Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 59-26 (2025) 151–156

The epistemic dimension in innovation diffusion and fair policy making ⋆

Eugenia Villa * **Camilla Quaresmini** * **Valentina Breschi** ** **Viola Schiaffonati** * **Mara Tanelli** *

* *Department of Electronics, Information and Bioengineering, Politecnico di Milano, Milano, 20133, Italy*
** *Department of Electrical Engineering, Eindhoven University of Technology, Eindhoven, MB 5600, The Netherlands*

Abstract: While many green technologies have clear environmental advantages, individual resistance and societal inertia can slow down their widespread adoption. Overcoming these barriers requires combining governmental incentives with social structures that promote diffusion through imitation. However, discrimination based on social identity can hinder this process. By formalizing the epistemic properties of a social environment through individual *credibility* and *reliability*, along with the concept of epistemic fairness, we analyze the impact of epistemic unfairness on innovation diffusion in open-loop and closed-loop scenarios. Our findings underscore the influence of epistemic factors on adoption dynamics and fair policy design, offering insights into solutions for a fairer dissemination of sustainable technologies.

Keywords: Algorithmic Fairness, Innovation Diffusion, Optimal Control, Fair Policy Design

1. INTRODUCTION

As social interactions shape individual predisposition to embrace new technologies (see, e.g., Jackson et al. (2008); Ravazzi et al. (2021)), it is paramount to systematically account for their impact when analyzing innovation diffusion and designing policies to promote the adoption of new sustainable technologies. At the same time, these social bonds are often heterogeneous, with social dictates and prejudices leading to differences in how others' opinions are perceived and trusted (Rogers et al., 2014). In turn, such heterogeneity can exacerbate biases, marginalizing societal sectors, and radicalizing opinions, especially when implementing nudging policies disregarding it.

In light of these considerations, we combine theories from social epistemology, i.e., the science that studies the mechanisms underlying the acquisition and validation of knowledge within a society (Fricker, 2007), and a simple, yet widely adopted, cascaded model (Granovetter, 1978) to make an initial step toward a formal representation of the *epistemic* dimension in innovation diffusion, i.e., in formalizing how individuals are perceived and trusted/distrusted in their roles of "knowers" based on their social attributes (Fricker, 2007). This perspective advances a broader notion of fairness in policy design by integrating individuals' epistemic standing within social systems, alongside classical distributive concerns. Specifically, we characterize innovation diffusion via a Linear Threshold Model (LTM) (Granovetter, 1978) over a weighted graph, describing adoption dynamics as a cascading process driven by personal attitudes and heterogeneous social connections.

Nonetheless, differently from the extensions of the LTM on weighted networks already proposed in the literature (see, e.g., Kempe et al. (2003); Cox et al. (2017)), we provide a novel, epistemic-based reinterpretation of such a cascade mode by linking the weights in our social network to the individual level of credibility on the considered innovation based on social features. As one's credibility can be affected by *epistemic biases*, in turn generating "credibility deficits" (i.e., individuals are systematically distrusted based on their attributes), this reinterpretation allows us to formalize the concept of *epistemic fairness* in innovation diffusion, enabling us to analyze the impact of epistemic biases on free and "forced" [1] innovation diffusion. The latter analysis focuses on optimal (in the linear quadratic sense (Bertsekas, 2012)) policies aimed at reducing individuals' resistance to the adoption of new technologies while reducing costs. Our analyses allow us to shed light on how epistemic biases impact the effectiveness of such promoting policies, indicating possible venues to reducing epistemic bias and achieving a more fair distribution of resources (Sinha et al., 2022; Tarekegne, 2020).

The paper is organized as follows. Section 2 introduces the LTM over weighted networks and its epistemic reinterpretation. We then formalize the concept of epistemic fairness in Section 3, according to which we evaluate the impact of epistemic biases in free (Section 4) and forced (Section 5) adoption dynamics. We present the results of a numerical analysis on the closed-loop impact of epistemic biases in Section 6, concluding the paper with some remarks and directions for future work.

⋆ The work was partially supported by PRIN project TECHIE (Cod. 2022KPHA24 CUP: D53D23001320006), MOST project founded by the European Union NextGenerationEU (PNRR – MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4-D.D. 1033 17/06/2022, CN00000023 SPOKE 5), MUSA project funded by the European Union NextGenerationEU (PNRR Mission 4 Component 2 Investment Line 1.5, ECS 00000037).

[1] When fostering interventions are put in place by a policymaker.

10.1016/j.ifacol.2025.12.026

2. AN EPISTEMICALLY GROUNDED MODEL FOR INNOVATION DIFFUSION

Notation The set of natural numbers (including zero) is denoted as $\mathbb{N}$ ($\mathbb{N}_0$), while the operator *or* is indicated as $\vee$. Given a set $\mathcal{A}$, its cardinality is denoted with $|\mathcal{A}|$, and its complementary is indicated as $\mathcal{A}^c$. Given a, $b \in \mathbb{R}$, $[a, b]$ denotes the set of real numbers between a and b, while $\{a, b\}$ denotes the set constituted by a and b only.

To model innovation diffusion, we rely on the Linear Threshold Model (LTM) (Granovetter, 1978), i.e., we postulate the adoption of a new technology over a set $\mathcal{V}$ of socially connected individuals (also referred to as *agents*) to be driven by their predisposition to adoption and the influence of neighbors. Therefore, we associate each agent with a binary state $a_x(t) \in \{0, 1\}$, with $a_x(t) = 1$ if the x-th agent is an adopter at time $t \in \mathbb{N}$ and $a_x(t) = 0$ otherwise, for all $x \in \mathcal{V}$. For these binary variables, we make the following assumption.

Assumption 1. *There exists a non-empty seed set* $S^\star(0) = \{x \in \mathcal{V} | a_x(0) = 1\}$ *of initial adopters.*

Therefore, as they are the initial adopters, the agents in $S^\star(0)$ are responsible for triggering the diffusion process within the considered social network. Differently from the standard LTM, the social network is here described as a *weighted*, strongly connected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{W})$, where $\mathcal{V}$ represents the individuals in the network, $\mathcal{E}$ characterizes the social connections between them, while each weight $w_{x,y}$ in the symmetric matrix $\mathcal{W} = \mathcal{W}^\top$ allow us to introduce the epistemic dimension into the model according to the concept of epistemic injustice theorized in Fricker (2007). In particular, we assume the following.

Assumption 2. *Each agent* $x \in \mathcal{V}$ *has a* credibility $\gamma_x \in [0, 1]$, *such that* $w_{x,y} = \gamma_x$ *for all* $y \neq x$, $y \in \mathcal{V}$.

The parameter $\gamma_x \in [0, 1]$ introduced in Assumption 2 represents how much an agent is *trusted* by others, allowing us to model an agent's adoption propensity as a combination of personal attitude, influence from neighbors, and the epistemic characterization of the environment.

For the time evolution of the binary variables $\{a_x(t)\}_{x \in \mathcal{V}}$, we maintain the same assumptions of the standard LTM.

Assumption 3. *Adoption is driven by the* relative popularity *of the new technology, and it is* irreversible *(once an adopter, an agent remains an adopter for all future times).*

Assumption 4. *The decisions of agents are influenced by their resistivity* $\rho_x \in [0, 1]$, *with* $x \in \mathcal{V}$.

Innovation diffusion is thus modeled by

$$a_x(t+1) = \begin{cases} 1 \text{ if } & a_x(t) = 1 \vee \dfrac{\sum_{y \in N_x^*(t)} \gamma_y}{\sum_{y \in N_x} \gamma_y} \geq \rho_x, \\ 0 \text{ otherwise,} & \end{cases} \tag{1}$$

for all $t \in \mathbb{N}_0$ and $x \in \mathcal{V}$, with $N_x = \{y \in \mathcal{V} | (x, y) \in \mathcal{E}\}$ being the set of neighbors of $x \in \mathcal{V}$, which influence the opinion of the x-th agent toward adoption, and $N_x^* \subseteq N_x$ is its set of neighboring adopters. In this way, agents become adopters if the influence of their adopting neighbors, weighted by their credibility values, exceeds their resistivity thresholds $\{\rho_x\}_{x \in \mathcal{V}}$.

Remark 1 (Limitation of (1))**.** *Even if simplifying the description of the adoption dynamics by assuming irreversibility and not considering stochastic elements is justifiable for short-term analyses, these simplifications lay the groundwork for alternative (stochastic) models.*

3. EPISTEMIC (UN)FAIRNESS & ITS IMPACT

Parallel to credibility, representing how much agents are trusted by their neighbors, we introduce *reliability*, namely

$$r_x \in [0, 1], \quad x \in \mathcal{V}. \tag{2}$$

Determined by quantifiable factors like education and experience, this additional individual feature indicates how much agents should be trusted by others without prejudice or, in other words, their potential capacity to influence peers based on their actual knowledge of the technology of interest. Though not explicitly defined in Fricker (2007), we use reliability to formalize the condition in which subjects are perceived as less credible than they should, thus suffering *epistemic injustice*. In particular, without bias, reliability equals credibility, i.e.,

$$\gamma_x = r_x, \quad \forall x \in \mathcal{V}. \tag{3}$$

On the contrary, in the presence of bias, individuals are subject to a credibility deficit $\Delta_x = r_x - \gamma_x$, $\Delta_x \in [0, r_x]$, for $x \in \mathcal{V}$, dictated by proxy factors like gender, ethnicity, or age, which reduces an agent's reliability up to their credibility. We can thus formalize *epistemic fairness* as follows.

Definition 3.1 (Epistemic Fairness)**.** *A social network* $\mathcal{G}$ *is said epistemically fair, i.e.,* $\gamma_x = r_x$, *and is thus denoted as EF(*$\mathcal{G}$*), if and only if*

$$\Delta_x = 0, \quad \forall x \in \mathcal{V}. \tag{4}$$

The condition in (4) ensures that agents are attributed the credibility they deserve in an ideal, free-from bias scenario. At the same time, in line with Fricker (2007), our definition of epistemic fairness relies on the simplifying assumption that credibility deficits $\{\Delta_x\}_{x \in \mathcal{V}}$ are only shaped by individual attributes, neglecting their relational dimensions. Note that, as the importance of the latter aspect is highlighted in empirical studies (see, e.g., Mahmoodi et al. (2018)), we will consider it in future work.

3.1 Investigating epistemic (un)fairness: useful results

In the footsteps of Acemoglu et al. (2011), we now analyze the effects of introducing the epistemic dimension (i.e., credibility and reliability) in the diffusion process modeled by (1). Toward this objective, we extend the definition of cohesive set provided in Acemoglu et al. (2011) as follows.

Definition 3.2 (Epistemically cohesive set)**.** *A set* $X \subseteq \mathcal{V}$ *is epistemically cohesive if*

$$\frac{\sum_{y \in (X \cap N_x)} \gamma_y}{\sum_{y \in N_x} \gamma_y} > 1 - \rho_x, \quad \forall x \in X. \tag{5}$$

Therefore, for a set $X \subseteq \mathcal{V}$ to be epistemically cohesive, the ratio between the credibility of neighboring agents belonging to X and the influence of all the neighboring agents has to be larger than $1 - \rho_x$ for all $x \in X$. The properties of cohesive sets in Acemoglu et al. (2011) also apply to the proposed epistemic extension.

Moreover, we introduce the set $\mathcal{S}(t)$ of agents switching to adoption at time $t \in \mathbb{N}$, namely

$$\mathcal{S}(t) = \{x \in \mathcal{V} | a_x(t-1) = 0 \wedge a_x(t) = 1\}, \tag{6}$$

and, accordingly, we provide the following definition of a fixed point in the innovation diffusion dynamics.

Definition 3.3 (Fixed point). *The set $\bar{S}$ is a fixed point for the dynamics dictated by* (1) *if*

$$S^\star(\tau) = \bar{S} \Rightarrow \mathcal{S}(t) = \emptyset, \quad \forall t > \tau \geq 0. \tag{7}$$

By following the same steps proposed by Acemoglu et al. (2011), we can thus arrive at a characterization of the final set of adopters, namely

$$S^{\star\star} = \{x \in \mathcal{V} | a_x(t) = 1, t \to +\infty\}, \tag{8}$$

for which the following result holds.

Corollary 3.4. *Let $\mathcal{M}$ be the largest cohesive subset of $\mathcal{V} \setminus S^\star(0)$. Then, $S^{\star\star} = \mathcal{M}^c$.*

The proof of this corollary follows the same steps carried out in Acemoglu et al. (2011), and it is thus omitted.

3.2 Epistemically fair vs *unfair settings in free evolution*

By leveraging Corollary 3.4, we focus on theoretically comparing $S^{\star\star}$ in an *epistemically fair* scenario, i.e., $\gamma_x = r_x$ for all $x \in \mathcal{V}$, and in an *epistemically unfair* one, namely when $\gamma_x = r_x - \Delta_x \leq r_x$ for all $x \in \mathcal{V}$. To this end, given a set $X \subseteq \mathcal{V}$, let us introduce

$$r_x^\alpha = \sum_{y \in N_x^\alpha} r_y, \quad r_x^\beta = \sum_{y \in N_x^\beta} r_y, \tag{9}$$

with $N_x^\alpha = X \cap N_x$ and $N_x^\beta = N_x \setminus X$ such that

$$N_x^\alpha \cap N_x^\beta = \emptyset, \quad N_x^\alpha \cup N_x^\beta = N_x.$$

By Definition 3.2, in an *epistemically fair* setting, the set X is epistemically cohesive if the following holds

$$\frac{r_x^\alpha}{r_x^\alpha + r_x^\beta} > 1 - \rho_x, \quad \forall x \in X. \tag{10}$$

Such characterization changes in an *epistemically unfair* since the agents' reliability has to be replaced with their credibility. In particular, one has that:

$$\sum_{y \in N_x^\alpha} \gamma_y = \sum_{y \in N_x^\alpha} r_y - \sum_{y \in N_x^\alpha} \Delta_y = r_x^\alpha - \Delta_x^\alpha, \tag{11a}$$

$$\sum_{y \in N_x^\beta} \gamma_y = \sum_{y \in N_x^\beta} r_y - \sum_{y \in N_x^\beta} \Delta_y = r_x^\beta - \Delta_x^\beta, \tag{11b}$$

with $\Delta_x^\alpha \in [0, r_x^\alpha)$ and $\Delta_x^\beta \in [0, r_x^\beta)$ are induced by the credibility deficit, for all $x \in \mathcal{V}$. Hence, with social bias, X is epistemically cohesive (see (5)) if

$$\frac{\sum_{y \in N_x^\alpha} \gamma_y}{\sum_{y \in N_x} \gamma_y} = \frac{r_x^\alpha - \Delta_x^\alpha}{r_x^\alpha - \Delta_x^\alpha + r_x^\beta - \Delta_x^\beta} > 1 - \rho_x, \ \forall x \in X. \tag{12}$$

These results translate into two conditions for the epistemic cohesiveness of X to be preserved even in the face of bias, as formalized in the following Lemmas.

Lemma 3.5. *Let $X \subseteq \mathcal{V}$ be a cohesive set in an* epistemically fair *scenario. Then, X is still cohesive in an* epistemically unfair *setting if*

$$\Delta_x^\beta r_x^\alpha > \Delta_x^\alpha r_x^\beta, \quad \forall x \in X. \tag{13}$$

Proof. The proof is in the Appendix. □

Lemma 3.6. *Let $X \subseteq \mathcal{V}$ be a cohesive set in an* epistemically unfair *scenario. Then, X is still cohesive in an* epistemically fair *setting if*

$$\Delta_x^\alpha r_x^\beta > \Delta_x^\beta r_x^\alpha, \quad \forall x \in X. \tag{14}$$

Proof. The proof unfolds as that of Lemma 3.5, and is thus omitted. □

These lemmas emphasize the impact of bias on cohesive sets and, consequently, on the final set of adopters (see Corollary 3.4), foreshadowing the results discussed in the numerical example in Section 4. Lemma 3.6 provides a sufficient condition to bound the set $S^{\star\star}$, applicable in both fair and unfair settings, indicating that the innovation will not reach $\mathcal{M}^{\text{fair}}$ when $\mathcal{M}^{\text{fair}} \cap S^\star(0) = \emptyset$, but it could also stop spreading before reaching $\mathcal{M}^{\text{fair}}$. Meanwhile, when the ratio of credibility deficit to reliability is equal for both adopters and non-adopters, i.e., $\Delta_x^\beta r_x^\alpha = \Delta_x^\alpha r_x^\beta, \ \forall x \in X$, the diffusion patterns in the unfair scenario align with those in the fair case, leading to the same final set of adopters. Therefore, the epistemic dimension only impacts diffusion when biases affect agents unevenly, with uniform biases nullifying their effects.

4. EPISTEMIC BIASES IN OPEN-LOOP

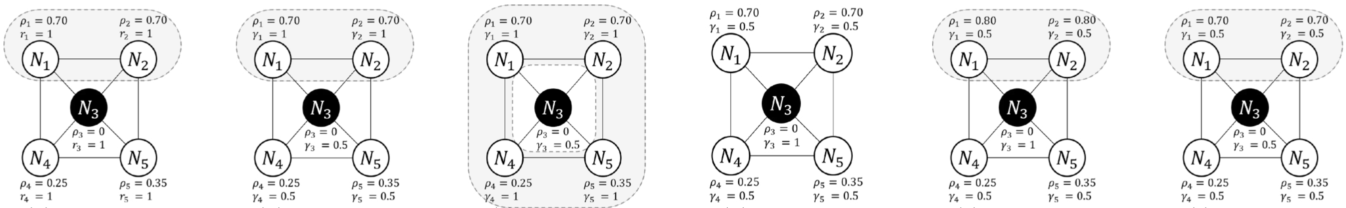
We now exemplify the theoretical results introduced in the previous section via a set of illustrative case studies. In particular, we consider the network with 5 agents, i.e., $\mathcal{V} = \{N_1, N_2, N_3, N_4, N_5\}$, shown in Fig. 1. Note that, in all cases, only agent N_3 belongs to the seed set, i.e., $S^\star(0) = \{N_3\}$. Such an agent is also the only one *directly* connected to all others.

Example 1 – nominal conditions (see Fig. 1(a)). All nodes are fully reliable and, thus, have the same influence on their neighbors. By Definition 3.2, the cohesive set of $\mathcal{V} \setminus S^\star(0)$ is unique and corresponds to $\mathcal{M}_1 = \{N_1, N_2\}$. By Corollary 3.4, this implies $S^{\star\star} = \mathcal{M}_1^c = \{N_3, N_4, N_5\}$.

Example 2 – credibility deficit on $\mathcal{M}_1^c$ (see Fig. 1(b)). With respect to the nominal case, now the agents N_3, N_4, and N_5 are affected by a deficit that halves their credibility with respect to the nominal value. Instead, $N_1, N_2 \in \mathcal{M}_1$ satisfy the condition of Lemma 3.5, as $\Delta_x^\beta r_x^\alpha = 1.5 > 0 = \Delta_x^\alpha r_x^\beta$. This still makes $\mathcal{M}_1$ an epistemically cohesive set. By checking all possible partitions of $\mathcal{V} \setminus S^\star(0)$, we find that the largest cohesive set $\mathcal{M}_2$ is (in this case) $\mathcal{M}_1$. Hence, the innovation does not spread to N_1 and N_2, leading to the same final adopters of the nominal case.

Example 3 – lower bound on the largest cohesive set (see Fig. 1(c)). By halving γ_3 with respect to its nominal value, we now assign a credibility deficit to agent N_3 only. In this scenario, $\mathcal{M}_1$ is still a cohesive set, since $\Delta_x^\beta r_x^\alpha = 1 > 0 = \Delta_x^\alpha r_x^\beta$, but it is not the only one. Indeed, also $\mathcal{M}_3 = \{N_1, N_2, N_4, N_5\} \supset \mathcal{M}_1$ is cohesive, actually representing the largest epistemically cohesive set in this scenario. This result confirms that, if the condition in Lemma 3.5 holds, we can only find a lower bound for the largest cohesive set. Moreover, innovation stops at the seed agent N_3 as the credibility deficit affects only $S(0)$.

Example 4 – credibility deficit on $\mathcal{M}_1$ (see Fig. 1(d)). If both $\mathcal{M}_1$ and $\mathcal{M}_3$ are affected by a credibility deficit, the sufficient condition in (13) does not hold. Indeed, $\Delta_x^\beta r_x^\alpha = 1 \not> 2 = \Delta_x^\alpha r_x^\beta$. This implies that $\mathcal{M}_1$ is not cohesive anymore. As a consequence, innovation can now also spread across N_1 and N_2. By further analyzing the network, it can be shown that there are no cohesive subsets in $\mathcal{V} \setminus S^\star(0)$, and that $\mathcal{M}_4 = \emptyset$. Therefore, the innovation spreads across the whole network thanks to the *positive impact* of credibility imbalances (especially when the credibility deficit affects all but the seed agents).



(a) Example 1 (b) Example 2 (c) Example 3 (d) Example 4 (e) Example 5 (f) Example 6

Fig. 1. Illustrative examples: filled nodes are initial adopters, while gray areas denote the largest cohesive sets. The fourth example does not feature a colored area since the associated largest cohesive set is the empty set.

Note that this result also rests on the values of individual resistivity, and hence, it might not hold any longer if they change.

Example 5 – counterexample with increased resistivity (see Fig. 1(e)). We now maintain the same epistemic conditions as in Fig. 1(d), but we increase the resistivity of N_1 and N_2. This implies that (13) does not hold. At the same time, analyzing all subsets of $\mathcal{V} \setminus S^\star(0)$, it can be seen that this scenario features a unique cohesive set $\mathcal{M}_5 = \{N_1, N_2\} = \mathcal{M}_1$. Hence innovation spreads once again only across the agents N_3, N_4, N_5.

Example 6 – uniform credibility deficit (see Fig. 1(f)). By setting uniform credibility deficits for all agents in the network, we unsurprisingly obtain that the largest cohesive set in the network $\mathcal{M}_6$ overlaps with $\mathcal{M}_1$, leading to the same outcome as the epistemically fair scenario.

All these results highlight that credibility deficits do not necessarily hinder the diffusion of an innovation. Nonetheless, its spread is slowed if only seed agents are affected by a credibility deficit. Meanwhile, when the credibility deficit affects seeds and/or non-seeds, diffusion is also driven by the other actors of the adoption model (e.g., individual resistivity and network topology) and, thus, the impact of credibility deficits cannot be judged a priori.

5. CREDIBILITY DEFICITS & CLOSED-LOOP INNOVATION DIFFUSION

Innovation diffusion models explicitly accounting for social biases can be key in designing resource allocation policies to nudge virtuous behaviors. Toward analyzing the effect of social bias (in the form of a credibility deficit) in closed-loop, we extend the epistemic-based LTM in (1) to include controlled inputs in the footsteps of Villa et al. (2023). Specifically, we make the following assumption.

Assumption 5. *External policies $\{u_x(t)\}_{x\in\mathcal{V}}$, with $t \in \mathbb{N}_0$, act on individual resistivity, i.e.,*

$$\rho_x^u(t+1) = \rho_x^u(t) + b_x u_x(t), \quad \forall x \in \mathcal{V},\ t > 0, \tag{15}$$

where $\rho_x^u(t) \in [0,1]$ and $\rho_x^u(0) = \rho_x$, with ρ_x featured in (1) and $b_x \in [-1, 0)$ for all $x \in \mathcal{V} \setminus S^\star(0)$.

Under this assumption, external actions are designed to modify adoption barriers in individual choices, reducing them if $u_x(t) \geq 0$ for all $t \in \mathbb{N}$ and $x \in \mathcal{V} \setminus S^\star(t)$. The effect of the external policy is modulated by b_x, capturing individual differences in policy responsiveness. In EV adoption, for instance, incentives like price discounts may drive adoption for some but have limited impact on others due to non-economic factors (Breschi et al., 2023; Mundaca and Samahita, 2020). Based on (15), the adoption mechanism in (1) is thus modified as

$$a_x(t+1) = 1 \text{ if } a_x(t) = 1 \vee \frac{\sum_{y\in N_x^\star(t)} \gamma_y}{\sum_{y\in N_x} \gamma_y} \geq \rho_x^u(t). \tag{16}$$

5.1 Epistemic-based optimal Policy Design

We now assume that a policymaker oversees the network $\mathcal{G}$ and can allocate unlimited resources [2] to influence agents toward virtuous behaviors. The corresponding resource allocation problem is here solved via standard (finite horizon) linear quadratic regulation (LQR) (Bertsekas, 2012), seeking optimality with respect to conventional goals of policymakers, i.e., minimization of costs and maximization of policy impact. Therefore, we assume that the policymaker designs interventions by solving

$$\min_{U} \sum_{x\in\mathcal{V}} \sum_{\tau=0}^{T-1} \left[\omega_x^\rho(t) e_x^2(\tau) + \omega^u u_x^2(\tau)\right] + \omega_x^\rho(T) e_x^2(T) \tag{17a}$$

$$\text{s.t.} \quad (15), (16), \quad \forall x \in \mathcal{V},$$

where $T > 0$ is a prefixed optimization horizon, $U = \{u_x(0), \ldots, u_x(T-1)\}_{x\in\mathcal{V}}$, $\omega_x^\rho(t) = \bar{\omega}^\rho(1 - a_x(t))$ so that tracking performance is not penalized for adopters, and $\bar{\omega}^\rho, \omega^u > 0$ are tunable by the policymaker to balance between conservative actions and effective promotion strategies. Meanwhile, $e_x(\tau)$ is the tracking error for given the set point $\bar{\rho}_x$ at time $\tau \in \mathbb{N}_0$, i.e.,

$$e_x(\tau) = \rho_x^u(\tau) - \bar{\rho}_x, \quad \tau = 0, \ldots, T, \quad x \in \mathcal{V}, \tag{17b}$$

while the last term in the loss is a terminal cost, enforcing error minimization at the end of the horizon. Note that, in line with (Villa et al., 2023), maximizing the spread of the innovation over the social network is sought by reducing individual resistivity toward the target set point

$$\bar{\rho}_x = \frac{\sum_{y\in N_x^\star(0)} \gamma_y}{\sum_{y\in N_x} \gamma_y}, \quad \forall x \in \mathcal{V} \setminus S^\star(0), \tag{18}$$

i.e., the minimum relative popularity of the innovation among an agent's neighbors required for the agent to become an adopter at $t = 0$. The reason for this choice of setpoint is two-fold. On the one hand, reducing individual resistance implies that less influence of neighboring adopters is required for one to become an adopter (thereby facilitating innovation diffusion through contagion). On the other hand, since $\bar{\rho}_x = \rho_x^u(0) \leq \rho_x^u(t)$, for all $x \in \mathcal{V}$ and all $t \in \mathbb{N}$, picking $\bar{\rho}_x$ allows us to contain the effort made by the policymaker, not demanding for excessive (and unneeded) reductions of individual resistivity.

As in Villa et al. (2023), (17) is solved in a *receding horizon fashion*, i.e., (17) is solved at each time $t \in \mathbb{N}_0$, but only the first optimal action is applied before proceeding to

[2] This assumption allows us to provide analytical results on the impact of credibility deficits in closed-loop. We will relax it in future works.

the next iteration and repeating the previous procedure. Standard arguments (see, e.g., Bertsekas (2012)) can be used to prove that $u_x(t)$ at time t is given by

$$u_x(t) = \kappa_x(t)(\rho^u_x(t) - \bar{\rho}_x), \tag{19}$$

with $\kappa_x(t) = \tilde{\kappa}_x(0|t) \in \mathbb{R}$ computed (backwards) through

$$\tilde{\kappa}_x(\tau|t) = -\frac{b_x\tilde{p}_x(\tau+1|t)}{\omega^u + \tilde{p}_x(\tau+1|t)b_x^2}, \tag{20a}$$

$$\tilde{p}_x(\tau|t) = \omega^\rho_x(t) + \tilde{p}_x(\tau+1|t) - \frac{b_x^2\tilde{p}_x^2(\tau+1|t)}{\omega^u + \tilde{p}_x(\tau+1|t)b_x^2}, \tag{20b}$$

for $\tau = T-1,\dots,0$, starting from $\tilde{p}_x(T|t) = \omega^\rho_x(T)$, for all $x \in \mathcal{V} \setminus S^\star(t)$. Note that, with this choice of the terminal cost-to-go, (19) computed as in (20) is the exact and unique solution of the strictly convex problem (17).

5.2 Epistemic dimension in closed-loop dynamics

By relying on the credibility deficit to characterize epistemic (un)fairness, we now analyze the impact of feedback-based policies designed according to (20). To this end, note that with unlimited resources, policies designed via (19) progressively lower resistivity thresholds, ensuring full network adoption through social contagion. Thus, in this closed-loop setting, the impact of a credibility deficit affects the *transient diffusion process* rather than the final adoption outcome. We hence analyze the former, by first noticing that $k_x(t)$ in (19) does not depend on the credibility deficit (see (20)), while the epistemic dimension affects the targets $\bar{\rho}_x$ to which the input is proportional (see (18)). Moreover, note that (19) can be equivalently recast as

$$u_x(t) = \underbrace{\kappa_x(t)(\rho^u_x(t) - 1)}_{=c_1(t)} + \kappa_x(t)\frac{\sum_{y\in(N^\star_x(0))^c}\gamma_y}{\sum_{y\in N_x}\gamma_y}, \tag{21}$$

where $c_1(t) \leq 0$ depends on the features of the control law for all $t \in \mathbb{N}_0$. Accordingly, the (nudging) input to an agent depends on individual predisposition and decreases as the credibility of their neighbors increases. As a result, agents with high resistance and low epistemic influence receive more resources, while those with low resistance and high influence require minimal intervention. Based on this result, we can now formalize the differences in input allocation and evolution between an epistemically fair and unfair closed-loop scenario as follows.

Lemma 5.1. *Given $x \in \mathcal{V} \setminus S^\star(t)$, assume that $\rho^u_x(t)$ and $a_x(t)$ are equal for the epistemically fair and unfair scenarios at a given time $t \in \mathbb{N}_0$. Then, the input fed to the agent x at time t and the reduction of the its resistivity at time $t+1$ are lessened by credibility deficits if*

$$\Delta^\beta_x r^\alpha_x > \Delta^\alpha_x r^\beta_x, \tag{22}$$

with Δ^α_x, r^α_x, Δ^β_x and r^β_x defined as in (11) and $N^\alpha_x = N_x \cap S^\star(0)$ and $N^\beta_x = N_x \cap (S^\star(0))^c$.

Proof. The proof can be found in the Appendix. □

6. CLOSED-LOOP NUMERICAL EXAMPLE

To empirically validate our formal results on the impact of social bias in closed-loop, we consider a population of $|\mathcal{V}| = 20$ individuals, each interacting with a subset of randomly assigned agents[3]. The agents in the seed set are also randomly selected, representing 10% of the total population, i.e., $|S^\star(0)| = 2$, while the others are assumed to verify $\rho^u_x(0) = 0.8$, $r_x = 1$ while $b_x = -1$ for all $x \in \mathcal{V} \setminus S^\star(0)$. For policy design, we further assume that $\bar{\omega}^\rho = \omega^u = 1$ and set $T = 10$ (see (17)). Under these conditions, we compare the three scenarios depicted in Fig. 2: $\gamma_x = \gamma_y$, $\forall x, y \in \mathcal{V}$ (all agents have the same credibility deficit); seeds are less credible than non-seeds, namely $\gamma_x < \gamma_y$, for all $x \in S^\star(0)$ and $y \in \mathcal{V} \setminus S^\star(0)$; seeds are more credible than non-seeds, i.e., $\gamma_x > \gamma_y$, for all $x \in S^\star(0)$ and $y \in \mathcal{V} \setminus S^\star(0)$. As all agents are equally (non-)credible, the first setting is ultimately an epistemically fair one, which is instead not the case for the other two. The attained results are quantitatively compared via the following indicator:

Table 1. Performance indexes

Scenario	$\bar{C}_x$	t^{**}
1	0.6954	7
2	0.7907	8
3	0.3448	6

$$\bar{C}_x = \frac{1}{N}\sum_{x\in\mathcal{V}}\sum_{t=1}^{T} u_x(t), \tag{23}$$

denoting the average individual policy cost, as well as the time t^{**} required to achieve full acceptance, i.e., $\{x \in \mathcal{V} | a_x(t^{**}) = 1\} \equiv \mathcal{V}$.

Table 1 confirms our theoretical expectations. Cost increases when only seeds face a credibility deficit, while adoption is slowed. Conversely, costs and adoption time decrease when credibility deficits affect non-seeds, a case in which epistemic biases might be beneficial in the cascaded processes. Meanwhile, Fig. 3 shows that higher inputs are required when non-seeds have greater influence. Indeed, when reducing seeds' influence (Scenario 2), agents must rely on mindset changes induced by external actions to embrace the new technology. Instead, when higher influence is attributed to adopters (Scenario 3), adoption spreads through imitation, requiring lower external efforts. These results align with Watts (2002), highlighting the double facet of credibility in innovation diffusion.

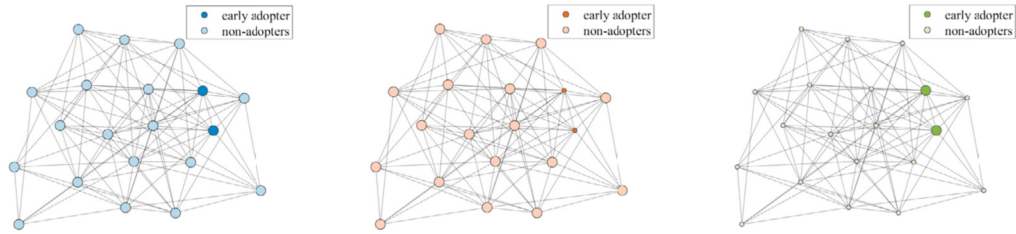
7. CONCLUSIONS

By incorporating individual credibility into an irreversible cascade model, we have analyzed the role of epistemic fairness in innovation diffusion and policy design. In particular, considering the epistemic dimension of the network enables us to recognize and mitigate forms of harm that remain invisible to classical models. This, in turn, supports the design of policies that are fair not only in distributive terms but also in addressing epistemic injustice. Our findings show that neglecting epistemic factors in nudging strategies can reinforce biases and lead to unfair outcomes, underscoring that this aspect can no longer be overlooked in policy design. Future research will refine the notion of credibility and the attribution of the related deficit, and study how to better adapt the diffusion model to encompass these epistemic concepts.

REFERENCES

Acemoglu, D., Ozdaglar, A., and Yildiz, E. (2011). Diffusion of innovations in social networks. In *2011 50th IEEE Conference on Decision and Control and European Control Conference*, 2329–2334. IEEE.

[3] The probability that an edge exists is set to 0.5.



(a) Scenario 1 (b) Scenario 2 (c) Scenario 3

Fig. 2. Social network *vs* epistemic scenarios: node colors and sizes indicate agents' initial statuses and credibility.

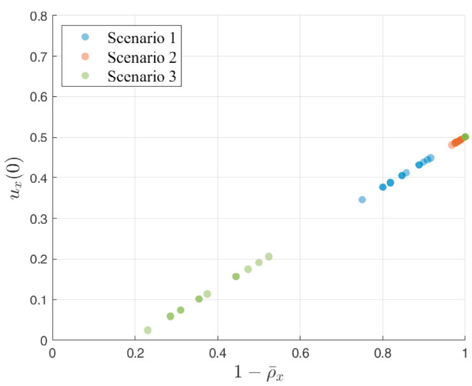


Fig. 3. Control actions *vs* $1-\bar{\rho}_x$ at $t=0$.

Bertsekas, D. (2012). *Dynamic programming and optimal control: Volume I*, volume 1. Athena scientific.

Breschi, V., Ravazzi, C., Strada, S., Dabbene, F., and Tanelli, M. (2023). Driving electric vehicles' mass adoption: An architecture for the design of human-centric policies to meet climate and societal goals. *Transportation Research Part A: Policy and Practice*, 171, 103651.

Cox, S., Horadam, K.J., and Rao, A. (2017). The spread of ideas in a weighted threshold network. In *Complex Networks & Their Applications V: Proceedings of the 5th International Workshop on Complex Networks and their Applications (COMPLEX NETWORKS 2016)*, 437–447. Springer.

Fricker, M. (2007). *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford University Press, New York.

Granovetter, M. (1978). Threshold models of collective behavior. *American Journal of Sociology*, 83(6), 1420–1443.

Jackson, M.O. et al. (2008). *Social and economic networks*, volume 3. Princeton university press Princeton.

Kempe, D., Kleinberg, J., and Tardos, É. (2003). Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, 137–146.

Mahmoodi, A., Bahrami, B., and Mehring, C. (2018). Reciprocity of social influence. *Nature Communications*, 9(1), 2474.

Mundaca, L. and Samahita, M. (2020). What drives home solar PV uptake? Subsidies, peer effects and visibility in Sweden. *Energy Research & Social Science*, 60, 101319.

Ravazzi, C., Dabbene, F., Lagoa, C., and Proskurnikov, A.V. (2021). Learning hidden influences in large-scale dynamical social networks: A data-driven sparsity-based approach, in memory of roberto tempo. *IEEE Control Systems Magazine*, 41(5), 61–103.

Rogers, E.M., Singhal, A., and Quinlan, M.M. (2014). Diffusion of innovations. In *An integrated approach to communication theory and research*, 432–448. Routledge.

Sinha, A., Balsalobre-Lorente, D., Zafar, M.W., and Saleem, M.M. (2022). Analyzing global inequality in access to energy: Developing policy framework by inequality decomposition. *Journal of environmental management*, 304, 114299.

Tarekegne, B. (2020). Just electrification: Imagining the justice dimensions of energy access and addressing energy poverty. *Energy Research & Social Science*, 70, 101639.

Villa, E., Breschi, V., Ravazzi, C., Dabbene, F., and Tanelli, M. (2023). Fostering the use of sharing mobility solutions via control-oriented policy design. *IFAC-PapersOnLine*, 56(2), 1–6.

Watts, D.J. (2002). A simple model of global cascades on random networks. *Proceedings of the National Academy of Sciences*, 99(9), 5766–5771.

PROOF OF LEMMA 3.5

To show that (13) guarantees that X is a cohesive set in an epistemically unfair scenario, thanks to (10), we only need to prove that there exists $\eta_x, \delta_x > 0$ such that

$$\frac{r_x^\alpha}{r_x^\alpha + r_x^\beta} + \frac{\eta_x}{\delta_x} = \frac{r_x^\alpha - \Delta_x^\alpha}{r_x^\alpha + r_x^\beta - (\Delta_x^\alpha + \Delta_x^\beta)} > 1 - \rho_x.$$

By imposing the first equality, we get

$$\delta_x = \frac{r_x^\alpha + r_x^\beta - (\Delta_x^\alpha + \Delta_x^\beta)}{r_x^\alpha + r_x^\beta}, \quad \eta_x = \frac{r_x^\alpha \Delta_x^\beta - r_x^\beta \Delta_x^\alpha}{(r_x^\alpha + r_x^\beta)^2},$$

with $\delta_x > 0$ by construction and $\eta_x > 0$ only if (13) is verified, concluding the proof.

PROOF OF LEMMA 5.1

From (21) it holds that

$$u_x(t) = \kappa_x(t)\left(\rho_x^u(t) - \nu_x^r\right) + \kappa_x(t)\left(\nu_x^r - \nu_x^\Delta\right),$$

where $\nu_x^r = r_x^\alpha/(r_x^\alpha + r_x^\beta)$, $\nu_x^\Delta = (r_x^\alpha - \Delta_x^\alpha)/(r_x^\alpha - \Delta_x^\alpha + r_x^\beta - \Delta_x^\beta)$. Since $\kappa_x(t) \geq 0$, for the credibility deficit to reduce the input given to the system, the following has to hold

$$\frac{r_x^\alpha\left(r_x^\alpha - \Delta_x^\alpha + r_x^\beta - \Delta_x^\beta\right) - \left(r_x^\alpha - \Delta_x^\alpha\right)\left(r_x^\alpha + r_x^\beta\right)}{\left(r_x^\alpha + r_x^\beta\right)\left(r_x^\alpha - \Delta_x^\alpha + r_x^\beta - \Delta_x^\beta\right)} < 0.$$

As the denominator on the ration above is positive, the input $u_x(t)$ is reduced when

$$r_x^\alpha\left(r_x^\alpha - \Delta_x^\alpha + r_x^\beta - \Delta_x^\beta\right) - \left(r_x^\alpha - \Delta_x^\alpha\right)\left(r_x^\alpha + r_x^\beta\right) < 0,$$

which, after straightforward manipulations, leads to the condition in (22). Our proof is concluded by plugging the expression for $u_x(t)$ obtained previously into (15).