

Optimal Policy Design for Repeated Decision-Making Under Social Influence

Chiara Ravazzi , *Senior Member, IEEE*, Valentina Breschi , *Member, IEEE*,
Paolo Frasca , *Senior Member, IEEE*, Fabrizio Dabbene , *Fellow, IEEE*,
and Mara Tanelli , *Senior Member, IEEE*

Abstract—In this article, we present a novel model to characterize individual tendencies in repeated decision-making scenarios, with the goal of designing model-based control strategies that promote virtuous choices amidst social and external influences. Our approach builds on the classical Friedkin and Johnsen model of social influence, extending it to include random factors (e.g., inherent variability in individual needs) and controllable external inputs. We explicitly account for the temporal separation between two processes that shape opinion dynamics: individual decision-making and social imitation. While individual decisions occur at regular, frequent intervals, the influence of social imitation unfolds over longer periods. The inclusion of random factors naturally leads to dynamics that do not converge in the classical sense. However, under specific conditions, we prove that opinions exhibit ergodic behavior. Building on this result, we propose a constrained asymptotic optimal control problem designed to foster, on average, social acceptance of a target action within a network. To address the transient dynamics of opinions, we reformulate this problem within a model predictive control (MPC) framework. Simulations highlight the significance of accounting for these transient effects in steering individuals toward virtuous choices while managing policy costs.

Index Terms—Behavioral sciences, closed loop systems, networked control systems, optimal control, social dynamics.

Received 31 July 2025; revised 19 December 2025; accepted 17 February 2026. Date of publication 23 February 2026; date of current version 30 July 2026. This work was supported in part by the European Union – Next Generation EU, Mission 4, Component 1, under the PRIN Project TECHIE: “A control and network-based approach for fostering the adoption of new technologies in the ecological transition” under Grant 2022KPHA24 CUP Master: D53D23001320006 and Grant CUP: B53D23002760006 and in part by PEPR project FORcasting Impacts of Mobility, BACKcasting Optimal Decisions (FORBAC), under Grant ANR-23-PEMO-0002. Recommended by Associate Editor A. A. Malikopoulos. (Corresponding author: Chiara Ravazzi.)

Chiara Ravazzi and Fabrizio Dabbene are with the Institute of Electronics, Computer and Telecommunication Engineering, National Research Council of Italy (CNR-IEIT), Politecnico di Torino, 10129 Torino, Italy (e-mail: chiara.ravazzi@cnr.it; fabrizio.dabbene@cnr.it).

Valentina Breschi is with the Department of Electrical Engineering, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands (e-mail: v.breschi@tue.nl).

Paolo Frasca is with CNRS, Inria, Grenoble INP, GIPSA-Lab, Univ. Grenoble Alpes, F-38000 Grenoble, France (e-mail: paolo.frasca@gipsa-lab.fr).

Mara Tanelli is with the Department of Electronics, Information and Bioengineering, Politecnico di Milano, 20133 Milano, Italy (e-mail: mara.tanelli@polimi.it).

Digital Object Identifier 10.1109/TAC.2026.3667054

I. INTRODUCTION

IN OUR daily routines, we often make repeated decisions. For example, each day we choose the mode of transportation for our commute, considering various factors such as our schedule, weather conditions, economic costs, and even ethical or social norms that influence our preference for specific mobility options. From this simple example, it is already clear that these decisions are heavily influenced by both the social environment one is immersed in and other external factors [1], as well as by our habits and the resistance to change them. Indeed, individuals often demonstrate stubborn behaviors, being hesitant to change their habits, even when it could benefit them in the long run [2].

Changes in one’s habits can nonetheless be fostered by policies that encourage specific virtuous actions, e.g., the daily adoption of sustainable mobility services like bike sharing. Given the interplay of different factors in shaping individual recurrent choices, these policies should account for i) the characteristics of individuals, ii) the complex relationships between them, as well as iii) the consequences of past policy decisions (as discussed in [3] and [4]).

In a related line of research, some recent works started investigating the influence of recommendations on the dynamics of social networks. Based on [5], where a feedback interaction between a recommendation system and a single user is introduced, [6], [7] integrate the dynamics of social interactions and the recommendation system within influence networks. Meanwhile, Sprenger et al. [8] leveraged an extended Friedkin and Johnsen model (see [2]) to introduce a model predictive control (MPC) strategy for user engagement maximization, hence designing a (nonpersonalized) recommendation policy that accounts for the evolution of opinions over time.

Contribution: In this work, our ultimate goal is to design effective nudging policies that favor virtuous actions within repetitive decision-making scenarios that account for the different players shaping the dynamics of opinions. To this end, we propose a new opinion dynamics model based on the Friedkin and Johnsen model, an opinion dynamics model widely recognized by the scientific community and experimentally validated for medium-size networks [9], [10]. Taking advantage of the ability of the model to capture the intricate interplay between individual behaviors and social influences, our model also incorporates *random factors* and *external* (controlled) policies that can shape opinion dynamics. The former represents the variability likely to

affect individual inclination over repetitive choices, as well as the random nature of social acceptance toward specific actions that a policymaker seeks to promote. The latter allows us to incorporate the effect of external policies in shaping personal choices. In the absence of external control actions, the proposed dynamics do not converge but persistently oscillate. Nonetheless, under suitable assumptions, we prove that the resulting stochastic process modeling individual opinions of the system is ergodic. In turn, this implies that, over a sufficiently long period, the time averages of inclinations converge to their expected values.

With this in mind, we propose two optimization-based control frameworks to foster the propagation and acceptance of desired actions within the network on average while containing policy efforts. Specifically, we initially design control actions only looking at the asymptotic behavior of average individual inclinations, then considering also transients by formulating an MPC problem to foster virtuous actions in repetitive settings. Compared with [8], in formulating the policy design problem, the recommender is not intended as a new node in the network able to influence others at each time. Instead, we intend the policymaker as an external influence able to act directly on the bias of individual inclination. This choice allows us i) to design *personalized* policies for each individual, and ii) to explicitly account for the fact that nudging/incentive policies often do not act directly on one's inclination, but rather aim at tearing down specific barriers that prevent a certain virtuous choice from being made (e.g., fidelity discounts for bike-sharing users). Furthermore, unlike prior literature, our policy design approach does not entail observing opinions reflecting inclinations toward acceptance. Instead, we are restricted to observing only their (binary) realizations. This choice, together with the introduction of a budget constraint, allows us to formulate the decision-making problem considering two core practical aspects in policy design, e.g., limited resources and access to limited information on individual opinions. It is worth pointing out that our approach still relies on the knowledge of individual biases, thus restricting the scenarios in which our strategy is directly applicable. Nonetheless, the policies proposed in the article are still valuable as theoretical benchmarks for other policies. Moreover, they are practically relevant, as individual biases can often be estimated from socio-demographic and economic characteristics collected through structured surveys (see, e.g., [11] and [12]). The impact of personalization is evaluated in our numerical example by comparing the proposed policies with a control action obtained by uniformly dividing the available budget across individuals, showcasing the benefit of tailor-made policy actions.

Outline: We introduce the proposed opinion dynamics model in Section II, characterizing its properties in the absence of control and in the presence of a persistent constant control action in Section III. Two constant control policies, respectively obtained by uniformly dividing the available budget across individuals and by optimizing an acceptance/cost tradeoff by looking at the asymptotic behavior of opinions only, are introduced in Section IV. These policies will serve as benchmarks for comparison with the model predictive controller presented in Section V, where we consider both the cases where an oracle provides prior information on the average inclinations and such information is

estimated from the available observations. The different strategies are then analyzed in Section VI on a numerical example. We conclude this article with some remarks and discussion on future developments.

Notation and useful facts: The sets of real and nonnegative integer numbers are denoted by $\mathbb{R}$ and $\mathbb{Z}_{\geq 0}$, respectively. Column vectors are indicated in boldface, while matrices are boldfaced capital letters. Unitary and zero (column) vectors are accordingly denoted as $\mathbf{1}$ and $\mathbf{0}$, respectively. A matrix $\mathbf{M} \in \mathbb{R}^{n \times n}$ is said row-stochastic if its entries are all nonnegative and $\mathbf{M}\mathbf{1} = \mathbf{1}$, while it is Schur stable if all its eigenvalues lay inside the unit circle. Moreover, $[\mathbf{M}]$ denotes a diagonal matrix whose entries are equal to the diagonal of $\mathbf{M}$. Given a random vector $\mathbf{v}$, $\mathbb{E}[\mathbf{v}]$ denotes its expected value.

II. NEW MODEL FOR REPETITIVE DECISION-MAKING SCENARIOS

We consider a social network represented by a directed graph¹ $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{P})$, consisting of a set $\mathcal{V}$ of agents and a set of edges $\mathcal{E} \in \mathcal{V} \times \mathcal{V}$ describing the (directed) influences among them, i.e., if $(v, w) \in \mathcal{E}$, we will say that v is potentially influenced by node $w \in \mathcal{V}$. The intensities of their potential interactions are encoded in a weight matrix $\mathbf{P} \in \mathbb{R}^{|\mathcal{V}| \times |\mathcal{V}|}$, that we assume to be row-stochastic and nonnegative, while the set of agents influencing node $v \in \mathcal{V}$ is denoted with the symbol $\mathcal{N}_v = \{w \in \mathcal{V} : P_{vw} > 0\}$.

Each agent in the network is associated with a pair $(x_v(t), y_v(t)) \in \mathcal{X} \times \mathcal{Y}$ of variables, where $x_v(t) \in \mathcal{X} = [0, 1]$ represents the individual *hidden* inclination to a specific choice while $y_v(t) \in \mathcal{Y} = \{0, 1\}$ is the effective realization of such choice at time $t \in \mathbb{N}$. Specifically, $x_v(t) \approx 0$ indicates opposition to making the target choice, and $x_v(t) \approx 1$ denotes support for it. Instead, $y_v(t) = 1$ means that the agent makes the target choice at time $t \in \mathbb{Z}$ and $y_v(t) = 0$ otherwise indicates that the agent has not made it.

We assume that each agent $v \in \mathcal{V}$ starts from an *initial belief* $x_v(0) = x_0$ that, at time step $t \in \mathbb{Z}_{\geq 0}$, is updated according to two possible mechanisms.

The hidden inclination of each agent $v \in \mathcal{V}$ evolves as a response to social interactions and external factors with probability $\alpha_v \in [0, 1]$, with the updated opinion becoming a convex combination of the opinions coming from the neighbors and the external influence. In particular, while the weight of the influence on $v \in \mathcal{V}$ from each neighbor $w \in \mathcal{N}_v$ is given by P_{vw} , the overall contribution of the neighbors is controlled by the parameter $\lambda_v \in [0, 1]$ and the updated opinion is given by

$$x_v(t+1) = \lambda_v \sum_{w \in \mathcal{N}_v} P_{vw} x_w(t) + (1 - \lambda_v) \eta_v(t) \quad (1)$$

where $\eta_v(t) \in [0, 1]$ encodes the effect of external influences and individual biases at time t . Instead, the agent's opinion is updated irrespective of social interaction with probability $1 - \alpha_v$, i.e., it becomes simply a convex combination of the previous individual

¹To avoid trivial cases, we assume that the graph has at least three nodes.

opinion and external influences/initial bias as

$$x_v(t+1) = \lambda_v x_v(t) + (1 - \lambda_v) \eta_v(t), \quad v \in \mathcal{V} \quad (2)$$

where $\lambda_v \in [0, 1]$ now has the role of a relative weight that quantifies the effect of the external influence and initial biases on the evolution of the opinion. It is worth remarking that the state variable $x_v(t)$ is interpreted as an individual propensity or inclination, which is a consequence of the modeling assumption (being a convex combination of quantities in $[0, 1]$), consistent with the semantics of the dynamics and with the structure of the bias terms.

In this dynamics, we decompose the last term $\eta_v(t)$ as follows:

$$\eta_v(t) = \eta_v^0 + \eta_v^{\text{nc}}(t) + u_v(t) \quad (3)$$

where $\eta_v^0 \in [0, 1]$ is a deterministic constant indicating the *individual bias* of the agent that, in the same spirit of the Friedkin and Johnsen model [2], steers individual opinion along with the social influence in the absence of external influences. Instead, the uncontrollable input $\eta_v^{\text{nc}}(t)$ represents possible (yet unpredictable) variations in the individual intrinsic bias (e.g., changes in the weather and, thus, the individual choice not to take a shared bike) and we will assume that $\eta_v^0 + \eta_v^{\text{nc}}(t)$ will be naturally in the range in $[0, 1]$ for all t and for all $v \in \mathcal{V}$. Meanwhile, the *control input*

$$u_v(t) \in [0, 1 - \eta_v^0 - \eta_v^{\text{nc}}(t)], \quad v \in \mathcal{V} \quad (4)$$

is what the policymaker can shape to nudge individual choices toward the target one, potentially in a personalized fashion. Note that these inputs could represent the normalized effects of incentive strategies or advertising campaigns (see, e.g., [11] for a preliminary approach to translate control inputs and actual policy actions). Throughout the remainder of this article, we will assume that policymakers have a limited budget β at their disposal to nudge agents' choices, i.e., the control input vector has to satisfy the following:

$$\mathbf{1}^\top \mathbf{u}(t) \leq \beta, \quad \forall t \in \mathbb{Z}_{\geq 0}. \quad (5)$$

By defining the vector $\mathbf{x}(t) \in [0, 1]^{|\mathcal{V}|}$ grouping the inclinations of all agents in the network and the vector $\boldsymbol{\eta}(t) \in [0, 1]^{|\mathcal{V}|}$ of all external factors and biases, the previous dynamics can be written in matrix form as

$$\mathbf{x}(t+1) = \mathbf{\Lambda} \mathbf{P}(t) \mathbf{x}(t) + (\mathbf{I} - \mathbf{\Lambda}) \boldsymbol{\eta}(t) \quad (6a)$$

where $\mathbf{P}(t) = \text{diag}(\mathbf{z}(t)) \mathbf{P} + (\mathbf{I} - \text{diag}(\mathbf{z}(t))) \mathbf{I}$ with

$$z_v(t) = \begin{cases} 1, & \text{w.p. } \alpha \\ 0, & \text{w.p. } 1 - \alpha \end{cases} \quad (6b)$$

and $\mathbf{\Lambda}$ is a diagonal matrix with elements equal (and equally ordered) to parameters $\{\lambda_v\}_{v \in \mathcal{V}}$.

Remark 1 (On the features of $\mathbf{P}$): If the bias $\eta_v(t)$ is persistently less than 1/2 for some agents $v \in \mathcal{V}$, the convex combination in the dynamics will tend to bring their opinions below the decision threshold, regardless of the influence of neighbors. Thus, the assumption that $\mathbf{P}$ has nonnegative elements does not imply that all trajectories converge to 1, as the presence of heterogeneous and potentially low values in the bias vector $\boldsymbol{\eta}(t)$,

combined with the random activation of social influence, leads to nontrivial equilibrium behavior.

Assumption 1: The uncontrollable inputs $\{\eta_v^{\text{nc}}(t)\}_{t \in \mathbb{Z}_{\geq 0}}$ are independent, identically distributed (i.i.d.) random vectors over time, have mutually independent components across agents, with zero mean and bounded support, so that, for every agent $v \in \mathcal{V}$ and every $t \in \mathbb{Z}_{\geq 0}$, the sum of the intrinsic bias and the uncontrollable disturbance remains a valid inclination, i.e., $\eta_v^0 + \eta_v^{\text{nc}}(t) \in [0, 1]$.

It is worth remarking that Assumption 1 is reasonable since η_v^0 represents a baseline inclination in $[0, 1]$, whereas $\eta_v^{\text{nc}}(t)$ captures small, zero-mean fluctuations around this baseline due to uncontrollable factors. Note that, assuming $\eta_v^{\text{nc}}(t)$ to be zero-mean implies that uncontrollable perturbations do not systematically push initial inclinations in one direction on average. Indeed, they solely introduce variability around the intrinsic bias η_v^0 . Meanwhile, requiring bounded support and $\eta_v^0 + \eta_v^{\text{nc}}(t) \in [0, 1]$ a.s. ensures that these fluctuations do not push an individual's inclination outside the range $[0, 1]$.

Moreover, the independence across nodes at each time step is not necessary, but from now on it is assumed for ease of tractability and could be relaxed in more general settings. The assumption that changes are not correlated among different individuals mimicks the fact that ultimately an external factor (e.g., an unexpected and temporary changes in one's daily routine) can have a different impact on one's choices (e.g., renting a shared bike).

Remark 2 (Dynamics and external factors): Assuming that the uncontrollable inputs are i.i.d. and independent of the state $\mathbf{x}(t)$ is crucial to prove the properties of the dynamics in (6) analyzed in Section III. Meanwhile, the assumption on the bounds is mainly exploited in policy design.

Finally, we assume that the measured acceptance variables of the target choice, grouped in the vector $\mathbf{y}(t) \in \{0, 1\}^{|\mathcal{V}|}$, are random variables as stated formally below.

Assumption 2 (Acceptance variables): Given $\mathbf{x}(t) \in [0, 1]$ for $t \in \mathbb{N}$, we assume the acceptance variables are independent random variables whose conditional distribution satisfies

$$\mathbb{P}(y_v(t) = 1 | \mathbf{x}(t)) = x_v(t) \quad (7a)$$

$$\mathbb{P}(y_v(t) = 0 | \mathbf{x}(t)) = 1 - x_v(t) \quad (7b)$$

for all $t \in \mathbb{Z}_{\geq 0}$.

Note that introducing these acceptance variables toward a target choice that a policy-maker intends to promote represents a situation in which only a noisy measure of individuals' opinions is available. In other words, the actual opinions of agents are not directly observable, and the available information is simply related to the manifested individual choices. It is worth remarking that the acceptance variables are independent across different $v \in \mathcal{V}$. Nonetheless, the framework could be extended to the case where the acceptance variables are correlated, provided that the condition $\mathbb{E}[\mathbf{y} | \mathbf{x}] = \mathbf{x}$ is preserved.

A. Novelty and Relation to Prior Works

The dynamics in (6a) is a novel extension of the classical Friedkin and Johnsen model [2] as well as its variations in

opinion dynamics [13]. The main distinguishing features of the model with respect to existing ones are: i) the presence of random factors, i.e., the random variable $\eta^{\text{nc}}(t)$, $t \in \mathbb{Z}_{\geq 0}$, modeling variability in individual inclinations due to external (uncontrollable) events; ii) the introduction of the dependence on a controlled input (through $u(t)$, $t \in \mathbb{Z}_{\geq 0}$), mimicking the impact of policy actions to nudge a target choice; iii) the probabilistic update of opinions, according to the parameter α in (6b).

Looking at how the controlled input impacts the evolution of opinions with respect to [8], in our model, the policymaker is not intended as a new node introduced in the network, able to influence individuals at each time with an equal action, but is meant as a (possibly personalized) external influence. In this sense, the control represents actions aimed at modifying the intrinsic bias of agent $v \in \mathcal{V}$, through, e.g., incentive strategies and information campaigns, rather than modifying their opinion directly.

Meanwhile, considering (6b) allows us to account for the fact that two processes, opinion updates related to social imitation and those related to individual decision-making, can act on different time scales. In fact, social influence or pressure resulting from interactions that individuals have with their neighbors generally takes longer to have a significant effect, e.g., several days, weeks, or even months. Instead, individual decision-making may occur more frequently and on a shorter time scale, e.g., decisions may be made on a daily basis. The parameter α in (6b) is used in our model to model the frequency of these two updating mechanisms.

Lastly, our choice of introducing the acceptance variables (see Assumption 2) allows us to distinguish our work from the literature related to robust and stochastic control [14]. Indeed, instead of designing controllers taking into account that the opinion dynamics are uncertain and not perfectly observable, our approach is based on the framework of iterated random systems. In turn, this allows us to leverage a rich ergodic theory that can be applied to study the system's long-run behavior and design nudging policies.

B. Illustrative Scenarios

Our model covers several aspects that customarily describe how people make repetitive decisions, e.g., daily choices of recycling garbage or using car/bike sharing services for daily commutes.

Indeed, one's choices are generally dictated by an intrinsic bias that is due to one's habits and/or socioeconomic status, which inherently makes one more or less inclined to repetitively make the same decision or adopt a certain service (see the discussion in [15]). The impact of such a bias can be mitigated by policies enacted by governments and stakeholders to favor new daily habits or facilitate access to a new service [16], e.g., charging fees to people not recycling or providing loyalty discounts for sharing services. In practice, these policies are nonetheless generally designed under budget constraints, capping investments that can be undertaken to nudge individuals to the target choice. This aspect is accounted for in our framework through the constraint in (5), in turn allowing us to consider

more realistic limited budget scenarios. At the same time, uncontrollable events can temporarily shape one's inclination, e.g., weather can either hinder or stimulate the use of bike-sharing services, depending on whether it is bad or good.

These external factors contribute to shaping individual opinions together with the social environment in which one is immersed. Indeed, in reality, acceptance of new habits or services is also driven by social dictates or homophily [17], with individual opinions being shaped by what is perceived as "good" or "bad" by the society as well as influenced by the benefits/downside of the targeted choice experiment by peers. Since it requires opinions on the targeted choice to form and shape as well as others make a certain choice or try services out, social influence is nonetheless likely to steer individual opinions on a different time scale with respect to the evolution of one inclination based on individual needs and irrespective of social dictates.

Ultimately, when analyzing and designing policy strategies, policymakers and stakeholders generally devise or have access to surveys (see, e.g., [18]) that allow them eventually to unveil the socioeconomic status and habits of an individual as well as the manifested acceptance/nonadoption of the targeted choice, but they generally do not give direct access to how one's opinion truly evolves. This is considered in our model through the introduction of the variable $y(t)$, $t \in \mathbb{Z}_{\geq 0}$, linked to the *hidden variables* thanks to Assumption 2.

III. ASYMPTOTIC ANALYSIS

We first present a theoretical analysis of the behavior of the dynamics (6a) in the absence of control, i.e., when $u(t) = 0$ for all $t \in \mathbb{Z}_{\geq 0}$. Despite this, we stress that two different sources of randomness have to be considered in this analysis. First, the social imitation mechanism acts in an intermittent way with a certain probability α . Second, the variation in the individual bias is still characterized by $\eta^{\text{nc}}(t)$, $t \in \mathbb{Z}_{\geq 0}$, that captures the intrinsic volatility of the conditions at each time instant. These two components ultimately make both the latent and acceptance variables continue to oscillate over time. In addition, as $y(t)$ are random variables for all $t \in \mathbb{Z}_{\geq 0}$, the acceptance variables introduce another source of uncertainty. Nonetheless, under Assumption 2 and an additional hypothesis on the topology of the influence network, we will show that both the latent and acceptance variables converge to a final limit profile.

Specifically, we make the following assumption on the topology of the influence network associated with P .

Assumption 3 (Network topology): For any node $v \in \mathcal{V}$, there exists a path from $v \in \mathcal{V}$ to a node $s \in \mathcal{V}$ such that $\lambda_s < 1$.

This assumption is common in opinion dynamics (see, e.g., [13]) and ensures that each node is connected (directly or indirectly) to a node with persistent bias (i.e., with $\lambda_s < 1$). Therefore, it allows us to avoid cases in which a node is completely isolated, and therefore not influenced by others, even indirectly, while being sufficient to guarantee that $\Lambda P(t)$ is, on average, Schur stable. At the same time, Assumption 3 is not particularly restrictive, as it encompasses all cases where the graph is strongly connected, as well as all cases in which all parameters are strictly smaller than 1. For instance, the special

case where $\lambda_v < 1$ for every $v \in \mathcal{V}$ simply implies that each user has some internal resistance to adopting virtuous behaviors.

By guaranteeing Schur stability on average of $\Lambda P(t)$, Assumption 3 ultimately allows the expected opinions dynamics to converge as shown in the following proposition.

Proposition 1 (Expected hidden inclinations' dynamics): Let $\mathbf{u}(t) = 0$ for all $t \in \mathbb{Z}_{\geq 0}$. Under Assumptions 1 and 3, the dynamics in (6) satisfy

$$\mathbb{E}[\mathbf{x}(t+1)] = \Lambda \bar{P} \mathbb{E}[\mathbf{x}(t)] + (\mathbf{I} - \Lambda) \boldsymbol{\eta}^0 \quad (8)$$

for any initial condition $\mathbf{x}(0) \in \mathbb{R}^{|\mathcal{V}|} \in [0, 1]$ and any $t \in \mathbb{Z}_{\geq 0}$. Moreover, $\lim_{t \rightarrow \infty} \mathbb{E}[\mathbf{x}(t)] = \mathbf{x}^*$ with

$$\mathbf{x}^* := \mathbb{E}[\mathbf{x}(\infty)] = (\mathbf{I} - \Lambda \bar{P})^{-1} (\mathbf{I} - \Lambda) \boldsymbol{\eta}^0 \quad (9)$$

and $\bar{P} := (\mathbf{I} - \text{diag}(\boldsymbol{\alpha})) \mathbf{I} + \text{diag}(\boldsymbol{\alpha}) P$.

Proof: By conditioning, from (2) we get

$$\begin{aligned} \mathbb{E}[\mathbf{x}(t+1)] &= \mathbb{E}[\mathbb{E}[\mathbf{x}(t+1)|\mathbf{x}(t)]] \\ &= \mathbb{E}[\mathbb{E}[\Lambda P(t) \mathbf{x}(t) | \mathbf{x}(t)] + (\mathbf{I} - \Lambda) \boldsymbol{\eta}^0] \\ &= \Lambda \bar{P} \mathbb{E}[\mathbf{x}(t)] + (\mathbf{I} - \Lambda) \boldsymbol{\eta}^0 \end{aligned}$$

where the second equality stems from Assumption 1 and $\bar{P} := (\mathbf{I} - \text{diag}(\boldsymbol{\alpha})) \mathbf{I} + \text{diag}(\boldsymbol{\alpha}) P$. Meanwhile, based on Assumption 3, it follows that $\Lambda \bar{P}$ is Schur stable, thus concluding the proof. $\square$

By relying on this result, as well as on Assumption 2, we can then formalize the following result on the acceptance variables.

Corollary 1 (Expected acceptance variables): Let $\mathbf{u}(t) = 0$ for all $t \in \mathbb{Z}_{\geq 0}$. Under Assumption 1–3 it holds that $\mathbb{E}[\mathbf{y}(t+1)] = \mathbb{E}[\mathbf{x}(t+1)]$ and

$$\lim_{t \rightarrow \infty} \mathbb{E}[\mathbf{y}(t)] = \mathbf{x}^* \quad (10)$$

for any initial condition $\mathbf{x}(0) \in \mathbb{R}^{|\mathcal{V}|} \in [0, 1]$.

Proof: The proof follows straightforwardly from Assumption 2 and Proposition 1. $\square$

These two results imply that opinions in a social system evolving according to (6a) are stable in expectation in the absence of externally controlled action. Moreover, they show that, in the absence of controlled inputs, a final opinion limit profile emerges that is a combination of the expected value of exogenous (uncontrollable) inputs, i.e., $\boldsymbol{\eta}_0(t)$ for $t \in \mathbb{Z}_{\geq 0}$, and the expected impact of social interactions.

After introducing these properties, we can now summarize the features of the dynamics in (6) as follows.

Given a stochastic process [19] $\{\mathbf{x}(t)\}_{t \in \mathbb{Z}_{\geq 0}}$, its time-average (also Cesàro or Polyak average in some contexts [20]) is denoted by

$$\bar{\mathbf{x}}(t) := \frac{1}{t} \sum_{\ell=0}^{t-1} \mathbf{x}(\ell). \quad (11)$$

Theorem 1 (Ergodicity of hidden inclinations): Consider the random process $\{\mathbf{x}(t)\}_{t \in \mathbb{Z}_{\geq 0}}$ defined in (6) and let $\mathbf{u}(t) = 0$, $\forall t \in \mathbb{Z}_{\geq 0}$. Under Assumptions 1 and 3, the following hold.

- 1) $\mathbf{x}(k)$ converges in distribution to random variables $\mathbf{x}_\infty$ and the distribution of $\mathbf{x}_\infty$ is the unique invariant distribution for (6).

- 2) The process is ergodic, i.e. $\bar{\mathbf{x}}(t) \rightarrow \mathbb{E}[\mathbf{x}_\infty]$ almost surely.
- 3) The limit random variables satisfy $\mathbb{E}[\mathbf{x}_\infty] = \mathbf{x}^*$.

Proof: Let us rewrite (6a) as

$$\mathbf{x}(t+1) = \mathbf{A}(t) \mathbf{x}(t) + \mathbf{b}(t) \quad (12)$$

with $\mathbf{A}(t) = \Lambda P(t)$ and $\mathbf{b}(t) = (\mathbf{I} - \Lambda) \boldsymbol{\eta}(t)$. As we assume that the controlled input is zero, then from Assumption 1 it holds that $\mathbb{E}[\mathbf{b}(t)] = (\mathbf{I} - \Lambda) \boldsymbol{\eta}^0$, which is constant. From Assumptions 1 and 3, we further obtain that $\mathbb{E}[\mathbf{A}(t)] = \Lambda \bar{P}$ is Schur stable. By direct application of [21, Thm. 1] we deduce that $\mathbf{x}(k)$ converges in distribution to a random variable $\mathbf{x}_\infty$, and the distribution of $\mathbf{x}_\infty$ is the unique invariant distribution for (12). The process (12) is ergodic and the limit random variable satisfies $\mathbb{E}[\mathbf{x}_\infty] = \mathbf{x}^*$. $\square$

The following theorem guarantees that the ergodicity of the latent opinions implies the ergodicity of the acceptance variables.

Theorem 2 (Convergence of acceptance variables): Consider the random processes $\{\mathbf{x}(t)\}_{t \in \mathbb{Z}_{\geq 0}}$ defined in (6a) and $\{\mathbf{y}(t)\}_{t \in \mathbb{Z}_{\geq 0}}$. Suppose that $\mathbf{u}(t) = 0$ for all $t \in \mathbb{Z}_{\geq 0}$. Then, under Assumptions 1–3, the following hold.

- 1) $\{\mathbf{y}(t)\}_{t \in \mathbb{Z}_{\geq 0}}$ converges in distribution to random variables $\mathbf{y}_\infty$ that is distributed as a Bernoulli distribution with parameter $\mathbf{x}_\infty$. In particular, $\mathbb{E}[\mathbf{y}_\infty] = \mathbf{x}^*$.
- 2) The process $\{\mathbf{y}(t)\}_{t \in \mathbb{Z}_{\geq 0}}$ is ergodic, i.e., $\bar{\mathbf{y}}(t) \rightarrow \mathbb{E}[\mathbf{y}_\infty]$ almost surely.

Proof: It is well known that $X_n \xrightarrow{D} X$ (convergence in distribution) if and only if for every bounded and continuous function $h : \mathbb{R} \rightarrow \mathbb{R}$, we have: $\mathbb{E}[h(X_n)] \rightarrow \mathbb{E}[h(X)]$. By exploiting Assumption 2, for every bounded and continuous function $h(\cdot)$, it holds that

$$\begin{aligned} \mathbb{E}[h(y_v(t))] &= \mathbb{E}[\mathbb{E}[h(y_v(t)) | x_v(t)]] \\ &= \mathbb{E}[h(1)x_v(t) + h(0)(1 - x_v(t))] \\ &= h(1)\mathbb{E}[x_v(t)] + h(0)\mathbb{E}[1 - x_v(t)]. \end{aligned}$$

Since $\mathbf{x}(t)$ converges to $\mathbf{x}_\infty$ in distribution (see Theorem 1), for the properties of $h(\cdot)$ we have

$$\begin{aligned} \lim_{t \rightarrow \infty} \mathbb{E}[h(y_v(t))] &= h(1)\mathbb{E}[x_v(\infty)] + h(0)\mathbb{E}[1 - x_v(\infty)] \\ &= \mathbb{E}[h(y_v(\infty))] \end{aligned}$$

in turn, implying that $\lim_{t \rightarrow \infty} \mathbb{E}[h(\mathbf{y}(t))] = \mathbb{E}[h(\mathbf{y}_\infty)]$. From this fact, we conclude that $\{\mathbf{y}(t)\}_{t \in \mathbb{Z}_{\geq 0}}$ converges in distribution to $\mathbf{y}_\infty$, which is distributed as a Bernoulli distribution with parameter $\mathbf{x}_\infty$.

To complete the proof, it remains to show that the process $\{\mathbf{y}(t)\}_{t \in \mathbb{Z}_{\geq 0}}$, where $\mathbf{y}(t) \sim \text{Bernoulli}(\mathbf{x}(t))$, is also ergodic, i.e.,

$$\lim_{t \rightarrow \infty} \frac{1}{t} \sum_{\ell=0}^{t-1} \mathbf{y}(\ell) = \mathbb{E}[\mathbf{y}_\infty] = \mathbf{x}^* \quad \text{almost surely.}$$

By the Chernoff and union bound [22], for any $\epsilon > 0$ the following holds:

$$\mathbb{P}\left(\left\|\frac{1}{t} \sum_{\ell=0}^{t-1} \mathbf{y}(\ell) - \frac{1}{t} \sum_{\ell=0}^{t-1} \mathbb{E}[\mathbf{y}(\ell)]\right\|_1 \geq \epsilon\right) \leq 2|\mathcal{V}| \exp\left(-\frac{\epsilon^2 t}{2|\mathcal{V}|}\right).$$

If we choose $\epsilon = t^{-1/2+\alpha}$ for any $\alpha > 0$, then

$$\sum_{t=1}^{\infty} 2|\mathcal{V}| \exp\left(-\frac{\epsilon_t^2 t}{2|\mathcal{V}|}\right) = \sum_{t=1}^{\infty} 2|\mathcal{V}| \exp\left(-\frac{t^{2\alpha}}{2|\mathcal{V}|}\right) < \infty$$

the Borel-Cantelli Lemma (see [23, Thm. 1.4.2]) implies that with probability one $\|\frac{1}{t} \sum_{\ell=0}^{t-1} \mathbf{y}(\ell) - \frac{1}{t} \sum_{\ell=0}^{t-1} \mathbb{E}[\mathbf{y}(\ell)]\|_1$ for all but finitely many values of $t \in \mathbb{N}$. Therefore, almost surely $\|\frac{1}{t} \sum_{\ell=0}^{t-1} \mathbf{y}(\ell) - \frac{1}{t} \sum_{\ell=0}^{t-1} \mathbb{E}[\mathbf{y}(\ell)]\|_1$ converges to zero as t goes to infinity. Finally, from triangular inequality and since $\mathbb{E}[\mathbf{y}(\ell)] = \mathbb{E}[\mathbf{x}(\ell)]$ for all $\ell \in \mathbb{N}$, we have

$$\begin{aligned} \left\| \frac{1}{t} \sum_{\ell=0}^{t-1} \mathbf{y}(\ell) - \mathbf{x}^* \right\|_1 &\leq \left\| \frac{1}{t} \sum_{\ell=0}^{t-1} \mathbf{y}(\ell) - \frac{1}{t} \sum_{\ell=0}^{t-1} \mathbb{E}[\mathbf{y}(\ell)] \right\|_1 \\ &\quad + \left\| \frac{1}{t} \sum_{\ell=0}^{t-1} \mathbb{E}[\mathbf{x}(\ell)] - \mathbf{x}^* \right\|_1. \end{aligned}$$

Since the first term on the right-hand side converges to zero almost surely as $t \rightarrow \infty$ from the arguments above and $\{\mathbf{x}(t)\}_{t \in \mathbb{Z}_{\geq 0}}$ is ergodic (see Theorem 1), we have that

$$\frac{1}{t} \sum_{\ell=0}^{t-1} \mathbf{y}(\ell) \rightarrow \mathbb{E}[\mathbf{x}_{\infty}] = \mathbb{E}[\mathbf{y}_{\infty}] = \mathbf{x}^* \text{ almost surely as } t \rightarrow \infty$$

and hence $\{\mathbf{y}(t)\}_{t \in \mathbb{Z}_{\geq 0}}$ is ergodic, thus concluding the proof. $\square$

Remark 3 (Open-loop properties and control actions): The properties summarized in Theorems 1–2 are still valid if one feeds the social system with a constant control input over time, i.e., $\mathbf{u}(t) = \mathbf{u}$ for all $t \in \mathbb{Z}_{\geq 0}$.

Based on this consideration, in the next section, we focus on the design of constant nudging policies, which preserve the ergodicity of both the hidden and manifested inclinations and are then used as a benchmark in Section VI against an (optimal) time-varying one.

IV. ASYMPTOTIC BEHAVIOR WITH CONSTANT CONTROL ACTION: MODEL-FREE VERSUS MODEL-BASED APPROACH

We now introduce two possible constant policies that the policymaker/stakeholder can enact to nudge a target choice. The first one is a *model-free* constant control policy (MFCCP), obtained by simply dividing the overall budget β [see (5)] in a uniform way. The second is an *optimal* policy designed by accounting for the asymptotic behavior of the dynamics in (6) and the fact that only the acceptance variables (see Assumption 2) are accessible. In both cases, one would require the (generally unrealistic) knowledge of the stochastic disturbance $\boldsymbol{\eta}^{\text{nc}}(t)$ to account for it by design while satisfying $\mathbf{x}(t) \in [0, 1]$ at all $t \in \mathbb{Z}_{\geq 0}$. Based on Assumption 1, we nonetheless have $\boldsymbol{\eta}^0 + \boldsymbol{\eta}^{\text{nc}}(t) \leq \boldsymbol{\eta}' \leq \mathbf{1}$ almost surely. Therefore, in designing our policies, we take a (simplistic and conservative) approach by imposing the designed policy to satisfy

$$\mathbf{u}(t) \in [0, 1 - \boldsymbol{\eta}'] \quad (13)$$

assuming $\boldsymbol{\eta}'$ to be known.

The most straightforward policy that could be deployed thus corresponds to a uniform distribution of the available resources

[see (5)], eventually saturated based on the individual worst-case bound. This policy can be formalized as follows and we will refer to it as MFCCP.

Policy 1 (MFCCP): The MFCCP is given by

$$u_v(t) = u_v^{\text{MF}} = \min \left\{ 1 - \eta'_v, \frac{\beta}{|\mathcal{V}|} \right\} \quad \forall v \in \mathcal{V}, \forall t \in \mathbb{Z}_{\geq 0}. \quad (14)$$

However, due to its nature, a constant policy like MFCCP is “blind” with respect to the imitation and influence effect due to interactions in the social network, which can instead be leveraged to nudge the target choice achieving the compromise generally sought by policymakers and stakeholders, i.e., maximizing the acceptance of a specific choice (i.e., maximizing the policy’s *efficiency*) while minimizing the costs of a policy. Indeed, disregarding social interactions under budget constraints may cause excessive (eventually unneeded) resource investments in some individuals that can be more “easily” influenced by others, while sufficient efforts are not invested in nudging other ones that instead would need more effort on the policymaker/stakeholder side to be convinced.

To account for social interactions in pursuing the compromise between the two opposite objectives of maximal efficiency and minimal cost, we introduce a personalized constant control policy, which we will refer to as model-based constant control policy (MBCCP). The latter is designed by leveraging the result in Theorem 2 and Remark 3 to optimally achieve a tradeoff between the asymptotic average acceptance of the target choice and policy effort, as follows.

Policy 2 (MBCCP): The MBCCP is obtained as

$$\begin{aligned} \mathbf{u}(t) = \mathbf{u}^{\text{MB}} &= \underset{\mathbf{u}}{\operatorname{argmin}} J_{\text{MB}}(\mathbf{u}) \\ \text{s.t. } \mathbf{1}^\top \mathbf{u} &\leq \beta, \mathbf{u} \in [0, 1 - \boldsymbol{\eta}'] \end{aligned} \quad (15a)$$

for all $t \in \mathbb{Z}_{\geq 0}$, with

$$J_{\text{MB}}(\mathbf{u}) := \mathbb{E} [\|\mathbf{1} - \mathbf{y}_{\infty}\|_{\mathbf{Q}}^2 + \|\mathbf{u}\|_{\mathbf{R}}^2] \quad (15b)$$

and $\mathbf{Q}, \mathbf{R}$ being user-defined positive definite matrices.

The policy design objective in (15b) can further be proven to be solely a function of the policy $\mathbf{u}$ to be designed and the initial individual bias $\boldsymbol{\eta}_0$ as formalized in the following result.

Proposition 2: Under Assumptions 1–3, the control objective in (15b) can be equivalently recast as

$$J_{\text{MB}}(\mathbf{u}) = \|\mathbf{u}\|_{\mathbf{W}}^2 + \mathbf{c}^\top \mathbf{u} + \gamma \quad (16a)$$

where

$$\mathbf{W} = \mathbf{R} + \mathbf{V}^\top (\mathbf{Q} - [\mathbf{Q}]) \mathbf{V} \quad (16b)$$

$$\mathbf{c}^\top = \mathbf{1}^\top ([\mathbf{Q}] - 2\mathbf{Q}) \mathbf{V} + 2\boldsymbol{\eta}_0^\top (\mathbf{Q} - [\mathbf{Q}]) \mathbf{V} \quad (16c)$$

$$\gamma = \mathbf{1}^\top \mathbf{Q} \mathbf{1} + \boldsymbol{\eta}_0^\top \mathbf{V}^\top (\mathbf{Q} - [\mathbf{Q}]) \mathbf{V} \boldsymbol{\eta}_0 + \mathbf{1}^\top ([\mathbf{Q}] - 2\mathbf{Q}) \mathbf{V} \boldsymbol{\eta}_0 \quad (16d)$$

with $\mathbf{V} = (\mathbf{I} - \boldsymbol{\Lambda} \bar{\mathbf{P}})^{-1} (\mathbf{I} - \boldsymbol{\Lambda})$.

Proof: Let us rewrite the first term in (15b) as follows:

$$\begin{aligned} \mathbb{E} [\|\mathbf{1} - \mathbf{y}_{\infty}\|_{\mathbf{Q}}^2] &= \mathbb{E} [\mathbb{E} [\|\mathbf{1} - \mathbf{y}_{\infty}\|_{\mathbf{Q}}^2 | \mathbf{x}_{\infty}]] \\ &= \|\mathbf{1} - \mathbf{x}^*\|_{\mathbf{Q}}^2 + \langle \mathbf{x}^*, \mathbf{1} - \mathbf{x}^* \rangle_{[\mathbf{Q}]} \end{aligned} \quad (17)$$

where the second equality follows from Theorem 2. Due to the presence of the constant input $\mathbf{u}$ to be optimized, similarly to (9), $\mathbf{x}^*$ now satisfies

$$\begin{aligned}\mathbf{x}^* &= \Lambda \bar{\mathbf{P}} \mathbf{x}^* + (\mathbf{I} - \Lambda)(\boldsymbol{\eta}_0 + \mathbf{u}) \\ \Rightarrow \mathbf{x}^* &= (\mathbf{I} - \Lambda \bar{\mathbf{P}})^{-1}(\mathbf{I} - \Lambda)(\boldsymbol{\eta}_0 + \mathbf{u}) = \mathbf{V}(\boldsymbol{\eta}_0 + \mathbf{u}).\end{aligned}$$

Putting the expression $\mathbf{x}^*$ into (15b) and grouping quadratic terms in $\mathbf{u}$, linear terms in $\mathbf{u}$ and those that are instead independent of the quantity to be optimized, the result straightforwardly follows. $\square$

This reformulation highlights that the MBCCP is the solution to a quadratic program (QP) with linear inequality constraints, that can be efficiently tackled with several well-known approaches, including interior point methods, active set, augmented Lagrangian, and conjugate gradient methods [24], to mention just a few.

Based on the reformulation of the loss provided in Proposition 2, it would also be possible to find explicit expressions for $\mathbf{u}_{\text{MB}}$ via the Karush–Kuhn–Tucker (KKT) conditions associated with (15), i.e.,

$$2\mathbf{W}\mathbf{u}^{\text{MB}} + \mathbf{c} + \nu_1 \mathbf{1} - \nu_2 + \nu_3 = 0 \quad (18a)$$

$$\nu_1 \geq 0, \nu_{2,v} \geq 0, \nu_{3,v} \geq 0 \quad \forall v \in \{1, \dots, |\mathcal{V}|\} \quad (18b)$$

$$\nu_1(\mathbf{1}^\top \mathbf{u}^{\text{MB}} - \beta) - \sum_{v=1}^{|\mathcal{V}|} \nu_{2,v} u_v^{\text{MB}} + \sum_{v=1}^{|\mathcal{V}|} \nu_{3,v} (u_v^{\text{MB}} - 1 + \eta'_v) = 0 \quad (18c)$$

$$\mathbf{1}^\top \mathbf{u}^{\text{MB}} - \beta \leq 0, -\mathbf{u}^{\text{MB}} \leq \mathbf{0}, \mathbf{u}^{\text{MB}} \leq \mathbf{1} - \boldsymbol{\eta}' \quad (18d)$$

where ν_1 is the Lagrange multiplier associated with the budget constraint, while ν_2 and ν_3 stack the multipliers associated with the lower and upper bounds on each component of $\mathbf{u}^{\text{MB}}$ dictated by (13), respectively. Accordingly, there are several possible scenarios for the (constant) value ultimately taken by $\mathbf{u}_{\text{MB}}$, depending on the combination of active constraints induced by the user-defined matrices $\mathbf{Q}$ and $\mathbf{R}$ [see (15b)], as well as the available budget β .

Remark 4 (MBCCP with inactive constraints): Whenever the constraints are not active [i.e., the conditions in (18d) are strict inequalities], the Lagrange multipliers have to be zero for (18b) and (18c) to be satisfied. This result ultimately leads to $\mathbf{u}^{\text{MB}} = -\frac{1}{2}\mathbf{W}^{-1}\mathbf{c}$. Given a fixed budget β , the previous result implies the following condition:

$$-\frac{1}{2} \begin{bmatrix} \mathbf{1}^\top \\ -\mathbf{I} \\ \mathbf{I} \end{bmatrix} \mathbf{W}(\mathbf{Q}, \mathbf{R})^{-1} \mathbf{c}(\mathbf{Q}) < \begin{bmatrix} \beta \\ \mathbf{0} \\ \mathbf{1} - \boldsymbol{\eta}' \end{bmatrix} \quad (19)$$

where we have explicitly highlighted the dependence of the weights $\mathbf{W}$ and $\mathbf{c}$ on the user-defined ones, i.e., $\mathbf{Q}$ and $\mathbf{R}$ [see (15b)]. Such a condition can hence be checked by the policymaker/stakeholder to understand whether the model-based constant policy they design allows them i) not to consume the overall budget, ii) to nudge everyone, and iii) to be careful with respect to external disturbances.

As already pointed out in Remark 3, whether the policymaker decides to adopt the MFCCP in (14) or the MBCCP, the stability in expectation of the opinion dynamics, ergodicity of both the hidden inclination and the manifest variables, as well as the convergence in the mean square sense of the dynamics, straightforwardly follow with minor adjustments (i.e., replacing $\boldsymbol{\eta}_0$ with $\boldsymbol{\eta}_0 + \mathbf{u}$) from Theorems 1–2.

Remark 5 (What the policymaker should know): To design the MBCCP and the predictive policies introduced in the next section, a policymaker should have access to (at least) an estimate of the parameters in (6) and the initial biases of each agent. Such estimates can be obtained from data by leveraging surveys (see, e.g., the approach adopted in [3] and [11]) and initial observations of individuals' responses to policy or communication campaigns. Future work will be devoted to explicitly accounting for inaccurate estimates of the opinion dynamics' parameters at the policy design stage.

V. ADAPTIVE POLICY DESIGN: A MODEL PREDICTIVE APPROACH

While the MBCCP strategy given by Policy 2 already seeks an optimal tradeoff between the efficiency and costs of a policy, it does so by restrictively looking at achieving an efficiency-cost balance in the long run and a steady-state scenario, neglecting the intertwined evolution of opinions and policies over time. We now make a step forward in overcoming this limitation, formulating the following policy design problem explicitly optimizing the efficiency-cost tradeoff accounting for how the acceptance of the target choice evolves over time under the influence of the nudging policy.

Policy 3 (Infinite-horizon policy with quadratic cost): Let the policymaker/stakeholder aiming to maximize the average realizations of unitary acceptance variables $\mathbf{y}(t)$ while minimizing the policy effort $\mathbf{u}(t)$, for all $t \in \mathbb{Z}_{\geq 0}$. A possible optimal policy to achieve this objective can be designed by solving

$$\begin{aligned}\min_{\mathbf{U}_\infty} & J(\mathbf{U}_\infty) \\ \text{s.t.} & \text{(6) and (7)} \quad \forall t \in \mathbb{Z}_{\geq 0} \\ & \mathbf{u}(t) \in [\mathbf{0}, \mathbf{1} - \boldsymbol{\eta}'] \quad \forall t \in \mathbb{Z}_{\geq 0} \\ & \mathbf{1}^\top \mathbf{u}(t) \leq \beta \quad \forall t \in \mathbb{Z}_{\geq 0}\end{aligned} \quad (20)$$

where $\mathbf{U}_\infty = \{\mathbf{u}(t)\}_{t \in \mathbb{Z}_{\geq 0}}$ comprises all the policy actions to be designed over time, and the infinite-horizon loss $J(\mathbf{U}_\infty)$ is given by

$$J(\mathbf{U}_\infty) = \sum_{t=0}^{+\infty} [\mathbb{E} \|\mathbf{1} - \mathbf{y}(t)\|_{\mathbf{Q}}^2 + \|\mathbf{u}(t)\|_{\mathbf{R}}^2] \quad (21)$$

where $\mathbf{Q}$ and $\mathbf{R}$ are positive definite matrices shaping the tradeoff between efficiency and control effort.

Note that this infinite policy formulation looks at maximizing manifested acceptance, i.e., the only information that can be accessed by the policymaker. Nonetheless, under Assumption 2, this quantity is still closely related to hidden inclinations variables $\mathbf{x}(t)$.

Remark 6 (Relation with MBCCP): The MBCCP (see Policy 2) corresponds to an infinite horizon policy with quadratic cost (similar to Policy 3) where the transient in the objective function is neglected.

As done for the MBCCP in Section IV, we first focus on the cost $J(U_\infty)$ and analyze its properties. To this end, let us introduce the following compact notation for the expected value of the hidden inclination variables:

$$\boldsymbol{\mu}(t) = \mathbb{E}[\mathbf{x}(t)]. \quad (22)$$

By exploiting this formalism, we can introduce and formalize the following result.

Proposition 3 (Equivalent cost): Under Assumption 2, the cost function in (21) is equivalent to

$$\begin{aligned} J_\infty(U_\infty) &= \sum_{t=0}^{+\infty} [\|\mathbf{1} - \boldsymbol{\mu}(t)\|_Q^2 + \|\mathbf{u}(t)\|_R^2 + \langle \boldsymbol{\mu}(t), \mathbf{1} - \boldsymbol{\mu}(t) \rangle_{[Q]}] . \end{aligned} \quad (23)$$

Proof: Let us focus on the first term of the cost in (21) only. The latter can be recast as

$$\mathbb{E}[\|\mathbf{1} - \mathbf{y}(t)\|_Q^2] = \mathbb{E}[\mathbb{E}[\|\mathbf{1} - \mathbf{y}(t)\|_Q^2 | \mathbf{x}(t)]] .$$

Thanks to Assumption 2, standard computations allow us to rewrite this term further as a function of $\boldsymbol{\mu}(t)$ in (22) as follows:

$$\begin{aligned} \mathbb{E}[\mathbb{E}[\|\mathbf{1} - \mathbf{y}(t)\|_Q^2 | \mathbf{x}(t)]] &= \|\mathbf{1} - \boldsymbol{\mu}(t)\|_Q^2 + \\ &\quad - \sum_{v \in \mathcal{V}} Q_{vv} (1 - \mu_v(t))^2 + \sum_{v \in \mathcal{V}} Q_{vv} \mathbb{E}[\mathbb{E}[(1 - y_v(t))^2 | \mathbf{x}(t)]] \\ &= \|\mathbf{1} - \boldsymbol{\mu}(t)\|_Q^2 - \sum_{v \in \mathcal{V}} Q_{vv} (1 - \mu_v(t))^2 \\ &\quad + \sum_{v \in \mathcal{V}} Q_{vv} \mathbb{E}[1 - \mu_v(t)] \\ &= \|\mathbf{1} - \boldsymbol{\mu}(t)\|_Q^2 + \langle \boldsymbol{\mu}(t), \mathbf{1} - \boldsymbol{\mu}(t) \rangle_{[Q]} . \end{aligned}$$

By replacing with this last expression the first term in (21), we straightforwardly get (23) and conclude the proof. $\square$

As shown in Proposition 3, the cost in (21) ultimately consists of three terms. The first is a quadratic function in the expected values of the agents' inclinations, the second is a quadratic form in the control input, and the third term allows us to equivalently recast the expected quadratic cost in a quadratic loss of expected values.

The reformulation of the cost in (23) further enables us to confine the infinite horizon cost in (21) within an interval, as summarized below.

Lemma 1 (Bound on the infinite-horizon cost): For any U_∞ satisfying the constraints in (20), the loss in (21) satisfies

$$\underline{J}_\infty(U_\infty) \leq J_\infty(U_\infty) \leq \bar{J}_\infty(U_\infty) \quad (24a)$$

where

$$\underline{J}_\infty(U_\infty) = \sum_{t=0}^{+\infty} [\|\mathbf{1} - \boldsymbol{\mu}(t)\|_Q^2 + \|\mathbf{u}(t)\|_R^2 - \|\mathbf{1}\|_{[Q]}^2] \quad (24b)$$

$$\bar{J}_\infty(U_\infty) = \sum_{t=0}^{+\infty} [\|\mathbf{1} - \boldsymbol{\mu}(t)\|_Q^2 + \|\mathbf{u}(t)\|_R^2 + \|\mathbf{1}\|_{[Q]}^2] . \quad (24c)$$

Proof: Let us focus on the last term in (23), verifying

$$\langle \boldsymbol{\mu}(t), \mathbf{1} - \boldsymbol{\mu}(t) \rangle_{[Q]} = (\boldsymbol{\mu}(t))' [Q] \mathbf{1} - \|\boldsymbol{\mu}(t)\|_{[Q]}^2$$

by definition, with $[Q]$ having all nonnegative entries since Q is positive definite. Since U_∞ satisfies the constraints imposed in (20), then $\boldsymbol{\mu}(t) \in [0, 1]^{|V|}$ for all $t \in \mathbb{Z}_{\geq 0}$. These two results imply that

$$0 \leq (\boldsymbol{\mu}(t))' [Q] \mathbf{1} \leq \|\mathbf{1}\|_{[Q]}^2 \quad \forall t \in \mathbb{Z}_{\geq 0} \quad (25)$$

and thus, the following upper bound holds:

$$\langle \boldsymbol{\mu}(t), \mathbf{1} - \boldsymbol{\mu}(t) \rangle_{[Q]} \leq (\boldsymbol{\mu}(t))' [Q] \mathbf{1} \leq \|\mathbf{1}\|_{[Q]}^2 \quad \forall t \in \mathbb{Z} .$$

The bound in (24c) straightforwardly follows. Meanwhile, the third term in (23) also satisfies

$$\langle \boldsymbol{\mu}(t), \mathbf{1} - \boldsymbol{\mu}(t) \rangle_{[Q]} \geq -\|\boldsymbol{\mu}\|_{[Q]}^2 \geq -\|\mathbf{1}\|_{[Q]}^2 \quad \forall t \in \mathbb{Z}_{\geq 0} ,$$

where the first inequality stems from the lower-bound in (25) while the last one holds thanks to the bounds on $\boldsymbol{\mu}(t)$ guaranteed by the assumption on U_∞ . The bound in (24b) easily follows, ending the proof. $\square$

This result shows that the cost function in (23) can be both lower and upper bounded by a quadratic cost in $\{\boldsymbol{\mu}(t)\}_{t \in \mathbb{Z}_{\geq 0}}$ and U_∞ except for a constant, that changes only in sign depending on whether we consider the lower or the upper bound.

While ideally guaranteeing an optimal efficiency-cost trade-off [according to (21)] over an infinite horizon, Policy 3 is knowingly intractable from a computational perspective even for small networks, because it entails the solution of an optimization problem with infinite constraints. Moreover, it might not truly fit the policymakers' and stakeholders' needs. Indeed, practical policies are likely never enacted for an infinite period of time but rather redesigned at times based on societal and technological changes (e.g., substantial technological advancements). Therefore, we now shift from infinite-horizon to finite-horizon policy design.

Remark 7 (Shift to finite-horizon policies): If $Q = [Q]$, Proposition 3 implies that the cost to be optimized is quadratic in U_∞ but only linear in $\{\boldsymbol{\mu}(t)\}_{t \in \mathbb{Z}_{\geq 0}}$. To avoid this different treatment of the controlled input and expected inclinations, we approximate the cost by considering only the quadratic terms in $\{\boldsymbol{\mu}(t)\}_{t \in \mathbb{Z}_{\geq 0}}$ and U_∞ common to (24b) and (24c) in going from an infinite-horizon to a finite-horizon policy.

A. Oracle Finite-Horizon MPC

We now step away from the impractical, infinite-horizon Policy 3 to a practical finite-horizon one, that still accounts that the individual inclinations are time-varying and exhibit an oscillating behavior. To this end, we initially assume that an *oracle* provides the policymaker/stakeholder with the expected values $\mu(t)$ of the hidden inclinations at all time instant $t \in \mathbb{Z}_{\geq 0}$. Under this assumption, by exploiting the result in Proposition 1, we cast the following finite-horizon policy design problem.

Policy 4 (Oracle finite-horizon policy): Given a finite horizon $T \in \mathbb{Z}_{\geq 0}$, with $T > 1$, the input $\mathbf{u}(t)$ to be fed to the social system to nudge individuals toward a target choice at time $t \in \mathbb{Z}_{\geq 0}$ can be designed by solving the following constrained QP:

$$\begin{aligned} \min_{\mathbf{U}_t} J_T(\mathbf{U}_t) \\ \text{s.t. } \mu(k+1|t) &= \Lambda \bar{\mathbf{P}} \mu(k|t) + (\mathbf{I} - \Lambda)(\eta_0 + \mathbf{u}(k|t)) \quad \forall k \in \mathcal{I}_T \\ \mathbf{u}(k|t) &\in [0, 1 - \eta'] \quad \forall k \in \mathcal{I}_T \\ \mathbf{1}^\top \mathbf{u}(k|t) &\leq \beta \quad \forall k \in \mathcal{I}_T \\ \mu(0|t) &= \mu(t) \\ \mu(T|t) &= (\mathbf{I} - \Lambda \bar{\mathbf{P}})^{-1}(\mathbf{I} - \Lambda)(\eta_0 + \mathbf{u}^{\text{MB}}) \end{aligned} \quad (26)$$

where $\mathcal{I}_T = [0, T-1]$, $\mathbf{u}^{\text{MB}}$ is the solution to (15), $\mu(k|t)$ is the vector of expected inclinations predicted according to the first set of constraints starting from the initial condition dictated by $\mu(t)$, and $\mathbf{U}_t = \{\mathbf{u}(k|t)\}_{k=0}^{T-1}$ is the set of inputs to be optimized. The cost $J_T(\mathbf{U}_t)$ to be minimized is instead

$$J_T(\mathbf{U}_t) = \sum_{k=0}^{T-1} [\|\mathbf{1} - \mu(k|t)\|_Q^2 + \|\mathbf{u}(k|t)\|_R^2] \quad (27)$$

according to Remark 7.

In principle, solving (26) results in a sequence of T inputs $\mathbf{U}_t^*$ that can be used to nudge individual over the user-defined finite horizon of length T . Nonetheless, enacting the policy in a recording horizon fashion, i.e., by setting $\mathbf{u}(t) = \mathbf{u}^*(0|t)$ and discarding the rest of $\mathbf{U}_t^*$, is advisable given the external uncontrollable factors temporarily influencing individual choices and our simplistic approach to account for them.

For the receding horizon oracle policy, we can prove the following formal result.

Theorem 3 (Oracle policy's properties): Let the initial expected inclination given by the oracle be $\mu(0) \in [0, 1]^V$ and assume that the optimal control problem in (26) is feasible at time $t = 0$. Then, the following properties hold for all positive definite $\mathbf{Q}$ and $\mathbf{R}$.

- i) The optimal control problem in (26) is recursively feasible $\forall t \geq 0$.
- ii) The expected inclinations asymptotically converge to the steady-state $\mu^{\text{MB}} = (\mathbf{I} - \Lambda \bar{\mathbf{P}})^{-1}(\mathbf{I} - \Lambda)(\eta_0 + \mathbf{u}^{\text{MB}})$, with $\mathbf{u}^{\text{MB}}$ being the solution of (15).

Proof: We first prove recursive feasibility [i.e., i)]. To this end, assume that (26) is feasible at $t \in \mathbb{Z}_{\geq 0}$ and denote the associated optimal control sequence and the predicted expected inclinations

as

$$\begin{aligned} \mathbf{U}_t^* &= \{\mathbf{u}^*(0|t), \mathbf{u}^*(1|t), \dots, \mathbf{u}^*(T-1|t)\} \\ \mathbf{M}_t^* &= \{\mu^*(0|t), \mu^*(1|t), \dots, \mu^*(T-1|t), \mu^{\text{MB}}\} \end{aligned}$$

respectively, where the last term in $\mathbf{M}_t^*$ is induced by the terminal constraint. Let us construct the candidate solution for the instance of (26) at time $t+1$ as

$$\mathbf{U}'_{t+1} = \{\mathbf{u}^*(1|t), \mathbf{u}^*(2|t), \dots, \mathbf{u}^*(T-1|t), \mathbf{u}^{\text{MB}}\} \quad (28)$$

where $\mathbf{u}^{\text{MB}}$ is the solution of (15), satisfying both the box constraint $\mathbf{u}^{\text{MB}} \in [0, 1 - \eta']$ and the budget constraint by construction. As the other elements of $\mathbf{U}'_{t+1}$ correspond to the last $T-1$ elements of $\mathbf{U}_t^*$, also satisfying both the budget constraint and the box constraints, the candidate input sequence is feasible. Let us now define the expected inclination sequence associated with $\mathbf{U}'_{t+1}$, i.e.,

$$\mathbf{M}'_{t+1} = \{\mu^*(1|t), \mu^*(2|t), \dots, \mu^{\text{MB}}, \mu'(T|t+1)\} \quad (29)$$

where the first $T-1$ elements are a feasible trajectory by construction (as they are the last T elements of $\mathbf{M}_t^*$). Meanwhile, since the $\mathbf{u}'(T-1|t+1) = \mathbf{u}^{\text{MB}}$ by construction and $\mu'(T-1|t+1) = \mu^{\text{MB}}$, then $\mu'(T|t+1) = \mu^{\text{MB}}$, thanks to the stationarity of the pair $(\mu^{\text{MB}}, \mathbf{u}^{\text{MB}})$. Therefore, $\mathbf{M}'_{t+1}$ is a feasible expected inclination sequence. Since this candidate can be defined for all $t \in \mathbb{Z}_{\geq 0}$, a feasible solution for (26) exists $\forall t \in \mathbb{Z}_{\geq 0}$ and property i) easily follows.

We now focus on property ii). To prove it, let us equivalently rewrite the cost in (26) as follows:

$$J_T(\mathbf{U}_t) = \sum_{k=0}^{T-1} \ell(\mu(k|t), \mu^{\text{MB}}, \mathbf{u}(k|t), \mathbf{u}^{\text{MB}})$$

where

$$\begin{aligned} \ell(\mu(k|t), \mu^{\text{MB}}, \mathbf{u}(k|t), \mathbf{u}^{\text{MB}}) &= \|\varepsilon_\mu(k|t) - \varepsilon_\mu^{\text{MB}}\|_Q^2 + \|\varepsilon_u(k|t)\|_R^2 \\ &+ 2(\varepsilon_\mu^{\text{MB}})^\top \mathbf{Q}(\varepsilon_\mu(k|t) - \varepsilon_\mu^{\text{MB}}) + 2(\mathbf{u}^{\text{MB}})^\top \mathbf{R}\varepsilon_u(k|t) \\ &+ \|\varepsilon_\mu^{\text{MB}}\|_Q^2 + \|\mathbf{u}^{\text{MB}}\|_R^2 \geq 0 \end{aligned}$$

with $\varepsilon_\mu(k|t) = \mathbf{1} - \mu(k|t)$, $\varepsilon_\mu^{\text{MB}} = \mathbf{1} - \mu^{\text{MB}}$ and $\varepsilon_u(k|t) = \mathbf{u}(k|t) - \mathbf{u}^{\text{MB}}$. Note that, $\ell(\mu(k|t), \mu^{\text{MB}}, \mathbf{u}(k|t), \mathbf{u}^{\text{MB}})$ satisfies the following:

$$\begin{aligned} \ell(\mu(k|t), \mu^{\text{MB}}, \mathbf{u}(k|t), \mathbf{u}^{\text{MB}}) &\geq \ell(\mu^{\text{MB}}, \mu^{\text{MB}}, \mathbf{u}^{\text{MB}}, \mathbf{u}^{\text{MB}}) \\ &= \|\varepsilon_\mu^{\text{MB}}\|_Q^2 + \|\mathbf{u}^{\text{MB}}\|_R^2 \end{aligned} \quad (30)$$

for all $k \in \mathcal{I}_T$, since the first two terms are always nonnegative and become zero (i.e., assume their minimal value) only when $\varepsilon_\mu(k|t) = \varepsilon_\mu^{\text{MB}}$. Moreover, let us compactly denote the cost associated with the optimal solution at a given time instant t and the one associated with the candidate solution as $J_{T|t}^* = J_T(\mathbf{U}_t^*)$ and $J'_{T|t} = J_T(\mathbf{U}'_t)$, respectively. Note that, by optimality, it is straightforward that $J_{T|t}^* \leq J'_{T|t}$ at all $t \in \mathbb{Z}_{\geq 0}$.

By considering the candidates in (28) and (29), note that the following holds:

$$\begin{aligned} J'_{T|t+1} &= J^*_{T|t} - \ell(\mu(0|t), \mu^{\text{MB}}, u(0|t), u^{\text{MB}}) \\ &\quad + \|\varepsilon_{\mu}^{\text{MB}}\|_Q^2 + \|u^{\text{MB}}\|_R^2 \\ &\leq J^*_{T|t} \end{aligned}$$

where the inequality follows from (30). Accordingly

$$J^*_{T|t+1} \leq J'_{T|t+1} \leq J^*_{T|t}$$

and hence $J^*_{T|t+1} - J^*_{T|t} \leq 0$, i.e., the optimal cost of (26) is nonincreasing. As a consequence, there exists a nonnegative scalar $J^*_{T|\infty}$ such that $\lim_{t \rightarrow \infty} J^*_{T|t} = J^*_{T|\infty}$. Therefore

$$\begin{aligned} J^*_{T|t+1} - J^*_{T|t} &\geq \ell(\mu(0|t), \mu^{\text{MB}}, u(0|t), u^{\text{MB}}) \\ &\quad - \|\varepsilon_{\mu}^{\text{MB}}\|_Q^2 - \|u^{\text{MB}}\|_R^2 \geq 0 \end{aligned}$$

where the first term tends to zero for $t \rightarrow \infty$, implying that $\lim_{t \rightarrow \infty} \mu(t) = \mu^{\text{MB}}$ and $\lim_{t \rightarrow \infty} u(t) = u^{\text{MB}}$, thus concluding the proof. $\square$

Remark 8: As well-known (see, e.g., [25]), the explicit solution to (26) is a piecewise affine (PWA) control law on the expected inclination, with the values of its gains depending on the active constraints at each time instant. Nonetheless, based on the KKT conditions for QP, one can also determine (if they are well defined) the conditions on the weights Q and R guaranteeing the constraints in (26) are always inactive (i.e., the bound characterizing them are never hit). In this case, the control law becomes a static feedback on the average inclinations. In this simpler scenario, our future research will be devoted to a formal analysis of the closed-loop properties of the resulting controlled system.

B. MPC With Estimated Mean Inclinations

All previous policy design approaches rely on the exact knowledge of the average inclination $\mu(t) \forall t \in \mathbb{Z}_{\geq 0}$, which is generally not directly accessible to the policymaker. Indeed (see Assumption 2), policymakers and stakeholders can generally access the acceptance variables $y(t)$, for $t \in \mathbb{Z}_{\geq 0}$. The latter can be used to define a practical estimate of the mean inclination over time as

$$\hat{\mu}(t) = \frac{1}{t} \sum_{\tau=0}^{t-1} y(\tau). \quad (31)$$

Despite the simplicity of this estimator, if the ergodicity of the system is preserved under the designed control action, the Cesàro time-average (31) converges to their expected value, namely

$$\mathbb{E}[\|\hat{\mu}(t) - \mu(t)\|_2^2] \rightarrow 0. \quad (32)$$

We can therefore substitute the oracle predictive control problem in (26) with the following:

$$\begin{aligned} \min_{U|t} J_T(U|t) \\ \text{s.t. } \mu(k+1|t) &= \Lambda \bar{P} \mu(k|t) + (I - \Lambda)(\eta_0 + u(k|t)) \quad \forall k \in \mathcal{I}_T \\ u(k|t) &\in [0, 1 - \eta'] \quad \forall k \in \mathcal{I}_T \end{aligned}$$

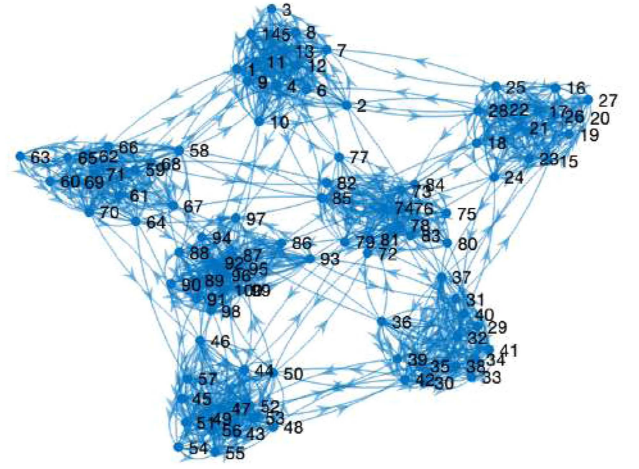


Fig. 1. Social network with 100 nodes considered in the numerical example.

$$\begin{aligned} 1^\top u(k|t) &\leq \beta \quad \forall k \in \mathcal{I}_T \\ \mu(0|t) &= \hat{\mu}(t) \\ \mu(T|t) &= (I - \Lambda \bar{P})^{-1} (I - \Lambda)(\eta_0 + u^{\text{MB}}) \end{aligned} \quad (33)$$

i.e., replace the “true” yet unknown average inclinations with their estimates, computed as in (31).

VI. NUMERICAL RESULTS

To illustrate the free evolution of the model proposed in Section II, as well as the effect of the control policies proposed in Sections IV and V, let us consider the randomly generated² modular graph consisting of $|\mathcal{V}| = 100$ nodes, grouped in 7 clusters depicted in Fig. 1, representing our social network. This set of nodes is split into two groups. One (smaller) group $\mathcal{V}_+ \subset \mathcal{V}$ of 20 nodes is favorably inclined toward a target choice of interest, having $x_v(0) = \eta_v^0$ drawn from a uniform distribution supported in $[0.45, 0.95]$, for all $v \in \mathcal{V}_+$. The remaining 80% of nodes are instead negatively inclined with respect to the target choice, with each of these 80 nodes having an initial inclination supported in $[0.05, 0.15]$ and noncontrollable noise $\eta_v^{\text{nc}}(t)$ sampled at random from a uniform distribution supported in $[-0.03, 0.03]$. Note that this implies that the considered population is mainly negatively biased with respect to the target choice.

We test the evolution of inclinations within this population according to the model proposed in Section II with $\alpha_v = \alpha$, for all $v \in \mathcal{V}$, in open and closed loop, in this second scenario adopting (and comparing) policies designed according to the strategies proposed in Sections IV–V.

With respect to the parameters in (6), the entries of the diagonal matrix Λ are drawn from a uniform distribution over the interval $[0, 1]$. The time-scale separation parameter α is instead chosen differently depending on whether we consider the free or the controlled evolution of (6). While in the first case, if fixed to

²We set the density of the edges $|\mathcal{E}|/|\mathcal{V}|^2$ to 0.05 and the probability of connections between nodes in different clusters to $\gamma = 0.9$.

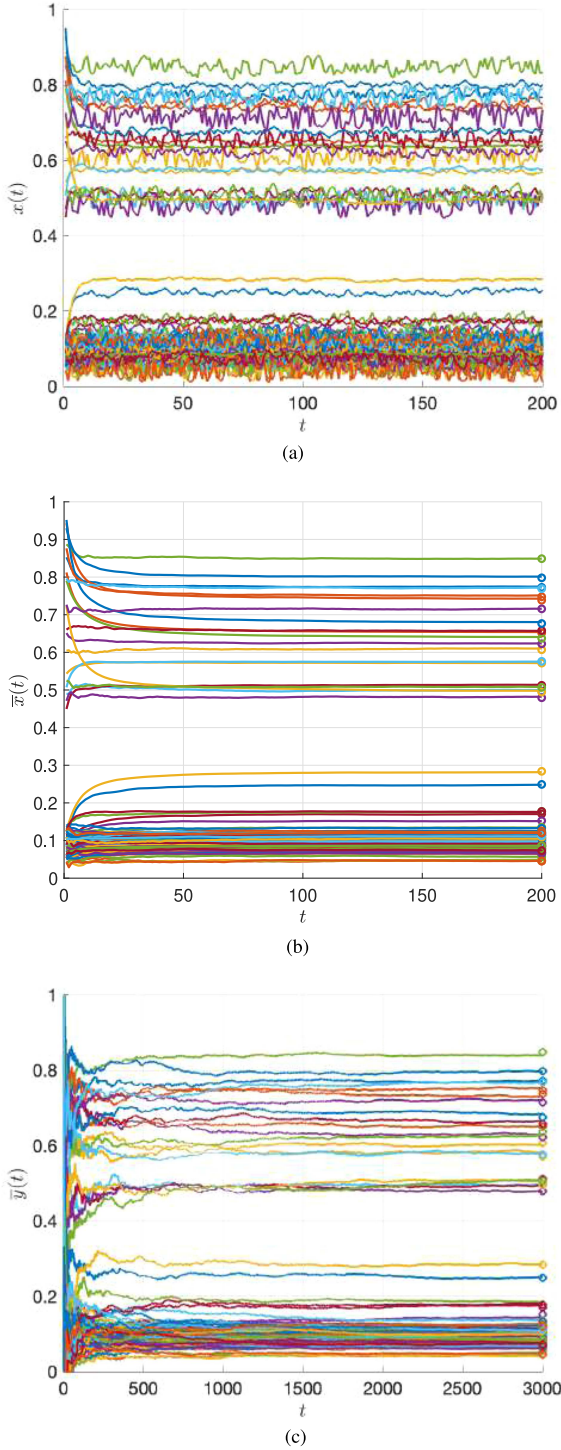


Fig. 2. Free evolution of the system: Evolution of hidden and manifested opinions and their time averages over time. Different colors represent different agents. (a) Hidden inclinations versus time. (b) Time average of hidden inclinations versus time. (c) Time average of manifested opinions versus time.

a (low) value, in the second one, it is varied to assess the impact of different time-scale separation parameters when closing the loop with the policies introduced in Sections IV–V.

Free evolution: Fig. 2 showcases the free evolution of the hidden opinions $x(t)$ characterized through dynamics in (6),

when $\alpha_v = 0.25$ for all $v \in \mathcal{V}$. These results confirm the theoretical results formalized in Theorems 1 and 2. Indeed, as the latent opinions $x(t)$ and the acceptance variables $y(t)$ continue to oscillate over time [see Fig. 2(a)], the corresponding time averages $\bar{x}(t)$ and $\bar{y}(t)$ converge to the same expected opinion profile, as shown in Fig. 2(b) and (c).

A. Performance Analysis of Control Policies

We now close the loop and compare the impact of nonpersonalized policies [i.e., MFCCP in (14)], personalized policies neglecting transients [see (15)], and the moving-horizon strategy with estimated mean inclinations presented in Section V-B against the free evolution of the social system. In our tests, we consider a simulation horizon of $T_{\text{sim}} = 20$ steps, impose a budget $\beta = 10$, and (as previously mentioned) vary the time-scale separation parameter $\alpha \in \{0.25, 0.5, 0.75, 1\}$. In designing both the MBCCP policy in (15) and the moving-horizon strategy with estimated mean inclinations [by time averaging, see (33)], we impose

$$Q = I, \quad R = rI. \quad (34)$$

In our closed-loop simulations, r is varied within the interval $(0, 5]$ to consider a spectrum of strategies with different penalization of policy costs. When employing the policy in (33), we set the prediction horizon T to 5.

a) Key performance indicators: To evaluate the cost-benefit tradeoff of the different policies (against the free evolution of the system), we consider two indexes.

The first indicator is defined as follows:

$$\Gamma_{T_{\text{sim}}} = \sum_{t=0}^{T_{\text{sim}}-1} \sum_{v \in \mathcal{V}} \mathbb{E} \|1 - y(t)\|_2^2 \quad (35)$$

represents the “social cost” of a policy, providing an indication of the effectiveness of each policy strategy at the end of the considered horizon. Specifically, the lower the social cost in (35) is, the more effective the control policy is in fostering the acceptance of the target choice. Note that this quantity can be computed for both the free and policy-induced evolution of the social system.

As an actual cost-benefit analysis cannot be carried out without considering the effort required by the policy-maker/stakeholder to enact it, the second indicator evaluates the policies’ “control cost,” and it is defined as

$$\Delta u_{T_{\text{sim}}} = \sum_{t=0}^{T_{\text{sim}}-1} \mathbb{E} \|u(t)\|_2^2. \quad (36)$$

Remark 9 (Evaluating (35)): The cost-benefit analysis we carry out relies on 20 Monte Carlo simulations for each value of r considered, changing the realization of the (uncontrollable) external factors affecting the social system. This choice allows us to compute the index in (35) by averaging over different instances of the problem.

b) Results and discussion: Fig. 3 shows the control cost (x -axis) against the social cost (y -axis) of the compared policies for r varying over a prespecified range. This choice

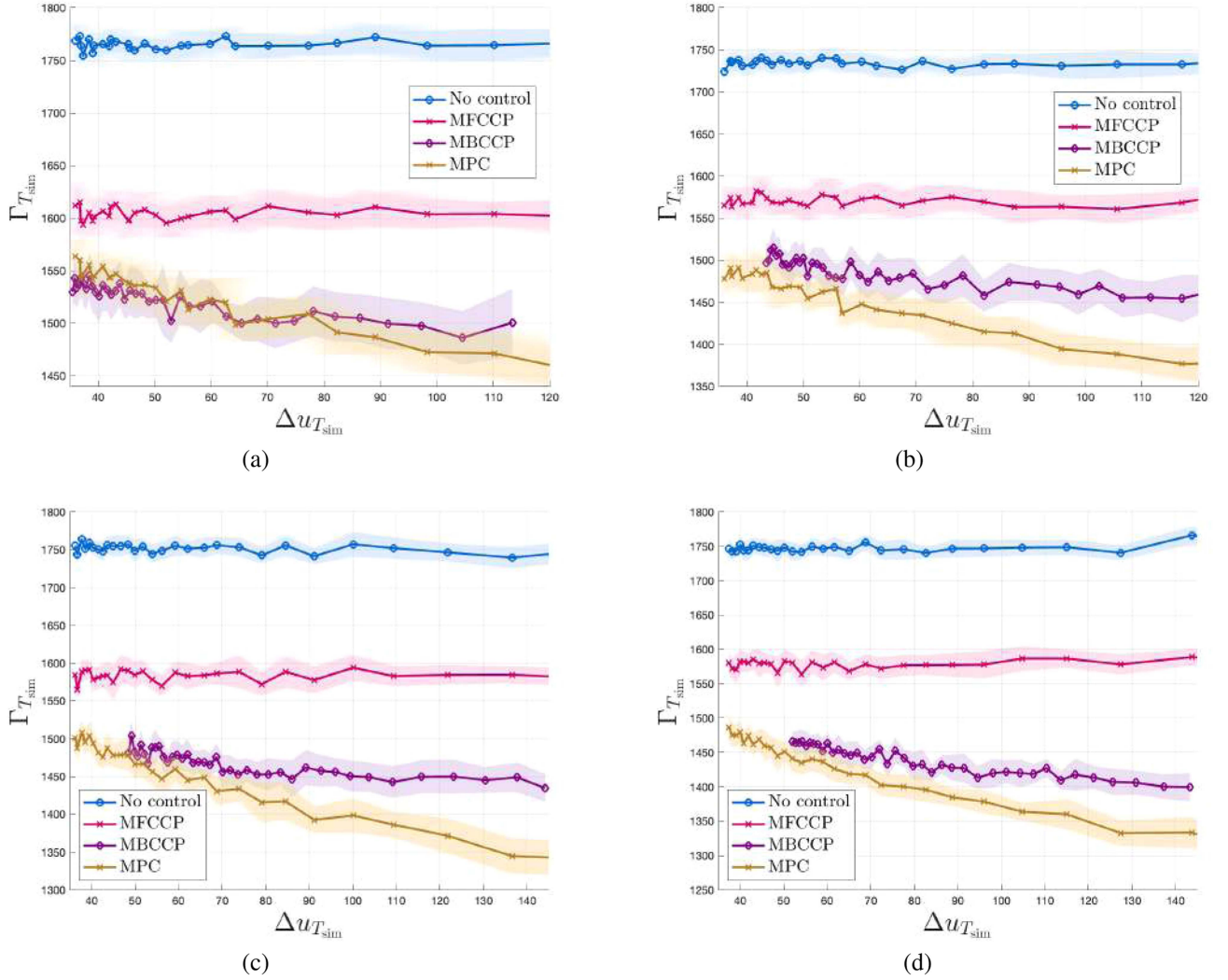


Fig. 3. Control cost versus social cost for different policies for modular graphs with intercluster connectivity $\gamma = 0.9$ for $r \in [10^{-12}, 5]$ and different values of α . Dotted lines indicate sample means, and shaded areas represent empirical standard deviations across 20 Monte Carlo simulations. The blue curve is used as a reference and represents the average social cost when *no control* is enacted. (a) $\alpha = 0.25$. (b) $\alpha = 0.5$. (c) $\alpha = 0.75$. (d) $\alpha = 1$.

allows us to make a fair comparison between the tested policy design strategies, since the control effort induced by the policies may differ even for the same value of the penalty r . The blue curve corresponds to the uncontrolled dynamics, hence the case where the system evolves freely without external (controlled) interventions. In this case, the control cost is naturally zero for this case, but including the free scenario in our comparison allows us to obtain a reference value for the social cost to be reduced when using fostering strategies. Moreover, it is worth pointing out that, although the free evolution (blue curve) and the outcome of the MFCCP strategy (magenta curve) are not affected by the choice of r , their value is not constant. Indeed, for each r , we consider 20 new realizations of the external factors (see Remark 9), differently affecting individual choices both in open-loop and with all the tested policies.

As expected, irrespective of the value taken by the time-scale separation parameter α , enacting either nonpersonalized policies

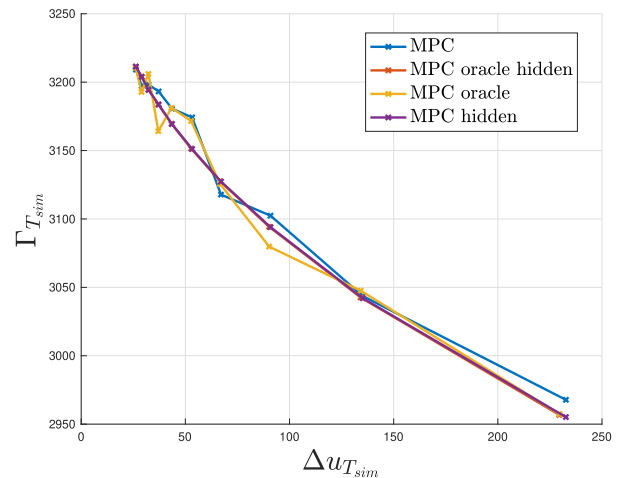


Fig. 4. Control cost versus social and “hidden” cost for the MPC policy.

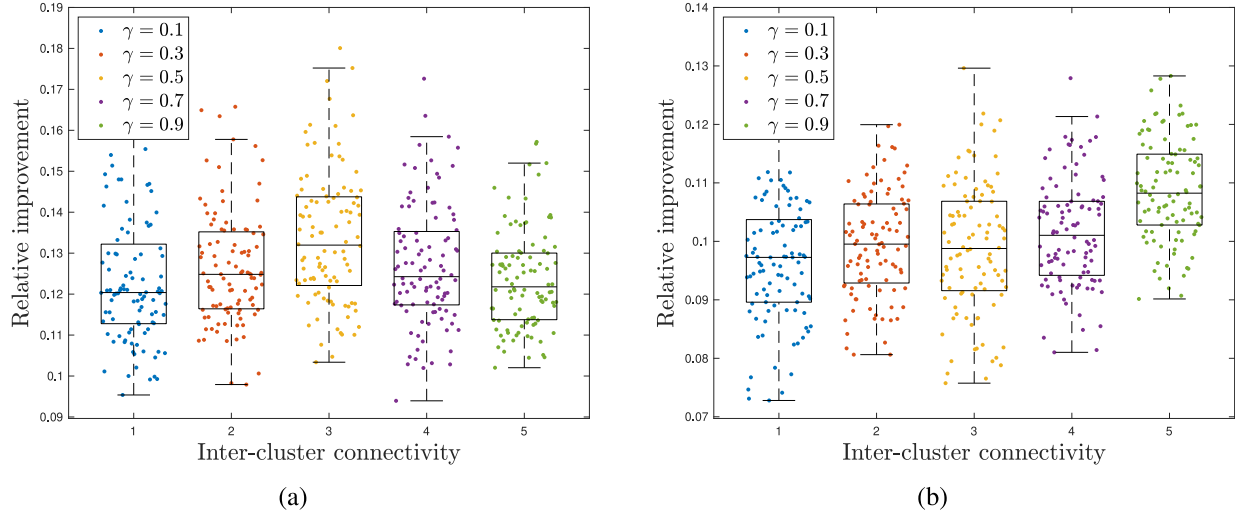


Fig. 5. Box and scatter plot showing the relative shift of social cost for 50 realizations of random modular graphs with 100 nodes for each value of the intercluster connectivity parameter γ for different values of α . (a) $\alpha = 0.25$. (b) $\alpha = 0.75$.

or personalized ones leads to a social cost that is lower than the free evolution of the system (blue curve), i.e., a more widespread acceptance of the target choice. At the same time, adopting the MFCCP strategy (magenta curve) only marginally improves the social cost with respect to the one attained with no enacted policy, as such strategy is blind to the social dynamics and solely relies on a uniform distribution of resources. In contrast, both the MBCCP (purple curve) and the MPC (yellow curve) approaches are clearly taking advantage of social imitation. Indeed, as the control cost increases, the social one decreases. This result indicates that a larger effort leads to a more widespread adoption of the target choice (and, hence, a greater social benefit if the latter aims to promote improved social welfare). Among the MBCCP strategy [see (15)] and the MPC-based one [see (33)], the latter achieves better results for the same control effort despite relying on an estimate of the hidden individual inclination. For both policies, but with a larger gain when considering the MPC-based one, it is also clear that the more frequent the social pressure (i.e., the higher α), the more effective the enacted policy. In fact, it can be seen that as α increases, the social cost is lowered for the same amount of control effort.

B. Comparison With Oracle Policy Under Perfect Information

We now assess the impact of using estimates of the hidden inclinations rather than the true ones by comparing the performance of two MPC schemes, one using (31) and the other (the so-called *oracle*) exploiting the actual values of the agents' inclinations. To this end, we set $\alpha_v = \alpha = 0.5$ for each $v \in \mathcal{V}$, λ is drawn from a uniform distribution over the interval $[0, 1]$, and r parameter is assumed to take values from a uniform grid in $[1, 5]$ with 10 points.

The social cost, as defined in (35), versus the control cost obtained with the two policies (averaged over 5 experiments) is displayed in Fig. 4. As can be seen, the behavior of the curves follows the same trend, with a slight loss in performance when

using the estimated opinions and a variability that is induced by the definition of $\Gamma_{T_{\text{sim}}}$. By considering the “hidden” social cost,³ i.e.,

$$\Gamma_{T_{\text{sim}}}^o = \sum_{t=0}^{T_{\text{sim}}-1} \sum_{v \in \mathcal{V}} \mathbb{E} \|1 - \mathbf{y}(t)\|_2^2 \quad (37)$$

for both policies, we notice that the performance degradation induced by not knowing the hidden inclination remains limited in expectation. In turn, this implies that the convergence to the stationary distribution of the Markov process (see Theorem 1) is fast and the approximation using the average mean inclinations does not excessively impact the outcomes of the MPC policies.

C. Impact of the Network Topology

By focusing on the MPC strategy only and by setting $r = 1$, we now analyze the impact of the network topology on the “relative improvement” in social cost achieved when using the MPC-based policy over adopting no policy at all, i.e.,

$$\Delta_{\Gamma_{T_{\text{sim}}}} = \frac{|\Gamma_{T_{\text{sim}}}^{\text{MPC}} - \Gamma_{T_{\text{sim}}}^{\text{ol}}|}{\Gamma_{T_{\text{sim}}}^{\text{ol}}} \quad (38)$$

where $\Gamma_{T_{\text{sim}}}^{\text{MPC}}$ and $\Gamma_{T_{\text{sim}}}^{\text{ol}}$ are the values of the social cost defined in (35) achieved with the MPC policy and in open-loop, respectively. To make this assessment, we consider new randomly generated modular graphs with 100 nodes (as before), varying the intercluster connectivity parameter γ . This choice allows us to compare the outcome of the policy for both homogeneous networks (i.e., for γ small) and clustered ones (corresponding to large values γ). Note that, for each value of γ , we consider 50 different realizations of the modular graph.

In contrast with the previous analysis, we fix Λ by imposing $\lambda_v = 0.9$ for all $v \in \mathcal{V}$ and we set $x_v(0) = 0.1$ for all $v \in \mathcal{V}$. This choice enables us to avoid confounding effects induced by

³This is computed by (17) and referred as “hidden” social cost, as it is not directly observable or computable in practice.

the polarization of the population, as well as differences in the impact of social interactions and external factors on the opinions' evolution.

Fig. 5 reports both a scatter plot and a box plot of the relative improvement in social cost [see (38)], considering both $\alpha = 0.25$ and $\alpha = 0.75$. It can be observed that when the time-scale separation parameter α is lower and, thus, social pressure is low, the network topology seems not to play a well-defined role in the outcome of the policy. Indeed, the best results are obtained when the intercluster connectivity has intermediate values. On the contrary, under higher social pressure ($\alpha = 0.75$), greater social impacts are obtained for higher values of connectivity. This result ultimately indicates, as expected, that the MPC-based policy benefits from the indirect effect of social contagion.

VII. CONCLUSION

This work presented a novel extension of the Friedkin and Johnsen model that characterizes the evolution of individual inclinations subject to social influence and to the effects of unpredictable phenomena and external (controlled and deterministic) interventions. An avenue of future investigation will be assessing the realism of our modeling choices through real-world data.

By relying on the proposed model and its ergodicity, we have then devised a set of policy design strategies, looking at the optimal balance between fostering the social acceptance of a target action on average and containing policy costs. Our numerical analysis spotlighted the differences, benefits, and drawbacks of the proposed strategies, setting the ground for further research into the use of these models for effective policy design in a social setting.

These strategies are designed considering a period-by-period budget constraint, rather than a finite-time horizon integral one. Indeed, in many situations, such as when devising awareness campaigns or incentive programs, policymakers can operate on recurring funding cycles, with the same resources being allocated at each time instant. Nonetheless, the proposed approach can be extended to consider an integral constraint over time, which is generally more appropriate when structural interventions (e.g., infrastructure improvements) are implemented. We refer to [26] for some preliminary results on the use of the integral budget constraint, while further investigations on this aspect in the framework proposed in this work are deferred to future research.

Moreover, the proposed model-based policies are conceived by relying on the knowledge the expected opinion dynamics, among which the individual bias and the average influence network. While these parameters cannot be directly observed in practice, they can be estimated from survey data, as in [3] and [11], or through a preliminary observation of the free dynamics (see, e.g., [13]). Future work will be devoted to accounting for the uncertainty of these estimates, to robustify policy design and further increase its practicality.

Future research will also be devoted to translating the proposed policies into the actual economic effort to be undertaken by policymakers, ultimately connecting the proposed framework with its practical application.

REFERENCES

- [1] N. E. Friedkin, "A formal theory of social power," *J. Math. Sociol.*, vol. 12, pp. 103–126, 1986.
- [2] N. E. Friedkin and E. C. Johnsen, "Social influence and opinions," *J. Math. Sociol.*, vol. 15, no. 3/4, pp. 193–206, 1990.
- [3] V. Breschi, C. Ravazzi, S. Strada, F. Dabbene, and M. Tanelli, "Driving electric vehicles' mass adoption: An architecture for the design of human-centric policies to meet climate and societal goals," *Transp. Res. Part A: Policy Pract.*, vol. 171, 2022, Art. no. 103651.
- [4] V. Breschi, C. Ravazzi, S. Strada, F. Dabbene, and M. Tanelli, "Fostering the mass adoption of electric vehicles: A network-based approach," *IEEE Trans. Control Netw. Syst.*, vol. 9, no. 4, pp. 1666–1678, Dec. 2022.
- [5] W. S. Rossi, J. W. Polderman, and P. Frasca, "The closed loop between opinion formation and personalized recommendations," *IEEE Trans. Control Netw. Syst.*, vol. 9, no. 3, pp. 1092–1103, Sep. 2022.
- [6] J. Castro, J. Lu, G. Zhang, Y. Dong, and L. Martínez, "Opinion dynamics-based group recommender systems," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 48, no. 12, pp. 2394–2406, Dec. 2018.
- [7] M. Goyal, D. Chatterjee, N. Karamchandani, and D. Manjunath, "Maintaining ferment," in *Proc. IEEE 58th Conf. Decis. Control*, 2019, pp. 5217–5222.
- [8] B. Sprenger, G. De Pasquale, R. Soloperto, J. Lygeros, and F. Dörfler, "Control strategies for recommendation systems in social networks," *IEEE Control Syst. Lett.*, vol. 8, pp. 634–639, 2024.
- [9] N. E. Friedkin and E. C. Johnsen, *Assessing the Model*. Cambridge, U.K.: Cambridge Univ. Press, 2011, pp. 80–112.
- [10] J. Dippong, W. Kalkhoff, and E. C. Johnsen, "Status, networks, and opinion change: An experimental investigation," *Social Psychol. Quart.*, vol. 80, no. 2, pp. 153–173, 2017.
- [11] E. Villa, V. Breschi, C. Ravazzi, M. Tanelli, and F. Dabbene, "Can control aid in attaining sustainable goals? An improved data-informed framework to promote shared mobility," *Control Eng. Pract.*, vol. 153, 2024, Art. no. 106106.
- [12] F. Kotthoff and T. Hamacher, "Calibrating agent-based models of innovation diffusion with gradients," *J. Artif. Societies Social Simul.*, vol. 25, no. 3, 2022, Art. no. 4. [Online]. Available: <http://jasss.soc.surrey.ac.uk/25/3/4.html>
- [13] C. Ravazzi, S. Hojjatinia, C. Lagoa, and F. Dabbene, "Ergodic opinion dynamics over networks: Learning influences from partial observations," *IEEE Trans. Autom. Control*, vol. 66, no. 6, pp. 2709–2723, Jun. 2021.
- [14] J. Fisher and R. Bhattacharya, "Linear quadratic regulation of systems with stochastic parameter uncertainties," *Automatica*, vol. 45, no. 12, pp. 2831–2841, 2009.
- [15] J. Korteling, G. Paradies, and J. Sassen-van Meer, "Cognitive bias and how to improve sustainable decision making," *Front. Psychol.*, vol. 14, pp. 872–876, 2023.
- [16] A. Jaffe, R. Newell, and R. Stavins, "Environmental policy and technological change," *Environ. Resource Econ.*, vol. 22, pp. 41–70, 2002.
- [17] M. McPherson, L. Smith-Lovin, and J. Cook, "Birds of a feather: Homophily in social networks," *Annu. Rev. Sociol.*, vol. 27, pp. 415–444, 2001.
- [18] D. Fiorello, L. Zani, P. Christidis, and E. N. Cawood, "EU survey on issues related to transport and mobility," Publications Office of the European Union, Luxembourg, Tech. Rep. LFNA-27334-EN-N (JRC96151), 2015, doi: [10.2791/48322](https://doi.org/10.2791/48322). [Online]. Available: <https://publications.jrc.ec.europa.eu/repository/handle/JRC96151>
- [19] J. L. Doob, *Stochastic Processes*. Hoboken, NJ, USA: Wiley, 1962.
- [20] B. Polyak and A. Juditsky, "Acceleration of stochastic approximation by averaging," *SIAM J. Control Optim.*, vol. 30, no. 4, pp. 838–855, 1992.
- [21] C. Ravazzi, P. Frasca, R. Tempo, and H. Ishii, "Ergodic randomized algorithms and dynamics over networks," *IEEE Trans. Control Netw. Syst.*, vol. 2, no. 1, pp. 78–87, Mar. 2015.
- [22] W. Hoeffding, "Probability inequalities for sums of bounded random variables," *J. Amer. Statist. Assoc.*, vol. 58, no. 301, pp. 13–30, 1963.
- [23] V. Borkar, *Probability Theory: An Advanced Course*. New York, NY, USA: Springer, 1995.
- [24] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [25] A. Bemporad, M. Morari, V. Dua, and E. Pistikopoulos, "The explicit solution of model predictive control via multiparametric quadratic programming," in *Proc. Amer. Control Conf.*, 2000, pp. 872–876.
- [26] L. Piccinin, V. Breschi, C. Ravazzi, F. Dabbene, and M. Tanelli, "Innovation diffusion dynamics toward long-term behavioral shifts," *IEEE Control Syst. Lett.*, vol. 9, pp. 661–666, 2025.



Chiara Ravazzi (Senior Member, IEEE) received the Ph.D. degree in mathematics for engineering sciences from Politecnico di Torino, Turin, Italy, in 2011.

She is currently a Senior Researcher with the CNR-IEIIT and an Adjunct Professor with Politecnico di Torino. In 2010, she was a Visiting Member with the Massachusetts Institute of Technology, Cambridge (LIDS) and she held Postdoctoral positions with Politecnico di Torino from 2011 to 2016. Her research interests include signal processing, optimization, and control of network systems.

Dr. Ravazzi served as an Associate Editor for *IEEE TRANSACTIONS ON SIGNAL PROCESSING* from 2019 to 2023, and she currently holds the same position for *IEEE Control Systems Letters* (since 2021) and the *European Journal of Control* (since 2023).



Valentina Breschi (Member, IEEE) received the B.Sc. degree in electronic and telecommunication engineering and the M.Sc. degree in electrical and automation engineering from the University of Florence, Florence, Italy, in 2011 and 2014, respectively, and the Ph.D. degree in control systems from IMT School for Advanced Studies Lucca, Lucca, Italy, in 2018.

She was a Visiting Scholar with the University of Michigan, USA. From 2018 to 2023, she was with Politecnico di Milano (Italy), first as a Post-

doctoral Researcher and then as a junior Assistant Professor from 2020 to 2023. In 2023, she joined the Department of Electrical Engineering, Eindhoven University of Technology, The Netherlands, as an Assistant Professor. Her research interests include data-driven control, system identification, meta learning, opinion dynamics, and human-centered policy design.



Paolo Frasca (Senior Member, IEEE) received the Ph.D. degree in mathematics for engineering sciences from Politecnico di Torino, Turin, Italy, in 2009.

After Postdoctoral appointments with the CNR-IAC, Rome, and in Torino, he has been an Assistant Professor with the University of Twente, Enschede, Netherlands, from 2013 to 2016. Since 2016, he has been a CNRS Researcher with GIPSA-lab, Grenoble, France. On the topics of his research interests, he has

(co)authored more than 60 journal and conference papers and the book *Introduction to Averaging Dynamics over Networks*, published by Springer. His research interests include the theory of network systems, with applications to transportation and social networks.

Dr. Frasca has served or is serving as Associate Editor in the Editorial Boards of several conferences and journals, including the *IEEE Control Systems Letters*.



Fabrizio Dabbene (Fellow, IEEE) received the Laurea degree and the Ph.D. degree in systems and computer engineering from Politecnico di Torino, Torino, Italy, in 1995 and 1999, respectively.

He is a Director of Research with IEIIT, National Research Council of Italy (CNR), Torino, where he is the Coordinator of the Information and Systems Engineering Group. He has held visiting and research positions with The University of Iowa, with Penn State University, and with the Russian Academy of Sciences, Institute of Control Science, Moscow.

On the topics of his research interests include, he has authored or coauthored more than 100 research papers and two books. His research interests include data-driven robust methods for systems and control, design of advanced guidance and navigation schemes for aerospace applications, and analysis of techno-social dynamical networks.

Dr. Dabbene was twice a recipient of the EurAgeng Outstanding Paper Award (2010 and 2024). He served as a Senior Editor for the *IEEE Control Systems Letters* (2018–2023) and as an Associate Editor for *Automatica* (2008–2014) and for *IEEE TRANSACTIONS ON AUTOMATIC CONTROL* (2008–2012), and he is currently serving as Senior Editor for *IEEE TRANSACTIONS ON CONTROL SYSTEMS TECHNOLOGY*. He was elected member of the Board of Governors (2014–2016) and served as the Vice-President for Publications (2015–2016). Since 2023, he has been serving as NMO representative for Italy at the International Federation of Automatic Control (IFAC).



Mara Tanelli (Senior Member, IEEE) received the M.Sc. degree in computer engineering and the Ph.D. degree in information engineering, from Politecnico di Milano, Milan, Italy, in 2003 and 2007, respectively, and the second M.Sc. degree in computer science from the University of Illinois at Chicago, Chicago, IL, USA.

She is a Full Professor of automatic control with the Politecnico di Milano. She has coauthored more than 150 scientific publications and more than 15 patents in the areas of her research interests.

Her research interests include active control of ground vehicles, energy management of electric vehicles, and “Smart Mobility.”

Since 2013, Dr. Tanelli has been part of the Conference Editorial Board of the IEEE CSS. She is currently AE for *IEEE TRANSACTIONS ON CONTROL SYSTEMS TECHNOLOGY* and *IEEE TRANSACTIONS ON HUMAN-MACHINE SYSTEMS*. Since 2020, she has been the Chair of the IEEE TC of Automotive Controls of the IEEE CSS.

Open Access funding provided by ‘Consiglio Nazionale delle Ricerche’ within the CRUI CARE Agreement