2024 European Control Conference (ECC)
June 25-28, 2024. Stockholm, Sweden

# On Modeling Collective Risk Perception via Opinion Dynamics

Lorenzo Zino, Francesca Giardini, Daniele Vilone, and Ming Cao

***Abstract*— Modeling the collective response to an emergency is a problem of paramount importance in social science and risk management. Here, we leverage the social-psychology literature to develop a mathematical model tailored to such a problem. In our model, a network of individuals revises their risk perception by processing information broadcast by the institution and shared by peers, accounting for heterogeneity in terms of individuals' trust in institutions, peers, and risk sensitivity. Analyzing the model, we establish that the temporal average opinions of the individuals converge to a steady state and, under some assumptions, we are able to analytically characterize such a steady state, shedding light on how the individuals' heterogeneous risk perception shapes the collective response.**

## I. Introduction

The development and analysis of mathematical models for social dynamics have witnessed an increasing interest in the systems and control community, providing novel theoretically-informed tools to understand and predict collective human behavior [1]–[8]. In particular, a key area of research focuses on studying opinion formation in social communities through the lens of opinion dynamics models [9]–[13], in which individuals' opinions evolve over time through a linear averaging process that accounts for the information exchanged with peers on a social network.

Concerning opinion formation, of particular interest is to understand the emergent behavior of a population during an emergency [14]. In this situation, it is crucial to have insights about how the individuals of a population collectively respond to the information that public authorities and institutions broadcast on the nature of the risk of the event under consideration in order to avoid underestimating the risk or, on the other extreme, emergence of panic reactions.

Despite the importance of such problem, the literature presents few mathematical models of opinion formation tailored specifically to such a scenario. On the one hand, classical mathematical models focus on an abstract representation of opinion dynamics [12]; on the other hand, social-psychological efforts are mostly concerned with unveiling the individual-level risk interpretation process [15]–[19], typically overlooking how such individual-level mechanism propagates at the population-level. In [20]–[22], different agent-based models tailored to capturing the emergence of collective risk perception about an emergency have been proposed and used to perform numerical simulations. However, the complexity of such models hinders rigorous analytical studies, calling for the development of new analytically-treatable mathematical models for collective risk perception.

Here, we fill in this gap by proposing a novel analytically-treatable model for collective risk perception, grounded in the theory of opinion dynamics [11], [23]. In particular, inspired by [21] and building on the social-psychology literature [15]–[19], we consider a network of interacting individuals who are forming their opinion on the risk of a given emergency. Specifically, individuals are exposed to two different sources of information: an evaluation of the risk which is officially broadcasted by the institutions and a local risk perception shared by peers on a dynamical social influence network [16], [17]. Consistent with the social-psychology literature on risk interpretation [18], individuals recursively revise their risk perception by processing these different information sources through their own risk sensitivity [19].

Our contribution is threefold. First, we propose the mathematical model for collective risk perception and we demonstrate that it can be cast as a generalized version of the well-known Friedkin–Johnsen opinion dynamic model [9] on a time-varying network. However, the complexity of the network formation process hinders its direct analysis using standard techniques [11], [23]. Second, we prove that, while individuals' opinions in general tend to keep oscillating, their temporal average converge under mild assumptions on the network structure. Third, under some assumptions, we analytically characterize the steady-state temporal average opinion, showing how the presence of individuals with high risk sensitivity could lead to overreactions and panic.

*Notation:* We denote the set of nonnegative and strictly positive integer numbers by $\mathbb{N}$ and $\mathbb{N}_+$, respectively. A vector $\boldsymbol{x}$ is denoted with bold lower-case font, with $i$th entry $x_i$ and $\boldsymbol{x}^\top$ denoting its transpose; a matrix $\boldsymbol{A}$ is denoted with bold upper-case font, with $j$th entry of the $i$th row $A_{ij}$. Given a stochastic event $E$, we denote its probability by $\mathbb{P}[E]$; given a random variable $x$, we denote its expectation by $\mathbb{E}[x]$.

## II. Model

We consider a population of $n \in \mathbb{N}_+$ individuals, denoted by the set $\mathcal{V} = \{1, \dots, n\}$. Individuals are connected through a time-invariant network $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ that captures social influence between the individuals. In particular, the directed edge $(i, j) \in \mathcal{E}$ if and only if $i$ can be influenced by the opinion of $j$. For any individual $i \in \mathcal{V}$, we denote by $\mathcal{N}_i := \{j : (i, j) \in \mathcal{E}\}$ the set of (out)-neighbors of $i$, that is,

L. Zino is with the Department of Electronics and Telecommunications, Politecnico di Torino, Italy (lorenzo.zino@polito.it). F. Giardini is with the Department of Sociology, Faculty of Behavioural and Social Sciences, University of Groningen, The Netherlands (f.Giardini@rug.nl). D. Vilone is with the Institute of Cognitive Sciences and Technologies, National Research Council, Rome, Italy (daniele.vilone@istc.cnr.it). M. Cao is with the Engineering and Technology Institute Groningen, University of Groningen, The Netherlands (m.cao@rug.nl).

the set of individuals who can directly influence the opinion of $i$, and by $d_i := |\mathcal{N}_i|$ the (out)-degree of the individual.

Each individual $i \in \mathcal{V}$ is characterized by an opinion $x_i(t) \in [0,1]$, which represents individual $i$'s *risk perception* on the emergency at discrete time $t \in \mathbb{N}$, with initial opinion $x_i(0) \in [0,1]$. Opinions are gathered into a vector $\boldsymbol{x}(t) = [x_1(t), \dots, x_n(t)]^\top$, which represents the state of the network at time $t$. Moreover, $i \in \mathcal{V}$ is characterized by three parameters: *risk sensitivity* $\rho_i \in \{-1, 0, +1\}$, *trust in institutions* $\tau_i \in [0,1]$, and *trust in peers* $\mu_i \in [0,1]$, with $\tau_i + \mu_i \leq 1$. Note that $1 - \tau_i - \mu_i$ can be interpreted as a measure of the stubbornness of the individual.

Opinions of the individuals evolve over time in accordance with observations from the social-psychology literature on risk interpretation, which provides evidence of the fact that individuals do not directly take the information broadcasted by the institution, but they process it using information from peers and their own risk sensitivity [15]–[19]. Grounded on such literature, we define a two-step update mechanism. First, the individuals gather information from the institutional source and from peers, and process such information, according to a weighted average dynamics, regulated by the parameters representing the individuals' trust in institutions and in peers, respectively. Second, the individuals revise their opinion by using such information gathered, and further processing it, based on their own risk sensitivity. Such a two-step mechanism is illustrated in Fig. 1.

In the following, we define these dynamics and explicitly derive the set of equations that governs the model. For simplicity, we denote the intermediate step of the opinion in the revision from $x_i(t)$ to $x_i(t+1)$ as $z_i(t)$.

### A. Step I: Information gathering

At each time step $t \in \mathbb{N}_+$, each individual $i \in \mathcal{V}$ receives information from the institutions about the nature of the risk. Specifically, the institution broadcasts a (constant) signal $\iota \in [0,1]$, which quantifies the *nature of the risk*. Such a signal should be interpreted as a normalized quantity, so that $\iota = 0$ means no risk and $\iota = 1$ corresponds to maximal risk.

At the same time, individuals share information with their peers, consistent with the observations from the social-psychological on risk management [16], [17]. Specifically, at each time-step $t \in \mathbb{N}_+$, each individual $i \in \mathcal{V}$ interacts with a peer $j$, selected uniformly at random in $\mathcal{N}_i$, independently of the past. The neighbor $j$ decides to share with $i$ their opinion with state-dependent probability equal to $f_j(x_j(t))$, where $f_j : [0,1] \to [0,1]$ is a function termed *sharing probability function* that maps the opinion of individual $j$ to their tendency to communicate it. This function captures the fact, well-known in the social-psychology literature, that people tend to transmit information that is in accordance with their risk perception [19].

To represent the information sharing process, we use a time-varying network $\mathcal{G}_t = (\mathcal{V}, \mathcal{E}_t)$. If at time $t \in \mathbb{N}$ individual $i$ interacts with $j$, and $j$ decides to share their opinion, then we add the link $(i,j)$ to the edge set $\mathcal{E}_t$. We define the adjacency matrix as a $n \times n$ time-varying matrix $\boldsymbol{A}(t)$, with off-diagonal entries $A_{ij}(t) = 1$ if $(i,j) \in \mathcal{E}_t$ and $A_{ij}(t) = 0$ otherwise. The diagonal entries are defined as $A_{ii}(t) = 1 - \sum_{j \in \mathcal{V} \setminus \{i\}} A_{ij}(t)$. Note that, at each time, exactly one entry per each row of $\boldsymbol{A}(t)$ is nonzero: this is the $j$th entry if $i$ receives information from $j$, or diagonal entry if $i$ does not receive information at time $t$.

Then, individual $i$ revises their opinion by averaging the current opinion $x_i(t)$ with the information received from the different sources of information ($\iota$ and, possibly, $x_j(t)$), with the weights given by the trust in institutions $\tau_i$ and in peers $\mu_i$, respectively, obtaining the following convex combination:

$$z_i(t) = (1 - \mu_i - \tau_i) x_i(t) + \mu_i \sum_{j \in \mathcal{V}} A_{ij}(t) x_j(t) + \tau_i \iota, \quad (1)$$

which reduces to $z_i(t) = (1 - \tau_i) x_i(t) + \tau_i \iota$, when no information is received from the network, i.e., if $A_{ii}(t) = 1$.

### B. Step II: Opinion processing through risk sensitivity

After having revised their opinion on the basis of the information gathered from external sources (institutions and peers), individuals further process their opinion through their own risk sensitivity. Specifically, following [21], we assume that each individual $i \in \mathcal{V}$ updates their opinion as

$$x_i(t+1) = \begin{cases} \frac{1}{2}(1 + z_i(t)) & \text{if } \rho_i = +1, \\ z_i(t) & \text{if } \rho_i = 0, \\ \frac{1}{2} z_i(t) & \text{if } \rho_i = -1, \end{cases} \quad (2)$$

which can be conveniently re-written as a linear combination:

$$x_i(t+1) = \big(1 - \frac{1}{2}|\rho_i|\big) z_i(t) + \frac{1}{4}|\rho_i|(1 + \rho_i). \quad (3)$$

We conclude this section by observing that the entire two-step opinion update mechanism can be cast in a compact form as the linear averaging dynamics on a (weighted) time-varying network, which is summarized in the following statement.

**Proposition 1.** *For each and every $i \in \mathcal{V}$, the opinion update mechanism reads*

$$x_i(t+1) = (1 - \lambda_i) \sum\nolimits_{j \in \mathcal{V}} \tilde{A}_{ij}(t) x_j(t) + \lambda_i u_i, \quad (4)$$

*where*

$$\tilde{A}_{ij}(t) = \begin{cases} \dfrac{\mu_i}{1 - \tau_i} A_{ij}(t) & \text{if } j \neq i, \\ 1 - \dfrac{\mu_i}{1 - \tau_i}\big(1 - A_{ii}(t)\big) & \text{if } j = i, \end{cases} \quad (5a)$$

$$\lambda_i = \frac{1}{2}|\rho_i|(1 - \tau_i) + \tau_i, \quad (5b)$$

$$u_i = \frac{\big(1 - \frac{1}{2}|\rho_i|\big)\tau_i \iota + \frac{1}{4}|\rho_i|(1 + \rho_i)}{\frac{1}{2}|\rho_i|(1 - \tau_i) + \tau_i}. \quad (5c)$$

*Proof.* By substituting Eq. (1) into Eq. (3), we obtain

$$\begin{aligned} x_i(t+1) &= \Big(1 - \tfrac{1}{2}|\rho_i|\Big)\Big((1 - \mu_i - \tau_i) x_i(t) \\ &\quad + \mu_i \textstyle\sum_{j \in \mathcal{V}} A_{ij}(t) x_j(t) + \tau_i \iota\Big) + \tfrac{1}{4}|\rho_i|(1 + \rho_i) \\ &= \Big(1 - \tfrac{1}{2}|\rho_i|\Big)\Big(\big(1 - \tau_i - \mu_i(1 - A_{ii}(t))\big) x_i(t) \\ &\quad + \mu_i \textstyle\sum_{j \in \mathcal{V} \setminus \{i\}} A_{ij}(t) x_j(t) + \tau_i \iota\Big) + \tfrac{1}{4}|\rho_i|(1 + \rho_i), \end{aligned} \quad (6)$$

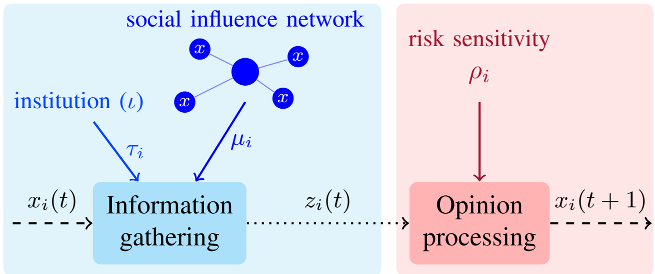
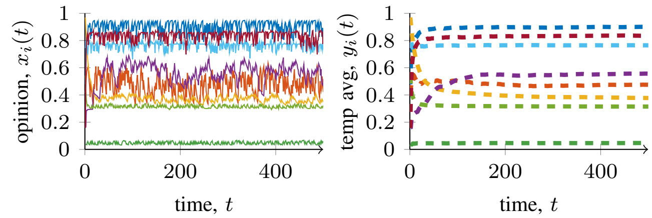



Fig. 1: Schematic of the two-step opinion update mechanism.




(a) Opinion (b) Temporal average opinion

Fig. 2: Temporal evolution of (a) opinions and (b) temporal average opinions for $n = 8$ individuals on a complete backbone network. Parameters $\tau_i$ and $\mu_i$ are selected uniformly at random in $[0, 1/2]$, $\rho_i$ in $\{-1, 0, +1\}$, and initial condition $x_i(0)$ in $[0, 1]$, for each $i \in \mathcal{V}$ independently of the others.

which, after simplification and proper-re-writing of the coefficients, yields Eqs. (4)–(5). $\square$

**Remark 1.** *From Proposition 1, we observe that Eq. (4) can be interpreted as a Friedkin–Johnsen opinion dynamics model on a time-varying network [9]. However, the complexity of the network formation process (which is inherently state-dependent) does not allow to directly apply classical results [11], [23], making the study of the model nontrivial.*

## III. Convergence Results

Here, we prove some general properties of the model to characterize the asymptotic behavior of the model. Specifically, we prove that, while individuals' opinion will tend to keep oscillating, their temporal average converges to a steady state. Before obtaining such a result, we start by observing that the model is always well-defined, that is, that the opinions will always remain within their domain.

**Lemma 1.** *The set $[0, 1]^n$ is positively invariant for the model in Eq. (4), that is, if $\boldsymbol{x}(0) \in [0, 1]^n$, then $\boldsymbol{x}(t) \in [0, 1]^n$, for all $t \in \mathbb{N}$.*

*Proof.* We proceed by induction. At $t = 0$, $x_i(0) \in [0, 1]$ for all $i \in \mathcal{V}$ by assumption. Now, assume that $x_i(t) \in [0, 1]$, for all $i \in \mathcal{V}$. Then, from Eq. (5a), all the entries of $\tilde{\boldsymbol{A}}$ are nonnegative and each row sums to 1. Hence, Eq. (4) states that $x_i(t + 1)$ is a convex combination of the states $x_j(t)$, and $u_i$. Hence $x_i(t+1) \geq \min\{\min_{j\in\mathcal{V}} x_j(t), u_i\} \geq 0$, being $u_i \geq \frac{\tau_i}{\tau_i+1}\iota \geq 0$; and $x_i(t+1) \leq \max\{\max_{j\in\mathcal{V}} x_j(t), u_i\} \leq 1$, being $u_i \leq 1 - \frac{\tau_i}{\tau_i+1}(1 - \iota) \leq 1$. $\square$

In general, the opinion of each node, $x_i(t)$, may not necessarily converge to a steady state value, but it can oscillate, due to the stochastic nature of the process that regulates the opinion exchange mechanism. See, e.g., the simulations in Fig. 2a. However, we can define the temporal average opinion of agent $i \in \mathcal{V}$ as $y_i(t) := \frac{1}{t+1}\sum_{s=0}^{t} x_i(s)$. From Fig. 2b, one can observe that the temporal average opinion vector $\boldsymbol{y}(t) = [y_1(t), \dots, y_n(t)]^\top$ seem to converge. This phenomenon resembles the emergent behavior of gossip consensus dynamics with stubborn agents [24], [25]. However, in our model, oscillations are due to heterogeneity in how individuals process information, rather than due to the presence of stubborn individuals. In the rest of this section, we will prove a convergence result to provide analytical support to such claim. We start by proving ergodicity.

**Proposition 2.** *The process $\boldsymbol{x}(t)$ with update mechanism in Eq. (4) is ergodic.*

*Proof.* The proof follows from the compact formulation in Eq. (4), which satisfies the assumptions in in [26, Th. 1]. $\square$

**Corollary 1.** *Since the process $\boldsymbol{x}(t)$ is ergodic, it holds that if the mean dynamics $\mathbb{E}[\boldsymbol{x}(t)]$ converges to a steady state $\bar{\boldsymbol{x}}$, then the temporal average opinion vector converges to the steady state of the mean dynamics, i.e., $\lim_{t\to\infty} \boldsymbol{y}(t) = \bar{\boldsymbol{x}}$.*

Based on Corollary 1, we study the mean dynamics, i.e., the evolution of $\mathbb{E}[\boldsymbol{x}(t)]$, in order to draw conclusions on the temporal average opinion. We start by explicitly deriving the update rule for the mean opinion dynamics.

**Proposition 3.** *For each and every $i \in \mathcal{V}$, the expected opinion evolves as*

$$\mathbb{E}[x_i(t+1)] = \big(1-\lambda_i\big) \sum\nolimits_{j\in\mathcal{V}} W_{ij}(\boldsymbol{x}(t))x_j(t) + \lambda_i u_i, \quad (7)$$

*with*

$$W_{ij}(\boldsymbol{x}(t)) = \begin{cases} \frac{\mu_i}{d_i(1-\tau_i)} f_j(x_j(t)) & \text{if } j \in \mathcal{N}_i, \\ 1 - \frac{\mu_i}{d_i(1-\tau_i)} \sum_{j\in\mathcal{N}_i} f_j(x_j(t)) & \text{if } j = i, \\ 0 & \text{otherwise}, \end{cases} \quad (8)$$

*and $\lambda_i$ and $u_i$ from Eq. (5b) and Eq. (5c), respectively.*

*Proof.* First, we compute the probability that $i$ receives information from $j \in \mathcal{N}_i$ at time $t$, as

$$\mathbb{P}[A_{ij}(t) = 1] = \mathbb{P}[i \text{ contacts } j]\mathbb{P}[j \text{ shares}] = \frac{f_j(x_j(t))}{d_i}. \quad (9)$$

Using Eq. (9), we compute the probability that $i$ receives information not only from the institution, but also from the network, at time $t$, as

$$\mathbb{P}\left[A_{ii}(t) = 0\right] = \sum_{j\in\mathcal{N}_i} \mathbb{P}[A_{ij}(t) = 1] = \frac{1}{d_i} \sum_{j\in\mathcal{N}_i} f_j(x_j(t)). \quad (10)$$

Hence, using Eq. (1), Eq. (9), and Eq. (10), we compute the expected value of the opinion of individual $i$ after the information exchange step, by conditioning on the values of the $i$th row of matrix $\boldsymbol{A}(t)$, as follows:

$$
\begin{aligned}
&\mathbb{E}[z_i(t)] = \mathbb{P}\left[A_{ii}(t)=1\right]\left[(1-\tau_i)x_i(t)+\tau_i\iota\right] \\
&+\sum_{j\in\mathcal{N}_i}\mathbb{P}\left[A_{ij}(t)=1\right]\left[(1-\mu_i-\tau_i)x_i(t)+\mu_i x_j(t)+\tau_i\iota\right] \\
&=\Big(1-\frac{1}{d_i}\sum\nolimits_{j\in\mathcal{N}_i} f_j(x_j(t))\Big)\left[(1-\tau_i)x_i(t)+\tau_i\iota\right] \\
&=\sum_{j\in\mathcal{N}_i}\frac{1}{d_i}f_j(x_j(t))\left[(1-\mu_i-\tau_i)x_i(t)+\mu_i x_j(t)+\tau_i\iota\right] \\
&=\Big(1-\frac{\mu_i}{d_i}\sum\nolimits_{j\in\mathcal{N}_i} f_j(x_j(t))-\tau_i\Big)x_i(t) \\
&\quad+\frac{\mu_i}{d_i}\sum\nolimits_{j\in\mathcal{N}_i} f_j(x_j(t))x_j(t)+\tau_i\iota.
\end{aligned}
\tag{11}
$$

Finally, we combine Eq. (11) and Eq. (3), obtaining an equation that determines the expected value of the opinion at time $t+1$, as a function of the current opinion of the individual, of their neighbors, and the model parameters:

$$
\begin{aligned}
\mathbb{E}[x_i(t+1)]=&\big(1-\frac{|\rho_i|}{2}\big)\big(1-\frac{\mu_i}{d_i}\sum_{j\in\mathcal{N}_i}f_j(x_j(t))-\tau_i\big)x_i(t) \\
&+\big(1-\frac{1}{2}|\rho_i|\big)\frac{\mu_i}{d_i}\sum\nolimits_{j\in\mathcal{N}_i} f_j(x_j(t))x_j(t) \\
&+\big(1-\frac{1}{2}|\rho_i|\big)\tau_i\iota+\frac{1}{4}|\rho_i|(1+\rho_i),
\end{aligned}
\tag{12}
$$

which can be re-written as Eq. (7), yielding the claim. □

Finally, we are ready to prove that, under some reasonable assumptions on the network of interactions and on the function $f$, the expected opinions and, ultimately, the temporal average opinions converge to a steady state.

**Assumption 1.** *Assume that the network $\mathcal{G}=(\mathcal{V},\mathcal{E})$ is strongly connected, $f_i(x)>0$ for all $x>0$ and $i\in\mathcal{V}$, $\iota>0$, and $\tau_i>0$, for all $i\in\mathcal{V}$.*

**Theorem 1.** *Under Assumption 1, the temporal average opinion vector $y_i(t):=\frac{1}{t+1}\sum_{s=0}^{t}x_i(s)$ under the opinion update in Eq. (4) converges almost surely to a steady state, i.e., $\lim_{t\to\infty}\boldsymbol{y}(t)=\bar{\boldsymbol{x}}\in[0,1]^n$.*

*Proof.* First of all, we observe that if $\tau_i=1$ or $\mu_i=0$, then an individual's opinion is not influenced by others, so $x_i(t)=x_i(0)$ for all $t\geq 0$, yielding the claim for individual $i$. Let now focus on the individuals with $\mu_i\neq 0$ and $\tau_i\neq 1$.

We start proving that, under Assumption 1, the mean dynamics of the ORE model $\mathbb{E}[x_i(t)]$ from Proposition 3 converges almost surely to a steady state, that is, $\lim_{t\to\infty}\mathbb{E}[x_i(t)]=\bar{x}_i\in[0,1]$. To obtain such convergence result, we consider the mean dynamics in Eq. (7), with the expression of $W_{ij}(x(t))$ reported in Eq. (8). First of all, we observe that, the update rule in Eq. (4) establishes a lower-bound on $x_i(t)$. In fact, since from Lemma 1, $x_i(t)\geq 0$, then we can further refine the bound by establishing that $x_i(t)\geq\lambda_i u_i\geq\frac{1}{2}\tau_i\iota$. We define the uniform bound $\alpha:=\min_{i\in\mathcal{V}}\frac{\mu_i}{d_i(1-\tau_i)}f_i(\frac{1}{2}\tau_i\iota)$. Under Assumption 1, we observe that $\frac{1}{2}\tau_i\iota>0$, which implies that also $f_i(\frac{1}{2}\tau_i\iota)>0$. Hence, $\alpha>0$. From Eq. (8), we observe that we can derive the following time-invariant bound on the weight for each link: $W_{ij}(\boldsymbol{x}(t))\geq\alpha$, for all $i\in\mathcal{V}$, $j\in\mathcal{N}_i$. Thus, the time-varying graph with weights $W$ is strongly connected, being $\mathcal{G}$ strongly connected. Hence, the mean dynamics in Eq. (7) is a Friedkin-Johnsen model on a strongly connected time-varying network, so $\mathbb{E}[\boldsymbol{x}(t)]$ converges [9], [11].

Finally, the fact that the mean dynamics $\mathbb{E}[\boldsymbol{x}(t)]$ converges almost surely to a steady state $\bar{\boldsymbol{x}}$ (proved in the above), combined with the fact that the process is ergodic (Proposition 2) implies that $\lim_{t\to\infty}\boldsymbol{y}(t)=\lim_{t\to\infty}\mathbb{E}[\boldsymbol{x}(t)]=\bar{\boldsymbol{x}}$ (Corollary 1), which yield the claim. □

## IV. Steady state characterization

In the previous section, we proved that, under some mild assumptions, the temporal average opinion of the individuals converges to a steady-state value. In general, the characterization of such a steady state is nontrivial due to the complexity of Eq. (7), which yields a system of $n$ coupled nonlinear recursive equations — one for each individual, where the inherent nonlinearity comes from the fact that the term $W_{ij}(\boldsymbol{x})$ (which couples the equations) is state-dependent. In this section, we consider a specific implementation of the model, for which we can analytically compute such a quantity, with a specific focus on the role of risk sensitivity. To perform such analysis, we make the following assumptions.

**Assumption 2.** *Let $\mathcal{G}$ be a complete network, that is, $\mathcal{N}_i=\mathcal{V}$, for all $i\in\mathcal{V}$. Moreover, let us assume that the parameters are uniform across the individuals, that is, $\tau_i=\tau\in(0,1]$ and $\mu_i=\mu$, and that the functions $f_i$ are uniform across the individuals and coincide with the identity function, that is, $f_i(x_i)=x_i$. We also assume $\iota>0$.*

In such a setting, we introduce the following notation. Let $\eta_+:=\frac{1}{n}|\{j:\rho_i=+1\}|$, $\eta_-:=\frac{1}{n}|\{j:\rho_i=-1\}|$, and $\eta_0:=\frac{1}{n}|\{j:\rho_i=0\}|$ be the fraction of population with high, low, and neutral risk perception, respectively. It clearly holds $\eta_++\eta_-+\eta_0=1$.

**Theorem 2.** *Under Assumption 2, the asymptotic value of the temporal average opinion of individual $i\in\mathcal{V}$ under the opinion update in Eq. (4) satisfies*

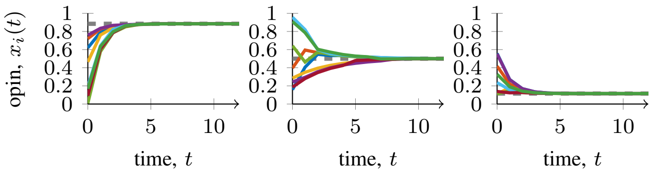
$$
\lim_{t\to\infty}y_i(t)=\begin{cases}\bar{y}_+ & \textit{if } \eta_i=+1\,,\\ \bar{y}_0 & \textit{if } \eta_i=0\,,\\ \bar{y}_- & \textit{if } \eta_i=-1\,,\end{cases}
\tag{13}
$$

*where $(\bar{y}_+,\bar{y}_0,\bar{y}_-)\in[0,1]^3$ is solution of*

$$
\begin{aligned}
\bar{y}_+=&\frac{1}{2}\Big(1-\mu\big(\eta_+\bar{y}_++\eta_0\bar{y}_0+\eta_-\bar{y}_-\big)-\tau\Big)\bar{y}_+ \\
&+\frac{1}{2}\mu\big(\eta_+\bar{y}_+^2+\eta_0\bar{y}_0^2+\eta_-\bar{y}_-^2\big)+\frac{1}{2}\tau\iota+\frac{1}{2},
\end{aligned}
\tag{14a}
$$

$$
\begin{aligned}
\bar{y}_0=&\Big(1-\mu\big(\eta_+\bar{y}_++\eta_0\bar{y}_0+\eta_-\bar{y}_-\big)-\tau\Big)\bar{y}_0 \\
&+\mu\big(\eta_+\bar{y}_+^2+\eta_0\bar{y}_0^2+\eta_-\bar{y}_-^2\big)+\tau\iota,
\end{aligned}
\tag{14b}
$$

$$
\begin{aligned}
\bar{y}_-=&\frac{1}{2}\Big(1-\mu\big(\eta_+\bar{y}_++\eta_0\bar{y}_0+\eta_-\bar{y}_-\big)-\tau\Big)\bar{y}_- \\
&+\frac{1}{2}\mu\big(\eta_+\bar{y}_+^2+\eta_0\bar{y}_0^2+\eta_-\bar{y}_-^2\big)+\frac{1}{2}\tau\iota.
\end{aligned}
\tag{14c}
$$




(a) $\eta_+ = 1$ (b) $\eta_0 = 1$ (c) $\eta_- = 1$

Fig. 3: Numerical simulation of the ORE model with $n = 8$ on a complete backbone network, with $\iota = 0.5$, $\tau_i = \mu_i = 0.3$, for all $i \in \mathcal{V}$, initial condition $x_i(0)$ selected uniformly at random in $[0, 1]$, for each $i \in \mathcal{V}$ independently of the others, and (a) $\rho_i = +1$, (b) $\rho_i = 0$, (c) $\rho_i = -1$, for all $i \in \mathcal{V}$. The gray dashed line is the predicted consensus (Proposition 4).

*Proof.* First, we observe that, according to Theorem 1, the temporal averages of individuals' opinion converge to a steady state $\bar{\boldsymbol{x}}$, which is the steady state of the mean dynamics. Then, we observe that ergodicity of the process guarantees that the steady states of the mean dynamics do not depend on the initial condition. Based on this observation, a symmetry argument can be used to guarantee that $\bar{x}_i = \bar{x}_j$ if $\rho_i = \rho_j$, being all the other parameters equal and the network fully connected, that is, Eq. (13) holds. At this stage, we observe that, at the equilibrium, under Assumption 2, the following two equalities hold true:

$$\frac{1}{d_i}\sum_{j\in\mathcal{N}_i} f(\bar{x}_j) = \frac{1}{n}\sum_{j\in\mathcal{V}} \bar{x}_j = \eta_+\bar{y}_+ + \eta_0\bar{y}_0 + \eta_-\bar{y}_- \tag{15}$$

and

$$\begin{aligned}\frac{1}{d_i}\sum_{j\in\mathcal{V}} f(\bar{x}_j)\bar{x}_j &= \frac{1}{n}\sum_{j\in\mathcal{N}_i} \bar{x}_j^2 = \frac{1}{n}\sum_{j:\rho_i=+1} \bar{y}_+^2 \\ &+\frac{1}{n}\sum_{j:\rho_i=0}\bar{y}_0^2 + \frac{1}{n}\sum_{j:\rho_i=-1}\bar{y}_-^2 = \eta_+\bar{y}_+^2 + \eta_0\bar{y}_0^2 + \eta_-\bar{y}_+^2.\end{aligned} \tag{16}$$

Finally, we write the equilibrium condition for the mean dynamics, starting from Eq. (12), and we substitute Eq. (15) and Eq. (16) into such expression, obtaining Eq. (14). □

Theorem 2 provides a powerful tool to characterize the steady-state temporal average opinion of the network. In general, given the parameter of the model, the solution of the three coupled quadratic equations in Eq. (14) can be easily computed using a numerical solver. On the other hand, determining the analytical solution may be, in general, challenging, due to the complexity of the equations. In the rest of this section, we use Theorem 2 to analytically characterize the steady-state temporal average opinion for some specific scenarios where analytical treatment is possible.

### A. Homogeneous population

First, we consider a homogeneous population where all the individuals have positive, neutral, or negative risk perception, i.e., setting $\eta_+ = 1$, $\eta_0 = 1$, or $\eta_- = 1$, respectively. In these scenarios, we prove almost sure convergence of the opinion of each individual to a consensus, which we characterize, with results confirmed by simulations in Fig. 3.

**Proposition 4.** *If Assumption 2 holds and the entire population has the same risk perception, then the ORE model in Eq. (4) almost surely converges to a consensus, that is,* $\lim_{t\to\infty} x_i(t) = x^*$ *with:*

$$\mathbb{E}[x^*] = \begin{cases} \iota + \dfrac{1-\iota}{1+\tau} & \text{if } \eta_+ = 1,\\ \iota & \text{if } \eta_0 = 1,\\ \iota - \dfrac{\iota}{1+\tau} & \text{if } \eta_- = 1. \end{cases} \tag{17}$$

*Proof.* First, we prove almost sure convergence using Theorem 3.3 from [27]. The proving argument involves the definition of an augmented network with an additional node (which we label as 0) with $\mu_0 = \tau_0 = 0$, and initial opinion

$$x_0(0) = \frac{(1-\frac{1}{2}|r|)\tau\iota + \frac{1}{4}|r|(1+r)}{\frac{1}{2}|r|(1-\tau)+\tau} \tag{18}$$

with $r = 1$ if $\eta_+ = 1$, $r = 0$ if $\eta_0 = 1$, and $r = -1$ if $\eta_- = 1$. Note that, being $\mu_0 = \tau_0 = 0$, then it holds true that $x_0(t) = x_0(0)$, for all $t \geq 0$. The entire model can be reformulated as a De Groot model on a time-varying (state-dependent) network [11], [23] with node 0 as a globally reachable node at every time $t$. Hence, Theorem 3.3 from [27] guarantees almost sure convergence to a consensus, which yields the first part of the claim. Since $x_0(0) = x_0(t)$, for all $t \geq 0$, necessarily the value of the expected consensus coincides with $x_0(t)$. Finally, by substituting $r \in \{+1, 0, -1\}$ into Eq. (18), we get Eq. (17). □

**Remark 2.** *From Proposition 4, we observe that, for uniform populations, the system converges to a consensus, whose expected value can be computed. In the absence of any risk perception biases, the consensus coincides with the actual information sent out by the institution* $x^* = \iota$. *Positive or negative risk perceptions would instead lead to an overestimation or a underestimation of the risk, respectively, as can be observed in Fig. 3.*

**Remark 3.** *Note that the trust in peers (i.e., parameter* $\mu$*) does not play a role in determining the asymptotic consensus state, but it may affect the speed of convergence. As a consequence, one could relax the assumption that such a quantity is homogeneous across the population in Assumption 2.*

### B. Role of heterogeneous risk sensitivity

Here, we investigate the role of individuals with high risk sensitivity in shaping the emergent behavior of the population. We consider a polarized scenario with half of the population with low risk sensitivity and half with high risk sensitivity, proving that individuals with high risk sensitivity would lead to an overestimation of the risk.

**Proposition 5.** *If Assumption 2 holds,* $\iota = 1/2$, $\eta_+ = \eta_- = 1/2$, *and* $\mu = 1 - \tau$ *then the temporal average opinion of each individual in the ORE model in Eq. (4) almost surely converges to a steady state with mean opinion* $< \bar{y}_i > := \frac{1}{n}\sum_{i\in\mathcal{V}} \bar{y}_i \geq 1/2$, *with strict inequality holding if* $\tau < 1$.

*Proof.* In this scenario, the equilibrium equations in Eq. (14) reduce to the following coupled quadratic equations:

$$\bar{y}_+ = \frac{1}{2}(1-\tau)\Big(\bar{y}_+ - \frac{1}{2}\bar{y}_- y_+ + \frac{1}{2}\bar{y}_-^2\Big) + \frac{1}{2}\tau + \frac{1}{2}, \tag{19a}$$

$$\bar{y}_- = \frac{1}{2}(1-\tau)\Big(\bar{y}_- - \frac{1}{2}\bar{y}_+ y_- + \frac{1}{2}\bar{y}_+^2\Big) + \frac{1}{2}\tau, \tag{19b}$$

where Eq. (14b) is omitted, being $\eta_0 = 0$. Let us define $\xi = \frac{\bar{y}_+ + \bar{y}_-}{2}$ and $\zeta = \frac{\bar{y}_+ - \bar{y}_-}{2}$ as the average and half-difference between the two mean opinions. We observe that the steady state with mean opinion $< \bar{y}_i > := \frac{1}{n}\sum_{i\in\mathcal{V}} \bar{y}_i = \xi$. Hence, the problem reduces to prove that $\xi > 1/2$. By computing the sum and the difference between the two equations in Eq. (19) and recalling the definition of $\xi$ and $\zeta$, we derive

$$\xi = \frac{1}{2}(1-\tau)(1-\xi)\xi + \frac{1}{2}(1-\tau)(\xi^2+\zeta^2) + \frac{1}{4}\tau + \frac{1}{4}, \tag{20a}$$

$$\zeta = \frac{1}{2}(1-\tau)(1-\xi)\zeta + \frac{1}{4}. \tag{20b}$$

From Eq. (20b), we explicitly compute $\xi = \frac{1-2\zeta(1+\tau)}{2\zeta(1-\tau)}$. Our objective is to verify that $\xi > 1/2$. A necessary and sufficient condition for having $\xi > 1/2$ is that $\xi > \frac{1}{2} \iff 1 - 2\zeta(1+\tau) > \zeta(1-\tau) \iff \zeta < \frac{1}{3\tau+1}$. To check this condition, we need to compute the solution of Eq. (20) for the variable $\zeta$. To this aim, we substitute the solution of $\xi$ into Eq. (20a), obtaining the third-order equation $\phi(\zeta) = 2(1-\tau)^3\zeta^3 + (\tau^2+4\tau+3)\zeta - (1+\tau) = 0$. It is straightforward to check that $\phi(\zeta)$ is monotonically increasing in $\zeta$ for any $\tau \in [0,1]$ (in fact $\phi'(\zeta) = 6(1-\tau)^3\zeta^2 + \tau^2 + 4\tau + 3 > 0$); that $\phi(0) < 0$ and $\phi(1) > 0$. Therefore, $\phi(\zeta) = 0$ has only one real solution, which lies in $[0,1]$. However, despite this solution can be analytically computed (being the unique real solution of a third-order equation), its complexity hinders the possibility to readily check whether it is less than $\frac{1}{3\tau+1}$. However, we can compute $\phi(\frac{1}{3\tau+1}) = \frac{1}{(3\tau+1)^3}(2(1-\tau)^3 + (\tau^2+4\tau+3)(1+3\tau)^2 - (1+3\tau)^3(1+\tau)) = \frac{4+6\tau+22\tau^2-14\tau^3-16\tau^4}{(3\tau+1)^3}$, which is strictly positive for any $\tau < 1$. Being $\phi(\zeta)$ strictly monotonically increasing, its unique zero must satisfy $\zeta < \frac{1}{3\tau+1}$, implying that $\xi > 1/2$, which yields the claim. □

## V. Conclusion

We proposed a model for collective risk perception grounded on the mathematical theory of opinion dynamics [9], [11], [23] and social-psychology literature [15], [16], [18], [19]. We proved convergence of the temporal average opinions on the risk of a given event, providing a characterization of the steady-state temporal average opinions, which gave us analytical insights into how individuals with high risk perception may lead to a collective overreactions.

Our results pave the way for several avenues of research. First, our theoretical analysis should be extended to investigate the speed of convergence of the temporal average opinions and their transient behavior, and generalize our characterization of the steady-state beyond the limitations of Assumption 2. Second, effort should be placed in extending the model to incorporate further real-world features, such as the presence of media which may bias the information provided by the institution. Third, validation and parametrization using experimental and survey data on risk perception is envisaged of our future research.

## References


[1] N. E. Friedkin, "The Problem of Social Control and Coordination of Complex Systems in Sociology: A Look at the Community Cleavage Problem," *IEEE Control Syst. Mag.*, vol. 35, no. 3, pp. 40–51, 2015.

[2] M. Ye *et al.*, "Collective patterns of social diffusion are shaped by individual inertia and trend-seeking," *Nat. Comm.*, vol. 12, 2021.

[3] L. Zino, M. Ye, and M. Cao, "A two-layer model for coevolving opinion dynamics and collective decision-making in complex social systems," *Chaos*, vol. 30, no. 8, p. 083107, 2020.

[4] ——, "Facilitating innovation diffusion in social networks using dynamic norms," *PNAS Nexus*, vol. 1, 2022.

[5] G. De Pasquale and M. E. Valcher, "A bandwagon bias based model for opinion dynamics: Intertwining between homophily and influence mechanisms," *Eur. J. Control*, vol. 68, p. 100675, 2022.

[6] H. Aghbolagh, M. Ye, L. Zino, Z. Chen, and M. Cao, "Coevolutionary dynamics of actions and opinions in social networks," *IEEE Trans. Autom. Control*, vol. 68, no. 12, pp. 7708–7723, 2023.

[7] A. Franci, M. Golubitsky, I. Stewart, A. Bizyaeva, and N. E. Leonard, "Breaking indecision in multiagent, multioption dynamics," *SIAM J. Appl. Dyn. Syst.*, vol. 22, no. 3, pp. 1780–1817, 2023.

[8] L. Wang *et al.*, "Consensus in concatenated opinion dynamics with stubborn agents," *IEEE Trans. Autom. Control.*, vol. 68, no. 7, pp. 4008–4023, Jul. 2023.

[9] N. E. Friedkin and E. C. Johnsen, "Social Influence and Opinions," *J. Math. Sociol.*, vol. 15, no. 3-4, pp. 193–206, 1990.

[10] G. Deffuant, D. Neau, F. Amblard, and G. Weisbuch, "Mixing beliefs among interacting agents," *Adv. Complex Syst.*, vol. 03, no. 01n04, pp. 87–98, Jan. 2000.

[11] A. V. Proskurnikov and R. Tempo, "A tutorial on modeling and analysis of dynamic social networks. Part I," *Annu. Rev. Control*, vol. 43, pp. 65–79, 2017.

[12] B. D. O. Anderson and M. Ye, "Recent advances in the modelling and analysis of opinion dynamics on influence networks," *Int. J. Autom. Comput.*, vol. 16, no. 2, pp. 129–149, Feb. 2019.

[13] C. Bernardo, C. Altafini, A. Proskurnikov, and F. Vasca, "Bounded confidence opinion dynamics: A survey," *Automatica*, vol. 159, p. 111302, 2024.

[14] M. K. Lindell, "Disaster studies," *Curr. Sociol.*, vol. 61, no. 5-6, pp. 797–825, May 2013.

[15] P. Slovic, "Perception of risk," *Science*, vol. 236, no. 4799, pp. 280–285, 1987.

[16] D. S. Mileti and L. Peek, "The social psychology of public response to warnings of a nuclear power plant accident," *J. Hazard. Mater.*, vol. 75, no. 2-3, pp. 181–194, 2000.

[17] C. W. Scherer and H. Cho, "A social network contagion theory of risk perception," *Risk Anal.*, vol. 23, no. 2, pp. 261–267, 2003.

[18] J. R. Eiser *et al.*, "Risk interpretation and action: A conceptual framework for responses to natural hazards," *Int. J. Disaster Risk Reduct.*, vol. 1, pp. 5–16, Oct. 2012.

[19] N. F. Popovic *et al.*, "Acute stress reduces the social amplification of risk perception," *Sci. Reps.*, vol. 10, no. 1, May 2020.

[20] M. Moussaïd, "Opinion formation and the collective dynamics of risk perception," *PLoS ONE*, vol. 8, no. 12, p. e84592, Dec. 2013.

[21] F. Giardini and D. Vilone, "Opinion dynamics and collective risk perception: An agent-based model of institutional and media communication about disasters," *J. Artif. Soc. Soc. Simul.*, vol. 24, no. 1, p. 4, 2021.

[22] K. C. Roy, S. Hasan, O. I. Abdul-Aziz, and P. Mozumder, "Understanding the influence of multiple information sources on risk perception dynamics and evacuation decisions: An agent-based modeling approach," *Int. J. Disaster Risk Reduct.*, vol. 82, p. 103328, 2022.

[23] A. V. Proskurnikov and R. Tempo, "A tutorial on modeling and analysis of dynamic social networks. Part II," *Annu. Rev. Control*, vol. 45, pp. 166–190, 2018.

[24] P. Frasca, C. Ravazzi, R. Tempo, and H. Ishii, "Gossips and prejudices: Ergodic randomized dynamics in social networks," in *Proc. IFAC NECSYS'13*, vol. 46, no. 27, 2013, pp. 212–219.

[25] D. Acemoğlu, G. Como, F. Fagnani, and A. Ozdaglar, "Opinion fluctuations and disagreement in social networks," *Math. Oper. Res.*, vol. 38, no. 1, pp. 1–27, 2013.

[26] 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, 2015.

[27] F. Fagnani and P. Frasca, *Introduction to Averaging Dynamics over Networks*. Cham, Switzerland: Springer, 2018.