; $\theta = 1$.

curves of economic cost vs. social cost, as we vary λ, for $t_{\max} = 1, 2, 4, 8$ years. These results have been obtained by running the multi-class model with the parameter θ, representing the increase of mortality due to ICU saturation, equal to 1, putting us in the most favorable conditions (i.e., in the presence of unlimited healthcare facilities) to decide to abandon the suppression strategy. Clearly, under the suppression strategy, the economic cost increases linearly with time, so for $t_{\max}$ large enough, this strategy becomes necessarily suboptimal[2].

Interestingly, curves shown in Fig. 4.9 can be split into two convex parts connected at the point where the population reaches natural herd immunity (the knee). The consequences of this behavior on the multi-objective function (4.7), for $\kappa_2 = 0$ and $\kappa_3 = 1$, which is linear with respect to trade-off factor κ_1, are illustrated in Fig. 4.10 for the case $t_{\max} = 4$ years. We observe that all points between B and C are not Pareto-efficient, hence cannot be optimal solutions for the optimization problem (4.7). The optimal strategy exhibits a phase transition with respect to κ_1: for small values of κ_1 (social cost much more important than economic cost), the best strategy is total suppression (point A), whereas for large κ_1 we end up operating beyond the herd immunity knee. Intermediate solutions between A and B also exist, but only for a very small, particular range of κ_1 values.

[2]It should be noticed, however, that a finite population model like ours is not adequate to describe a system running for more than, say, a few years.

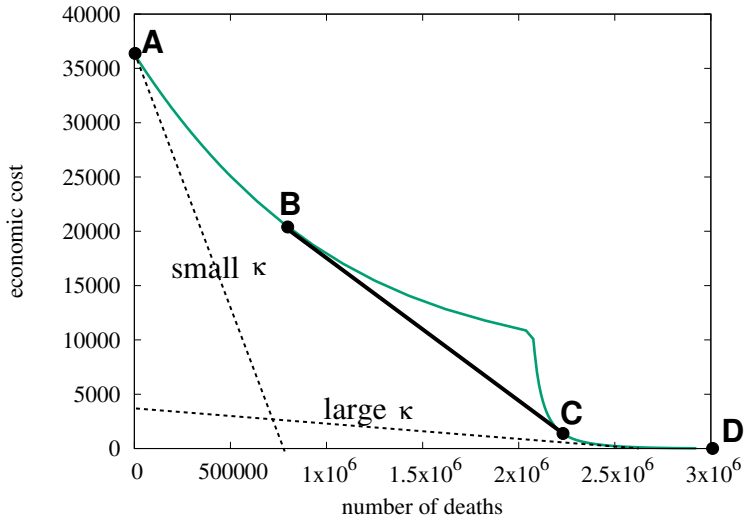


Fig. 4.10 Pareto frontier of the multi-objective function (4.7) in the case $t_{\max} = 4$ years; $\theta = 1$.

The particular value of κ_1 at which the phase transition occurs, in addition to the time horizon $t_{\max}$, depends crucially on the exponent α, as one can intuitively understand from Fig. 4.11, which shows economic and social costs as a function of the controlled rate λ, for fixed $t_{\max} = 4$ years, and different values of the exponent in the economic cost $\alpha = 1,2,3$: while the social cost is the same for all α, the economic cost depends dramatically on α. Note that $\alpha = 1$ is the extreme case for the validity of Proposition 3.

Proper values of α to be used in the model are difficult to set. However, the general conclusion remains the same: unless one considers considerably long (but unlikely to be significant) time horizons, the best option always appears to be the minimization of λ. With the parameters of COVID-19, and in particular, for the delta variant, the opposite 'let it rip' strategy in which one tries to achieve the natural herd immunity (while still controlling λ to avoid ICU saturation) produces an unreasonable social cost in terms of deaths. Some countries (like the UK) initially considered this option at the onset of the pandemic but quickly switched back to the suppression strategy after a few months.

Another reason why the 'let it rip' strategy considered so far is perilous is that it relies on the assumption that recovered people are immune forever, i.e., recovery rate $\mu = 0$. In the case of COVID-19, natural immunity is progressively lost over time, so reinfections are possible about six months after recovery. Even

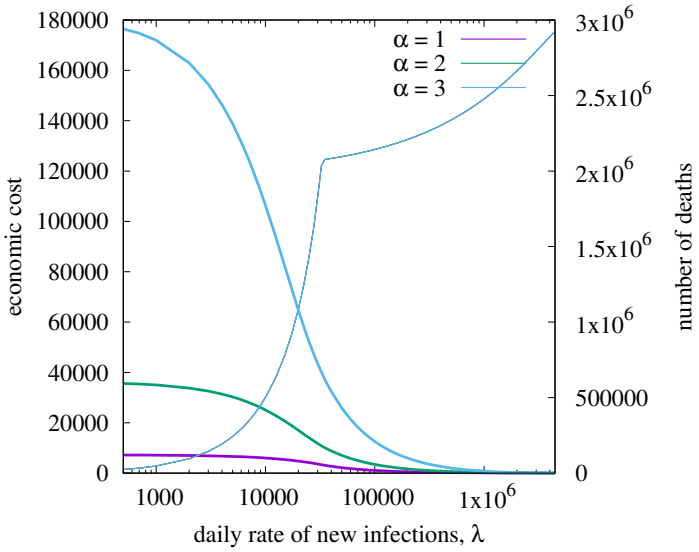


Fig. 4.11 Economic and social costs as a function of controlled rate λ, for fixed $t_{\max} = 4$ years and different values of α.

assuming that reinfected people are much less likely to develop a severe form of the disease, we expect a significantly higher social cost when $\mu > 0$. This observation is confirmed by results in Fig. 4.12, showing economic and social costs for $t_{\max} = 4$ years, mortality reduction after the first exposure $q_{\texttt{post}} = 10$, and different values of the average sojourn time in the immune state, equal to 6 months (as estimated for COVID-19), 1 year, 2 years, in addition to the optimistic hypothesis $\mu = 0$.

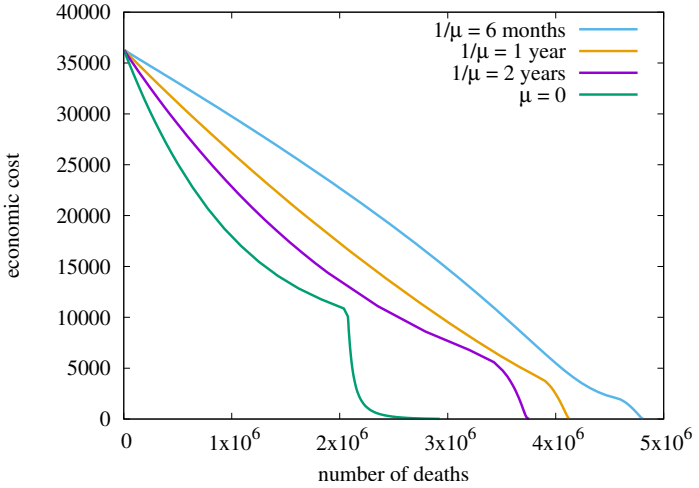


Fig. 4.12 Parametric curves of economic cost vs social cost, as we vary λ, for fixed $t_{\max} = 4$ years and different values of μ. Mortality rate after reinfection is reduced by $q_{\texttt{post}} = 10$.

Note that at the beginning of the pandemic, the decision on which strategy to choose was particularly difficult, as the characteristics of the virus were unknown and the effect of the vaccinations was unclear. Of course, it was also unknown which mutations of the original virus would emerge and when they would have replaced the original strain. In later sections, we will take a closer look at these two fundamental factors that were crucial to the development of the pandemic after the first year.

4.6 Control on Hospitalizations and ICU Occupancy

Recall that, according to the HT strategy, the control variable $\rho(t)$ is directly related to the instantaneous numbers $H(t)$ and $T(t)$ of patients who are currently hospitalized or under intensive treatment, respectively. Many countries have widely adopted this strategy, being particularly simple to implement.

4.6.1 Stability Analysis

We assume that Hospitals and ICUs have a maximum capacity $\widehat{H}$ and $\widehat{T}$, correspondingly. A maximum level of restrictions $\rho_{\max}$ is applied whenever either $H(t)$ exceeds $H_{\max}$ (with $H_{\max} \leq \widehat{H}$), or $T(t)$ exceeds $T_{\max}$ ($T_{\max} \leq \widehat{T}$). When $H(t) < H_{\max}$ and $T(t) < T_{\max}$, we assume that two control functions $\rho_H : \mathbb{R}^+ \to [1,\infty)$ and $\rho_T : \mathbb{R}^+ \to [1,\infty)$ provide two different levels of restrictions, the larger (i.e. stricter) of which is actually applied: $\rho := \max\{\rho_H \circ H, \rho_T \circ T\}$.

Assumption 2. *Let* $\rho_H \in C^1[0, H_{\max}]$, $\rho_T \in C^1[0, T_{\max}]$ *such that* $\rho_H(0) = \rho_T(0) = 1$, $\rho_H(H_{\max}) = \rho_T(T_{\max}) = \rho_{\max}$, *with* $\inf_{x\in(0,H_{\max})} \dot{\rho}_H(x) > 0$ *and* $\inf_{x\in(0,T_{\max})} \dot{\rho}_T(x) > 0$.

To analyze the system stability under the above type of control, we first assume $S_{r,p}(t) \approx N f_{r,p}$. We will later extend the analysis to the general case through a time-scale separation approach. Under the assumption $S_{r,p}(t) \approx N f_{r,p}$ we have that the total number of infected 'edges' is governed by (4.8).

Proposition 5 (Stationary solutions)**.** *Under the assumption $S_{r,p}(t) \approx N f_{r,p}$ and Assumption 2 the stationary solutions satisfy:*

$$H^* = \frac{\gamma}{\phi}\widetilde{I}^* \frac{\mathbb{E}[r\, p_{r,p}^{IH}]}{\mathbb{E}[r^2]}, \quad T^* = \frac{\gamma}{\tau}\widetilde{I}^* \frac{\mathbb{E}[r\, p_{r,p}^{IH}\, p_{r,p}^{HT}]}{\mathbb{E}[r^2]}. \tag{4.15}$$

Proof. From the definition we have $I_{r,p}(t) = \widetilde{I}(t)\frac{r f_{r,p}}{\mathbb{E}[r^2]}$, $I(t) = \widetilde{I}(t)\frac{\mathbb{E}[r]}{\mathbb{E}[r^2]}$. It should be noted that at equilibrium necessarily $\rho^*(t) = \mathscr{R}_0$ for all t and, by monotonicity of ρ_H and ρ_T, we have one of the following cases:

- $H^* = \rho_H^{-1}(\mathscr{R}_0)$, and $T^* \le \rho_T^{-1}(\mathscr{R}_0)$;
- $T^* = \rho_T^{-1}(\mathscr{R}_0)$, $H^* \le \rho_H^{-1}(\mathscr{R}_0)$.

Hence,

$$\widetilde{I}^* = \min\left(\rho_H^{-1}(\mathscr{R}_0)\frac{\phi}{\gamma}\frac{\mathbb{E}[r^2]}{\mathbb{E}[r\, p_{r,p}^{IH}]}, \rho_T^{-1}(\mathscr{R}_0)\frac{\tau}{\gamma}\frac{\mathbb{E}[r^2]}{\mathbb{E}[r\, p_{r,p}^{IH}\, p_{r,p}^{HT}]}\right).$$

Now, from (4.2), we obtain detailed equilibrium points:

$$I_{r,p}^* = \widetilde{I}^* \frac{r f_{r,p}}{\mathbb{E}[r^2]}, H_{r,p}^* = \frac{\gamma}{\phi} I_{r,p}^*\, p_{r,p}^{IH}, T_{r,p}^* = \frac{\phi}{\tau} H_{r,p}^*\, p_{r,p}^{HT}$$

Therefore, summing over (r,p), we get corresponding equilibria for the total number of people hospitalized or under intensive therapy as given by (4.15). □

Theorem 6 (Stability analysis)**.** *Let ρ_H and ρ_T satisfy Assumption 2 and H^* and T^* be stationary solutions as given in Proposition 5. If at least one of the following conditions is satisfied:*

- $\rho_H(H^*) > \rho_T(T^*)$
- $\rho_T(T^*) \ge \rho_H(H^*)$ *and* $\phi + \tau \ge \frac{T^* \dot{\rho}_T(T^*)\gamma}{\mathscr{R}_0}$

then the system is locally stable.

Proof. Let us consider small perturbations around the equilibrium point $\widetilde{I}^*$: $\widetilde{I}(t) = \widetilde{I}^* + \widetilde{\eta}(t)$ with $\widetilde{\eta}(t) \ll \widetilde{I}^*$.

We will assume that $0 < H^* < H_{\max}$, and $0 < T^* < T_{\max}$. From Assumption 2, by denoting with $\alpha_H^* = \dot{\rho}(H^*)$ and $\alpha_T^* = \dot{\rho}(T^*)$ we have the following cases.

1. If $\rho_H(H^*) > \rho_T(T^*)$ by continuity we get that $\rho(t) = \rho_H(H(t)) > \rho_T(T(t))$ and assuming initial conditions $H(0) = H^*$, $T(0) = T^*$, after some algebra we get the Laplace transform of $\eta(t)$:

$$\mathscr{L}\{\eta(t)\} = \frac{\eta(0)(s+\phi)}{s(s+\phi) + \frac{H^* \alpha_H^* \phi\gamma}{\mathscr{R}_0}} \tag{4.16}$$

   In this case, the system is always stable for any value of parameters ϕ, γ,$\mathscr{R}_0$, since the real part of the poles of (4.16) is always negative. As we increase the amplitude of coefficient $\frac{H^* \alpha_H^* \phi\gamma}{R_0}$, the real part of the dominating pole moves from 0 to $-\phi$.

2. If $\rho_T(T^*) > \rho_H(H^*)$ then, by continuity, we have $\rho(t) = \rho_T(T(t)) > \rho_H(H(t))$ and, by first-order analysis and computing the Laplace transform, we get

$$\mathscr{L}\{\eta(t)\} = \frac{\eta(0)(s+\phi)(s+\tau)}{s(s+\phi)(s+\tau) + \frac{T^* \alpha_T^* \tau\phi\gamma}{\mathscr{R}_0}}$$

   The system may be unstable since we obtain in the denominator a third-order equation whose complex solutions can fall in the positive half-plane. In particular, the system is stable when:

$$\phi + \tau \geq \frac{T^* \alpha_T^* \gamma}{\mathscr{R}_0} \tag{4.17}$$

   while it becomes unstable otherwise. Indeed, pure imaginary solutions $s = i\omega$ are roots of the above third order equation when $\omega = \sqrt{\tau\phi}$, while relation (4.17) is satisfied with equality.

□

Theorem 6 provides conditions guaranteeing the local stability of the system. In particular, it is worth remarking that once $H_{\max} \leq \widehat{H}$ has been fixed, condition $\rho_H(H^*) > \rho_T(T^*)$ can always be achieved by arranging a sufficiently large number of available intensive therapy facilities. Indeed, even when $\mathscr{R}_0$ is not perfectly known, it is sufficient to guarantee:

$$\rho_H^{-1}(y)\frac{\phi}{\gamma}\frac{\mathbb{E}[r^2]}{\mathbb{E}[r\, p_{r,p}^{IH}]} < \rho_T^{-1}(y)\frac{\tau}{\gamma}\frac{\mathbb{E}[r^2]}{\mathbb{E}[r\, p_{r,p}^{IH}\, p_{r,p}^{HT}]}$$

for every $\rho_{\min} < y < \rho_{\max}$, i.e. $\frac{\rho_T^{-1}(y)}{\rho_H^{-1}(y)} > \frac{\phi}{\tau}\frac{\mathbb{E}[r\, p_{r,p}^{IH}\, p_{r,p}^{HT}]}{\mathbb{E}[r\, p_{r,p}^{IH}]}$. Observe that the above constraint can be met if:

$$T_{\max} > \frac{\phi}{\tau}\frac{\mathbb{E}[r\, p_{r,p}^{IH}\, p_{r,p}^{HT}]}{\mathbb{E}[r\, p_{r,p}^{IH}]} H_{\max} \tag{4.18}$$

by adopting controllers satisfying: $\rho_H(xH_{\max}) \geq \rho_T(xT_{\max})\ \forall\ 0 \leq x \leq 1$.

When the number of intensive therapies is, instead, under-dimensioned, we have $\rho_H(H^*) > \rho_T(T^*)$, and the system stability essentially depends on the average time spent in hospitals and ICU, through the sum $\phi + \tau$ of transitions rates out of compartments H, T (both are equally important).

Assuming $S_{r,p}(t)$ are almost constant on the time scale on which the stability is studied, the analysis can be extended by replacing the basic reproduction number $\mathscr{R}_0$ with the effective reproduction number $\mathscr{R}(t)$. In this way, the evolution of the total number of infected edges (4.14) becomes formally identical to (4.8).

4.6.2 Sensitivity Analysis - The impact of $H_{\max}$ and $T_{\max}$

In the same settings as in Section 4.5.3, we start analyzing the impact of the maximum tollerable levels of hospitalized and ICU patients $H_{\max}$ and $T_{\max}$ on the system dynamics. Both the implemented controllers are linear.

In all the cases we have set the capacities as $\widehat{H} = 50000$ and $T_{\max} = \widehat{T}$. The choice $T_{\max} = \widehat{T}$ is justified by our previous analysis, according to which the maximization of ratio $T_{\max}/H_{\max}$ favors system local stability around the equilibrium point. Note that our choice of parameters guarantees local stability also in cases in which the tightest control at the equilibrium point is exerted by intensive therapy occupancy Finally note that since in our scenario $\frac{\phi}{\tau}\frac{\mathbb{E}[r\, p_{r,p}^{IH}\, p_{r,p}^{HT}]}{\mathbb{E}[r\, p_{r,p}^{IH}]} = 0.331$, we should enforce $T_{\max}/H_{\max} > 0.331$ to guarantee that at the equilibrium point, the tighter control is exerted by hospitalizations.

Figure 4.13 reports some results. First, we have fixed $T_{\max} = 10000$ and we let $H_{\max}$ vary. In particular we have chosen: $H_{\max} = 20000$ (top left plot), $H_{\max} = 30000$ (top right plot), $H_{\max} = 50000$ (bottom left plot).

Only the first choice for $H_{\max}$ satisfies condition (4.18). Note that by reducing $H_{\max}$, we significantly reduce oscillations since the control on hospitalization becomes reactive. Periods in which the tightest control is exerted by hospital-

Table 4.2 costs and deaths [$k \equiv 10^3$]

$T_{\max}$	$H_{\max}$	cost ($\alpha = 1$)	cost ($\alpha = 2$)	cost ($\alpha = 3$)	deaths
$5k$	$10k$	$2.03k$	13.0 k	102 k	13.2 k
$10k$	$10k$	$2.03k$	13.0 k	101 k	13.2 k
$10k$	$20k$	$1.94k$	11.1 k	71.3 k	25.3 k
$10k$	$30k$	$1.92k$	10.9 k	66.6 k	35.9 k
$10k$	$50k$	$2.12k$	14.6 k	115 k	42.2 k
$20k$	$40k$	$1.88k$	10.2 k	59.3 k	49.0 k

izations/intensive therapy occupancy are highlighted in the figures. In no cases, saturation of intensive treatment facilities is observed. Table 4.2 complements the previous figure by reporting economic costs (with economic cost exponent $\alpha = 1,2,3$) and deaths for all scenarios. In general, more conservative choices of $H_{\max}$ lead to significant reductions in the number of deaths, and in some cases also in the economic cost, as an effect of the reduction of oscillations.

We have also tested, reporting results in Table 4.2), situations in which $T_{\max}/H_{\max}$ is kept fixed equal to two (so to guarantee the satisfaction of condition (4.18), while $T_{\max}$ is set respectively to 5000, 10000 and 20000. Note that we obtain different trade-offs between economic cost and number of deaths. In general, by increasing $T_{\max}$, we reduce the economic cost and increase the number of deaths. Evolution of metrics for the case $T_{\max} = 5000$, $H_{\max} = 10000$ is shown in Figure 4.13 (bottom right plot). In this case, contrarily to the case $T_{\max} = 10000$ and $H_{\max} = 20000$, intensive therapy control exerts the tightest control for a given short period.

At last, Table 4.2 reports results for the case $T_{\max} = 10000$, $H_{\max} = 10000$. Observe that the performance of this last case is almost indistinguishable from the case $T_{\max} = 5000$, $H_{\max} = 10000$ (which requires just half of the intensive therapy facilities) both in terms of deaths and economic cost.

In conclusion, in our scenario keeping the ratio $T_{\max}/H_{\max} \approx 2$ appears to be the best choice, as it guarantees that the tightest control is essentially always exerted by hospitalizations in dynamic conditions. Then $T_{\max}$ (and consequently $H_{\max}$) should be chosen instead to achieve the desired trade-off between deaths and economic cost (as previously observed, deaths are more sensitive to parameters than economic costs). In our analysis, we have neglected the costs

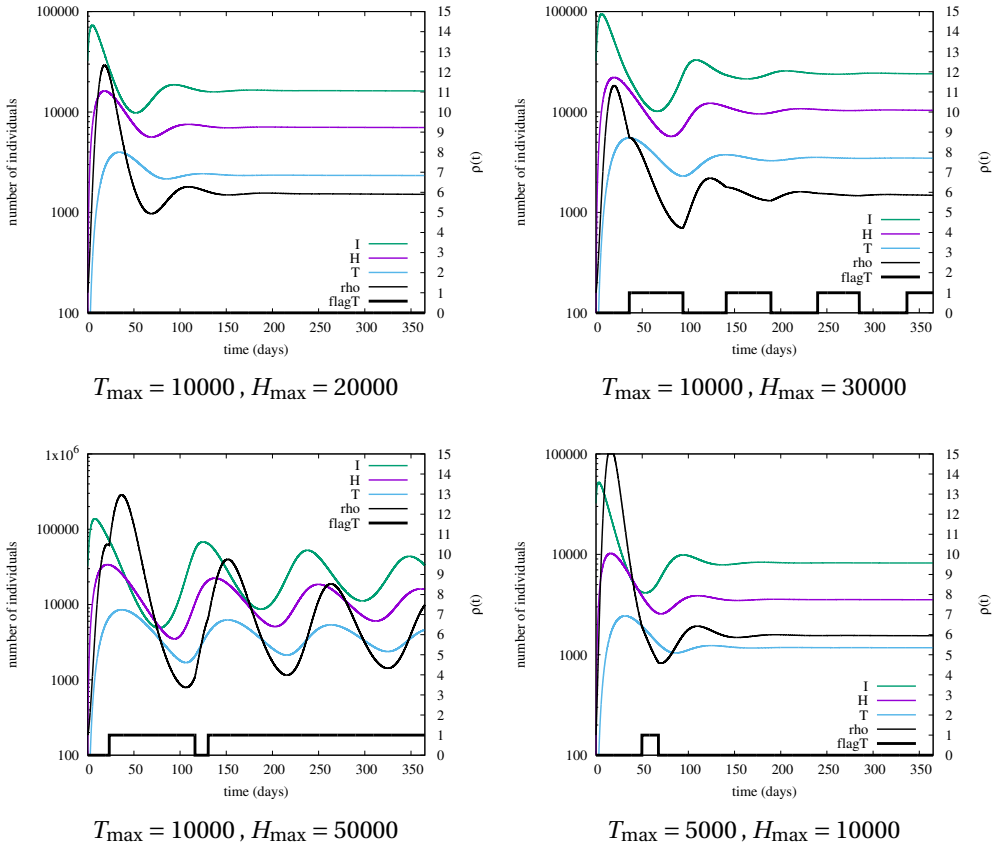


Fig. 4.13 Evolution of $I(t), T(t)$ (left y axes) and $\rho(t)$ (right y axes), for different combinations of $T_{max} = \widehat{T}$ and H_{max}, and fixed $\widehat{H} = 100000$. The rectangles at the bottom of the plots indicate periods in which $\rho(t)$ is determined by $T(t)$.

related to the creation/maintenance of sanitary facilities (which are typically small with respect to general economic costs due to restrictions) to limit the number of free parameters. However, extending the model to include such costs would be relatively immediate.

4.7 Experiments in a Comprehensive Scenario

Our numerical results are obtained in a reference scenario roughly inspired by the actual evolution of COVID-19 in Italy during a period of 3 years, starting from the onset of the virus at the beginning of 2020. During this period, the dynamics of COVID-19 in Italy (and similarly in other European countries) have been characterized by three main phases, each spanning about one year:

1. **First phase**: in this phase the most dangerous strains of the virus, e.g., the alpha and delta mutations propagated in the absence of pharmaceutical interventions (vaccines), causing the majority of all deaths attributed to COVID-19.

2. **Second phase**: since the beginning of 2021, vaccines started to be massively distributed to the population, and almost all individuals (excluding no-vax people) completed the vaccination cycle (by receiving one or two doses) by the end of the second year.

3. **Third phase**: since the beginning of 2022, with the onset of the omicron variant, less dangerous but more virulent strains became prevalent, substituting the initial strains. Vaccines originally developed for the alpha and delta mutations also protected people against the omicron variant, though with reduced efficacy.

To capture the above dynamics, we made some simplifying approximations to limit the model complexity: we assume that a single variant (strain 1), with basic reproduction number $R_0^1 = 6$, propagates during the first 2 phases, after which a new variant (strain 2) appears with higher $R_0^2 = 12$ and reduced mortality (by factor q_{21} with respect to the mortality of strain 1, for each class of people).

The parameters of our reference scenario are summarized in Table 4.3 and their choice is justified in Appendix D. Although our model and parameters can only roughly describe the actual dynamics of COVID-19 in Italy, they provide a realistic scenario in which different strategies to contain the virus can be compared. We stratify the population based on the $f_{r,p}$ distribution calculated for Italy. Transition probabilities between compartments I,H,T,D satisfying constraint (4.3) are set for simplicity as follows: $p_{r,p}^{IH} = p_{r,p}^{HT} = \hat{p}_{r,p}^{TD} = p^{1/3}$.

Strain 1 starts at time 0 with 1 initially infected individual. Similarly, strain 2 starts at time t_2 with 1 initially infected individual. We consider the economic cost function: $\mathfrak{C}(\rho) = (\rho - 1)^\alpha$ which satisfies the assumptions of Proposition 3 for $\alpha \geq 1$, and allows us to explore the impact of costs caused by more substantial non-pharmaceutical interventions by varying the single parameter α. We emphasize that the resulting scenario is not specific to Italy: similar assumptions and parameters could describe equally well, at a high level, the dynamics of COVID-19 in other mid-size European countries or a single US state with comparable popu-

Table 4.3 Parameters of reference scenario

Symbol	**Value**	**Description**
N	60 million	total population size
N^{novax}	6 million	no-vax population size
$t_{\max}$	3 years	time-horizon
t_V	1 year	time at which vaccinations start
t_2	2 years	time at which strain 2 appears
R_0^1	6	basic reproduction number of strain 1
R_0^2	12	basic reproduction number of strain 2
$1/\gamma$	8 days	average sojourn time in state I
$1/\phi$	16 days	average sojourn time in state H
$1/\tau$	16 days	average sojourn time in state T
VE_1^2	0.9	vaccine efficacy against strain 1
VE_2^2	0.7	vaccine efficacy against strain 2
q_{21}	5	mortality reduction of strain 2 vs strain 1
$q_{\texttt{post}}$	20	mortality reduction after exposure to virus/vaccine
$\rho_{\max}$	15	maximum transmissibility reduction
$T_{\max}$	20,000	ICU control parameter
$H_{\max}$	40,000	Hospitalizations control parameter
$\widehat{T}$	20,000	ICU capacity
$\widehat{H}$	50,000	Hospitalizations capacity
θ	10	mortality increase due to ICU saturation
α	2	exponent of economic cost

lation size. At last, while each of the first two phases lasted approximately one year, in our analysis to have a complete view of the potential impact of different control approaches, we have also considered cases in which no effective treatments have been available for several years. When the epidemic spread out at the beginning of 2020, and the first decisions had to be made, no one could predict how long it would have taken to have effective vaccines/treatments available. In the following, for the sake of simplicity, we neglect the term associated with the healthcare system stress by taking into consideration only social (deaths) and economic costs, this corresponds to set $\kappa_2 = 0$.

In the following two subsections, we first examine the interplay between mobility restrictions, enforced through our two Rate and HT controls, and vaccination prioritization schemes (MSF and MVF) hence considering the first two years of our reference scenario. Then, we consider the complete three-year scenario.

4.7.1 Mobility Restrictions and Vaccinations

We first consider the 'first' and 'second phase' of our reference scenario, considering the joint impact of vaccination policies and control strategies during the first two years of the pandemic. Recall from Section 4.4.2 that we focus on two extreme vaccine prioritization policies: Most Vulnerable First (MVF) and Most Social First (MSF).

We will consider a single type of vaccine to be administered in two doses separated by a variable interval of Δ days. In this way, we can address an issue raised in some countries, e.g., the UK, when vaccines started to be available for mass distribution, i.e., whether it is better to follow the recommended protocol ($\Delta = 21$ days) or to give one dose to the largest possible population, before administering the second dose. The latter policy, which aims at partially immunizing a vast portion of the population, corresponds to choosing $\Delta = 135$ days. In our investigation, we assume the vaccination rate to be constant and such that the entire population can receive two doses after 9 months (270 days).

No vaccine is available during the first year (first phase). To better compare our two control strategies, we initially start the system at the equilibrium point (I^*, H^*, T^*), disregarding the transient needed to reach such equilibrium[3]. Under the HT strategy, we assume that control is always determined by the occupation of regular hospitals, rather than ICU, by adequately setting the ratio $T_{\max}/H_{\max}$. Moreover, note that the parameters of the HT strategy can be tuned to achieve the desired number I^* of infected people at the beginning of the pandemic. This allows us to compare the trade-offs achievable by our two control policies.

Given the current understanding of COVID-19 vaccines, one limitation of the approach is the uncertainty surrounding the specific efficacy of different vaccines and their effectiveness against emerging variants. Vaccine efficacy can vary depending on age, underlying health conditions, and individual immune response. Additionally, the duration of vaccine-induced protection and the potential for waning immunity over time are still being studied. As a result, the parameters related to vaccine prioritization, such as the efficacy rates and the duration of protection, are subject to a range of values rather than precise estimates. The lack of comprehensive knowledge about these parameters restricts the ability to

[3] A comprehensive analysis of the complete scenario also comprising the initial transient will be presented later in Section 4.7.2.

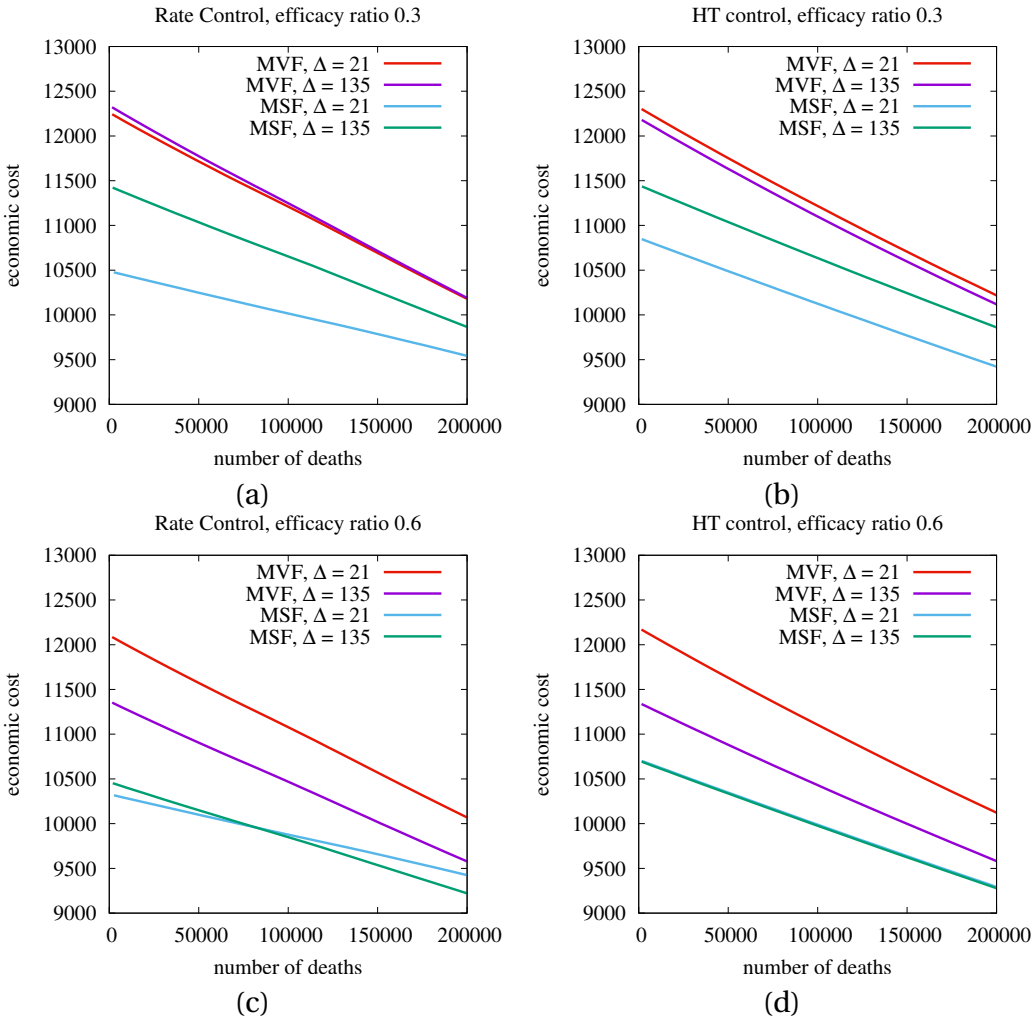


Fig. 4.14 Impact of vaccination policies and control strategies on deaths and economic cost. All individuals are vaccinated in 270 days.

determine an optimal vaccination strategy with certainty. Therefore, the study may need to consider a range of plausible values for vaccine-related parameters and perform sensitivity analyses to assess the robustness of the results under different scenarios. Given the considerations above, we introduce variability in the efficacy ratio between the first and second doses of the vaccine. Specifically, we examine two different values for this ratio, denoted as $\mathrm{VE}^1/\mathrm{VE}^2$, namely 0.3 and 0.6. Meanwhile, we keep the efficacy of the second dose fixed at $\mathrm{VE}^2 = 0.9$.

By incorporating this range of values for the efficacy ratio, we account for the uncertainty surrounding the relative effectiveness of the two vaccine doses.

The achievable trade-offs between economic cost and number of deaths, measured at the end of the second year, are shown in plots (a),(b),(c), and (d) of Fig. 4.14, for the four combinations arising from the two considered control policies and the two considered efficacy ratios (see plot titles). Each plot contains four curves related to the four combinations of vaccination policies (MSF vs. MVF, $\Delta = 21$ vs. $\Delta = 135$).

Several observations are in order. First, the MSF policy (green and blue curves) generally outperforms MVF (red and purple curves). This fact is not trivial and depends crucially on the extent of the negative correlation between r and p in the population distribution $f_{r,p}$. Note that the MSF policy is hardly implementable in practice. Indeed, only the MVF policy has been deployed in many countries, by simple age prioritization, except for special categories of workers (e.g., healthcare workers) who have also received the vaccine in advance due to their exposition to the virus. Second, as expected, the efficacy ratio of 0.6 leads to better outcomes than the efficacy ratio of 0.3. In particular, delaying the distribution of the second dose ($\Delta = 135$) is not advisable if the first dose is relatively ineffective (efficacy ratio 0.3). Third, the impact of different control strategies is fairly small, with rate control slightly outperforming HT control. The best possible trade-offs, i.e., the lowest possible curves, are generated by the rate control, MSF, and a properly tuned Δ (note the crossing between blue and green curves on plot Fig. 4.14(c)). The effect of the two control strategies, combined with different vaccination policies, can be better understood by looking at temporal dynamics shown in Fig. 4.15 for rate and HT control. In both cases, we assume an initial number of infected people $I^* = 32,000$ (corresponding to $\lambda_C = 4,000$) while restricting ourselves to an efficacy ratio of 0.6.

The evolution of $D(t)$, $I(t)$, $T(t)$, $\rho(t)$ in Fig. 4.15 is shown by curves of different colors, respectively red, green, blue, and black. Thick (thin) lines correspond to MVF (MSF). Solid (dashed) lines correspond to $\Delta = 21$ ($\Delta = 135$). Let us start with the simpler case of rate control in Fig. 4.15. Here, $I(t)$ is maintained constant through the entire period of two years. When vaccinations start (day 365), two extreme behaviors for $\rho(t)$ arise, as expected, by MSF with $\Delta = 21$ (thin dashed black line) and MVF with $\Delta = 21$ (thick dashed black line), with the other curves

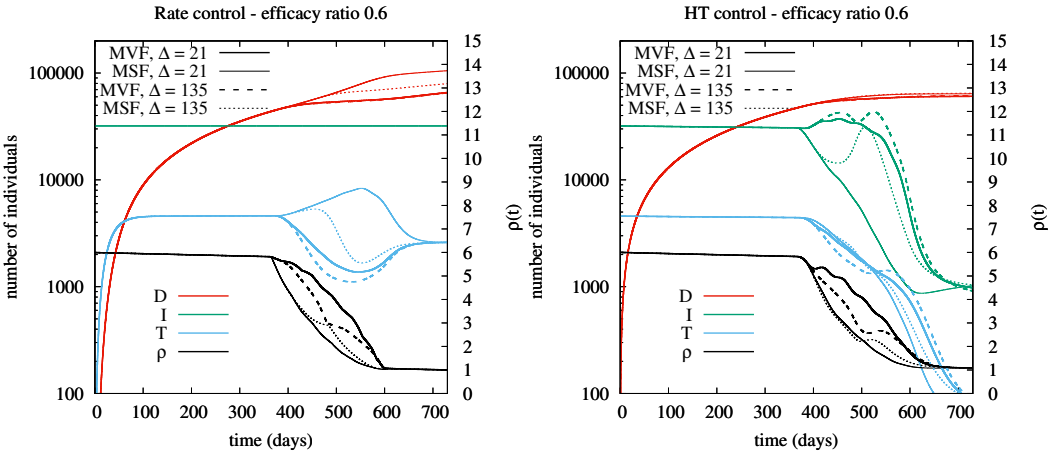


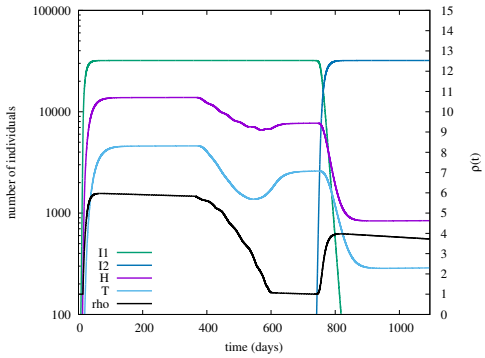
Fig. 4.15 Evolution of $I(t), D(t), T(t), \rho(t)$ in the case of Rate control (left) and HT control (right), efficacy ratio 0.6, and different vaccination policies (different line styles of the same color).

(related to $\Delta = 135$) lying in between these two. MSF with $\Delta = 21$ allows us to release social restrictions more quickly, lowering the economic cost at the expense of more deaths. The case of HT control in Fig. 4.15 is more complex, since here $I(t)$ is not constant and, in fact, decreases drastically during the second year thanks to the self-adaptive nature of HT control.

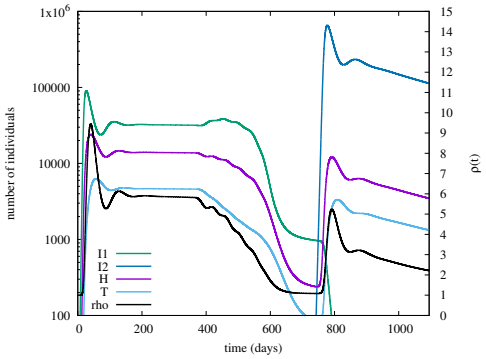
The fact that better trade-offs are achieved by the not self-adaptive rate control at the end of the second year may appear counter-intuitive. Note, however, that such better trade-offs are only possible under a carefully tuned MSF policy, and they are thus hardly achievable in practice. At last, observe that in a more realistic setting, one might not arbitrarily choose the rate of new infections. For example, if one cannot operate below $\lambda_C = 4,000$, from Fig. 4.15, the best option would likely be MVF, which produces significantly fewer deaths at the expense of a tolerable and largely justifiable increase of the economic cost. Interestingly, in this case, $\Delta = 135$ would produce a significantly lower penalty in the economic cost with respect to $\Delta = 21$ while generating an almost identical number of deaths.

4.7.2 Considering all Three Phases

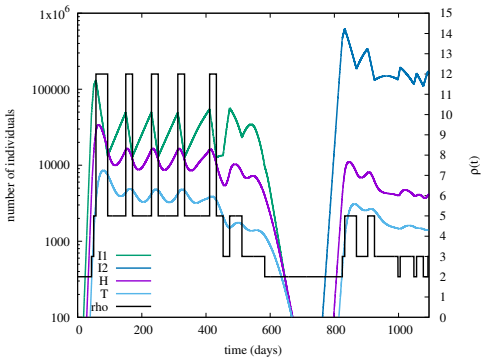
At last, we consider a scenario that includes all three epidemic phases over three years, as described in Section 4.7. The MVF-$\Delta = 21$ vaccination policy was chosen because many countries have actually adopted this policy. The ratio between the efficacy of the first and second dose was set at 0.6.



(a) Rate Control, $\lambda_C = 4000$.



(b) HT control: $T_{\max} = \widehat{T} = 20000$; $H_{\max} = 40000$, $\widehat{H} = 100000$.



(c) Imperfect HT control.

Fig. 4.16 Evolution of $I(t)$, $H(t)$, $T(t)$ (left y axes) and $\rho(t)$ (right y axes) in the comprehensive scenario.

Figure 4.16(a) and Figure 4.16(b) report the evolution of the metrics, respectively, for the case in which the control is on the rate of new infected ($\lambda_C = 4,000$) and the HT (with $T_{\max} = 20000$ and $H_{\max} = 40000$). Parameters have been set so that the two controls operate around approximately the same operational point during the first year.

Rate control appears more reactive in the early phase of the epidemic. As already observed, due to its intrinsic delay in the control ring, HT control exhibits some initial oscillations, which are not observable when rate control is applied. Therefore, it should not be surprising that rate control leads to better performance indices at the end of the first year, as shown in Table 4.4. Note that costs are expressed in arbitrary units, while deaths are expressed in thousands. However, when the second variant starts spreading, the rate-control strategy may overreact, forcing the system to work in over-restricted conditions for quite a long time (note

Table 4.4 Control strategies in a comprehensive scenario: economic cost and deaths

	1st year				three years			
	Cost			Deaths	Cost			Deaths
α	1	2	3		1	2	3	
rate	1.68	8.18	39.9	41.1	3.35	13.5	58.0	70.5
HT	1.76	9.22	50.2	45.8	3.19	13.4	64.6	78.8
IHT	2.02	15.5	146	40.6	3.71	22.1	187	68.7

that at the end of the three-year period, rate-control is far from being completely relieved). Instead, HT control can automatically adjust its operational point as an effect of the mutated environmental conditions, i.e., a smaller intrinsic lethality index of the variant and a significant fraction of vaccinated individuals who are protected against severe outcomes.

We remark that these strategies, which tightly and precisely control either the infection rate or the hospitalization/ICU occupancy, are hardly implementable. However, they provide valuable insights. To shed light on more practical controls, we examine an implementable rough version of the HT control, denoted as Imperfect HT (IHT). Figure 4.16(c) shows the evolution of the epidemic when the IHT strategy is adopted. In this case, the control dynamically selects the current alert level from the following finite set *green, white, yellow, orange, red, purple*. A different set of non-pharmaceutical restrictions corresponds to every alert level, determining a corresponding value of $\rho(t) \in \{1,2,3,5,12,15\}$ (note that intermediate values of $\rho(t)$ corresponding to different alert levels, do not need to be perfectly known). Every week a simple threshold mechanism is implemented to establish the current alert level for the following week, with normalized thresholds (with respect to $H_{\max}$ or $T_{\max}$) set respectively to $\{0.01,0.1,0.2,0.4,1.0\}$. Any alert level must be maintained for at least three weeks before it can be decreased. Despite the behavior of IHT does not significantly deviate from HT, a high extra economic cost is paid for the effect of unavoidable oscillations between consecutive alert levels, especially for large values of α.

4.8 Concluding Remarks

Our research draws on the lessons we have learned from the COVID-19 pandemic. It takes a comprehensive approach to address the challenges of effective planning

and implementation of measures in large communities. The aim is to find a balance between minimizing economic costs and reducing the number of fatalities, a crucial task in pandemic management.

We have developed a sophisticated multi-class model that takes into account the complex relationship between mortality rates and the different risk exposures of population segments. In fact, it is important to note, that we have observed a negative correlation between a person's age and their level of risk exposure to the virus. Moreover, this correlation is not uniform across countries.

In this chapter, we have found that control strategies based on monitoring infection rates or assessing the burden on the healthcare system can be subject to significant instability. This instability becomes particularly notable when one considers that the key system parameters are often highly noisy.

In line with real-world responses to the COVID-19 pandemic, our research provides insights into the practical effectiveness of strategies that control infection rates and hospitalizations. We have also conducted a comparative analysis of vaccination policies, examining different approaches and variations in dosing intervals, inspired by the experiences of countries such as the UK and Italy.

Chapter 5

Conclusion

In this thesis, we have looked at processes over networks, considering interactions between individuals, whether through online contacts, reading each other's posts in an online social network, or through physical interactions in the context of the spread of an epidemic. More broadly, we have presented tools for modeling and analyzing social phenomena, highlighting the connections between different subjects. We have seen how such tools can come from different disciplines (see the Ising model for ferromagnetic materials) and can be used in the social domain.

The first step we took was to describe online interactions in Chapter 2. Online social networks have recently played an increasingly important role in opinion formation, and understanding the mechanisms underlying this modern communication paradigm requires the development of new, flexible frameworks. We have developed an opinion model specifically tailored to these types of interactions, with a particular focus on the interplay between regular users and influencers. We also characterize influential people on online platforms by grounding our design decisions on data from real-world online social networks. Similar to other recent work in the literature, we considered content personalization in a flexible and tunable way. We have shown how content personalization reinforces inequality by favoring structurally advantaged individuals and, in most cases, preventing some influencers from remaining visible on the platform, which could potentially hinder the diversity of opinions in the population. In addition, even in structurally balanced conditions, personalization can lead to the emergence of echo chambers in which users' opinions are radicalized by the influencer's point of view, even

on topics that are not the influencer's reference topic. Or unstable situations in which a single person hegemonizes the online scene.

The popularity evolution as a function of the consistency parameter is an interesting aspect that should be given further attention. An influencer can gain an advantage over the others by simply adjusting the number of posts he publishes in its reference direction. Note that this behavior is highly dependent on the underlying distribution of regular users and requires careful evaluation. A possible future research direction is to consider an influencer with time-varying consistency that aims to increase its popularity (e.g., using low consistency values, see Fig. 2.11a), and understand to what extent this behavior may give the influencer an advantage. We would also like to mention that, even if the modeling framework is quite complicated, many hypotheses have been put forward. It is true that we have observed that an influencer's popularity increases the more they post, but does this phenomenon ever reach saturation? Users, and generally individuals, have a limited budget of attention, so it may not be a wise decision to overload them with posts. However, one needs to understand if the platform is doing such a pruning of posts and how influencers who flood the network are treated, as there are pages that publish hundreds of posts per day.

The model presented in Chapter 2 is one of the first attempts to faithfully describe the complexity of online interactions but, again, comes with some limitations. Users are considered passive entities and influencers are stubborn agents. Moreover, homophily is the primary driver of user interaction, as we did not consider any other explicit relationship structure. Nonetheless, despite the simplifying assumptions, the emergent behavior of the model proved rich enough to show the effects of content personalization and shed light on the dynamics of influencer popularity. In Chapter 3, we addressed one of these limitations by allowing influencers to modify their opinions to maximize their impact on a user population. We formalized this problem of maximizing online social impact over an online social network. In the case of a single influencer over the network, we characterized the optimal strategy thanks to a trellis-like structure that can efficiently solve the problem to be solved efficiently. We found that maximizing the influence at each time step is not always optimal. It seems to be advantageous to first group users together and then exert influence on the group. Inspired by this experiment, we developed a competitive game that we extended for the case of closed-loop interactions (post-feedback-popularity-filtering). We showed that

a disadvantaged influencer should compromise her target opinion more in order to get closer to the mass of the population so that she can exert a greater degree of persuasion on them. We found that an influencer (the advantaged one) tends to monopolize attention and attract virtually all users to her opinion when her popularity develops rapidly.

Competition in social networks has proven to be a challenging and rather unexplored area of research. One aspect that we have only touched on, but which is probably of great interest, is the evolution of the popularity of influencers who deal with the same topics. An interesting question is to what extent success on social media platforms is determined by competence and what role randomness plays. Do posts go viral by chance? Both empirical and modeling efforts are required to answer such questions.

Finally, in Chapter 4, we shifted our focus from opinion dynamics to epidemics and presented a SIR-like model that considers COVID-19 as a use case. We have defined a comprehensive approach to address the challenges of effective planning and implementation of countermeasures at the country level in the event of a pandemic. The aim is to find a balance between minimizing economic costs and reducing deaths, a crucial endeavor in pandemic management. At the heart of our approach lies our multi-class model, which takes into account the complex relationship between mortality rates and the different risk exposures of different population segments. For example, it considers the negative correlation that often exists between an individual's age and their level of risk exposure to the virus. We have developed a data-driven approach that makes it possible to construct country-specific distributions using publicly available data.

Our analysis has shown that epidemic outbreak control strategies based on monitoring infection rates or assessing the burden on the healthcare system can be subject to considerable instability. We assessed stability under different types of delays and identified the conditions that lead to instability. Our framework allows us to examine the interplay of such containment strategies with vaccine prioritization. It is important to emphasize that there is no one-size-fits-all approach to epidemic control. What is appropriate and effective depends on the context, objectives, and available resources. Our approach recognizes this and does not claim to offer a universally superior solution. Instead, it takes a pragmatic perspective. We prioritize the simplicity and ease of implementation of our measures.

We aim to provide a toolkit of strategies that can be considered alongside more complex models and provide policymakers with a range of options. We conducted a comparative analysis of vaccination policy, looking at different approaches and variations in dosing intervals, drawing on the experience of countries such as the UK and Italy. This comparative approach allows us to examine the potential impact of different vaccination strategies in different epidemiological contexts and provide valuable guidance to policy makers.

Apart from the technical contributions, I would like to spend a few words about the challenges ahead. Our society is undergoing profound change, driven by technological breakthroughs, new forms of interaction, and also geographical and political instability. We need to understand these changes and the impact they can have on all aspects of our lives. This will enable us to better adapt to the new habits and needs in the areas of education, work, health, and social interaction, to name a few. To be prepared for these challenges, we must work together to develop social models to have an interpretive key for the new society.

References

[1] Erasmo Recami. L'articolo di ettore majorana su "il valore delle leggi statistiche nella fisica e nelle scienze sociali" (ettore majorana's article on "the value of statistical laws in physics and in social sciences"), 2007.

[2] Zaheer Allam and Zaynah A. Dhunny. On big data, artificial intelligence and smart cities. *Cities*, 89, 2019.

[3] Alain Barrat, Marc Barthélemy, and Alessandro Vespignani. *Dynamical Processes on Complex Networks*. Cambridge University Press, 2008.

[4] Franz Utermohlen. Mean field theory solution of the ising model, 2018.

[5] Charles Kittel. *Elementary statistical physics*. Courier Corporation, 2004.

[6] Giuseppe Grosso and Giuseppe Pastori Parravicini. *Chapter 17 - Magnetic Ordering in Crystals*. Academic Press, 2014.

[7] KATARZYNA SZNAJD-WERON and JÓZEF SZNAJD. Opinion evolution in closed community. *International Journal of Modern Physics C*, 11(06), 2000.

[8] F. Slanina and H. Lavicka. Analytical results for the sznajd model of opinion formation. *The European Physical Journal B - Condensed Matter*, 35(2), 2003.

[9] Matteo Cinelli, Gianmarco De Francisci Morales, Alessandro Galeazzi, Walter Quattrociocchi, and Michele Starnini. The echo chamber effect on social media. *Proceedings of the National Academy of Sciences*, 118(9), 2021.

[10] Zhilan Feng Fred Brauer, Carlos Castillo-Chavez. *Mathematical Epidemiology*. Springer Berlin Heidelberg, 2008.

[11] Miller McPherson, Lynn Smith-Lovin, and James M Cook. Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27(1), 2001.

[12] Francesco Fallucchi, Marco Faravelli, and Simone Quercia. Fair allocation of scarce medical resources in the time of covid-19: What do people think? *Journal of Medical Ethics*, 47(1), 2021.

[13] Franco Galante, Luca Vassio, Michele Garetto, and Emilio Leonardi. Modeling communication asymmetry and content personalization in online social networks. *Online Social Networks and Media*, 37-38, 2023.

[14] Brooke Auxier and Monica Anderson. Social media use in 2021. *Pew Research Center*, 1, 2021.

[15] Solomon E. Asch. Opinions and social pressure. *Scientific American*, 193(5), 1955.

[16] John R. P. French. A formal theory of social power. *Psychological Review*, 63(3), 1956.

[17] Leon Festinger. A theory of social comparison processes. *Human Relations*, 7(2), 1954.

[18] Loretta Mastroeni, Pierluigi Vellucci, and Maurizio Naldi. Agent-based models for opinion formation: A bibliographic survey. *IEEE Access*, 7, 2019.

[19] Morris H. Degroot. Reaching a consensus. *Journal of the American Statistical Association*, 69(345), 1974.

[20] Noah E. Friedkin and Eugene C. Johnsen. Social influence and opinions. *The Journal of Mathematical Sociology*, 15(3–4), 1990.

[21] Pranav Dandekar, Ashish Goel, and David T. Lee. Biased assimilation, homophily, and the dynamics of polarization. *Proceedings of the National Academy of Sciences*, 110(15), 2013.

[22] Rainer Hegselmann and Ulrich Krause. Opinion dynamics and bounded confidence: Models, analysis and simulation. *Journal of Artificial Societies and Social Simulation*, 5, 2002.

[23] Guillaume Deffuant, David Neau, Frederic Amblard, and Gérard Weisbuch. Mixing beliefs among interacting agents. *Advances in Complex Systems*, 03(01n04), 2000.

[24] Jan Lorenz. Continuous opinion dynamics under bounded confidence: A survey. *International Journal of Modern Physics C*, 18(12), 2007.

[25] PETER CLIFFORD and AIDAN SUDBURY. A model for spatial conflict. *Biometrika*, 60(3), 1973.

[26] Richard A. Holley and Thomas M. Liggett. Ergodic theorems for weakly interacting infinite systems and the voter model. *The Annals of Probability*, 3(4), 1975.

[27] Petter Holme and M. E. J. Newman. Nonequilibrium phase transition in the coevolution of networks and opinions. *Physical Review E*, 74(5), 2006.

[28] Richard Durrett, James P. Gleeson, Alun L. Lloyd, Peter J. Mucha, Feng Shi, David Sivakoff, Joshua E. S. Socolar, and Chris Varghese. Graph fission in an evolving voter model. *Proceedings of the National Academy of Sciences*, 109(10), 2012.

[29] Cecilia Nardini, Balázs Kozma, and Alain Barrat. Who's talking first? consensus or lack thereof in coevolving opinion formation models. *Physical Review Letters*, 100(15), 2008.

[30] Boris L. Granovsky and Neal Madras. The noisy voter model. *Stochastic Processes and their Applications*, 55(1), 1995.

[31] E. Ben-Naim. Opinion dynamics: Rise and fall of political parties. *Europhysics Letters*, 69(55), 2005.

[32] G. Toscani. Kinetic models of opinion formation, 2006.

[33] Claudio Castellano, Santo Fortunato, and Vittorio Loreto. Statistical physics of social dynamics. *Reviews of Modern Physics*, 81(2), 2009.

[34] J. T. Cox. Coalescing random walks and voter model consensus times on the torus in zd. *The Annals of Probability*, 17(4), 1989.

[35] L. Frachebourg and P. L. Krapivsky. Exact results for kinetics of catalytic reactions. *Physical Review E*, 53(4), 1996.

[36] Krzysztof Suchecki, Víctor M. Eguíluz, and Maxi San Miguel. Voter model dynamics in complex networks: Role of dimensionality, disorder, and degree distribution. *Physical Review E*, 72(3), 2005.

[37] V. Sood and S. Redner. Voter model on heterogeneous graphs. *Physical Review Letters*, 94(17), 2005.

[38] Ercan Yildiz, Asuman Ozdaglar, Daron Acemoglu, Amin Saberi, and Anna Scaglione. Binary opinion dynamics with stubborn agents. *ACM Transactions on Economics and Computation*, 1(4), 2013.

[39] Carlo Michele Valensise, Matteo Cinelli, and Walter Quattrociocchi. The dynamics of online polarization, 2022.

[40] Antonio F. Peralta, Matteo Neri, János Kertész, and Gerardo Iñiguez. Effect of algorithmic bias and network structure on coexistence, consensus, and polarization of opinions. *Physical Review E*, 104(4), 2021.

[41] Nicola Perra and Luis E. C. Rocha. Modelling opinion dynamics in the age of algorithmic personalisation. *Scientific Reports*, 9(1), 2019.

[42] Antonio F. Peralta, János Kertész, and Gerardo Iñiguez. Opinion dynamics in social networks: From models to data, 2022.

[43] Santo Fortunato and Claudio Castellano. Scaling and universality in proportional elections. *Physical Review Letters*, 99, 2007.

[44] Christopher A. Bail, Lisa P. Argyle, Taylor W. Brown, John P. Bumpus, Haohan Chen, M. B. Fallin Hunzaker, Jaemin Lee, Marcus Mann, Friedolin Merhout, and Alexander Volfovsky. Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences*, 115(37), 2018.

[45] T. J. H. Morgan, L. E. Rendell, M. Ehn, W. Hoppitt, and K. N. Laland. The evolutionary basis of human social learning. *Proceedings of the Royal Society B: Biological Sciences*, 279(1729), 2011.

[46] Juan Fernández-Gracia, Krzysztof Suchecki, José J. Ramasco, Maxi San Miguel, and Víctor M. Eguíluz. Is the voter model a model for voters? *Physical Review Letters*, 112, 2014.

[47] Pablo Barberá, John T. Jost, Jonathan Nagler, Joshua A. Tucker, and Richard Bonneau. Tweeting from left to right: Is online political communication more than an echo chamber? *Psychological Science*, 26(10), 2015.

[48] Abir De, Sourangshu Bhattacharya, Parantapa Bhattacharya, Niloy Ganguly, and Soumen Chakrabarti. Learning linear influence models in social networks from transient opinion dynamics. *ACM Trans. Web*, 13(3), 2019.

[49] Corrado Monti, Giuseppe Manco, Cigdem Aslay, and Francesco Bonchi. Learning ideological embeddings from information cascades. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*. Association for Computing Machinery, 2021.

[50] Eytan Bakshy, Solomon Messing, and Lada A. Adamic. Exposure to ideologically diverse news and opinion on facebook. *Science*, 348(6239), 2015.

[51] Sinan Aral, Lev Muchnik, and Arun Sundararajan. Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proceedings of the National Academy of Sciences*, 106(51), 2009.

[52] Hannes Risken. *Fokker-Planck equation*, pages 63–95. Springer, 1996.

[53] Robin Cohen, Alan Tsang, Krishna Vaidyanathan, and Haotian Zhang. Analyzing opinion dynamics in online social networks. *Big Data & Information Analytics*, 1(4), 2016.

[54] Franco Galante, Michele Garetto, and Emilio Leonardi. Competition of influencers: A model for maximizing online social impact. In *Proceedings of the 16th ACM Web Science Conference*. Association for Computing Machinery, 2024.

[55] Claudio Castellano, Santo Fortunato, and Vittorio Loreto. Statistical physics of social dynamics. *Reviews of Modern Physics*, 81(2), 2009.

[56] K. Sridhar Moorthy. Using game theory to model competition. *Journal of Marketing Research*, 22(3), 1985.

[57] Morris H. Degroot. Reaching a consensus. *Journal of the American Statistical Association*, 69, 1974.

[58] Noah Friedkin and Eugene Johnsen. Social influence networks and opinion change. *Advances in Group Processes*, 16, 1999.

[59] Guillaume Deffuant, David Neau, Frederic Amblard, and Gérard Weisbuch. Mixing beliefs among interacting agents. *Advances in Complex Systems*, 03(01n04), 2000.

[60] Hegselmann Rainer and Ulrich Krause. Opinion dynamics and bounded confidence: Models, analysis and simulation. *Journal of Artificial Societies and Social Simulation*, 5(3), 2002.

[61] Hossein Noorazar. Recent advances in opinion propagation dynamics: a 2020 survey. *The European Physical Journal Plus*, 135, 2020.

[62] Carmela Bernardo, Claudio Altafini, Anton Proskurnikov, and Francesco Vasca. Bounded confidence opinion dynamics: A survey. *Automatica*, 159, 2024.

[63] David Kempe, Jon Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. In *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2003.

[64] Wei Chen, Alex Collins, Rachel Cummings, Te Ke, Zhenming Liu, David Rincon, Xiaorui Sun, Yajun Wang, Wei Wei, and Yifei Yuan. Influence maximization in social networks when negative opinions may emerge and propagate. In *Proceedings of the 2011 SIAM International Conference on Data Mining (SDM)*, pages 379–390, 2011.

[65] Wei Lu, Wei Chen, and Laks V. S. Lakshmanan. From competition to complementarity: comparative influence diffusion and maximization. *Proc. VLDB Endow.*, 9, 2015.

[66] Heather Z. Brooks and Mason A. Porter. A model for the influence of media on the ideology of content in online social networks. *Physical Review Research*, 2, 2020.

[67] Rainer Hegselmann and Ulrich Krause. Opinion dynamics under the influence of radical groups, charismatic leaders, and other constant signals: A simple unifying model. *Networks and Heterogeneous Media*, 10, 2015.

[68] Anahita Mirtabatabaei, Peng Jia, and Francesco Bullo. Eulerian opinion dynamics with bounded confidence and exogenous inputs. *SIAM Journal on Applied Dynamical Systems*, 13(1), 2014.

[69] Suttida Wongkaew, Marco Caponigro, and Alfio Borzì. On the control through leadership of the hegselmann–krause opinion formation model. *Mathematical Models and Methods in Applied Sciences*, 25(03), 2015.

[70] Rainer Hegselmann, Stefan König, Sascha Kurz, Christoph Niemann, and Jörg Rambau. Optimal Opinion Control: The Campaign Problem. *Journal of Artificial Societies and Social Simulation*, 18(3), 2015.

[71] Sascha Kurz. Optimal control of the freezing time in the hegselmann–krause dynamics. *Journal of Difference Equations and Applications*, 21, 2014.

[72] Antonia Masucci and Alonso Silva. Strategic resource allocation for competitive influence in social networks. *2014 52nd Annual Allerton Conference on Communication, Control, and Computing, Allerton 2014*, 2014.

[73] Swapnil Dhamal, Walid Ben-Ameur, Tijani Chahed, and Eitan Altman. Optimal investment strategies for competing camps in a social network: A broad framework. *IEEE Transactions on Network Science and Engineering*, 6, 2019.

[74] Vineeth Varma, Morarescu Irinel-Constantin, Samson Lasaulce, and Samuel Martin. Opinion dynamics aware marketing strategies in duopolies. In *IEEE 56th Annual Conference on Decision and Control (CDC)*, 2017.

[75] Vineeth S. Varma, Samson Lasaulce, Julien Mounthanyvong, and Irinel-Constantin Morărescu. Allocating marketing resources over social networks: A long-term analysis. *IEEE Control Systems Letters*, 3(4), 2019.

[76] Mohak Goyal and D. Manjunath. Opinion control competition in a social network. In *2020 International Conference on COMmunication Systems & NETworkS (COMSNETS)*, 2020.

[77] Franco Galante, Chiara Ravazzi, Michele Garetto, and Emilio Leonardi. Planning interventions in a controlled pandemic: the covid-19 case. *IEEE Transactions on Network Science and Engineering*, 2023.

[78] Laura Matrajt, Julia Eaton, Tiffany Leung, and Elizabeth R. Brown. Vaccine optimization for covid-19: Who to vaccinate first? *Science Advances*, 7(6), 2021.

[79] Chadi M Saad-Roy, Shaun E Morris, C Jessica E Metcalf, Michael J Mina, Robert E Baker, Jeremy Farrar, Edward C Holmes, Oliver G Pybus, Andrea L Graham, Simon A Levin, Bryan T Grenfell, and Christina E Wagner. Epidemiological and evolutionary considerations of sars-cov-2 vaccine dosing regimes. *Science (New York, N.Y.)*, 372(6540), 2021.

[80] Joshua R. Goldstein, Thomas Cassidy, and Kenneth W. Wachter. Vaccinating the oldest against COVID-19 saves both the most lives and most years of life. *PNAS*, 118(11), 2021.

[81] Jan Medlock and Alison P. Galvani. Optimizing Influenza Vaccine Distribution. *Science*, 325(5948), 2009.

[82] Frank G Sandmann, Nicholas G Davies, Anna Vassall, W John Edmunds, Mark Jit, and Centre for the Mathematical Modelling of Infectious Diseases COVID-19 working group. The potential health and economic value of sars-cov-2 vaccination alongside physical distancing in the uk: a transmission model-based future scenario analysis and economic evaluation. *The Lancet Infectious Diseases*, 21(7), 2021.

[83] Michel Monod, Alexander Blenkinsop, Xiaoyue Xi, David Hebert, Seth Bershan, Savanna Tietze, Marc Baguelin, Valerie C Bradley, Yen Ting Chen, Helen Coupland, Sarah Filippi, Jennifer Ish-Horowicz, Matthew McManus, Thomas Mellan, Axel Gandy, Martin Hutchinson, Helen JT Unwin, Sabine L van Elsland, Michael A C Vollmer, Samir B Weber, ... [and 14 others], and Imperial College COVID-19 Response Team. Age groups that sustain resurging covid-19 epidemics in the united states. *Science*, 371(6536), 2021.

[84] Peter Jentsch, Madhur Anand, and Chris T. Bauch. Prioritising COVID-19 vaccination in changing social and epidemiological landscapes: a mathematical modelling study. *The Lancet Infectious Deseases*, 21, 2021.

[85] William Ogilvy Kermack and A. G. McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 115(772), 1927.

[86] Dave Osthus, Kyle S. Hickmann, Petruţa C. Caragea, Dave Higdon, and Sara Y. Del Valle. Forecasting seasonal influenza with a state-space SIR model. *The Annals of Applied Statistics*, 11(1), 2017.

[87] D. Aldila, N. Nuraini, and E. Soewono. Optimal control problem in preventing of swine flu disease transmission. *Applied Mathematical Sciences*, 8, 2014.

[88] Herbert W. Hethcote. The mathematics of infectious diseases. *SIAM Review*, 42(4), 2000.

[89] Juliana Tolles and ThaiBinh Luong. Modeling epidemics with compartmental models. *JAMA*, 323(24), 2020.

[90] Dina Mistry, Maria Litvinova, Ana Piontti, Matteo Chinazzi, Laura Fumanelli, Marcelo Ferreira da Costa Gomes, Syed Haque, Quan-Hui Liu, Kunpeng Mu, Xinyue Xiong, M. Halloran, Ira Longini, Stefano Merler, Marco Ajelli, and Alessandro Vespignani. Inferring high-resolution human mixing patterns for disease modeling. *Nature Communications*, 12, 2021.

[91] Qianying Lin, Shi Zhao, Daozhou Gao, Yijun Lou, Shu Yang, Salihu S. Musa, Maggie H. Wang, Yongli Cai, Weiming Wang, Lin Yang, and Daihai He. A conceptual model for the coronavirus disease 2019 (covid-19) outbreak in wuhan, china with individual reaction and governmental action. *International Journal of Infectious Diseases*, 93, 2020.

[92] Cleo Anastassopoulou, Lucia Russo, Athanasios Tsakris, and Constantinos Siettos. Data-based analysis, modelling and forecasting of the covid-19 outbreak. *PLOS ONE*, 15(3), 2020.

[93] F. Casella. Can the covid-19 epidemic be controlled on the basis of daily test reports? *IEEE Control Systems Letters*, 5(3), 2021.

[94] Giulia Giordano, Franco Blanchini, Raffaele Bruno, Patrizio Colaneri, Alessandro Di Filippo, Angela Di Matteo, and Marta Colaneri. Modelling the covid-19 epidemic and implementation of population-wide interventions in italy. *Nature Medicine*, 26(6), 2020.

[95] Kate M. Bubar, Kyle Reinholt, Stephen M. Kissler, Marc Lipsitch, Sarah Cobey, Yonatan H. Grad, and Daniel B. Larremore. Model-informed covid-19 vaccine prioritization strategies by age and serostatus. *Science*, 371(6532), 2021.

[96] Jeffrey P Townsend, Hayley B Hassler, Zheng Wang, Sayaka Miura, Jaiveer Singh, Sudhir Kumar, Nancy H Ruddle, Alison P Galvani, and Alex Dornburg. The durability of immunity against reinfection by sars-cov-2: a comparative evolutionary study. *The Lancet Microbe*, 2(12), 2021.

[97] Howard M. Taylor. Some models in epidemic control. *Mathematical Biosciences*, 3, 1968.

[98] Andris Abakuks. An optimal isolation policy for an epidemic. *Journal of Applied Probability*, 10(2), 1973.

[99] R. Morton and K. H. Wickwire. On the optimal control of a deterministic epidemic. *Advances in Applied Probability*, 6(4), 1974.

[100] K.H. Wickwire. Optimal isolation policies for deterministic and stochastic epidemics. *Mathematical Biosciences*, 26(3), 1975.

[101] Suresh P. Sethi and Preston W. Staats. Optimal control of some simple deterministic epidemic models. *The Journal of the Operational Research Society*, 29(2), 1978.

[102] Eric Hansen and Troy Day. Optimal control of epidemics with limited resources. *Journal of Mathematical Biology*, 62, 2011.

[103] Thomas Kruse and Philipp Strack. Optimal Control of an Epidemic through Social Distancing. Cowles Foundation Discussion Papers 2229, Cowles Foundation for Research in Economics, Yale University, 2020.

[104] Horst Behncke. Optimal control of deterministic epidemics. *Optimal Control Applications and Methods*, 21(6), 2000.

[105] Luca Bolzoni, Elena Bonacini, Cinzia Soresina, and Maria Groppi. Time-optimal control strategies in sir epidemic models. *Mathematical Biosciences*, 292, 2017.

[106] John R. Birge, Ozan Candogan, and Yiding Feng. Controlling epidemic spread: Reducing economic losses with targeted closures. *Management Science*, 68(5), 2022.

[107] P.E. Paré, C L. Beck, and T. Modeling, estimation, and analysis of epidemics over networks: an overview. *Annual Reviews in Control*, 50, 2020.

[108] Tom Britton, Maria Deijfen, and Andreas N Lagerås. Optimal control of vaccination dynamics during an influenza epidemic. *Journal of Mathematical Biology*, 78(1-2), 2019.

[109] Zhilan Feng and Xue-Qiao Zhao. Optimal vaccination and treatment in a multi-group epidemic model. *Journal of Mathematical Biology*, 67(6–7), 2013.

[110] Jia Zhang, Meng Li, and Zhien Ma. Optimal control of a vaccination model with vaccination age. *Mathematical Biosciences and Engineering*, 11(5), 2014.

[111] F. Altarelli. Containing epidemic outbreaks by message-passing techniques. *Physical Review X*, 4, 2014.

[112] A. Galeotti, B. Golub, and S. Goyal. Strategic immunization and group structure. *American Economic Journal: Microeconomics*, 5(2), 2013.

[113] Tom Britton, Frank Ball, and Pieter Trapman. A mathematical model reveals the influence of population heterogeneity on herd immunity to sars-cov-2. *Science*, 369(6505), 2019.

[114] Daron Acemoglu, Victor Chernozhukov, Iván Werning, and Michael D. Whinston. Optimal targeted lockdowns in a multi-group sir model. Technical Report 27102, National Bureau of Economic Research, 2020.

[115] Rajasree Bhattacharyya, Sayan Mukherjee, and Anjan Mukhopadhyay. Modelling the covid-19 pandemic in india: A compartmental model study considering heterogeneity and mobility. *Chaos, Solitons & Fractals*, 146, 2021.

[116] Rajiv Chowdhury, Kevin Heng, Md Shafiur Rahman Shawon, Gabriel Goh, Daisy Okonofua, Carolina Ochoa-Rosales, Valentina Gonzalez-Jaramillo, Abbas Bhuiya, Daniel Reidpath, Shamini Prathapan, et al. Dynamic interventions to control covid-19 pandemic: A multivariate prediction modelling study comparing 16 worldwide countries. *European Journal of Epidemiology*, 35(5), 2020.

[117] Xin Li, Meng Liu, Yiqing Deng, Zhongjie Sun, Lu Xiao, and Ye Feng. Epidemiological and dynamical analysis of covid-19 transmission in shenzhen china. *Annals of Translational Medicine*, 8(21), 2020.

[118] Rene Carmona and Peiqi Wang. Finite state mean field games with major and minor players, 2016.

[119] Alexander Aurell, René Carmona, Gökçe Dayanikli, and Mathieu Laurière. Optimal incentives to mitigate epidemics: A stackelberg mean field game approach. *SIAM Journal on Control and Optimization*, 60(2), 2022.

[120] Romuald Elie, Emma Hubert, and Gabriel Turinici. Contact rate epidemic control of COVID-19: an equilibrium view. *Mathematical Modelling of Natural Phenomena*, 15, 2020.

[121] Arthur Charpentier, Romuald Elie, Mathieu Laurière, and Viet Chi Tran. Covid-19 pandemic control: balancing detection policy and lockdown intervention under icu sustainability, 2020.

[122] Romualdo Pastor-Satorras, Claudio Castellano, Piet Van Mieghem, and Alessandro Vespignani. Epidemic processes in complex networks. *Reviews of Modern Physics*, 87(3), 2015.

[123] Mark EJ Newman. Spread of epidemic disease on networks. *Physical Review E*, 66(1), 2002.

[124] Matt J Keeling and Ken T D Eames. Networks and epidemic models. *Journal of The Royal Society Interface*, 2(4), 2005.

[125] Laura Matrajt, Tom Britton, M. Elizabeth Halloran, and Ira M. Longini. One versus two doses: What is the best use of vaccine in an influenza pandemic? *Epidemics*, 13, 2015.

[126] Xiao-Jie Li, Cong Li, and Xiang Li. Minimizing social cost of vaccinating network sis epidemics. *IEEE Transactions on Network Science and Engineering*, 5(4), 2018.

[127] Markus Kantner and Thomas Koprucki. Beyond just "flattening the curve": Optimal control of epidemics with purely non-pharmaceutical interventions. *Journal of Mathematics in Industry*, 10, 2020.

[128] Leonardo Cianfanelli, Francesca Parise, Daron Acemoglu, Giacomo Como, and Asuman Ozdaglar. Lockdown interventions in sir models: Is the reproduction number the right control variable? In *2021 60th IEEE Conference on Decision and Control (CDC)*, 2021.

[129] Mingtao Xia, Lucas Böttcher, and Tom Chou. Controlling epidemics through optimal allocation of test kits and vaccine doses across networks. *IEEE Transactions on Network Science and Engineering*, 9(3), 2022.

[130] Ashish R. Hota, Jaydeep Godbole, and Philip E. Paré. A closed-loop framework for inference, prediction, and control of sir epidemics on networks. *IEEE Transactions on Network Science and Engineering*, 8(3), 2021.

[131] Luca Vassio, Michele Garetto, Carla Chiasserini, and Emilio Leonardi. Temporal dynamics of posts and user engagement of influencers on facebook

and instagram. In *Proceedings of the 2021 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*. Association for Computing Machinery, 2022.

[132] Martino Trevisan, Luca Vassio, and Danilo Giordano. Debate on online social networks at the time of covid-19: An italian case study. *Online Social Networks and Media*, 23, 2021.

[133] Carlo Signorelli and Anna Odone. Age-specific COVID-19 case-fatality rate: no evidence of changes over time. *International Journal of Public Health*, 65(8), 2020.

[134] Selene Ghisolfi, Ingvild Almås, Justin C Sandefur, Tillman von Carnap, Jesse Heitner, and Tessa Bold. Predicted COVID-19 fatality rates based on age, sex, comorbidities and health system capacity. *BMJ Global Health*, 5(9), 2020.

[135] Ian C. Marschner. Estimating age-specific COVID-19 fatality risk and time to death by comparing population diagnosis and death patterns: Australian data. *BMC Medical Research Methodology*, 21(1), 2021.

[136] World Health Organization. Life tables by country (ghe: Life tables), 2020. data retrieved from Global Health Observatory, https://www.who.int/data/gho/data/indicators/indicator-details/GHO/gho-ghe-life-tables-by-country.

[137] Ying Liu and Joacim Rocklöv. The reproductive number of the delta variant of SARS-CoV-2 is far higher compared to the ancestral SARS-CoV-2 virus. *Journal of Travel Medicine*, 28(7), 2021.

[138] Kevin Linka, Mathias Peirlinck, and Ellen Kuhl. The reproduction number of COVID-19 and its correlation with public health interventions. *Computational Mechanics*, 66(4), 2020.

[139] Nick Shryane, Maria Pampaka, Andrea Lisette Aparicio Castro, Shazaad Ahmad, Mark Elliot, Ji Hye Kim, Jennifer Murphy, Wendy Olsen, Diego Andres Perez Ruiz, and Arkadiusz Wiśniowski. Length of stay in ICU of covid-19 patients in england, march - may 2020. *International Journal of Population Data Science*, 5(4), 2021.

[140] Eleanor M. Rees, Emily S. Nightingale, Yalda Jafari, Naomi R. Waterlow, Samuel Clifford, Carl A. B. Pearson, CMMID Working Group, Thibaut Jombart, Simon R. Procter, and Gwenan M. Knight. COVID-19 length of hospital stay: a systematic review and data synthesis. *BMC Medicine*, 18(1), 2020.

[141] Andrew William Byrne, David McEvoy, Aine B Collins, Kevin Hunt, Miriam Casey, Ann Barber, Francis Butler, John Griffin, Elizabeth A Lane, Conor McAloon, Kirsty O'Brien, Patrick Wall, Kieran A Walsh, and Simon J More. Inferred duration of infectious period of SARS-CoV-2: rapid scoping review and analysis of available evidence for asymptomatic and symptomatic COVID-19 cases. *BMJ Open*, 10(8), 2020.

[142] Governo Italiano. Report vaccini anti covid-19. https://www.governo.it/it/cscovid19/report-vaccini/. Accessed Jul. 17, 2023.

[143] Yuan Liu, Yangyang Yu, Yanji Zhao, and Daihai He. Reduction in the infection fatality rate of omicron variant compared with previous variants in south africa. *International Journal of Infectious Diseases*, 120, 2022.

[144] Kazem Rahmani, Rasoul Shavaleh, Mahtab Forouhi, Hamideh Feiz Disfani, Mostafa Kamandi, Rozita Khatamian Oskooi, Molood Foogerdi, Moslem Soltani, Maryam Rahchamani, Mohammad Mohaddespour, and Mostafa Dianatinasab. The effectiveness of COVID-19 vaccines in reducing the incidence, hospitalization, and mortality from COVID-19: A systematic review and meta-analysis. *Frontiers in Public Health*, 10, 2022.

[145] ISTAT. Rapporto annuale 2020. La situazione del Paese. Presented in Palazzo Montecitorio, 3rd July 2020, Roma, 2020.

[146] L. Freddi. Optimal control of the transmission rate in compartmental epidemics. *Mathematical Control & Related Fields*, 12(1), 2022.

[147] Greg Kaplan, Benjamin Moll, and Giovanni L Violante. The great lockdown and the big stimulus: Tracing the pandemic possibility frontier for the u.s. Working Paper 27794, National Bureau of Economic Research, 2020.

[148] Andreas Kasis, Stelios Timotheou, Nima Monshizadeh, and Marios Polycarpou. Optimal intervention strategies to mitigate the covid-19 pandemic effects. *Scientific Reports*, 12(1), 2022.

[149] Jing Yuan, Meng Li, Guodong Lv, and Zhi-Kai Lu. Monitoring transmissibility and mortality of covid-19 in europe. *International Journal of Infectious Diseases*, 95, 2020.

[150] Office for National Statistics. Deaths at home increased by a third in 2020, while deaths in hospitals fell except for COVID-19. https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/articles/deathsathomeincreasedbyathirdin2020whiledeathsinhospitalsfellexceptforcovid19/2021-05-07, 2021.

Appendix A

Description of the Dataset

We collected data from real online social networks to support the hypotheses of our model and compare emergent behaviors. We focus on two popular social networks: Facebook (FB) and Instagram (IG). Facebook has long been the most popular social media application, while Instagram has undergone a surge in popularity in recent years.

In Facebook and Instagram, a *profile*, i.e., a social network user, can be followed by other profiles, i.e., its *followers*. A profile with a large number of followers is also called an *influencer* - we consider profiles with more than ten thousand followers as influencers. Influencers post content (i.e., *posts*) that the profile's followers and anyone registered on the platform can see, *like* and comment. Note that when we use the term influencer, we do not only mean individuals but also groups, soccer teams, newspapers, or companies.

To get the list of such popular profiles, we exploited the online analytics platform hypeauditor.com for IG, and www.socialbakers.com and www.pubblicodelirio.it for FB. We restricted the analysis to influencers with at least 10,000 followers on June 1, 2021. The obtained 649 influencers are the same as our previous paper [131] and are publicly available.[1] In this work, we are interested in the posts of influencers and their temporal sequence. For each monitored influencer, we downloaded all the data related to the posts published between January 1, 2016, and June 1, 2021, using the CrowdTangle tool and its API[2]. CrowdTangle

[1] https://mplanestore.polito.it:5001/sharing/P4WnRClQn
[2] https://github.com/CrowdTangle/API

is a content discovery and social analytics tool owned by Meta and available to researchers and analysts worldwide to support research, subject to a partnership agreement. Finally, we stored the data, which takes around 110 GB of disk space, on a Hadoop-based cluster, and we used PySpark for scalable processing.

A.1 Details on Post Classification

We developed a classifier that can categorize posts according to a particular set of subjects, similar to what we have done in our previous work [132]. First, we arbitrarily identified a subset of topics that sufficiently characterize the discussions on the monitored profiles. Specifically, these topics are *sports*, *politics*, *food and cooking*, *music*, and *pandemics*, which are intentionally loose and relatively uncorrelated to each other. We developed a keyword classifier to classify the posts. For each topic, we manually defined a list of representative keywords. For example, if we consider *pandemic*, we search for words like COVID, pandemic, and coronavirus in Italian (and commonly used terms in other languages). We search for the topic-specific terms in the text corpus of the post, and if we find a match, we mark the post as belonging to the topic. Notice that since keywords of various topics may be present in the same corpus, we can flag a message as discussing *multiple* topics. In this work, we discard posts marked as multiple and only consider posts associated with a single topic.

We are not interested in classifying all posts by an influencer, first because our list of topics does not cover all possible ones, and second, because we only need a large enough subsample of posts to make some statistical considerations. Conversely, it is of utmost importance that the accuracy of the classifier is high since misclassified posts could lead to wrong conclusions about the distribution among the available topics. Therefore, we manually validate the accuracy of our methodology for topic detection, as described in the following paragraph.

A.2 Classifier Precision Evaluation

We empirically evaluated the accuracy of the classifier by taking a random subsample of the labeled posts, i.e., 100 posts for each topic for a total of 500 messages,

and manually classifying them. To this end, we defined a lower and upper bound for accuracy. Indeed, even for a human being, it is challenging to univocally classify posts based on their content. Therefore, we defined three possible states for each classification decision: t correct classification, f incorrect classification, and *ncc* standing for not completely correct (indicating that the assigned topic is related to the post but may not be the main topic of the post or the classification of the post is difficult). Given this state subdivision, the precision bounds are as follows:

$$P_L = \frac{N_t}{N_t + N_f + N_{ncc}} \tag{A.1}$$

$$P_U = \frac{N_t + N_{ncc}}{N_t + N_f + N_{ncc}} \tag{A.2}$$

We refined our term selection for each topic to improve precision based on this analysis .[3] The classifier's precision is subject-dependent but was consistently above 80% considering the upper bound defined in (A.1). The classification is particularly effective in the case of *politics* and *pandemic*, where the precision goes above 90%. Table A.1 summarises the bounds on precision achieved by the procedure described above. These results are sufficient to use the classification to support our modeling assumptions.

Table A.1 Per-topic Precision

Topic	**Precision l.b.**	**Precision u.b.**
Sports	76.9	83.2
Politics	87.0	94.4
Music	53.4	84.5
Food	65.5	82.4
Pandemic	76.6	93.1

The average percentage of messages classified is 27.8% for all influencers in the dataset. Considering the final classifier and the analyzed dataset, we automatically flagged about one million posts[4] with at least one topic. Of these, only 6.7% of the posts were flagged with multiple labels, indicating the message dealt with more than one topic. We decided to consider in the rest of the work only influencers

[3] We make the final list of terms available at https://mplanestore.polito.it:5001/sharing/0wD5oU6xr.

[4] 1167963 posts were tagged with at least one label.

for whom it was possible to classify more than a thousand posts in the observed period. At the end of this filtering process, we could keep 237 influencers for whom the average posts' classification percentage is 53.2%.

The dataset used contains a subset of Italian politicians. To check the correctness of the labeling procedure, we checked whether the derived reference topic for all politicians was *politics*. It turned out that two politicians did not have *politics* as reference: Vincenzo De Luca had *pandemic*, and Renata Briano had *food*. However, this is entirely understandable as the latter runs a food blog and the former was known for his firm and frequent statements on the pandemic situation during the COVID-19 pandemic.

Appendix B

Proofs of Theorems 1 and 2

B.1 Proof of Theorem 1

Let us start assuming $k_i(0) > 0 \; \forall i$. In such a case we denote with $i_0 = \arg\max_i k_i(1)$ and with $K := k_{i_0}(1)$. First, we show that the problem:

$$k_i(y_i) - c\,y_i = 0, \quad \text{with } y_i \in [0,1] \;\; \forall i \tag{B.1}$$

admits a solution for any $c \geq K$. Indeed by choosing $c \geq K$ we have that necessarily $k_i(1) \leq k_{i_0}(1) \leq c \cdot 1 \; \forall i$ while $k_i(0) > c \cdot 0 = 0$; therefore a zero $z_i(c)$ must exist for every i. This zero is unique as a consequence of the concavity of $k_i(\cdot)$. The set of zeros $z_i(c)_i$ provides a solution of (B.1). Now to get a solution to the original problem (2.13) we need to show that there exist a c such that $\{z_i(c)\}_i$ are normalized. Observe that for $c = K$ by construction $z_{i_0}(K) = 1$ while $0 < z_i(K) \leq 1$ for $i \neq i_0$, therefore $\sum_i z_i(K) > 1$. Now, due to the monotonicity and concavity of $k_i(\cdot)$, $z_i(c)$ is by construction decreasing with respect to c, moreover $z_i(c) \to 0$ as $c \to \infty$ $\forall i$, therefore since $\sum_i z_i(\cdot)$ is a continuous function of its argument, there will necessarily be a c_0 in correspondence of which $\sum_i z_i(c_0) = 1$. In the case in which $k_i(0) = 0$, observe that 0 is a solution of (B.1) for any c, i.e. $z_i(c) = 0$. Moreover, for any $c \geq K$ a second zero may exist. For example, by construction, $z_{i_0}(K) = \{0,1\}$. Therefore for $c = K$, as before, we can always choose as set of zeros $\{z_i(K)\}_i$, such that $z_i(K) = 0$ if $k_i(0) = 0$, and $i \neq i_0$, $z_{i_0}(K) = 1$. By construction $\sum_i z_i(K) \geq 1$. In particular $\sum_i z_i(K) > 1$ is there exists a i such that $k_i(0) > 0$. In this latter case, by increasing c all the non null zeros decrease, therefore, as before, there will

necessarily be a c_0 in correspondence of which $\sum_i z_i(c_0) = 1$. □

B.2 Proof of Theorem 2

We first show that $||\bar{\pi}^{(1)} - \bar{\pi}^{(2)}||_{L_\infty} = \max_i |\bar{\pi}_i^{(1)} - \bar{\pi}_i^{(2)}| = ||\mathscr{G}(F_1(x,z)) - \mathscr{G}(F_2(x,z))||_{L_\infty} \leq M||F_1(x) - F_2(x)||_{L_\infty}$; then we show that we can always enforce: $||F_1(x,z) - F_2(x,z)||_{L_\infty} = ||\mathscr{H}(\bar{\pi}^{(1)}) - \mathscr{H}(\bar{\pi}^{(2)})|| \leq 1/(2M)||\bar{\pi}^{(1)} - \bar{\pi}^{(2)}||_{L_\infty}$ by properly choosing $\omega(\cdot,\cdot)$ and $\theta(\cdot)$. Therefore, we can conclude that $||\mathscr{H} \circ \mathscr{G}(F_1(x,z)) - \mathscr{H} \circ \mathscr{G}(F_2(x,z))|| \leq 1/(2M)||(\mathscr{G}(F_1(x)) - (\mathscr{G}(F_2(x))|| \leq M/(2M)||F_1(x,z) - F_2(x,z)|| = 1/2||F_1(x,z) - F_2(x,z)||$.

First note that $||F_1(x,z) - F_2(x,z)||_{L_\infty} = \sup_x |F_1(x,z) - F_2(x,z)|$ coincides with the Kolmogorov distance between the two distributions.

Let us denote with

$$k_i(y, F_1(x,z)) = \lambda f^{(i)} \int\int \theta(|x - x_i|)\rho(\bar{\pi}_i, |x - x_i|)\mathrm{d}F_1(x,z),$$

and similarly for $k_i(y, F_2(x,z))$ we assume that:

$$\sup_{y \in [0,1], i} |k_i(y, F_1(x,z)) - k_i(y, F_2(x,z))| := \Delta K(F_1, F_2) \leq a||F_1(x,z) - F_2(x,z)||_{L_\infty} \qquad a \in \mathbb{R}^+$$

and

$$\frac{\mathrm{d}k_i(y, F_1(x,z))}{\mathrm{d}y}|_{y=0} < \max_i k_i(1, F_1(x,z)) \qquad \frac{\mathrm{d}k_i(y, F_2(x,z))}{\mathrm{d}y}|_{y=0} < \max_i k_i(1, F_2(x,z)) \; \forall i.$$

Without lack of generality we assume $\max_i k_i(1, F_1(x)) \geq \max_i k_i(1, F_2(x))$. Let the pair $(\bar{\pi}^{(1)} = \{\bar{\pi}_i^{(1)}\}_i, c_1)$ be the solution of

$$k_i(y_i, F_1(x,z)) - c y_i = 0 \quad \text{s.t} \sum_i y_i = 1, y_i \geq 0, \; \forall i$$

now let $(\{\hat{p}_i^{(2)}\}_i)$ the non necessarily normalized solution of

$$k_i(y_i, F_2(x,z)) - c_1 y_i = 0 \quad \text{s.t } y_i \geq 0, \; \forall i.$$

by means of elementary geometric considerations we can bound:

$$|\bar{\pi}_i^{(1)} - \hat{p}_i^{(2)}| \leq \frac{\Delta K(F_1, F_2)}{c_1 - h_1}$$

where $h_1 = \frac{\mathrm{d}k_i(y,F_1(x))}{\mathrm{d}y}\,|_{y=\min(\bar{\pi}_i^{(2)},\widehat{p}_i^{(2)})} \leq \frac{\mathrm{d}k_i(y,F_1(x))}{\mathrm{d}y}\,|_{y=0}$. We recall that by construction (see proof of Theorem 1) we have $c_1 > \max_i k_i(1,F_1(x)))$.

Denoting with $|\widehat{p}^{(2)}| = \sum_i \widehat{p}_i^{(2)}$, we have

$$1 - \sum_i |\widehat{p}_i^{(2)} - \bar{\pi}_i^{(1)}| \leq |\widehat{p}^{(2)}| \leq 1 + \sum_i |\widehat{p}_i^{(2)} - \bar{\pi}_i^{(1)}|$$

Now denoted with $(\{\bar{\pi}_i^{(2)}\}_i, c_2)$ the solution of

$$k_i(y_i, F_2(x,z)) - c\,y_i = 0 \quad \text{s.t} \sum_i y_i = 1, y_i \geq 0,\ \forall i$$

we have, by construction, that:

$$\frac{1}{\max(1,|\widehat{p}^{(2)}|)} < \frac{c_1}{c_2} < \frac{1}{\min(1,|\widehat{p}^{(2)}|)}$$

and therefore, exploiting again elementary geometrical arguments, we can bound:

$$|\widehat{p}_i^{(2)} - \bar{\pi}_i^{(2)}| \leq \left| \left(\frac{c_1 - h_2}{c_2 - h_2} - 1 \right) \widehat{p}_i^{(2)} \right|$$

where $h_2 = \frac{dk_i(y,F_2(x,z))}{dy}\,|_{y=\min(\widehat{p}_i^{(2)},\bar{\pi}_i^{(2)})} = \frac{dk_i(y,F_2(x,z))}{dy}\,|_{y=0}$. Putting everything together, we have proved that:

$$\max_i |||\bar{\pi}_i^{(1)} - \bar{\pi}_i^{(2)}|| = ||\mathscr{G}(F_1(x,z)) - \mathscr{G}(F_2(x,z))||_{L_\infty} \leq M||F_1(x,z) - F_2(x,z)||_{L_\infty}$$

To conclude the proof, first note that by properly choosing $\rho(\cdot,\cdot)$ and $\theta(\cdot)$ we can assume $v_x(x,z)$ and $\sigma_x^2(x,z)$ to depend sufficiently smoothly on $\bar{\pi}$, i.e. $\forall \varepsilon > 0$ we can assume:

$$\sup_x \left|\left| v_x^{(1)}(x,z) - v_x^{(2)}(x,z) \right|\right|_{L_\infty} \leq \varepsilon ||\bar{\pi}^{(1)} - \bar{\pi}^{(2)}||_{L_\infty} \qquad \forall z,$$

$$\sup_x \left|\left| \sigma_x^{2,(1)}(x,z) - \sigma_x^{2,(2)}(x,z) \right|\right|_{L_\infty} \leq \varepsilon ||\bar{\pi}^{(1)} - \bar{\pi}^{(2)}||_{L_\infty} \qquad \forall z,$$

and

$$\sup_x \left|\left| \frac{\partial \sigma_x^{2,(1)}(x,z)}{\partial x} - \frac{\partial \sigma_x^{2,(2)}(x,z)}{\partial x} \right|\right|_{L_\infty} \leq \varepsilon ||\bar{\pi}^{(1)} - \bar{\pi}^{(2)}||_{L_\infty} \qquad \forall z.$$

Then observe that the solution of the Fokker-Planck equation given in (2.10) on a compact interval (and so also its primitive) depends smoothly on function $v_x(x,z)$, function $\sigma_x^2(x,z)$ and its first derivative, as long as $\inf_{x,z}\sigma_x^2(x,z)$ is bounded away from zero. As a final remark note that the set of weakly-increasing functions $F(x)$, such that $F(a)=0$ and $F(b)=1$ equipped with the L_∞-norm forms a closed set in a complete metric space. □

Appendix C

Sensitivity Analysis of the Case Study

For the *Case Study* in Section 2.7.2, we were able to measure or define some characteristic parameters (e.g., $f^{(i)}$ and $x_{Endgovernment}$). However, others were unknown, and even if there are methods to measure, for example, user's opinions (see Section 2.1), this is far beyond the scope of the present work. We could not measure the updating weights α, β, the *feedback* function θ, and the users' opinion on $Politics$. Nevertheless, it is sufficient to assume that a) Conte and Salvini have opposing views over *Politics*, i.e., $x^{(Salvini)} < 0.5$ and $x^{(Conte)} > 0.5$ and that b) Salvini has a more extreme viewpoint, i.e., $|x^{(Salvini)} - 0.5| > |x^{(Conte)} - 0.5|$ c) the population is not too *volatile* in their viewpoints, i.e., the *degree of stubbornness* is not too low c) the *feedback* function θ decreases quickly with the opinion distance (exponential), motivated by the fact that about *politics* people do not easily like viewpoints too far from theirs, we were able to obtain the results in Figure 2.18. These are obtained with the parameters specified in Table 2.3.

To justify our choices, we run a *sensitivity analysis* on the parameters we did not measure, defining the scenarios in Table C.1. Figure C.1 presents the results. It is evident that the increase in the normalized popularity π_{Conte} does not depend on these parameters and is due to the timing (see Fig. 2.17) and the skewness in the population distribution, see Figure 2.16. Note that in Figure C.1 there is a negative *offset* that tends to underestimate Conte's normalized popularity. This is because we use the measured rate $f^{(Conte)} = 0.108$, which puts him at quite a disadvantage compared to Salvini, who has a much higher publication frequency.

Table C.1 Definition of the scenarios for the sensitivity analysis

Scenario	α	β	θ param.	$x_0^{(Conte)}$	$x_0^{(Salvini)}$
1	0.25	0.708	8.0	0.74	0.01
2	0.25	0.708	8.0	0.74	0.05
3	0.25	0.708	8.0	0.78	0.01
4	0.25	0.708	8.0	0.78	0.05
5	0.45	0.475	8.0	0.74	0.01
6	0.45	0.475	8.0	0.74	0.05
7	0.45	0.475	8.0	0.78	0.01
8	0.45	0.475	8.0	0.78	0.05
9	0.25	0.708	8.25	0.74	0.01
10	0.25	0.708	8.25	0.74	0.05
11	0.25	0.708	8.25	0.78	0.01
12	0.25	0.708	8.25	0.78	0.05
13	0.45	0.475	8.25	0.74	0.01
14	0.45	0.475	8.25	0.74	0.05
15	0.45	0.475	8.25	0.78	0.01
16	0.45	0.475	8.25	0.78	0.05
17	0.25	0.708	8.5	0.74	0.01
18	0.25	0.708	8.5	0.74	0.05
19	0.25	0.708	8.5	0.78	0.01
20	0.25	0.708	8.5	0.78	0.05
21	0.45	0.475	8.5	0.74	0.01
22	0.45	0.475	8.5	0.74	0.05
23	0.45	0.475	8.5	0.78	0.01
24	0.45	0.475	8.5	0.78	0.05

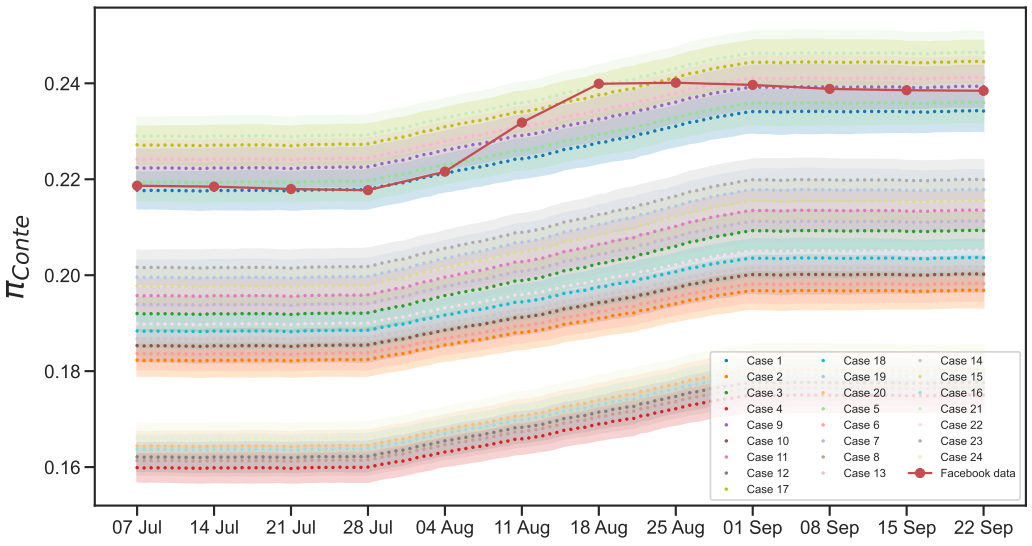


Fig. C.1 Sensitivity analysis for the Figure 2.18 considering the 24 different scenarios (in terms of parameters set) defined in Table C.1.

Appendix D

Data-Driven Modeling Framework

D.1 Modeling Population Heterogeneity

D.1.1 A Data-Driven Approach

The model proposed in this work assumes the knowledge of some characteristics of the population involved in the epidemiological process. At the country level, populations present notable differences in their features, for instance, in terms of age distribution, overall health condition, or the daily contacts among individuals, which in turn depend on the country's customs and, more broadly, on the country's wealth and employment conditions. The heterogeneity in the population may play a role in the disease transmission process. Indeed, different contact patterns among individuals may result in a faster virus outbreak. To account for this heterogeneity, we characterized the population through a distribution function $f_{r,p}$ (equivalently $f_{r,a}$, where a is the age, assuming a one-to-one correspondence between the age a and p, see Section D.1.4) whose parameters represent the average daily contacts of an individual with other individuals in the population r and the probability of dying p as a consequence of the contraction of COVID-19. The parameters on which these variables depend are numerous (i.e., gender, occupation, individual habits, and pre-existing medical conditions, to name a few); a precise population characterization is outside this work's scope. We considered the individual's age as the main factor on which the average daily contacts r and the probability of death p depend. Thus, we stratified the population by age and

derived country-specific distributions exploiting the contacts patterns presented in [90] and the case fatality rate from [133], which have both been obtained for various age classes. The construction procedure of the function $f_{r,p}$ is discussed in depth in the following paragraphs.

D.1.2 Chance of Death due to COVID-19

One of the characterizing parameters of our population is the probability of dying after infection p, for which it is crucial to have an estimation. To this end, we considered the Case Fatality Rate (CFR); see Table I for the values we used, which refer to Italy. The CFR represents the proportion of deaths due to a specific disease compared to the total number of people diagnosed with that disease in a certain period. The data we leveraged come from [133] in which CFR values have been obtained for 10-year-wide age classes considering Italian COVID-19 epidemiological data by 18th August 2020.

An observation at this point is in order, the case fatality rate refers to *confirmed* cases. In contrast, another popular indicator: the Infection Fatality Rate (IFR), considers the *actual* number of infected cases. Since COVID-19 can give rise to infections with mild or no symptoms, the number of actual cases might be significantly higher than the confirmed cases. For these reasons, it has been debated that the CFR overestimates the probability of death. However, the actual number of cases is unavailable, and the studies considering the IFR relied on estimations. In this direction, it is worth mentioning the work of Ghisolfi et al. [134] in which values for the IFR have been extrapolated accounting for age, gender, comorbidities, and health system capacity by discriminating between low-income and high-income regions. Considering the controversy around the CFR and IFR indicators, we decided to employ the CFR values to obtain an estimate for the mortality probability.

Moreover, for the characterization of our synthetic population, we have not considered other comorbidities (i.e., seropositive status, pre-existing medical conditions) nor gender differences (see [134] for such differentiated values). Again, the scope of this work is not to faithfully describe the population but to capture some quantitative differences in population classes and retain the significant traits of a heterogeneous population. Many studies confirm the strong correlation

between age and CFR (or IFR) [133][134][135]. The latter study considers data from 25th January to 10th December 2020 for the State of Victoria, Australia. In this period, the state experienced two waves of the virus, and by the end of the data series, the infection was eradicated from the Victorian population [135]. This observation made the writers conclude that their estimate of the CFR is not spoiled by the underestimation bias, which is customary for an ongoing outbreak. We report the observed CFR data in Table II. Even though the values for the CFR vary slightly among the cited studies, the trend is clear: older individuals are more *fragile*, and the CFR drastically increases with age.

Age Class	CFR
0-19	0.1
20-29	0.1
30-39	0.3
40-49	0.9
50-59	2.8
60-69	10.9
70-79	26.7
≥ 80	34.6

Table I - Italy

Age Class	CFR
0-9	0
10-19	0
20-29	0.02
30-39	0.06
40-49	0.04
50-59	0.63
60-69	2.16
70-79	14.41
80-89	31.90
≥ 90	40.03

Table II - Australia

The data from [133] have been assumed as representative of the mortality probability for all the countries considered for the $f_{r,p}$ distribution construction. It is a simplification justified by the absence of an extensive and precise study concerning the case fatality rate at a country level. Figure D.1 represents these data on the Case Fatality Rate as a function of age. The linear fitting of the empirical points is also shown in the figure, which allows the expansion of the number of age classes considered in the synthetic population.

D.1.3 Daily Number of Contacts

Another crucial piece of information for describing a population in epidemiological terms is the daily number of contacts of an individual. Indeed, an infection can occur when an infected individual encounters a healthy individual. The more

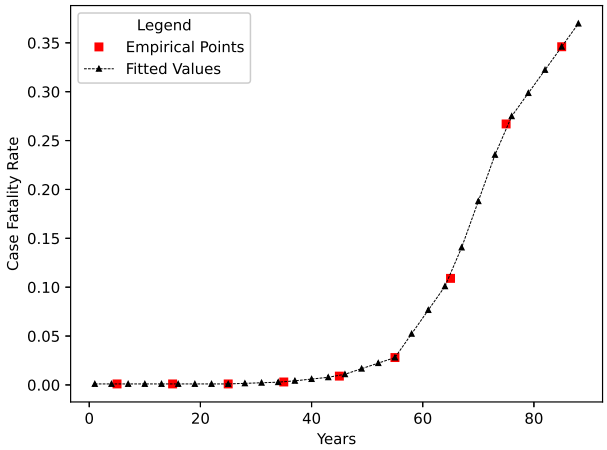


Fig. D.1 Values for the Case Fatality Rate for the various age classes from [133]. The graph shows the empirical points (in red) and the linearly fitted values of CFR in order to expand the number of available age classes for our population. The strong correlation between age and the probability of dying is well depicted by this graph.

individual-to-individual contacts, the higher the chance of an infection. In [90], age-specific contact matrices have been generated from detailed census and survey data on key demographic features for 35 countries. Four different settings are considered (household, school, workplace, and community), producing as many specific contact matrices. A linear combination with appropriate weights of these matrices produces an overall contact matrix $\{M_{ij}\}$ which indicates the daily per capita number of contacts an individual of age i has with individuals of age j (the interested reader is referred to [90] for more details). For this work, the entire contact matrix M is not necessary. We are only interested in the average number of contacts of an individual of age i regardless of the age of the individual encountered. This value is easily computed from M by summing over the columns of the matrix: $\mathbb{E}[r_i] = \sum_j M_{ij}$. In [90] for each country considered in the study, the number of people belonging to 1-year-wide age classes is also provided, allowing for the definition of the age distribution $d_a(\text{age})$ of the population. Figure D.2 reports the age distribution for Italy truncated at 84 years, together with the average daily contacts for each age class.

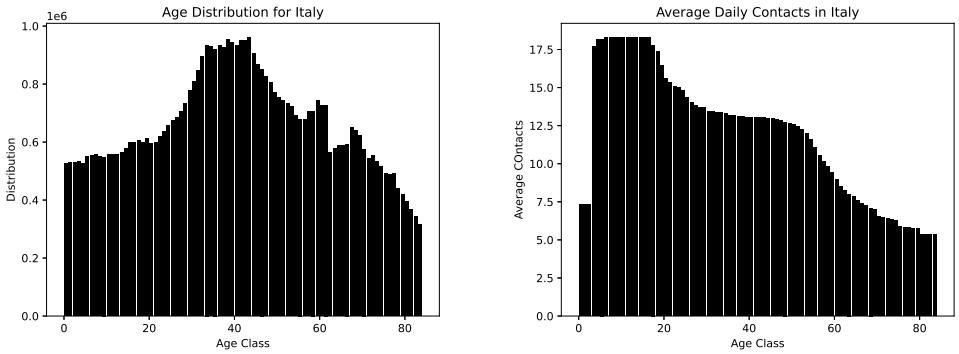


Fig. D.2 Characterization of the Italian population in terms of the age distribution and the daily average contacts among individuals as reported in [90].

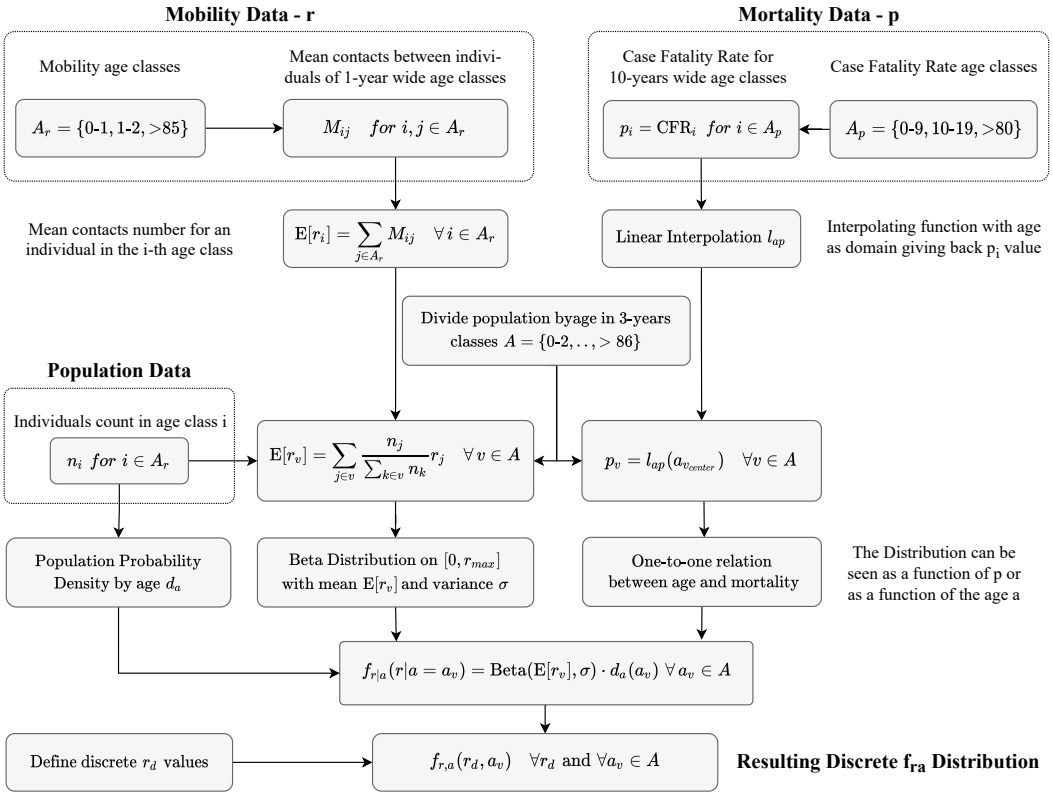
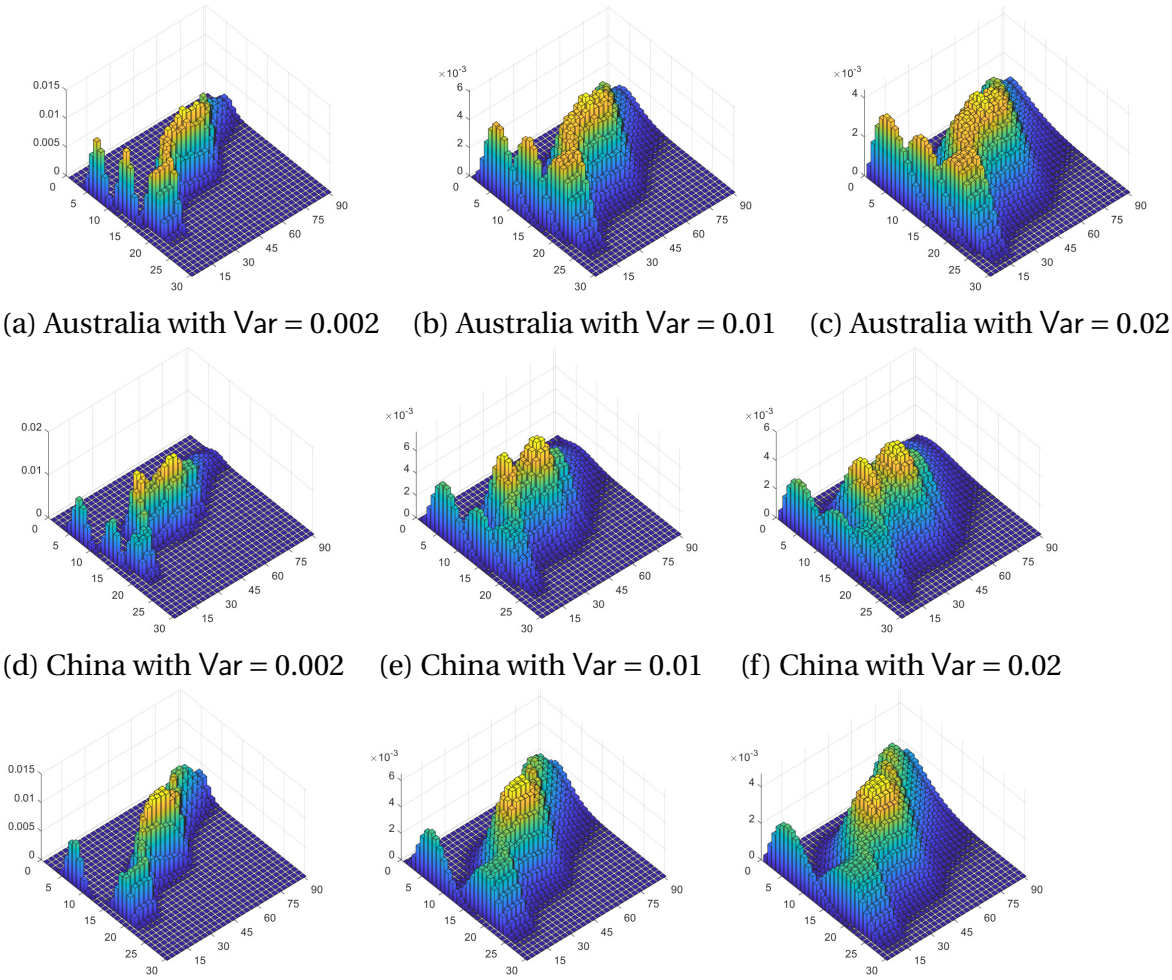


Fig. D.3 Pictorial representation of the construction's procedure for the synthetic distribution $f_{r,a}$. Starting from empirical data a discrete $f_{r,a}$ distribution is obtained for arbitrary values of the risk exposure r and age classes a.



(a) Australia with Var = 0.002 (b) Australia with Var = 0.01 (c) Australia with Var = 0.02

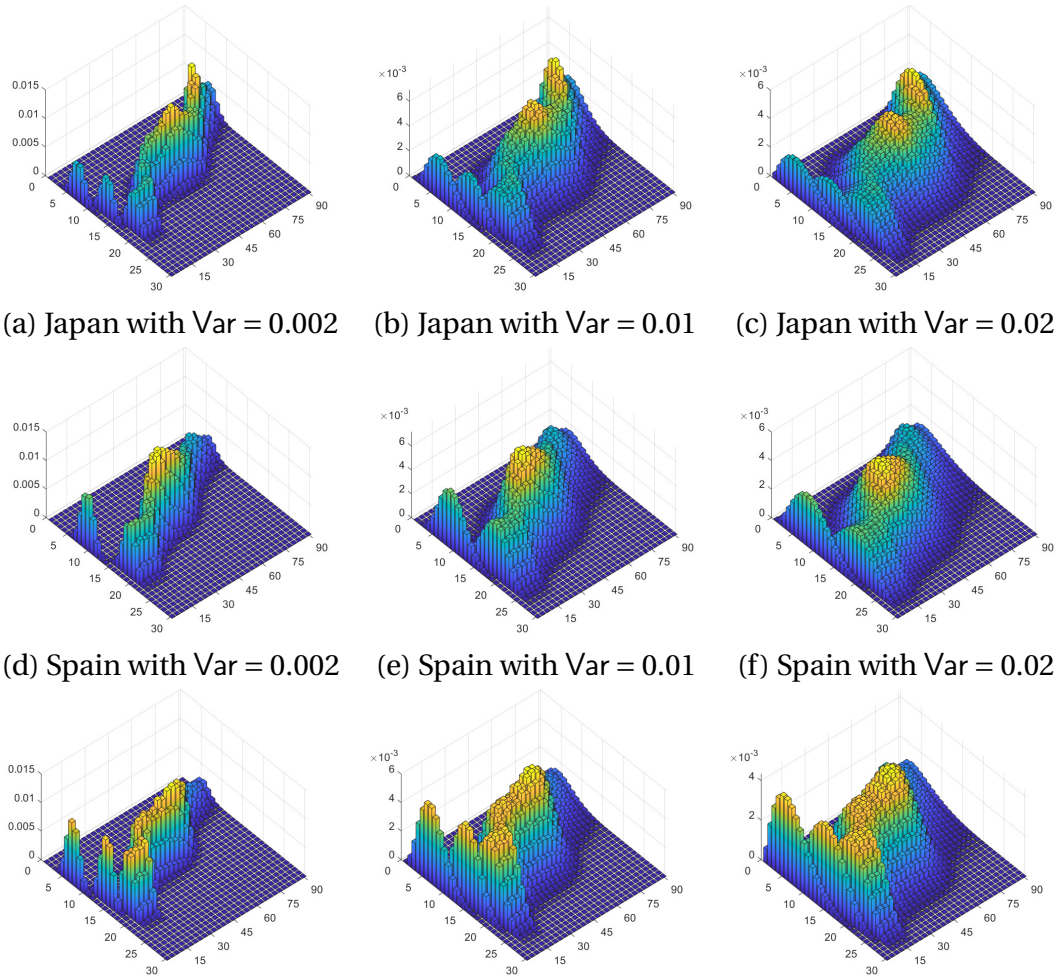
(d) China with Var = 0.002 (e) China with Var = 0.01 (f) China with Var = 0.02

(g) Italy with Var = 0.002 (h) Italy with Var = 0.01 (i) Italy with Var = 0.02

Fig. D.4 $f_{r,a}$ distributions for the countries that have been considered: Australia, China, Italy. The age bins are $A = \{0\text{-}2, 3\text{-}5, .., \geq 87\}$, and the daily number of contacts $r \in [0, r_{max} = 30]$.

D.1.4 Construction of the Distribution

The quantities reported in the previous paragraphs are the essential ingredients for constructing the $f_{r,p}$ distribution function. Indeed, we start from the knowledge of the age distribution d_a, the average daily contacts $\mathbb{E}[r_a]$, and the case fatality rate p_a, both given for certain age classes a. The flowchart reported in Figure D.3 exemplifies the procedure followed to construct the $f_{r,p}$ distribution, highlighting at what point we used which data. First, we partitioned the population into uniform age classes, each with a width equal to three years $A = \{0-2, .., \geq 87\}$. It



(a) Japan with Var = 0.002 (b) Japan with Var = 0.01 (c) Japan with Var = 0.02

(d) Spain with Var = 0.002 (e) Spain with Var = 0.01 (f) Spain with Var = 0.02

(g) USA with Var = 0.002 (h) USA with Var = 0.01 (i) USA with Var = 0.02

Fig. D.5 $f_{r,a}$ distributions for the countries that have been considered: Japan, Spain and the United States. The age bins are $A = \{0\text{-}2, 3\text{-}5, .., \geq 87\}$, and the daily number of contacts $r \in [0, r_{max} = 30]$.

must be observed that the raw data refer to age bins of different sizes: one-year width for the average contacts as well as the population's composition and ten-year width for the case fatality rate. They must be manipulated to refer to the same age classes $v \in A$. Thus, we averaged the mean daily contacts, weighting the average value by the relative frequency of the sub-bin of the wider class v. Starting from $\mathbb{E}[r_i]$, for $i \in A_r = \{0-1, 1-2, .., > 85\}$ we obtained:

$$\mathbb{E}[r_v] = \sum_{j \in v} \frac{n_j}{\sum_{k \in v} n_k} r_j \quad \forall v \in A = \{0-2, .., \geq 87\} \tag{D.1}$$

where n_j indicates the number of individuals in the age class $j \in A_r$. Similarly, we have to adapt the CFR defined for ten-year wide classes $A_p = \{0-9, .., > 80\}$ to the new three-year bins. For this purpose, the empirical values of the CFR have been linearly interpolated. This allows finding a CFR (p) value for each $v \in A$.

Given that we do not consider individual traits such as gender, ethnicity, or pre-existing medical conditions, we consider the death probability p as a fixed value, given the age class $v \in A$. This value has some variability because of the aspects just discussed and individuals' intrinsic differences. Nevertheless, the mortality probability p does not directly affect the system's dynamical behavior; it just determines the number of individuals entering the "death" D state, so its contribution matters on average. Note that in our setting, it is equivalent to talking about $f_{r,p}$ and $f_{r,a}$ due to the one-to-one relationship between an age class a and the corresponding probability of death p_a. On the other hand, the number of daily contacts r enters directly into the epidemiological system's dynamical rule (see Equations 4.2. Assuming r constant over a specific age interval would be a too substantial simplification. It is known that individuals of the same age have very different mobility patterns due to, for instance, their occupation, their lifestyle, and their medical condition. To account for this variability, we spread the individual's daily contacts around their mean value according to a Beta distribution with a given variance σ. We choose the support of the Beta distribution to be $[0, r_{max}]$. It is clear that $r \geq 0$, at the same time, we decided to fix an upper bound for the number of contacts assuming that just a negligible number of individuals would fall outside this interval. The average value of contacts $\mathbb{E}[r_v]$ fixes one of the Beta distribution parameters. The other parameter characterizing the Beta distribution appears to be a free variable for which we selected different values to obtain sufficiently different variance values. The resulting distributions have been fed as input to the dynamical model giving qualitatively the same results. Since the desired $f_{r,a}$ distribution needs to be discrete. We defined a set of equispaced r values for which we computed the discrete probability by integrating the Beta distribution in the interval of interest. This procedure allows obtaining a discrete bi-dimensional distribution for 3-year-wide age bins (each corresponding to a certain mortality probability p) and equispaced values of r in the interval of definition $[0, r_{max}]$.

D.1.5 Estimation of Years of Life Lost

One of the metrics employed to compare the proposed vaccination strategies consists of evaluating the years of life lost (YLL) due to deaths attributable to the virus. It is a measure of premature mortality that not only considers the chance of death from the virus but also the age of the deceased individuals. In order to perform such an estimation, country-specific demographic data are needed for all countries considered to construct the synthetic populations. The World Health Organization (WHO) collects health-related statistics in the Global Health Observatory [136], gathering information regarding all 194 WHO member states. In particular, it provides the life expectation stratified by age $\mathbb{E}[L_a]$, pivotal to compute the years of life lost. $\mathbb{E}[L_a]$ is provided for 5-year-wide age bins except for the first bin, which is partitioned into the individuals younger and older than one, and the last one, which groups all people older than 85. We made an interpolation to obtain a $\mathbb{E}[L_a]$ value for each of the age classes of our synthetic populations. We linearly interpolated the empirical data of the Global Health Observatory, providing the life expectation for each three-year-wide age class. To present some of the differences that might be present in $\mathbb{E}[L_a]$ considering different countries, we report the WHO's life expectation values for Italy (left) and China (right) in Table III.

D.2 Disease Specific Parameters Choice

The parameters specified in Table IV refer to the specific behavior of the COVID-19 pandemic. This section discusses the choice of such parameters, providing the necessary references that justify them.

In our reference scenario, we considered two phases of the pandemic, which is, of course, a first simplification. In the first phase of our scenario, there was more than one variant on the national territory. Indeed, there were several different variants with slightly different basic reproduction numbers. For simplicity, since our work does not aim to represent the pandemic's evolution accurately but rather to provide a data-driven demonstration of the proposed model, we considered only one reproduction number for the period, roughly equivalent to that of the Delta variant but also compatible with earlier variants. The Delta variant was

Age Class	Life Expectation
<1	82.97
1-4	82.20
5-9	78.24
10-14	73.27
15-19	68.30
20-24	63.36
25-29	58.44
30-34	53.52
35-39	48.61
40-44	43.73
45-49	38.90
50-54	34.15
55-59	29.53
60-64	25.04
65-69	20.73
70-74	16.63
75-79	12.80
80-84	9.24
≥ 85	6.21

Age Class	Life Expectation
<1	77.43
1-4	76.96
5-9	73.04
10-14	68.11
15-19	63.18
20-24	58.28
25-29	53.44
30-34	48.60
35-39	43.79
40-44	39.05
45-49	34.39
50-54	29.79
55-59	25.33
60-64	21.06
65-69	17.01
70-74	13.24
75-79	9.95
80-84	7.06
≥ 85	4.76

Table III: Life expectation - Italy (left) and China (right).

first identified in India in October 2020 and has a basic reproduction number that ranges from 3.2 to 8 [137]. Indeed, the value chosen for strain 1 is not far from the values of the basic reproduction number in the early outbreak, which averaged 4.22 when considering several European countries, and was estimated to be 6.33 in Germany and 5.88 in the Netherlands, while in Italy it was 4.25 [138]. The value of 6 chosen for strain 1 is roughly between the values for the early outbreak and the later Delta variant. For strain 2, "The Omicron variant has an average basic reproduction number of 9.5 and a range from 5.5 to 24" [138]. These values motivate us to choose a basic reproduction number of 12 for the second strain in our reference scenario since its occurrence corresponds approximately to the onset of the Omicron variant in Italy.

Regarding the average length of stay in the different states, it is reported in [139] that the time spent in the ICU was 18.4 days (before 25th March) and 15.4 days (after 7th April), which is why we chose 16 days. This choice is also consistent with

the study in [140], which examined several studies and found that the median length of total hospital stay (Length of Stay) ranged from 5 to 29 days. "Most studies (43/52) reported LoS for total hospitalization only, with four studies reporting LoS for ICU only, and five studies reporting both." [140]. Regarding the length of stay in the ICU in this study, the median ranged from 5 to 19 days [140]. We used an average sojourn time of 16 days for both hospitalizations and ICU. Table 4 of [141] reports the infectious period (IP) for symptomatic cases from several studies, ranging from 3 to 20 days. The infectious period seems to be about a week. We set it at eight days in our scenario.

On the website of the Italian Government [142], a report regarding vaccinations is available, as of 17th July 2023, 90,25% of the population over 12 has completed the vaccinal cycle, justifying the 10% of no-vax individuals considered in our scenario.

To what concerns the mortality reduction of strain 2 compared to strain 1 and the mortality reduction associated with acquired/natural immunity: In [143], it is found that in South Africa, the infection fatality ratio was reduced by 78.7%, which approximately corresponds to a factor 5 decrease, as used in our scenario. For the reduction in mortality, we assumed the effect of vaccination and natural immunity are comparable. In [144] is reported that the COVID-19-related mortality of the Pooled Vaccine Effectiveness (PVE) was 92%, corresponding to a hazard ratio of 0.08, which we optimistically associated with a factor of reduction of 20.

As for the number of "regular" hospital and intensive care beds, we rely on the ISTAT (Italian Institute of Statistics) report [145]. The number of "regular" beds is not easy to determine since there are different types of hospital beds, depending on the expected length of stay and the medical service to be provided. This number has declined over the last 30 years and is now about 200000 units. We took a conservative approach in our analysis and assumed $\hat{H} = 50000$ for two reasons: First, the above number concerns all types of hospital beds; second, not all beds can be reserved for COVID-19 patients; we considered a quarter of the total number a reasonable choice. Regarding ICU, the report mentioned earlier [145] states that the number of ICU beds is, on average, 15.1 for every 100000 individuals, which is about 10000 beds. Note that while we have set $\hat{T}$ to 20000, we have considered $T_{max} = 10000$ in most of our simulations. We argue that it was also interesting to investigate scenarios with a larger capacity.

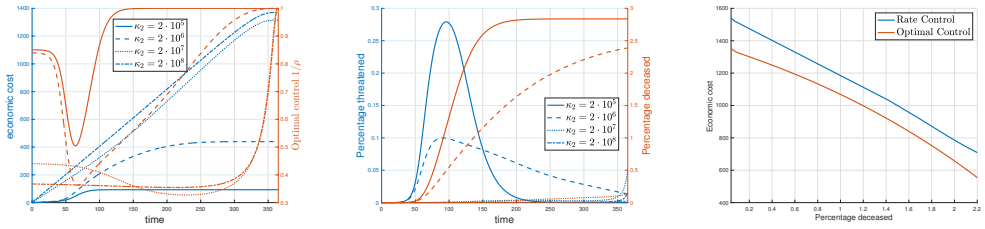
Appendix E

Comparative Analysis with Optimal Control

E.1 Our Study and the Optimal Control Framework

Our study takes a different perspective than the optimal control approach. We deliberately examine *simple* control strategies to provide practical insights and guidelines for decision-makers who may not have access to sophisticated optimization techniques or detailed knowledge of the underlying epidemic mathematical laws (e.g., parameters).

Optimal control approaches in epidemiology have certain drawbacks worth mentioning. These approaches often lack closed-form analytical solutions, requiring numerical methods for their implementation [146, 147]. While numerical solutions can provide valuable insights, they can be challenging to interpret and translate into practical control measures. This lack of interpretability hampers the ability to univocally understand the implications and consequences of the obtained optimal control strategies. Another drawback is that optimal control solutions may strongly depend on parameters, leading to abrupt changes in the optimal interventions. These "phase transitions" can make implementing and managing the control measures in practice difficult. Minor changes in the parameters or system conditions may result in significant shifts in the optimal strategies. Furthermore, optimal control approaches often rely on detailed and precise knowledge of the system dynamics, including accurate parameter values



(a) (b) (c)

Fig. E.1 (a) Economic cost and control effort via optimal control (b) Percentage of threatened and deceased individuals via optimal control implementation. (c) Comparison of optimal control and Rate control strategies

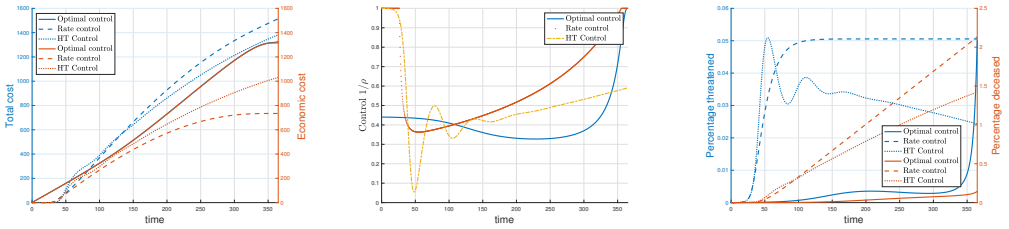
and functional forms of the underlying equations. In real-world scenarios, such detailed information may be unavailable or subject to significant uncertainties. In the same spirit as [128], we will mainly focus on two control feedback strategies based on controlling the rate of new infections or maintaining the occupation of healthcare facilities below a given level, and we evaluate the economic cost of non-pharmaceutical interventions and the social cost in terms of number of deaths. For the sake of simplicity and analytical tractability, we consider an "ideal" scenario in which the system operates near the equilibrium point, where the effective reproduction number equals one (we provide local stability results). This regimen appears to be a desirable condition whereby the number of infected individuals, and thus those requiring intensive treatment, is maintained at a sustainable level, even over long periods, while applying minimal durable mobility restrictions.

E.2 Comparative Analysis

In this section, we perform a comparative analysis of the proposed model against optimal control and homogeneous models to assess its effectiveness and advantages in addressing the research problem.

In this section, we compare the Rate and HT controllers with optimal control during the first phase of the pandemic in which vaccines are still unavailable. All experiments refer to the single-class version of the model in (4.2).

In [148], various government intervention strategies are compared against a specific percentage of the deceased population while employing different control



(a) (b) (c)

Fig. E.2 (a) Comparison of different control strategies: (a) Overall costs and economic costs; (b) Control function. (c) Percentage of threatened and deceased individuals.

policies. We replicate similar experiments and present the numerical solutions obtained via the optimal control approach within a time horizon $t_{\max} = 365$ days. We fix $\kappa_1 = 200$, $\kappa_3 = 20$, and $\zeta = 4$, and we let κ_2 to vary from 10^5 to 10^8.

The considered economic cost corresponds to the teal curve depicted in Figure 4.7. Such cost has been chosen non-convex, on purpose, to put the Rate control strategy in the most challenging conditions (indeed Proposition 3 and 4 do not hold). Observe that, with our choice of parameters, the first term of the objective function in (4.7) is typically small with respect to the third, and therefore the choice of κ_2 becomes fundamental to determine the proper trade-off between economic cost and death/ICU occupancy, where the last two metrics are highly correlated, since by tightly controlling ICU occupancy, we exert tight control on the deaths and vice versa.

The average sojourn time in state I has been set equal to 8 days and that in states H and T to 16 days. Therefore $\gamma = 1/8$, $\phi = \tau = 1/16$. The transition probabilities between compartments I, H, T, and D, satisfying constraint (2), are set for simplicity as follows: $p^{IH} = p^{HT} = \hat{p}^{TD} = p^{1/3}$ with overall mortality rate $p = 0.01$. The assumed value for the basic reproduction number has been fixed to $\mathscr{R}_0 = 3$, as reported in [149], and the healthcare capacity parameter to $3.33 \cdot 10^{-3}$.

In Figure E.1(a), the economic cost (in blue) and the control (in red) are shown as a function of time. For lower values of the parameter ($\kappa_2 \in [10^5, 10^6]$), the control measures are moderate and remain relatively constant for a brief period of approximately 50 days. After this initial phase, the control is tighter, reaching its maximum level of restriction. During this period, stringent measures are implemented to contain the epidemic effectively. Subsequently, as the situation

improves or specific goals are achieved, the control is gradually relaxed, allowing for a more lenient approach to managing the epidemic. This sequential pattern of moderate-tightened-relaxed control measures aims to strike a balance between mitigating the spread of the disease and minimizing the socioeconomic impact on the population. The stringent initial intervention effectively disrupts the early exponential growth of the epidemic, leading to a dampened peak number of infections. In both previous cases at the end of the observation window, i.e., for $t \approx t_{\max}$, the population reaches herd immunity. As we increase the value of κ_2, as expected, the control becomes more stringent and is kept constant for most of the time (around 350 days). At the same time, the healthcare system experiences less stress, and the number of deaths decreases at the expense of higher economic costs. Notably, when setting κ_2 above $2 \cdot 10^7$, we can confidently guarantee that the peak number of patients requiring intensive care remains below around 30000 (see Figure E.1(b) where blue curves refer to ICU occupancy and red curves to cumulative deaths). This behavior highlights the importance of appropriately calibrating control parameters to achieve optimal outcomes in managing the epidemic and preventing overwhelming pressure on the healthcare infrastructure. Observe that at the end of the observation window, i.e., for $t \approx t_{\max}$, the control is always completely released, i.e., $\rho(t_{\max}) = 1$. This effect is a by-product of the optimal control approach, which does not account for what happens when $t > t_{\max}$. Indeed, as t approaches $t_{\max}$, releasing the control leads to an instantaneous reduction of the economic cost, while, due to the delay, the resulting increase in ICUs and deaths is negligible (as it will take place after $t_{\max}$).

The analysis in Figure E.1(c) highlights the trade-offs between economic cost and human lives achieved by optimal control and Rate control, respectively. The curves have been obtained by varying parameter $\kappa_2 \in [10^5, 10^8]$ for optimal control, $\lambda \in [1000, 700000]$ for Rate control. We have disregarded the healthcare stress cost, using the total number of deaths as a proxy of it. It should be noted that the optimal control strategy proves to be the most effective, outperforming the rate control strategy. However, if we fix the number of deaths, for example, to 0.2% (i.e., 100000 deaths), the rate control strategy exhibits only a slightly worse economic cost. The difference between the economic cost curves of the two strategies is not substantial, with a modest 7% increase obtained by the rate control strategy. Despite the increase in economic costs resulting from the rate control strategy, the

difference is relatively small, indicating that both strategies remain competitive in managing the epidemic.

It is worth remarking that approximately the same value of the overall objective function in (4.7), which takes into account economic cost, deaths, and healthcare stress in the optimal control strategy, can be achieved through fairly different approaches. Figure E.2 provides a comparison of the optimal control strategy with $\kappa_2 = 2 \cdot 10^7$, the Rate control with $\lambda = 120000$, and the linear HT control with $\rho_H(H) = \min\left(15, \frac{H_0}{H_0 - H}\right)$ and $\rho_T(T) = \min\left(15, \frac{T_0}{T_0 - T}\right)$, with and $H_0 = 480000$ and $T_0 = 300000$.

While the overall cost for the three strategies is approximately the same, the different components of the cost are significantly different. For what concerns the economic cost, the optimal control strategy appears to be the least favorable, resulting in the highest economic burden compared to the other strategies (the economic cost for the optimal control strategy is hardly distinguishable from the overall cost). On the other hand, the rate control strategy is the most efficient in minimizing economic costs, offering a more economically sustainable approach. The HT control strategy falls in an intermediate position, achieving a balance between cost-effectiveness and epidemic management. Regarding the number of deaths, the optimal control strategy demonstrates its strength, resulting in the lowest fatality rate among the three strategies. It effectively minimizes the loss of life during the epidemic. Conversely, the rate control strategy shows the highest number of deaths, indicating that this approach is less effective in preventing fatalities. The HT control strategy lies in between, offering an intermediate level of protection against the loss of life compared to the other two strategies. In conclusion, adopting the overall cost as the unique driver for the choice of $\rho(t)$ turns out to be not particularly appealing to decision-makers because it does not allow them to exert direct control on the different components of the cost.

Appendix F

Miscellaneous

F.1 Model Extension - Adding Direct Transitions

To extend the base model, it is possible to consider that infected people die without being hospitalized or undergoing intensive treatment. This requires specifying the probabilities $p_{r,p}^{ID}$ and $p_{r,p}^{HD}$ associated with direct transitions from state I (state H) to state D, respectively, representing the occurrence of premature death events. Previous probabilities $p_{r,p}^{IH}$ and $p_{r,p}^{HT}$ are then redefined as transition probabilities conditioned to the event that such premature deaths do not occur. The new model is depicted in Figure F.1 and the modified system is:

$$
\begin{aligned}
\dot{S}_{r,p}(t) &= -\sigma(t)\left(\sum_{r',p'} r' I_{r',p'}(t)\right)\frac{r S_{r,p}(t)}{\sum_{r',p'} r' N f_{r',p'}} + \mu M(t) \\
\dot{I}_{r,p}(t) &= \sigma(t)\left(\sum_{r',p'} r' I_{r',p'}(t)\right)\frac{r S_{r,p}(t)}{\sum_{r',p'} r' N f_{r',p'}} - \gamma I_{r,p}(t) \\
\dot{H}_{r,p}(t) &= \gamma(1-p_{r,p}^{ID})p_{r,p}^{IH} I_{r,p}(t) - \phi H_{r,p}(t) \\
\dot{T}_{r,p}(t) &= \phi(1-p_{r,p}^{HD})p_{r,p}^{HT} H_{r,p}(t) - \tau T_{r,p}(t) \\
\dot{D}_{r,p}(t) &= \tau p_{r,p}^{TD}(t) T_{r,p}(t) + \gamma p_{r,p}^{ID} I_{r,p}(t) + \phi p_{r,p}^{HD} H_{r,p}(t) \\
\dot{M}_{r,p}(t) &= \gamma(1-p_{r,p}^{ID})(1-p_{r,p}^{IH}) I_{r,p}(t) \\
&\quad + \phi(1-p_{r,p}^{HD})(1-p_{r,p}^{HT}) H_{r,p}(t) \\
&\quad + \tau(1-p_{r,p}^{TD}(t)) T_{r,p}(t) \quad - \mu M(t)
\end{aligned}
\tag{F.1}
$$

Note that in this case, we must ensure that the overall death probability p satisfies (for $T(t) \leq \widehat{T}$)

$$p_{r,p}^{ID} + (1 - p_{r,p}^{ID}) p_{r,p}^{IH} \left[p_{r,p}^{HD} + (1 - p_{r,p}^{HD}) p_{r,p}^{HT} \hat{p}_{r,p}^{TD} \right] = p \tag{F.2}$$

which replaces (4.3).

As a final remark, we point out that while this extension is straightforward, it introduces the need for additional parameters that could be difficult to be determined. Moreover, models such as the SIDARTHE [94], discussed above and explicitly developed for COVID-19, do not consider such transitions. Lastly, there is evidence that the majority of deaths due to COVID-19 have occurred in hospitals or care homes [150]. Therefore, we employed the first version of the model in our simulations.

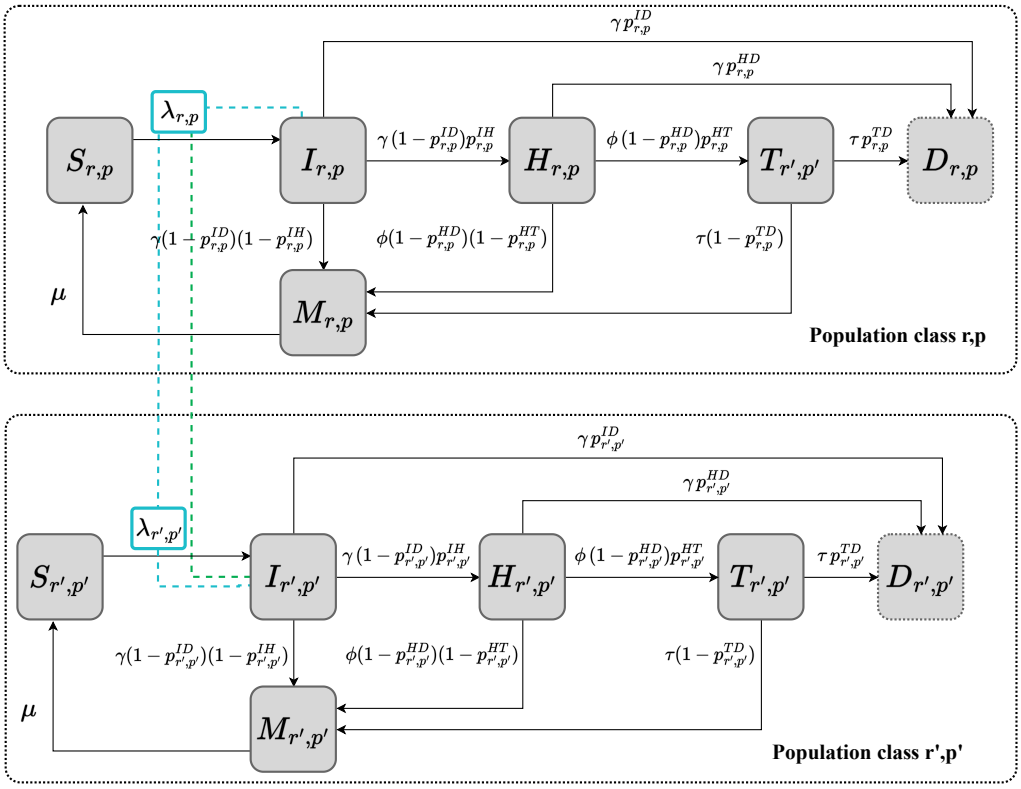


Fig. F.1 Schematic representation of the proposed model without vaccination adding transitions from the *Infected* (I) and *Hospitalized* stets to the *Death* (D) state, as detailed by the system of equations (F.1). This model is rather similar to that in Figure 4.4, only two transitions for each $\{r, p\}$ individual's class have been added.

F.2 Motivating Heterogeneity - Rate Control

F.2.1 Comparison with a Single-Class Model

In this section, we question the importance of introducing in the model a stratification based on distribution $f_{r,p}$, given that, at least when $S_{r,p}(t) \approx N f_{r,p}$, our multi-class (stratified) model is equivalent to a single-class (non-stratified) model with a transmission rate modified by a factor $\beta = \frac{\sum_{r,p} r^2 f_{r,p}}{\sum_{r,p} r f_{r,p}}$.

We now show that, besides being necessary to evaluate prioritized vaccination strategies, our stratification is fundamental also to compute the cost of pure non-pharmaceutical interventions. To this purpose, we consider a simple SIR model under perfect control of new infections, i.e., where $\rho(t) = \max\left\{1, \frac{\lambda_U(t-\varepsilon)}{\lambda}\right\}$ with $\varepsilon = 1/100$ day.

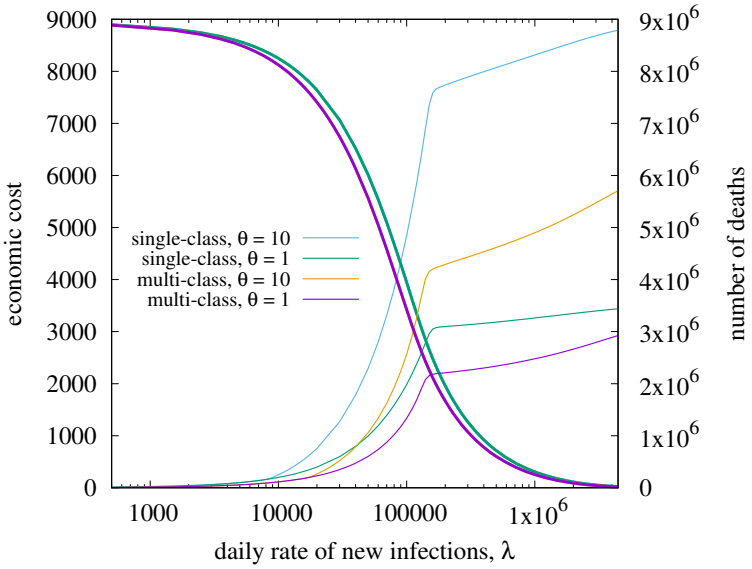


Fig. F.2 Comparison between single-class and multi-class models, and impact of parameter θ, on economic and social costs, as a function of controlled rate λ of new infections.

In Fig. F.2, we compare the economic and social costs derived by our model with those computed by a single-class model in which the contact rate and mortality of all individuals are set equal to mean values $\mathbb{E}[r]$ and $\mathbb{E}[p]$, respectively. In this experiment, the time horizon is $t_{\max} = 1$ year.

We also compare the case in which saturation of intensive therapy capacity does not affect mortality ($\theta = 1$) and the case where mortality is severely increased

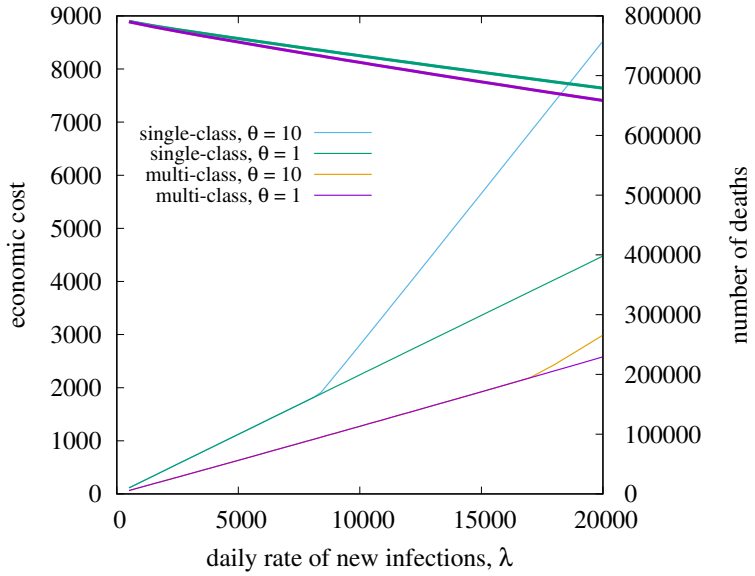


Fig. F.3 Zoom of Fig. F.2 on the more reasonable control regime of small λ.

by ICU saturation ($\theta = 10$). Fig. F.2 uses a log horizontal scale to encompass also large values of λ, corresponding to an almost uncontrolled epidemics, i.e., the attempt to quickly reach the naturally-acquired herd immunity (which occurs at the 'knees' of the curves for the number of deaths). Notice that such a strategy to achieve herd immunity produces a dramatic number of deaths (2 million out of a population of 60 million) even in the most favorable case (multi-class model with $\theta = 1$). We remark that the case fatality rate varies by time and location, and its measurement is affected by well-known biases exacerbated during the COVID-19 pandemic. Hence, the mortality in the numerical results might be overestimated, as the case fatality rate has been retrieved from data statistics that include both deaths *due to* or *with* COVID-19.

Fig. F.3 reports identical results, but zooming in on the more reasonable control strategy in which small λ values are enforced. The economic cost (left y axes) is approximately the same under single/multi-class models. This observation depends on the fact that, under the considered distribution $f_{r,p}$, $\mathbb{E}[r] = 12.1$ and $\mathbb{E}[r^2]/\mathbb{E}[r] = 13.5$, are very close to each other. Moreover, the economic cost does not depend on θ, as expected. In contrast, the predicted number of deaths (right y axes) is highly diverse across different models and considered values of θ. Even with $\theta = 1$, the single-class model predicts a much larger number of deaths. This can be explained by the fact that under any realistic distribution $f_{r,p}$ (see Fig.

D.4 and Fig. D.5), contact rate r and mortality p are negatively correlated, such that most vulnerable individuals (elderly) have lower contact rate. Therefore, the single-class model, in which all individuals are identical, is more pessimistic in terms of deaths. More significant discrepancies between single- and multi-class models are observed with $\theta = 10$, since after saturation of ICU, occurring at the "bifurcation" points appearing in Fig. F.3, saturation effect of mortality probability (Eq. (4.4)) occurs in the multi-class scenario for some disadvantaged classes, but not in the single-class model.

F.3 Additional Proofs

Here, we present the proofs of Corollary 2 and 3, for completeness, we report again the statement of the corollaries. The two corollaries explore two interesting cases of delay distributions.

Corollary 4 (Exponential delay distribution)**.** *If* $\mathfrak{f}_d(\tau) = u(\tau)\delta e^{-\delta(\tau)}$*, then the system is always (locally) stable.*

Proof. Consider an exponential delay distribution of parameter δ (mean $1/\delta$). We obtain

$$\mathscr{L}\{\eta(t)\} = \frac{\eta(0)}{s + \frac{\gamma\delta}{s+\delta}} \tag{F.3}$$

The poles of (F.3) are the roots of the second-order equation $s^2 + s\delta + \gamma\delta$. Since the real part of both roots is negative for any δ, the system is always stable. For $\frac{1}{\delta} > \frac{1}{4\gamma}$ it exhibits dumped oscillations, otherwise it exhibits an exponential decay. □

Corollary 5 (Shifted exponential delay distribution)**.** *Let* $\mathfrak{f}_d(\tau) = u(\tau - d)\delta e^{-\delta(\tau-d)}$*. For any given* $\delta > 0$*, there exists a critical delay* $d^* = \frac{1}{\gamma} f(\delta)$*, such that the system is (locally) stable if* $d < d^*$*, otherwise the system is unstable. As* δ *grows from 0 to* ∞*,* d^* *grows from* $1/\gamma$ *to* $\pi/(2\gamma)$*.*

Proof. Consider the shifted exponential distribution: $\mathfrak{f}_d(\tau) = u(\tau - d)\delta e^{-\delta(\tau-d)}$. In this case we have

$$\mathscr{L}\{\eta(t)\} = \frac{\eta(0)}{s + \frac{\gamma\delta}{s+\delta} e^{-sd}} \tag{F.4}$$

whose poles $z = b + i\omega$ satisfy the equations:

$$\begin{cases} b^2 - \omega^2 + b\delta + \gamma\delta e^{-bd}\cos(\omega d) = 0 \\ 2b\omega + \omega\delta - \gamma\delta e^{-bd}\sin(\omega d) = 0 \end{cases}. \tag{F.5}$$

At the critical point, $b = 0$, and above equations reduce to:

$$\begin{cases} \omega^2 = \gamma\delta\cos(\omega d) \\ \omega = \gamma\sin(\omega d) \end{cases}. \tag{F.6}$$

For given δ, we can solve the above two equations in the unknowns $\omega^\star$, $d^\star$, obtaining:

$$\omega^\star = \sqrt{\frac{-\delta^2 + \sqrt{\delta^4 + 4\gamma^2\delta^2}}{2}}, \quad d^\star = \frac{\arcsin(\frac{\omega}{\gamma})}{\omega}.$$

□

Fig. F.4 shows the critical values $d^\star$ and $\tau^\star = 2\pi/\omega^\star$ as a function of $1/\delta$ (in the plot all quantities are normalized by the average sojourn time $1/\gamma$ in the infectious state).

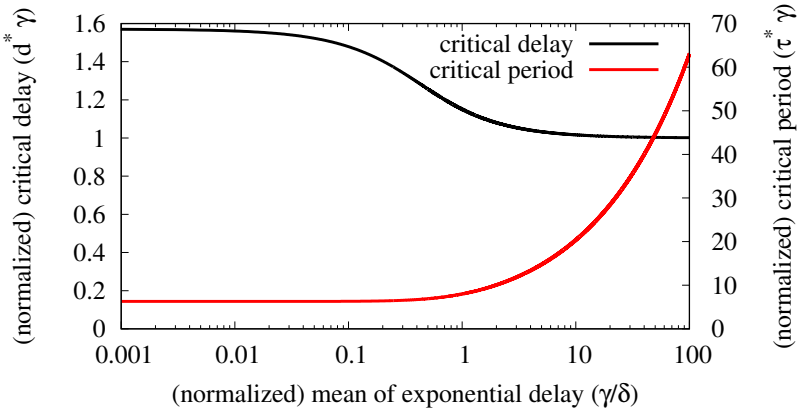


Fig. F.4 Critical values $d^\star$ (left y axes) and $\tau^\star$ (right y axes) as function of $1/\delta$. All quantities are normalized by $1/\gamma$.

Interestingly, the critical value $d^\star$ of the shift decreases as we increase the average $1/\delta$ of the exponential distribution. For example, with $\delta = \gamma$ we have $d^\star\gamma \approx 1.15$, in constrast to $d^\star \approx 1.57$ with deterministic delay.