A Dynamic Population Model of Strategic Interaction and Migration under Epidemic Risk

Ezzat Elokda, Saverio Bolognani and Ashish R. Hota

Abstract— In this paper, we show how a dynamic population game can model the strategic interaction and migration decisions made by a large population of agents in response to epidemic prevalence. Specifically, we consider a modified susceptible-asymptomatic-infected-recovered (SAIR) epidemic model over multiple zones. Agents choose whether to activate (i.e., interact with others), how many other agents to interact with, and which zone to move to in a time-scale which is comparable with the epidemic evolution. We define and analyze the notion of equilibrium in this game, and investigate the transient behavior of the epidemic spread in a range of numerical case studies, providing insights on the effects of the agents' degree of future awareness, strategic migration decisions, as well as different levels of lockdown and other interventions. One of our key findings is that the strategic behavior of agents plays an important role in the progression of the epidemic and can be exploited in order to design suitable epidemic control measures.

I. INTRODUCTION

Infectious diseases or epidemics spread through society by exploiting social interactions. As the disease becomes more prevalent, individuals and authorities take various measures to prevent its spread. These measures include restricting large gatherings and travel, imposing lock-downs, and others. Furthermore, individuals voluntarily reduce their social interaction and even migrate to safer locations in a strategic and non-myopic manner [1], [2] which plays a significant role in epidemic evolution. Therefore, it is critical to rigorously investigate the interplay between individual behavior, emerging network structure and epidemic evolution.

Mathematical modeling of epidemic evolution on networks has a rich history; see [3]–[5] for comprehensive reviews. Recent work has explored epidemic mitigation strategies based on centralized resource allocation [6], [7] as well as decentralized or game-theoretic protection strategies [8]–[10]. Several papers have considered dynamic or evolutionary game-theoretic approaches in the context of networked epidemics [11]–[13]. Evolution of network topology and epidemic states in a comparable time-scale have recently been studied in the framework of activity-driven networks [14], [15]. Game-theoretic decision-making in this framework was recently studied in [16] where the authors consider myopic bounded rational agents who decide whether to activate or not as a function of the current prevalence of the epidemic.

E. Elokda and S. Bolognani are with ETH Zurich, 8092 Zurich, Switzerland. {elokdae,bsaverio}@ethz.ch

A. R. Hota is with the Department of Electrical Engineering, Indian Institute of Technology Kharagpur, India. ahota@ee.iitkgp.ac.in

To the best of our knowledge, there have been few rigorous game-theoretic formulations that model agents that

- decide their degree of activation, and consequently influence the resulting network topology,
- decide whether to migrate to different locations, and
- maximize both current and long-run future pay-off

in the same time-scale as epidemic evolution. In this work, we present a framework to address the above research gap.

Motivated by the presence of asymptomatic carriers in COVID-19 [17], [18], we build upon the SAIR epidemic model studied in [19], [20]. We consider a large population regime where the state of an individual agent is characterized by its infection state and location or zone. At discrete time instants, each agent decides its degree of activation and next location. The agent is then paired randomly with other agents, and its infection state evolves following an augmented SAIR epidemic model which also takes into account *unknowingly recovered* agents as described in Section II. Agents decide strategically in order to maximize a discounted infinite horizon expected reward which is a function of the aggregate infection state, the zonal distribution of the agents, and the policy followed by the population.

In a departure from the conventional assumption of a static population distribution in the classical population game setting [21], epidemic evolution leads to a dynamically evolving population state which makes the analysis particularly challenging. Previous works that study dynamic strategic interactions in large populations include anonymous sequential games [22]–[24] and mean field games [25]–[29], which are often difficult to apply in practical models. We utilize the recently developed framework of *dynamic population games* [30], in which the authors show a reduction of the dynamic setting to a static population game setting [21]. This simplifies the analysis in comparison to the existing approaches, and is particularly useful for epidemic models. As a consequence of the reduction, standard evolutionary models [21] can be adapted for the coupled dynamics of the agents' states and strategic decision making, which evolve on the same time-scale.

The paper is structured as follows: the dynamic population model is presented in Section II, and its stationary equilibria are analyzed in Section III. The evolutionary update of agents' policies is modeled in Section IV. Numerical experiments reported in Section V provide compelling insights into agents' behavior, effects of lockdown measures and strategic mobility patterns. For instance, we observe that if recovered agents are exempt from lockdown measures, it leads to an increased level of activity by susceptible and asymptomatic

agents as well without having much impact on the peak and total infections. In other words, strategic behavior does not lead to a higher infection level and the social welfare improves due to higher activity levels.

II. MODEL

We consider a homogeneous population of agents, each corresponding to an individual. The state and the dynamics of this population are described by the following elements.

A. States

We augment the SAIR epidemic model to distinguish between recovered agents who are aware of being recovered and those who are recovered, but unaware of ever being infected. Specifically, each agent can be in one of the following infection states or compartments: *Susceptible* (S), *infected Asymptomatic* (A), *Infected symptomatic* (I), *Recovered* (R), *Unknowingly recovered* (U) (corresponding to agents that have recovered without showing symptoms). Agents in states R and U are immune from further infection. Moreover, each agent resides in one of Z zones.

Formally, we define the state of each agent as $(s, z) \in \mathcal{S} \times \mathcal{Z}$, where $\mathcal{S} = \{\, \text{S}, \text{A}, \text{I}, \text{R}, \text{U} \,\}$ and $\mathcal{Z} = \{\, 1, \dots, Z \,\}$. The state distribution is $d \in \mathcal{D} = \Delta(\mathcal{S} \times \mathcal{Z})$, where $d[s, z]$ denotes the proportion of agents with infection state s residing in zone z.

B. Actions and policies

We consider a dynamic environment that evolves in discrete-time (e.g., each time interval representing a day). At each time step, each agent strategically chooses:

- its activation degree $a \in \mathcal{A} = \{\, 0, 1, \dots, a_{\max} \,\}$, which denotes the number of other agents it chooses to interact with ($a = 0$ signifies no activation), and
- the zone $\tilde{z} \in \mathcal{Z}$ where to move (at the end of the day) and reside for the next day.

The combined action is denoted $(a, \tilde{z}) \in \mathcal{A} \times \mathcal{Z}$.

A (Markovian) policy is denoted by $\pi : \mathcal{S} \times \mathcal{Z} \to \Delta(\mathcal{A} \times \mathcal{Z})$, and it maps an agent's state $(s, z) \in \mathcal{S} \times \mathcal{Z}$ to a randomization over the actions $(a, \tilde{z}) \in \mathcal{A} \times \mathcal{Z}$. The set of all possible policies is denoted Π. Explicitly, $\pi[a, \tilde{z} \mid s, z]$ is the probability that an agent plays $(a, \tilde{z})$ when in state (s, z).

All agents are homogeneous and follow the same policy π. Policies need to be consistent with the information structure of the problem: agents that have never shown symptoms act in the same way, i.e. $\pi[\cdot \mid \text{S}, z] = \pi[\cdot \mid \text{A}, z] = \pi[\cdot \mid \text{U}, z]$.

C. State transitions

In order to study the evolution of the epidemic and the behavior of the agents, we now derive a dynamic model of the evolution of state distribution d when the agents adopt a policy π. We denote the policy-state pair (π, d) as the *social state*. The state of each agent changes at every time step according to transition probabilities encoded by the *stochastic matrix*

$$
\begin{aligned}
P[s^+, z^+ \mid s, z](\pi, d) \\
= \sum_{a, \tilde{z}} \pi[a, \tilde{z} \mid s, z]\, p[s^+, z^+ \mid s, z, a, \tilde{z}](\pi, d), \quad (1)
\end{aligned}
$$

where $p[s^+, z^+ \mid s, z, a, \tilde{z}](\pi, d)$ denotes the probability distribution over the next state when an agent in infection state s and zone z chooses action $(a, \tilde{z})$ in social state (π, d).

Note that the Markov chain $P[s^+, z^+ \mid s, z](\pi, d)$ is not time-homogeneous, since the transition probabilities are parametrized by the time-varying social state (π, d). However, it becomes time-homogeneous if (π, d) is stationary (see the equilibrium analysis in Section III).

The crucial element of (1) is the state transition function

$$
\begin{aligned}
p[s^+, z^+ \mid s, z, a, \tilde{z}](\pi, d) \\
:= \mathbb{P}[s^+ \mid s, z, a](\pi, d)\, \mathbb{P}[z^+ \mid z, \tilde{z}](\pi, d), \quad (2)
\end{aligned}
$$

specified hereafter. The term $\mathbb{P}[z^+ \mid z, \tilde{z}](\pi, d)$ corresponds to zone transitions, which we assume to be deterministic and independent of (π, d) and thus defined as

$$
\mathbb{P}[z^+ \mid z, \tilde{z}](\pi, d) = \begin{cases} 1 & \text{if } z^+ = \tilde{z}, \\ 0 & \text{otherwise.} \end{cases}
$$

In order to derive the infection state transition probabilities $\mathbb{P}[s^+ \mid s, z, a](\pi, d)$ (which are schematically represented in Figure 1), we combine the transition rules of the (augmented) SAIR model with the specific activation actions as follows.

- At a given time, an agent in state s in zone z chooses its activation degree a according to policy π. Then, it is paired randomly with a other individuals in zone z with the probability of being connected with another agent being proportional to the activation degree of the target agent (analogous to the configuration model [31]).
- Once the network is formed, a susceptible agent becomes asymptomatic with probability $\beta_\text{A} \in [0, 1]$ per each asymptomatic neighbor and with probability $\beta_\text{I} \in [0, 1]$ per each infected neighbor independently.
- An asymptomatic agent becomes (symptomatically) infected with probability $\delta_\text{A}^\text{I} \in (0, 1]$,[1] and recovers without being aware of it with probability $\delta_\text{A}^\text{U} \in [0, 1]$.
- An infected agent recovers with probability $\delta_\text{I}^\text{R} \in (0, 1]$.
- An individual in state U becomes aware of its recovery with probability $\delta_\text{U}^\text{R} \in [0, 1]$ (for example via serological tests on the population). Once recovered (knowingly or unknowingly), an agent has developed immunity and does not get infected again.
- The network thus formed gets discarded at the next time step and the process repeats.

We now formally state the above model of epidemic evolution. Note that with the exception of the transition from state S to A, all other state transition probabilities are defined via exogenous parameters and do not depend on the social

---

[1]This represents both the aggravation of the illness and the result of testing in the asymptomatic population.
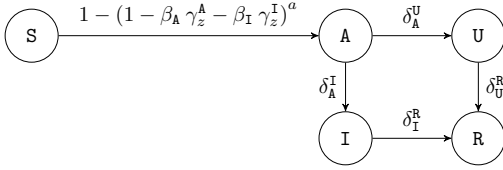
Fig. 1: Infection state transition diagram

state. Let us then compute the probability of interacting with asymptomatic or infected agents. We define the total amount of activity in zone z, which is determined both by the mass of active agents and by their degrees of activation, as

$$e_z(\pi, d) = \sum_s d[s, z] \sum_{a, \tilde{z}} a\, \pi[a, \tilde{z} \mid s, z].$$

Similarly, the amount of activity by *asymptomatic agents* in zone z is given by

$$e_z^{\mathtt{A}}(\pi, d) = d[\mathtt{A}, z] \sum_{a, \tilde{z}} a\, \pi[a, \tilde{z} \mid \mathtt{A}, z],$$

and the amount of activity by *infected agents* in zone z is

$$e_z^{\mathtt{I}}(\pi, d) = d[\mathtt{I}, z] \sum_{a, \tilde{z}} a\, \pi[a, \tilde{z} \mid \mathtt{I}, z].$$

In order to consider the event of encountering no agents to connect with when the amount of activity in the zone $e_z(\pi, d)$ is minute, we introduce a small constant amount $\epsilon > 0$ of fictitious activation that does not belong to any of the agents. Consequently in zone z, the probability of not interacting with any agents, the probability of a randomly chosen agent being asymptomatic and the probability of a randomly chosen agent being symptomatic are, respectively,

$$\gamma_z^{\emptyset}(\pi, d) = \frac{\epsilon}{e_z(\pi, d) + \epsilon}, \tag{3}$$

$$\gamma_z^{\mathtt{A}}(\pi, d) = \frac{e_z^{\mathtt{A}}(\pi, d)}{e_z(\pi, d) + \epsilon}, \quad \gamma_z^{\mathtt{I}}(\pi, d) = \frac{e_z^{\mathtt{I}}(\pi, d)}{e_z(\pi, d) + \epsilon}. \tag{4}$$

As a result, the probability of a susceptible agent to *not* get infected upon activation is

$$\mathbb{P}[s^+ = \mathtt{S} \mid s = \mathtt{S}, z, a](\pi, d)$$
$$= \left(1 - \beta_{\mathtt{A}}\, \gamma_z^{\mathtt{A}}(\pi, d) - \beta_{\mathtt{I}}\, \gamma_z^{\mathtt{I}}(\pi, d)\right)^a.$$

It is easy to see that when a susceptible agent does not interact with any other agent (i.e., $a = 0$), it remains susceptible.[2] When it participates in exactly one interaction ($a = 1$) in z, the probability that its neighbor is asymptomatic (respectively, infected) is $\gamma_z^{\mathtt{A}}(\pi, d)$ (respectively, $\gamma_z^{\mathtt{I}}(\pi, d)$). When it draws $a > 0$ independent agents to interact with, it must not get infected in any of the interactions to remain susceptible, and this occurs with the specified probability.

The remaining infection state transition probabilities follow directly from the model as:

$$\mathbb{P}[s^+ = \mathtt{A} \mid s = \mathtt{S}, z, a](\pi, d) =$$
$$1 - \mathbb{P}[s^+ = \mathtt{S} \mid s = \mathtt{S}, z, a](\pi, d)$$

[2]We define $0^0 = 1$ for the edge case if $1 - \beta_{\mathtt{A}}\gamma_z^{\mathtt{A}}(\pi, d) - \beta_{\mathtt{I}}\gamma_z^{\mathtt{I}}(\pi, d) = 0$.

$$\mathbb{P}[s^+ = \mathtt{I} \mid s = \mathtt{A}, z, a](\pi, d) = \delta_{\mathtt{A}}^{\mathtt{I}},$$
$$\mathbb{P}[s^+ = \mathtt{U} \mid s = \mathtt{A}, z, a](\pi, d) = \delta_{\mathtt{A}}^{\mathtt{U}},$$
$$\mathbb{P}[s^+ = \mathtt{A} \mid s = \mathtt{A}, z, a](\pi, d) = 1 - \delta_{\mathtt{A}}^{\mathtt{I}} - \delta_{\mathtt{A}}^{\mathtt{U}},$$
$$\mathbb{P}[s^+ = \mathtt{R} \mid s = \mathtt{I}, z, a](\pi, d) = \delta_{\mathtt{I}}^{\mathtt{R}},$$
$$\mathbb{P}[s^+ = \mathtt{I} \mid s = \mathtt{I}, z, a](\pi, d) = 1 - \delta_{\mathtt{I}}^{\mathtt{R}},$$
$$\mathbb{P}[s^+ = \mathtt{R} \mid s = \mathtt{U}, z, a](\pi, d) = \delta_{\mathtt{U}}^{\mathtt{R}}$$
$$\mathbb{P}[s^+ = \mathtt{U} \mid s = \mathtt{U}, z, a](\pi, d) = 1 - \delta_{\mathtt{U}}^{\mathtt{R}}$$
$$\mathbb{P}[s^+ = \mathtt{R} \mid s = \mathtt{R}, z, a](\pi, d) = 1.$$

These expressions completely specify $\mathbb{P}[s^+ \mid s, z, a](\pi, d)$ (Figure 1) and therefore the state transition function (2).

D. Rewards

Each agent's own state (s, z) and action $(a, \tilde{z})$ yields an immediate reward for the agent, composed of a reward $r_a[s, z, a]$ for their activation decision, a reward $r_m[s, z, \tilde{z}]$ for their migration decision, and a reward $r_d[s]$ for how the agent's health is affected by the disease. Formally,

$$r[s, z, a, \tilde{z}] := r_{\text{act}}[s, z, a] + r_{\text{mig}}[s, z, \tilde{z}] + r_{\text{dis}}[s]. \tag{5}$$

The activation reward is defined as

$$r_{\text{act}}[s, z, a] := o[a] - c[s, z, a],$$

where $o[a] \in \mathbb{R}_+$ denotes the social benefit of interacting with a other agents and is assumed to be non-decreasing in a with $o[0] = 0$, and $c[s, z, a] \in \mathbb{R}_+$ denotes the cost imposed by the authority in zone z to discourage activity. We assume that $c[s, z, a]$ are non-decreasing in a and satisfy

$$c[\mathtt{I}, z, a] \geq c[\mathtt{S}, z, a] = c[\mathtt{A}, z, a] = c[\mathtt{U}, z, a] \geq c[\mathtt{R}, z, a]$$

element-wise, since lockdown measures can be more stringent against individuals showing symptoms and more benign for individuals who are known to be immune.

The migration reward $r_{\text{mig}}[s, z, \tilde{z}]$ encodes the non-negative cost of migrating to a new zone. In this work, we define

$$r_{\text{mig}}[s, z, \tilde{z}] := \begin{cases} 0 & \text{if } \tilde{z} = z, \\ -c_{\text{mig}} & \text{if } \tilde{z} \neq z. \end{cases}$$

However, one may consider a richer cost function that incorporates specific travel restrictions between zones, and is a function of the infection state (i.e., different cost for symptomatic and immune agents).

The third term in (5) encodes the cost of being ill:

$$r_{\text{dis}}[s] := \begin{cases} -c_{\text{dis}} & \text{if } s = \mathtt{I}, \\ 0 & \text{otherwise.} \end{cases}$$

E. Agents' Strategic Decisions

We now introduce the strategic decision-making process of the agents. The immediate expected reward of an agent in state (s, z) when it follows policy π is

$$R[s, z](\pi) = \sum_{a, \tilde{z}} \pi[a, \tilde{z} \mid s, z]\, r[s, z, a, \tilde{z}],$$

with $r[s, z, a, \tilde{z}]$ as defined in (5). The expected discounted infinite horizon reward of an agent in state (s, z) with discount factor $\alpha \in [0, 1)$ following the homogeneous policy π is recursively defined as

$$V[s, z](\pi, d) = R[s, z](\pi)$$
$$+ \alpha \sum_{s^+, z^+} P[s^+, z^+ \mid s, z](\pi, d) \, V[s^+, z^+](\pi, d),$$

or, equivalently in vector form,

$$V(\pi, d) = (I - \alpha \, P(\pi, d))^{-1} \, R(\pi). \qquad (6)$$

Note that (6) is continuous in the social state (π, d), as $I - \alpha \, P(\pi, d)$ is guaranteed to be invertible for $\alpha \in [0, 1)$.

While an agent can compute the expected discounted reward $V(\pi, d)$ at a given social state (π, d), the policy π may not be optimal for the agent. Thus, we assume that each agent chooses its current action $(a, \tilde{z})$ in order to maximize

$$Q[s, z, a, \tilde{z}](\pi, d) := r[s, z, a, \tilde{z}]$$
$$+ \alpha \sum_{s^+, z^+} p[s^+, z^+ \mid s, z, a, \tilde{z}](\pi, d) \, V[s^+, z^+](\pi, d), \quad (7)$$

i.e., the agent is cognizant of the immediate reward and the effect of its action on their future state; however, it assesses the future reward based on a stationarity assumption on the social state (π, d). In other words, the agent chooses its action to maximize a single-stage deviation from the homogeneous policy π [32, Section 2.7]. Intuitively, this corresponds to assuming that the rest of the population is acting rationally at this moment, and that its own actions are not going to affect the social state significantly. This is a sensible working approximation in practice, and becomes exact at the equilibrium as discussed in the following section.

III. EQUILIBRIUM ANALYSIS

We start by introducing the notion of best response based on the single-stage deviation reward defined in (7).

Definition 1 (Best Response). *The best response of an agent in state (s, z) at the social state (π, d) is the set valued correspondence $B_{s,z} : \Pi \times \mathcal{D} \rightrightarrows \Delta(\mathcal{A} \times \mathcal{Z})$ given by*

$$B_{s,z}(\pi, d) \in \left\{ \sigma \in \Delta(\mathcal{A} \times \mathcal{Z}) : \forall \sigma' \in \Delta(\mathcal{A} \times \mathcal{Z}) \right.$$
$$\left. \sum_{a, \tilde{z}} (\sigma[a, \tilde{z}] - \sigma'[a, \tilde{z}]) \, Q[s, z, a, \tilde{z}](\pi, d) \geq 0 \right\}. \quad (8)$$

The above notion of best response is from the perspective of an individual agent in state (s, z) when all other agents are following the homogeneous policy π and their states are distributed as per d. The agent will choose any distribution σ over the actions which maximizes its expected single-stage deviation reward Q at the current state (s, z).

Consequently, a social state (π, d) is stationary when agents in all states are playing their best response when they follow the policy π, and additionally, the state distribution d is stationary under this policy. Thus, we have the following definition of a stationary equilibrium.

Definition 2 (Stationary equilibrium). *A stationary equilibrium is a social state $(\pi, d) \in \Pi \times \mathcal{D}$ which satisfies*

$$\pi[\cdot \mid s, z] \in B_{s,z}(\pi, d), \quad \forall (s, z) \in \mathcal{S} \times \mathcal{Z}, \qquad \text{(SE.1)}$$
$$d = d \, P(\pi, d). \qquad \text{(SE.2)}$$

Thus, at the equilibrium, the state Markov chain $P(\pi, d)$ (1) is time-homogeneous, and the agents behave optimally in the corresponding Markov decision process [32].

Theorem 1 ([30]). *A stationary equilibrium (π, d) for the proposed dynamic population game is guaranteed to exist.*

We refer to [30] for the details of the proof, which relies on a fixed-point argument. The proof requires that the state transition and reward functions are continuous in the social state (π, d), which is satisfied in our model.

The following proposition shows that the stationary distribution d does not have any asymptomatic or infected agents (i.e., eventually, the epidemic will die out). The final stationary distribution is however not unique.

Proposition 1. *Let $r_{\mathrm{act}}^*[s, z] := \max_a r_{\mathrm{act}}[s, z, a]$ be the maximum activation reward in state s and zone z, and let $\mathcal{A}_{s,z}^*$ be the set of activation levels that achieve such reward. Let $\bar{r}_{\mathrm{act}}[s] := \max_z r_{\mathrm{act}}^*[s, z]$ be the maximum activation reward in state s across all zones, and let $\delta_{\mathrm{U}}^{\mathrm{R}}, \alpha, c_{\mathrm{mig}} > 0$[3]. Let us define the subset of zones*

$$\bar{\mathcal{Z}}_s = \operatorname*{argmax}_z r_{\mathrm{act}}^*[s, z],$$

$$\mathcal{Z}_s^0 = \left\{ z \mid \bar{r}_{\mathrm{act}}[s] - r_{\mathrm{act}}^*[s, z] > \frac{1 - \alpha}{\alpha} c_{\mathrm{mig}} \right\}.$$

Then, any social state (π, d) satisfying

$$\begin{cases} d[\mathtt{A}, z] = d[\mathtt{I}, z] = d[\mathtt{U}, z] = 0 & \forall z, \\ d[s, z \in \mathcal{Z}_s^0] = 0 & \forall s, \\ \pi[a \notin \mathcal{A}_{s,z}^*, \tilde{z} \mid s, z] = 0 & \forall \tilde{z}, s, z, \\ \pi[a, \tilde{z} \notin \bar{\mathcal{Z}}_s \mid s, z \in \mathcal{Z}_s^0] = 0 & \forall a, s, \\ \pi[a, \tilde{z} \neq z \mid s, z \notin \mathcal{Z}_s^0] = 0 & \forall a, s. \end{cases}$$

is a stationary equilibrium.

Proposition 1 (whose proof is in the Appendix) also shows how stationary equilibria can be computed without solving a fixed-point problem, and directly identifies some dominant strategies for the agents. The identification of dominant strategies is extremely insightful for the design of interventions (e.g., lockdown measure for the different compartments). For example, it is possible to verify that

$$\mathcal{A}_{\mathtt{R},z}^* = \operatorname*{argmax}_a \, r_{\mathrm{act}}[\mathtt{R}, z, a] = \operatorname*{argmax}_a \, (o[a] - c[\mathtt{R}, z, a])$$

corresponds to the dominant activation strategies for agents in $\mathtt{R}$ (*knowingly* immune agents, including those that tested positive to a serological test). Notice that the activation caused by immune agents appears in the denominator of the

---

[3] An analogous result can be obtained when $\delta_{\mathrm{U}}^{\mathrm{R}} = 0$, in which case $d[\mathtt{U}, z]$ does not need to be zero. The case $\alpha = 0$ is easily derived by interpreting $(1 - \alpha)/\alpha$ as $+\infty$.

probability that a generic agent interacts with an infectious agent – see (4) – therefore looser lockdown measures for recovered agents ($c[\mathtt{R}, z, a] = 0$) can be used to reduce the spreading as shown in simulations in Section V.

IV. Social State Dynamics

Different stationary equilibria correspond to drastically different outcomes in terms of the impact of the epidemic on the population. For this reason, it is interesting to investigate the *transient behavior* leading to the equilibrium, which is determined by the state dynamics from Section II-C, i.e.,

$$d^+ = d \, P(\pi, d),$$

and the way in which agents update their policies. For this second part, we get inspiration from the *evolutionary dynamic* models in classical population games [21], and more precisely from the *perturbed best response dynamics*.

We assume that the agents are not perfectly rational, with the bounded rationality factor $\lambda \in [0, \infty)$. When they are making a decision on which action to play, they follow the *logit choice* function [21, Section 6.2], given by

$$\tilde{\pi}[a, \tilde{z} \mid s, z](\pi, d) = \frac{\exp\left(\lambda \, Q[s, z, a, \tilde{z}](\pi, d)\right)}{\sum_{a', \tilde{z}'} \exp\left(\lambda \, Q[s, z, a', \tilde{z}'](\pi, d)\right)}.$$

For $\lambda = 0$, it results in a uniform distribution over all the actions. At the limit $\lambda \to \infty$, we recover the perfect best response. At finite value of λ, it assigns higher probabilities to actions with higher payoffs, modeling the effect of concurrent reasons (unrelated to the epidemic) that affect the agents' strategic behavior.

In order to model the fact that agents update their policies gradually, we consider the discrete-time update

$$\pi^+[\cdot \mid s, z] = (1 - \eta) \, \pi[\cdot \mid s, z] + \eta \, \tilde{\pi}[\cdot \mid s, z],$$

where $\eta \in (0, 1]$ is a parameter that controls the rate of policy change: for $\eta < 1$, agents have *inertia* in their decision making, while for $\eta = 1$ agents promptly update their action decision to the perturbed best response $\tilde{\pi}$. Note that this update model leads to a perturbed version of the equilibrium policy π at the rest points, rather than the exact policy [21].

V. Numerical Case Studies

We present a select number of case studies to showcase how our model can be insightful for the analysis of different phenomena in epidemic spreading:

- the effect of agents' strategic activation decisions on the spread of the epidemic,
- the impact of lockdown interventions on both the containment of the epidemic and the economic welfare of the population, and
- the effect of strategic migration decisions on how the epidemic spreads across multiple locations.

For this purpose, we consider an infectious epidemic characterized by $\beta_{\mathtt{A}} = \beta_{\mathtt{I}} = 0.2$, $\delta_{\mathtt{A}}^{\mathtt{I}} = \delta_{\mathtt{A}}^{\mathtt{U}} = 0.08$, and $\delta_{\mathtt{I}}^{\mathtt{R}} = 0.04$. The agents can activate up to degree $a_{\max} = 6$, and the activation reward is linear in the activation degree, with a unit reward for maximum activation $o[a_{\max}] = 1$. The illness

is quite severe, with a discomfort cost $c_{\mathrm{dis}} = 10$. Initially, we let the agents choose an activation degree uniformly at random and do not plan any move. The agents are highly rational ($\lambda = 10$) and unless otherwise stated, we consider that they update their decisions with an inertia $\eta = 0.2$. We consider both a single zone and a two zone setting, and denote the zones by Z1 and Z2. In all cases, the epidemic starts in Z1 with a proportion of 2% of that zone's population asymptomatic (A), and 1% infected (I).

We further consider that authorities can enforce lockdown regulations through the parameter $a_{\mathrm{lock}}[s, z]$, which represents the *maximum allowed activation degree* and can differ between zones and for agents in different infection states. Lockdown is implemented by setting $c[s, z, a] = 0$ if $a \le a_{\mathrm{lock}}[s, z]$, and $c[s, z, a] = 3 \, o[a]$ otherwise. Regardless of the lockdown measures, we always assume that the discomfort of the illness is sufficient to prevent symptomatically infected agents from activating. As a consequence, the main threat of the epidemic is due to the presence of asymptomatically infected agents in the population.

A. Lockdown measures

We first investigate the single zone scenario, with a focus on how agents of different cognitive ability react under various lockdown measures, and the resulting effects on the epidemic spread. Figure 2 shows an example with lockdown degree $a_{\mathrm{lock}} = 2$ and three cases. In both cases (2a) and (2b), the lockdown is enforced on the whole population, and the cases differ in the agents' cognitive level of the future. In (2a), the agents are completely myopic. Notice how they simply adhere to the lockdown degree[4], which aligns with standard epidemic models. Farsighted agents (2b), on the other hand, behave considerably differently under the same lockdown measures. We observe that they actively adjust their activation decisions in response to the epidemic threat, and volunteer to limit their activity beyond the lockdown requirement at peak infection times. This aligns with the intuitive notion of individuals *acting responsibly* under the epidemic threat, and our model demonstrates that this is indeed in their best interest. The reduction in activity levels leads to a less severe epidemic spread, with both a considerably smaller total and peak number of infections.

Case (2c) also considers farsighted agents, but this time *recovered* agents are exempt from lockdown and thus activate at the maximum degree, as per their dominant strategy. This leads to a significant reduction of the total amount of infections. In fact, due the prevalence of activity of immune agents, it becomes less likely for a susceptible agent to encounter an infected agent. Consequently, susceptible, asymptomatic and unknowingly recovered agents too increase their level of activation. Nevertheless, the peak infection remains being largely unchanged.

This insight is further explored in Figure 3, which shows the effect of different lockdown measures on three main

---

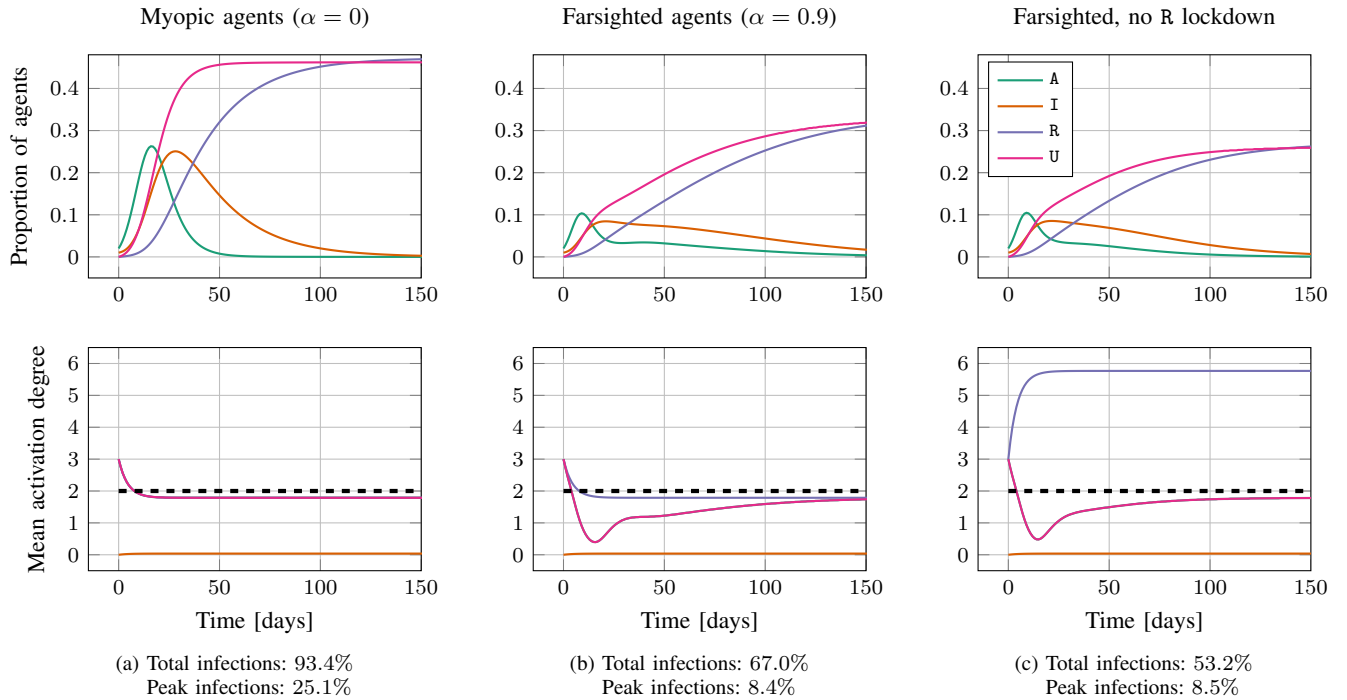[4] A slight deviation to the expected degree of activation is due to the agents' bounded rationality.

Fig. 2: The effect of the agents' farsightedness and of the exclusion of immune agents from lockdown measures on the epidemic.
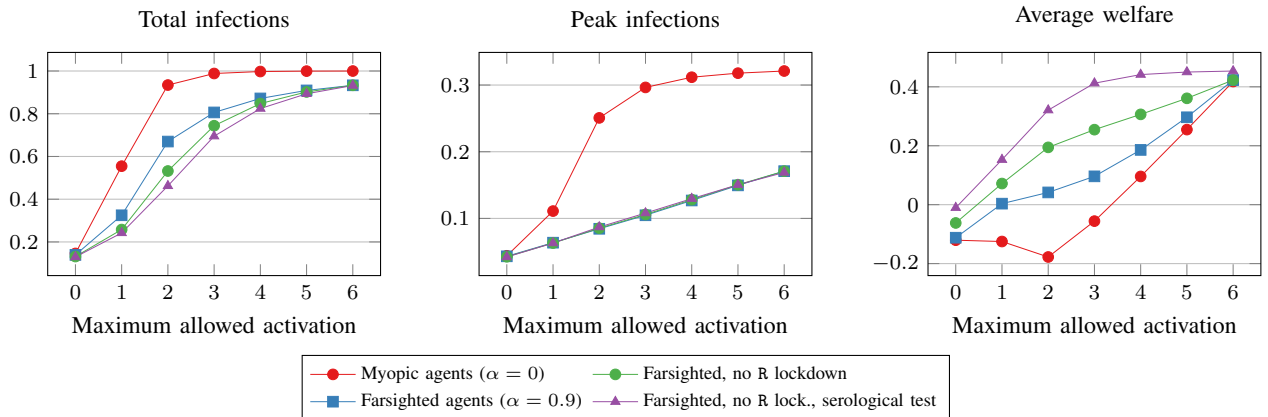

Fig. 3: The effect of the severity of the lockdown measures on three key epidemic indicators, for different scenarios.

performance indices: the *total infections* ($R + U$ at the end of the epidemic), the *peak infections* (highest value of I[5]), and the *average welfare* (mean reward (5) in the population over the duration of the epidemic). We depict the same cases considered in Figure 2 and perform a parameter sweep over the strictness of the lockdown a_{lock}. Additionally, we showcase the effect of performing *serological tests* to increase the amount of knowingly immune agents, at a testing rate $\delta_U^R = 0.05$. We observe that the myopic agent model is rather pessimistic in all the performance metrics. For farsighted agents, the exemption of recovered agents and serological tests lead to significant improvements in the average welfare,

mostly because the knowingly immune agents are able to achieve their maximum activation reward without increasing the threat of the epidemic. On the contrary, as discussed before, these measures reduce the probability of infection and ultimately improve the final outcome (total infections).

B. Strategic migration

We showcase the effect of strategic migration in a setting with two zones, with zone Z1 initially holding 90% of the total population, and zone Z2 initially infection free. In both zones, (knowingly) recovered agents are exempt from lockdown, and the lockdown restrictions for the other agents are different in the zones. Namely, Z1 has a looser lockdown, with a maximum allowed activation degree of 4, whereas Z2 only allows a maximum of 2. The migration cost is $c_{mig} = 2$,

---

[5]We do not consider infected but asymptomatic agents since they do not contribute to the load on medical facilities.
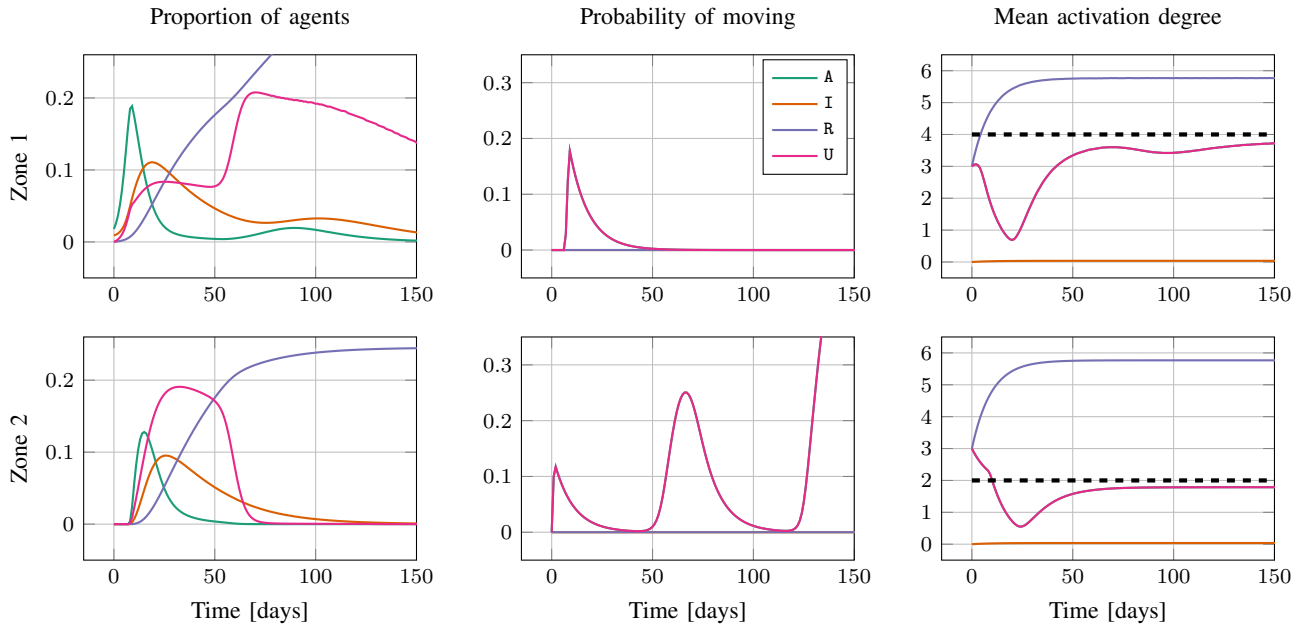
Fig. 4: Effect of strategic migration/activation decisions on the epidemic spreading between two zones with different lockdown measures.

and the agents are farsighted with $\alpha = 0.9$. Note that with these parameters, susceptible agents will want to move to Z1 when the epidemic is not prevalent, as per Proposition 1[6]. The inertia in the policy update is $\eta = 0.1$.

Additionally, both zones perform serological testing at rate $\delta_U^R = 0.01$. The resulting epidemic spread is displayed in Figure 4. First, notice how the proportion of unknowingly recovered agents decays, in contrast to Figure 2, in which no serological testing is performed.

We now focus on the strategic migration behavior of agents who are either susceptible to infection or think they are (S, A, U), as the epidemic unfolds. Note that symptomatically infected and recovered agents never move since they are immune to the threat of the epidemic, and the activation costs are the same for them in both zones. Initially, the epidemic risk is still small in Z1, and the occupants of Z2 start moving there to benefit from the more lenient lockdown measure. This trend soon reverses, however, with the rise of infections in Z1: the strategic agents elect to move to the zone with stricter lockdown to escape the epidemic risk. Since a proportion of the movers are asymptomatically infected, this leads to an outbreak of the epidemic in Z2 as well, with lower, but significant, peak infections than Z1. Eventually, once the infections in Z1 has decreased sufficiently (at approximately day 50), some Z2 residents move to Z1 again, initiating a *second wave of infections* in Z1. At the end of the epidemic, all the remaining agents move from Z2 as per their dominant strategy.

VI. CONCLUSIONS AND FUTURE WORKS

In this paper, we propose a model of strategic behavior at both individual and societal level based on first princi-

ples, demonstrate its potential to explain complex activation and migration patterns, and to guide the design of more effective epidemic control measures. We characterize the stationary equilibria in the proposed dynamic population game setting and illustrate how a better understanding of the emergent behavior can be leveraged to design effective mitigation strategies. For instance, we show that withdrawing restrictions on recovered agents leads to higher levels of activity by susceptible agents without increasing the peak and total infection levels. This observation provides a rigorous justification for conducting large-scale serological testing and letting individuals with immunity to interact freely for significantly improving the welfare of the society.

A natural extension of this model includes further control measures such as vaccinations and the generalization to a heterogeneous population game that would allow, for example, to consider agents with different "home zones." The problem of designing optimal intervention strategies that incorporates the strategic response of agents to such interventions is a challenging avenue for future research.

APPENDIX

Proof of Proposition 1

The first property follows the fact that states A, I, and U are *transient* in the infection state Markov chain (see Figure 1), and evolve independently of the agents' migrations. We omit the formal steps because of space limitations.

We now observe that, given a distribution that is only supported on the two compartments S and R, no transition between compartments of the extended SAIR model is possible. The reward for agents in both these compartments then becomes independent from the actions of others, and the set $\mathcal{A}_{s,z}^*$ defines their *dominant activation strategies*.

---

[6]Here, $r_{\text{act}}^*[\mathtt{S}, \mathtt{Z2}] = \frac{1}{3}$ and $\bar{r}_{\text{act}}[\mathtt{S}] = \frac{2}{3}$, therefore $\mathtt{Z2} \in \mathcal{Z}_\mathtt{S}^0$.

We now consider migration strategies of the agents at the equilibrium $(\boldsymbol{\pi}, \boldsymbol{d})$. We first consider an agent in infection state $s \in \{\mathtt{S}, \mathtt{R}\}$ residing in zone $z \in \bar{\mathcal{Z}}_s$. In this zone, the activation reward is maximum among all other zones for the agent in state s. Since the infection state remains unchanged, and migration is costly ($c_{\mathrm{mig}} > 0$), migrating to a different zone does not lead to a beneficial single-stage deviation for the agent.

We now consider an agent in zone $z \in \mathcal{Z}_s^n := \mathcal{Z} \setminus \{\bar{\mathcal{Z}}_s \cup \mathcal{Z}_s^0\}$. According to policy $\boldsymbol{\pi}$, the agent does not migrate to a different zone. Consequently, the value function for the agent is

$$V[s, z](\boldsymbol{\pi}, \boldsymbol{d}) = \frac{r_{\mathrm{act}}^*[s, z]}{1 - \alpha}.$$

Now, consider a single-stage deviation where the agent moves to location $z' \in \bar{\mathcal{Z}}_s$ and chooses activation degree $a' \in \mathcal{A}_{s,z}^*$. Consequently, using the definition of $\mathcal{Z}_s^n$,

$$Q[s, z, a', z'] = r_{\mathrm{act}}^*[s, z] - c_{\mathrm{mig}} + \alpha \frac{\bar{r}_{\mathrm{act}}[s]}{1 - \alpha} < V[s, z](\boldsymbol{\pi}, \boldsymbol{d}).$$

In other words, an agent in a zone in $\mathcal{Z}_s^n$ does not find it beneficial to migrate anywhere else in a single-stage deviation.

It remains to show that an agent in zone $z \in \mathcal{Z}_s^0$ finds it beneficial to move to a zone in $\bar{\mathcal{Z}}_s$ and consequently, we must have $d[s, z] = 0$ for $z \in \mathcal{Z}_s^0$. Under policy $\boldsymbol{\pi}$, we have

$$\boldsymbol{\pi}[a \notin \mathcal{A}_{s,z}^*, z' \notin \bar{\mathcal{Z}}_s | s, z] = 0.$$

Consequently, the value function for the agent is

$$V[s, z](\boldsymbol{\pi}, \boldsymbol{d}) = r_{\mathrm{act}}^*[s, z] - c_{\mathrm{mig}} + \alpha \frac{\bar{r}_{\mathrm{act}}[s]}{1 - \alpha} > \frac{r_{\mathrm{act}}^*[s, z]}{1 - \alpha};$$

in other words, the policy $\boldsymbol{\pi}$ yields a higher value compared to any policy that does not include migration. Similarly, one can show that migrating to a zone in $\mathcal{Z}_s^n$ instead does not yield a beneficial single-stage deviation since $r_{\mathrm{act}}^*[s, z'] < \bar{r}_{\mathrm{act}}[s]$ for any zone $z' \in \mathcal{Z}_s^n$. $\qquad\square$

REFERENCES

[1] M. Pachetti, B. Marini, F. Giudici, F. Benedetti, S. Angeletti, M. Ciccozzi, C. Masciovecchio, R. Ippodrino, and D. Zella, "Impact of lockdown on COVID-19 case fatality rate and viral mutations spread in 7 countries in europe and north america," *Journal of Translational Medicine*, vol. 18, no. 1, pp. 1–7, 2020.

[2] Google. (2020) Covid-19 community mobility reports. Accessed: 2020-10-26. [Online]. Available: https://www.google.com/covid19/mobility/

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

[4] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, "Epidemic processes in complex networks," *Reviews of Modern Physics*, vol. 87, no. 3, p. 925, 2015.

[5] P. E. Paré, C. L. Beck, and T. Başar, "Modeling, estimation, and analysis of epidemics over networks: An overview," *Annual Reviews in Control (To appear)*, 2020.

[6] V. M. Preciado, M. Zargham, C. Enyioha, A. Jadbabaie, and G. J. Pappas, "Optimal resource allocation for network protection against spreading processes," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 1, pp. 99–108, 2014.

[7] A. R. Hota, J. Godbole, P. Bhariya, and P. E. Paré, "A closed-loop framework for inference, prediction and control of SIR epidemics on networks," *arXiv:2006.16185*, 2020.

[8] S. Trajanovski, Y. Hayel, E. Altman, H. Wang, and P. Van Mieghem, "Decentralized protection strategies against SIS epidemics in networks," *IEEE Transactions on Control of Network Systems*, vol. 2, no. 4, pp. 406–419, 2015.

[9] A. R. Hota and S. Sundaram, "Game-theoretic vaccination against networked SIS epidemics and impacts of human decision-making," *IEEE Transactions on Control of Network Systems*, vol. 6, no. 4, pp. 1461–1472, 2019.

[10] T. C. Reluga, "Game theory of social distancing in response to an epidemic," *PLoS Comput Biol*, vol. 6, no. 5, p. e1000793, 2010.

[11] K. Paarporn, C. Eksin, J. S. Weitz, and J. S. Shamma, "Networked SIS epidemics with awareness," *IEEE Transactions on Computational Social Systems*, vol. 4, no. 3, pp. 93–103, 2017.

[12] C. Eksin, J. S. Shamma, and J. S. Weitz, "Disease dynamics on a network game: A little empathy goes a long way," *Scientific Reports*, vol. 7, p. 44122, 2017.

[13] Y. Huang and Q. Zhu, "A differential game approach to decentralized virus-resistant weight adaptation policy over complex networks," *IEEE Transactions on Control of Network Systems*, vol. 7, no. 2, pp. 944–955, 2020.

[14] L. Zino, A. Rizzo, and M. Porfiri, "An analytical framework for the study of epidemic models on activity driven networks," *Journal of Complex Networks*, vol. 5, no. 6, pp. 924–952, 2017.

[15] ——, "Analysis and control of epidemics in temporal networks with self-excitement and behavioral changes," *European Journal of Control*, vol. 54, pp. 1 – 11, 2020.

[16] A. R. Hota, T. Sneh, and K. Gupta, "Impacts of game-theoretic activation on epidemic spread over dynamical networks," *arXiv:2011.00445*, 2020.

[17] Z. Hu, C. Song, C. Xu, G. Jin, Y. Chen, X. Xu, H. Ma, W. Chen, Y. Lin, Y. Zheng *et al.*, "Clinical characteristics of 24 asymptomatic infections with COVID-19 screened among close contacts in Nanjing, China," *Science China Life Sciences*, vol. 63, no. 5, pp. 706–711, 2020.

[18] W. Pang, "Public health policy: COVID-19 epidemic and SEIR model with asymptomatic viral carriers," *arXiv:2004.06311*, 2020.

[19] S. Ansumali, S. Kaushal, A. Kumar, M. K. Prakash, and M. Vidyasagar, "Modelling a pandemic with asymptomatic patients, impact of lockdown and herd immunity, with applications to sars-cov-2," *Annual Reviews in Control*, 2020.

[20] R. H. Chisholm, P. T. Campbell, Y. Wu, S. Y. Tong, J. McVernon, and N. Geard, "Implications of asymptomatic carriers for infectious disease transmission and control," *Royal Society open science*, vol. 5, no. 2, p. 172341, 2018.

[21] W. H. Sandholm, *Population games and evolutionary dynamics*. MIT Press, 2010.

[22] B. Jovanovic and R. W. Rosenthal, "Anonymous sequential games," *Journal of Mathematical Economics*, vol. 17, no. 1, pp. 77–87, 1988.

[23] G. Y. Weintraub, L. Benkard, and B. Van Roy, "Oblivious equilibrium: A mean field approximation for large-scale dynamic games," *Advances in neural information processing systems*, vol. 18, pp. 1489–1496, 2005.

[24] S. Adlakha, R. Johari, and G. Y. Weintraub, "Equilibria of dynamic games with many players: Existence, approximation, and market structure," *Journal of Economic Theory*, vol. 156, pp. 269–316, 2015.

[25] M. Huang, R. P. Malhamé, P. E. Caines *et al.*, "Large population stochastic dynamic games: closed-loop mckean-vlasov systems and the nash certainty equivalence principle," *Communications in Information & Systems*, vol. 6, no. 3, pp. 221–252, 2006.

[26] J.-M. Lasry and P.-L. Lions, "Mean field games," *Japanese journal of mathematics*, vol. 2, no. 1, pp. 229–260, 2007.

[27] D. A. Gomes, J. Mohr, and R. R. Souza, "Discrete time, finite state space mean field games," *Journal de mathématiques pures et appliquées*, vol. 93, no. 3, pp. 308–328, 2010.

[28] J. Doncel, N. Gast, and B. Gaujal, "Mean-field games with explicit interactions," *Preprint hal-01277098, February*, 2016.

[29] B. A. Neumann, "Stationary equilibria of mean field games with finite state and action space," *Dynamic Games and Applications*, 2020.

[30] E. Elokda, S. Bolognani, A. Censi, E. Frazzoli, and F. Dörfler, "Dynamic population games," 2021, Working Paper.

[31] M. Newman, *Networks: An Introduction*. Oxford University Press, 2010.

[32] J. Filar and K. Vrieze, *Competitive Markov decision processes*. Springer Science & Business Media, 2012.