POLITECNICO DI TORINO
Repository ISTITUZIONALE

# Opinion Dynamics on Correlated Subjects in Social Networks

*Original*
Opinion Dynamics on Correlated Subjects in Social Networks / Nordio, Alessandro; Tarable, Alberto; Chiasserini, Carla Fabiana; Leonardi, Emilio. - In: IEEE TRANSACTIONS ON NETWORK SCIENCE AND ENGINEERING. - ISSN 2327-4697. - STAMPA. - 74:3(2020), pp. 1901-1912. [10.1109/TNSE.2019.2956861]

*Availability:*
This version is available at: 11583/2767812 since: 2020-09-03T14:12:30Z

*Publisher:*
IEEE

*Published*
DOI:10.1109/TNSE.2019.2956861

*Terms of use:*

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

*Publisher copyright*
IEEE postprint/Author's Accepted Manuscript

©2020 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collecting works, for resale or lists, or reuse of any copyrighted component of this work in other works.

(Article begins on next page)

12 October 2024

# Opinion Dynamics on Correlated Subjects in Social Networks

Alessandro Nordio, *Member, IEEE,* Alberto Tarable, *Member, IEEE,* Carla Fabiana Chiasserini, *Fellow, IEEE,* Emilio Leonardi, *Senior Member, IEEE*

**Abstract**—Understanding the evolution of collective beliefs is of critical importance to get insights on the political trends as well as on social tastes and opinions. In particular, it is pivotal to develop analytical models that can predict the beliefs dynamics and capture the interdependence of opinions on different subjects. In this paper we tackle this issue also accounting for the individual endogenous process of opinion evolution, as well as repulsive interactions between individuals' opinions that may arise in the presence of an adversarial attitude of the individuals. Using a mean field approach, we characterize the time evolution of opinions of a large population of individuals through a multidimensional Fokker-Planck equation, and we identify the conditions under which stability holds. Finally, we derive the steady-state opinion distribution as a function of the individuals' personality and of the existing social interactions. Our numerical results show interesting dynamics in the collective beliefs of different social communities, and they highlight the effect of correlated subjects as well as of individuals with an adversarial attitude.

**Index Terms**—Social networks, belief dynamics, opinion dynamics, mean-field approach, Fokker-Planck equation, stability.

## 1 Introduction

An increasing deal of attention has recently been devoted to the understanding and the analysis of collective social belief dynamics over social networks [1], [2], [3], [4]. This interest has been stimulated by the growing awareness of the fundamental role that social networks and media may play in the formation and diffusion of opinions/beliefs. For example, it is widely recognized that social media have played a pivotal role in several recent political events, such as the "Arabian Spring" or the last US presidential campaign. Moreover, the availability of a large amount of social data generated by users has attracted the interest of companies and government agencies, which envision opportunities for exploiting such data to get important real-time insights on evolution of trends, tastes, and opinions in the society.

Several experimental approaches, based on sentiment analysis [5], have been proposed for a timely analysis of social dynamics. Furthermore, several analytic frameworks have been developed with the goal of understanding and predicting dominant belief dynamics. These models aim at providing important insights on the dynamics of social interactions, as well as possible explanatory mechanisms for the emergence of strong collective opinions. Additionally, they have also been used to devise possible efficient strategies to influence social beliefs.

The existing models can be coarsely partitioned into two classes:

- Discrete models, in which a discrete variable is associated to every individual corresponding to a node on a graph, and represents the current belief/position of each individual, e.g., favorable, contrary, neutral, with respect to the considered subject. This body of work also includes studies such as [6], [7] where the naming game is used to model phenomena such as opinion dynamics in a population of agents. The social interactions are represented by the graph edges and the state of a node changes for effect of the interactions with its neighbors, i.e., the state of a node is a deterministic/stochastic function of the states of its neighboring nodes. Several different mechanisms, such as the Voter Model [8], Bootstrap-percolation [9], and Linear Threshold Models [10], have been proposed. The dynamics of the process terminates when the system reaches a globally consistent configuration.
- Continuous models, in which the opinion of individuals on a particular subject is described by means of a continuous variable, whose value is adapted as a result of social interactions with individuals having different opinions [11], [12], [13], [14], [15], [16], [17], [18]. Also in this case, social interactions between individuals are typically modeled by using (static or dynamic) graphs, which reflect the structure of the society and describe how individuals interact.

All of the above pieces of work have considered that the beliefs of an individual depend on her social interaction and vary for effect of pairwise "attractive" forces. In particular, a sub-class of continuous models that have attained considerable popularity, considers the so-called bounded confidence, according to which interactions between individuals are effective only if their beliefs are sufficiently close to each other [13], [14], [15], [16], [17], [19], [20], [21], [22]. Furthermore, all previous models represent the evolution of the individuals' opinions about a specific subject, neglecting how beliefs on different, yet correlated, topics may vary over time. This is essentially equivalent to assume the evolution of opinions on different subjects to be independent. Unfortunately, things are much more involved, and opinions on correlated topics exhibit complex inter-dependencies. Consider

- *A. Nordio and A. Tarable are with IEIIT-CNR (Institute of Electronics, Telecommunications and Information Engineering of the National Research Council of Italy), Italy, email: firstname.lastname@ieiit.cnr.it.*
- *C. F. Chiasserini and E. Leonardi are with DET, Politecnico di Torino, Torino, Italy, email: firstname.lastname@polito.it, and are associate researchers with IEIIT-CNR.*

for example the following situation: A group of people discuss about two correlated subjects, e.g., fish (in general) and salmon (in particular) as a part of diet. A person disliking fish also dislikes salmon. If the influence process changes the individual's attitude toward fish, say promoting fish as a healthy part of a diet, then such a person may change her food preferences in favor of salmon as well.

So far, only few pieces of work [23], [24], [25], [26], [27] have tackled the dynamics of opinions on multiple, correlated subjects. In most of such pieces of work, opinions are represented as vector-valued variables, evolving over a multidimensional space in which every axis represents a different subject. In particular, in [23], [24], [25] the evolution of opinions on different axes exhibit a weak dependency, i.e., two users interact only if their Euclidean distance between opinions does not exceed a prefixed threshold. The pieces of work in [26], [27], instead, propose a linear multidimensional model that explicitly accounts for the interdependence of opinions on various topics, and they provide conditions for both stability and convergence [27].

*Our contribution and methodology.* In this paper, we move a step forward with respect to the existing work. First, we generalize the model in [26], [27], introducing a noise component, which represents the individual endogenous process of opinion/belief evolution. Second, we enhance the model by accounting for possible *repulsive* interactions due to adversion between individuals. Using such a model and adopting a mean field approach holding for large population of users, we characterize the evolution of opinions on correlated subjects through a multidimensional Fokker-Planck equation. We derive ergodicity conditions (i.e., conditions for the existence of a unique stationary solution), and, under mild assumptions, we obtain a *closed-form* expression for the stationary distribution of individuals' opinions. We remark that the stability analysis in the presence of an adversarial individuals' attitude is much less obvious than in the traditional models (such as [18]) where only attractive forces were considered. Finally, we provide novel, efficient numerical techniques to analyze both the steady state and the transient solution of the Fokker-Planck (FP) equation, and show interesting effects of opinions dynamics and correlated topics in some relevant scenarios.

### 1.1 Paper organization

The paper is organized as follows. In Sect. 2 we introduce the model and we derive the Fokker-Planck equation based on mean-field approach. In Sect. 3 we develop a methodology for the solution of the FP equation. In Sect. 4 we analyse the stability of the system depending on the different parameters. Sect. 5 presents an analysis of the steady-state regime. in Sect. 6 we summarize the mathematical tools we have used in the paper. Numerical results are reported in Sect 7. Finally we draw some conclusions in Sect. 8.

### 1.2 Notation

Boldface uppercase and lowercase letters denote matrices ad column vectors, respectively. $\mathbf{I}_n$ is the identity matrix of size $n$ and the transpose of the generic matrix $\mathbf{A}$ is denoted by $\mathbf{A}^\mathsf{T}$. The notation $\mathbf{A} = \{a(i,j)\}$ is sometimes used to define a matrix $\mathbf{A}$ whose $(i,j)$-th element is $a_{i,j}$. Similarly, $\mathbf{a} = \mathrm{cat}\{\mathbf{a}(i)\}$ indicates that the column vector $\mathbf{a}$ is obtained by concatenating the column vectors $\mathbf{a}_i$. The Laplace transform of the function $f(x)$ is denoted by $\hat{f}(s)$. Finally, the symbol $\otimes$ denotes the Kronecker product.

## 2 System model

Consider a set of agents $\mathcal{U}$, with cardinality $U$, with agent $i$ exhibiting personality $P_i \in \mathcal{P}$. The agent's personality accounts for her interests and habits, e.g., the social networks to which she has subscribed or the forums in which she participates. We consider that agents have opinions on $N$ different topics and that an opinion formed on one subject is influenced by the opinions on some of the other subjects, i.e., topics are interdependent [27], [28], [29]. We define $\mathbf{C}$ as the coupling matrix, with $c_{mn}$ expressing the entanglement of subject $m$ on subject $n$. The opinions that agent $i \in \mathcal{U}$ has on the different subjects is represented by a vector of size $N$, denoted by $\mathbf{x}_i(t) \in \mathcal{X}^N$, which evolves over continuous time, $t \in \mathbb{R}_+$. We define the prejudice vector $\mathbf{u}(P_i)$ as the a-priori $N$-dimensional belief that agent $i$ has on the different subjects; also the prejudice depends on the agent's personality. We represent through a graph the existence and the intensity of social relationships between users, which depend on the personality of the agents and on the similarity between the agents' beliefs. The actual influence that agents exert on each other then depends on their opportunity to interact, as well as on their sensitivity to others' beliefs.

As a result, the evolution of agent $i$'s belief over time can be represented as:

$$\mathbf{x}_i(t+\mathrm{d}t) = \mathbf{x}_i(t) + \mathbf{C}\mathbf{d}_{\mathbf{x},i}(t) \tag{1}$$

where $\mathbf{x}_i(t)$ denotes the belief of agent $i$ on the $N$ topics at the current time instant, $\mathbf{d}_{\mathbf{x},i}(t)$ accounts for the variation of agent $i$ opinions in the time interval $[t, t+\mathrm{d}t]$, and $\mathbf{C}$ accounts for the influence of the opinion on one topic on the opinions on other topics. The quantity $\mathbf{d}_{\mathbf{x},i}(t)$ is given by

$$\begin{aligned}\mathbf{d}_{\mathbf{x},i} \;=\; & \frac{1-\alpha(P_i)}{U-1}\sum_{\substack{j\in\mathcal{U}\\ j\neq i}} \underbrace{\zeta\left(P_i,P_j\right)\left[\mathbf{x}_j(t)-\mathbf{x}_i(t)\right]\,\mathrm{d}t}_{(a)} \\ & + \underbrace{\alpha(P_i)\left[\mathbf{u}(P_i)-\mathbf{x}_i(t)\right]\,\mathrm{d}t}_{(b)} + \underbrace{\sigma\,\mathrm{d}\mathbf{w}_i(t)}_{(c)}\,.\end{aligned} \tag{2}$$

The meaning of the terms on the right hand side of the above expression is as follows.

- The first term $(a)$ represents the interaction of agent $i$ with all other agents in $\mathcal{U}$. In particular,
  - $\alpha(P_i) \in (0,1]$ indicates how insensitive $i$ is to other agents' beliefs, which, as also discussed in [30], plays an important role in opinion dynamics. This parameter will also be referred to as the agent's level of stubbornness. When $\alpha(P_i) \to 1$, the agent becomes completely insensitive to others' beliefs (stubborn). Instead, as $\alpha(P_i)$ decreases, the agent is more inclined to accept others' beliefs and is less conditioned by her own prejudice. For brevity, in the following we denote $\bar{\alpha}(P_i) = 1-\alpha(P_i)$;
  - $\zeta(P_i,P_j)$ represents the presence and the strength of interactions between agents $i$ and $j$ (hereinafter also referred to as mutual influence). It is a function of both agents' personality and defines the structure of the social graph [30]. Note that the interactions between agents do not depend on the proximity of their opinions, i.e., they are independent of $\mathbf{x}_j(t)-\mathbf{x}_i(t)$. Also, whenever $\zeta(P_i,P_j) = 0$, the

two agents do not influence each other, i.e., they never interact. Finally, it is fair to assume that each element of $\zeta(P_i, P_j)$ is upper bounded by a constant and is continuous with respect to its first and second arguments.

- The second term $(b)$ represents the tendency of an agent to retain her prejudice.
- The third term $(c)$ accounts for the endogenous process of the belief evolution within each agent. Such process is modeled as an i.i.d. standard Brownian motion with zero drift and scale parameter $\sigma^2$ [4].

We remark that $\mathbf{x}_i(t+\mathrm{d}t)$, i.e., the belief of agent $i$ at time $t+\mathrm{d}t$, depends on her personality $P_i$ and the current agent's belief. In other words, the temporal evolution of agents' beliefs $\{\mathbf{x}_i(t),\ i \in \mathcal{U}\}$ is Markovian over $\mathcal{Y}^U$, where $\mathcal{Y} = \mathcal{P} \times \mathcal{X}^N$ is an $(N+1)$-dimensional continuous space. Furthermore, in the following we assume that $\alpha(P_i)$, $\zeta(P_i, P_j)$, and each element of $\mathbf{u}(P_i)$ are continuous with respect to their arguments. For ease of notation, we denote $\mathbf{Z}(P_1, P_2) = \zeta(P_i, P_j)\mathbf{C}$; thus, by replacing (2) in (1), the latter can be rewritten as

$$\begin{aligned}\mathbf{x}_i(t+\mathrm{d}t) = \mathbf{x}_i(t) + \frac{1-\alpha(P_i)}{U-1}\sum_{j\in\mathcal{U}/\{i\}} \mathbf{Z}\left(P_i,P_j\right)\left[\mathbf{x}_j(t)-\mathbf{x}_i(t)\right]\,\mathrm{d}t \\ +\alpha(P_i)\mathbf{C}\left[\mathbf{u}(P_i)-\mathbf{x}_i(t)\right]\,\mathrm{d}t + \sigma\mathbf{C}\,\mathrm{d}\mathbf{w}_i(t)\,. \quad (3)\end{aligned}$$

### 2.1 From the discrete to the continuous model

We now extend the above model to the continuous case by using the mean-field theory. We leverage on the procedure presented in [18] and apply it to the multi-subject scenario. More in detail, we define the empirical probability measure, $\rho^{(U)}(\,\mathrm{d}p,\,\mathrm{d}\mathbf{x}, t)$ over $\mathcal{Y}$ at time $t$, as:

$$\rho^{(U)}(\,\mathrm{d}p,\,\mathrm{d}\mathbf{x}, t) = \frac{1}{U}\sum_{i\in\mathcal{U}} \delta_{(P_i,\mathbf{x}_i(t))}(\,\mathrm{d}p,\,\mathrm{d}\mathbf{x})\,. \quad (4)$$

In the above expression, $\delta_{(P_i,\mathbf{x}_i(t))}(\,\mathrm{d}p,\,\mathrm{d}\mathbf{x})$ is the Dirac measure centered at $(P_i, \mathbf{x}_i(t))$, i.e., $\delta_{(P_i,\mathbf{x}_i(t))}(\,\mathrm{d}p,\,\mathrm{d}\mathbf{x})$ represents the mass probability associated with opinion $\mathbf{x}_i(t)$ of agent $i$, which has personality $P_i$. Note that in (4) agents are seen as particles in the continuous space $\mathcal{Y}$, moving along the opinion axis $\mathbf{x}$. As shown in [18], by applying the mean-field theory [31], [32], as $U \to \infty$, $\rho^{(U)}(\,\mathrm{d}p,\,\mathrm{d}\mathbf{x}, t)$ converges in law to the asymptotic distribution $\rho(p, \mathbf{x}, t)$, provided that $\rho^{(U)}(\,\mathrm{d}p,\,\mathrm{d}\mathbf{x}, 0)$ converges in law to $\rho(p, \mathbf{x}, 0)$. Moreover, $\rho(p, \mathbf{x}, t)$ can be obtained from the following non-linear Fokker-Planck (FP) equation [31], [32]:

$$\begin{aligned}\frac{\partial}{\partial t}\rho(p,\mathbf{x},t) = -\sum_{n=1}^{N}\frac{\partial}{\partial x_n}\left[\mu_n(p,\mathbf{x},t)\rho(p,\mathbf{x},t)\right] \\ +\frac{1}{2}\sum_{m,n=1}^{N} D_{mn}\frac{\partial^2}{\partial x_m \partial x_n}\rho(p,\mathbf{x},t) \quad (5)\end{aligned}$$

where $D_{mn}$ is the $(m,n)$-th entry of the diffusion tensor $\mathbf{D} = \sigma^2\mathbf{C}\mathbf{C}^{\mathsf{T}}$, and $\mu_n(p, \mathbf{x}, t)$ is defined as the component of the instantaneous speed along axis $x_n$ of a generic agent whose personality and opinion at time $t$ are equal to $p$ and $\mathbf{x}$. The instantaneous speed is given by:

$$\begin{aligned}\boldsymbol{\mu}(p,\mathbf{x},t) &= \bar{\alpha}(p)\int_{\mathbf{y}\in\mathcal{X}^N}\int_{\mathcal{P}} \mathbf{Z}(p,q)(\mathbf{y}-\mathbf{x})\rho(q,\mathbf{y},t)\,\mathrm{d}^N\mathbf{y}\,\mathrm{d}q \\ &\quad + \alpha(p)\mathbf{C}[\mathbf{u}(p)-\mathbf{x}] \\ &= \bar{\alpha}(p)\left[\boldsymbol{\gamma}(p,t) - \boldsymbol{\Gamma}(p)\mathbf{x}\right] + \alpha(p)\mathbf{C}[\mathbf{u}(p)-\mathbf{x}] \\ &= -\boldsymbol{\Xi}(p)\mathbf{x} + \boldsymbol{\phi}(p,t) \quad (6)\end{aligned}$$

where we defined

$$\boldsymbol{\Gamma}(p) \triangleq \iint_{\mathbf{y},q} \mathbf{Z}(p,q)\rho(q,\mathbf{y},t)\,\mathrm{d}^N\mathbf{y}\,\mathrm{d}q \overset{(a)}{=} \int_q \mathbf{Z}(p,q)\rho_0(q)\,\mathrm{d}q, \quad (7)$$

where in $(a)$ we wrote $\rho(q,\mathbf{y},t) = \rho(\mathbf{y},t|q)\rho_t(q)$ and exploited the fact that by definition $\int_{\mathbf{y}} \rho(\mathbf{y},t|q)\,\mathrm{d}^N\mathbf{y} = 1$. Since the distribution of the agents' personality at time $t$, $\rho_t(q)$, does not depend on $t$, we have: $\rho_t(q) = \rho_0(q)$. Furthermore,

$$\boldsymbol{\gamma}(p,t) \triangleq \iint_{\mathbf{y},q} \mathbf{Z}(p,q)\mathbf{y}\rho(q,\mathbf{y},t)\,\mathrm{d}^N\mathbf{y}\,\mathrm{d}q, \quad (8)$$

$$\boldsymbol{\Xi}(p) \triangleq \bar{\alpha}(p)\boldsymbol{\Gamma}(p) + \alpha(p)\mathbf{C}, \quad (9)$$

$$\boldsymbol{\phi}(p,t) \triangleq \bar{\alpha}(p)\boldsymbol{\gamma}(p,t) + \alpha(p)\mathbf{C}\mathbf{u}(p) \quad (10)$$

and we considered a zero-drift Brownian motion process $\mathbf{w}(t)$. In the following, we analyze the system dynamics by solving the above FP equation for $\rho(p, \mathbf{x}, t)$ so as to obtain the distribution of agents over $\mathcal{Y}$.

## 3 Solution of the Fokker-Planck (FP) equation

In this section, we solve the $N$-dimensional FP equation, which describes an $N$-dimensional Ornstein-Uhlenbeck (OU) [33] random process. Considering a general initial density $\rho(p, \mathbf{x}, 0) = \rho_0(\mathbf{x}|p)\rho_0(p)$, we obtain the solution of the FP equation $\rho(p, \mathbf{x}, t)$ as shown in Appendix A in the Supplemental Material:

$$\rho(p,\mathbf{x},t) = \rho_0(p)\int_{\mathbf{y}} \mathcal{G}\left(\mathbf{x},\mathbf{m}(p,\mathbf{y},t),\boldsymbol{\Sigma}(p,t)\right)\rho_0(\mathbf{y}|p)\,\mathrm{d}^N\mathbf{y} \quad (11)$$

where $\mathcal{G}\left(\mathbf{x}, \mathbf{m}(p,\mathbf{y},t), \boldsymbol{\Sigma}(p,t)\right)$ is the pdf of the Gaussian multivariate distribution with covariance

$$\boldsymbol{\Sigma}(p,t) \triangleq \int_0^t \mathrm{e}^{-\boldsymbol{\Xi}(p)\tau}\mathbf{D}\mathrm{e}^{-\boldsymbol{\Xi}(p)^{\mathsf{T}}\tau}\,\mathrm{d}\tau \quad (12)$$

and mean

$$\begin{aligned}\mathbf{m}(p,\mathbf{y},t) &\triangleq \mathrm{e}^{-\boldsymbol{\Xi}(p)t}\left[\mathbf{y} + \int_0^t \mathrm{e}^{\boldsymbol{\Xi}(p)\tau}\boldsymbol{\phi}(p,\tau)\,\mathrm{d}\tau\right] \\ &\overset{(a)}{=} \mathrm{e}^{-\boldsymbol{\Xi}(p)t}\mathbf{y} + \alpha(p)\boldsymbol{\Xi}^{-1}(p)\left(\mathbf{I}_N - \mathrm{e}^{-\boldsymbol{\Xi}(p)t}\right)\mathbf{C}\mathbf{u}(p) \\ &\quad + \bar{\alpha}(p)\int_0^t \mathrm{e}^{-\boldsymbol{\Xi}(p)(t-\tau)}\boldsymbol{\gamma}(p,\tau)\,\mathrm{d}\tau \quad (13)\end{aligned}$$

where in (a) we used the definition of $\boldsymbol{\phi}(p,\tau)$ provided in (10). However, notice that $\mathbf{m}(p,\mathbf{y},t)$ in (13) is a function of $\boldsymbol{\gamma}(p,t)$, which in turn is a function of $\rho(p,\mathbf{x},t)$ (see (8)). As such, we have to impose a self-consistency condition; precisely, replacing (11)

in (8), we obtain

$$\begin{aligned}\boldsymbol{\gamma}(p,t) &= \int_{\mathbf{y}}\int_{q} \mathbf{Z}(p,q)\mathbf{y}\rho(q,\mathbf{y},t)\,\mathrm{d}^N\mathbf{y}\,\mathrm{d}q \\ &= \int_{\mathbf{y}}\int_{q} \mathbf{Z}(p,q)\mathbf{y}\rho_0(q)\int_{\mathbf{z}}\mathcal{G}\left(\mathbf{y},\mathbf{m}(q,\mathbf{z},t),\boldsymbol{\Sigma}(q,t)\right) \\ &\qquad \cdot\rho_0(\mathbf{z}|q)\,\mathrm{d}^N\mathbf{z}\,\mathrm{d}^N\mathbf{y}\,\mathrm{d}q \\ &= \int_{q}\mathbf{Z}(p,q)\rho_0(q)\int_{\mathbf{z}}\mathbf{m}(q,\mathbf{z},t)\rho_0(\mathbf{z}|q)\,\mathrm{d}^N\mathbf{z}\,\mathrm{d}q \\ &= \boldsymbol{\gamma}_0(p,t)+\boldsymbol{\gamma}_1(p,t)+\int_{q}\mathbf{Z}(p,q)\rho_0(q)\bar{\alpha}(q) \\ &\qquad \cdot\int_0^t \mathrm{e}^{-\boldsymbol{\Xi}(q)(t-\tau)}\boldsymbol{\gamma}(q,\tau)\,\mathrm{d}\tau\,\mathrm{d}q \end{aligned} \tag{14}$$

where we used (13) and defined for brevity:

$$\boldsymbol{\gamma}_0(p,t) \triangleq \int_{q}\mathbf{Z}(p,q)\rho_0(q)\mathrm{e}^{-\boldsymbol{\Xi}(q)t}\int_{\mathbf{y}}\mathbf{y}\rho_0(\mathbf{y}|q)\,\mathrm{d}^N\mathbf{y}\,\mathrm{d}q \tag{15}$$

$$\begin{aligned}\boldsymbol{\gamma}_1(p,t) &\triangleq \int_{q}\mathbf{Z}(p,q)\rho_0(q)\alpha(q)\boldsymbol{\Xi}^{-1}(q) \\ &\qquad \cdot\left(\mathbf{I}_N-\mathrm{e}^{-\boldsymbol{\Xi}(q)t}\right)\mathbf{C}\mathbf{u}(q)\,\mathrm{d}q \end{aligned} \tag{16}$$

Interestingly, (14) is a linear Volterra equation of the second kind [34]. We can take over time the Laplace transform of (14) and get

$$\begin{aligned}\widehat{\boldsymbol{\gamma}}(p,s) &= \widehat{\boldsymbol{\gamma}}_0(p,s)+\widehat{\boldsymbol{\gamma}}_1(p,s) \\ &\quad +\int_{q}\mathbf{Z}(p,q)\rho_0(q)\bar{\alpha}(q)\mathbf{X}^{-1}(s,q)\widehat{\boldsymbol{\gamma}}(q,s)\,\mathrm{d}q \end{aligned} \tag{17}$$

where $\mathbf{X}(s,q)=s\mathbf{I}_N+\boldsymbol{\Xi}(q)$

$$\widehat{\boldsymbol{\gamma}}_0(p,s) = \int_{q}\mathbf{Z}(p,q)\rho_0(q)\mathbf{X}^{-1}(s,q)\int_{\mathbf{y}}\mathbf{y}\rho_0(\mathbf{y}|q)\,\mathrm{d}^N\mathbf{y}\,\mathrm{d}q \tag{18}$$

$$\widehat{\boldsymbol{\gamma}}_1(p,s) = \int_{q}\mathbf{Z}(p,q)\rho_0(q)\alpha(q)s^{-1}\mathbf{X}^{-1}(s,q)\mathbf{C}\mathbf{u}(q)\,\mathrm{d}q \tag{19}$$

Eq. (17) is a non-homogeneous integral equation, whose solution is unique if and only if the associated homogeneous equation has no nonzero solutions. If the solution is unique, then it gives the solution of the FP through (11).

**Remark**: By restricting the definition of $\boldsymbol{\gamma}(p,t)$ over an arbitrary compact domain $\mathcal{P}\times[0,T]$, (14) can be rewritten in an operational form as $(\mathcal{I}-\mathcal{A})[\boldsymbol{\gamma}(p,t)] = \boldsymbol{\gamma}_0(p,t)+\boldsymbol{\gamma}_1(p,t)$ where the operator $\mathcal{A}[\boldsymbol{\gamma}(p,t)] = \int_q \mathbf{Z}(p,q)\rho_0(q)\bar{\alpha}(q)\int_0^t \mathrm{e}^{-\boldsymbol{\Xi}(q)(t-\tau)}\boldsymbol{\gamma}(q,\tau)\,\mathrm{d}\tau\,\mathrm{d}q$. Note that, as an immediate consequence of the structure of Volterra equations over compact domains, we have $\|\mathcal{A}^n\|<1$ for sufficiently large $n$ [35, cap. 2. pp. 50–51]. Thus, $(\mathcal{I}-\mathcal{A})[\cdot]$ is invertible (i.e., the solution is unique) and an expression for $\boldsymbol{\gamma}(p,t)$ can be obtained as $\boldsymbol{\gamma}(p,t) = (\mathcal{I}-\mathcal{A})^{-1}[(\boldsymbol{\gamma}_0(p,t)+\boldsymbol{\gamma}_1(p,t))] = \sum_n \mathcal{A}^n[(\boldsymbol{\gamma}_0(p,t)+\boldsymbol{\gamma}_1(p,t))]$.

Provided that the integral equation (17) admits a unique solution, in general it is still rather challenging to explicitly find it. In the following, we will particularize our analysis to two cases in which it is possible to find an explicit analytical expression for such as solution.

### 3.1 The case of $\zeta(p,q)$ in product form

We recall that $\mathbf{Z}(p,q)=\zeta(p,q)\mathbf{C}$. If $\zeta(p,q)=\zeta_1(p)\zeta_2(q)$, then from (8) we have

$$\begin{aligned}\boldsymbol{\gamma}(p,t) &= \zeta_1(p)\int_{\mathbf{y}}\int_{q}\zeta_2(q)\mathbf{C}\mathbf{y}\rho(q,\mathbf{y},t)\,\mathrm{d}^N\mathbf{y}\,\mathrm{d}q\,. \\ &= \zeta_1(p)\mathbf{b}(t) \end{aligned} \tag{20}$$

Taking the Laplace transform of (20), we obtain $\widehat{\boldsymbol{\gamma}}(p,s) = \zeta_1(p)\widehat{\mathbf{b}}(s)$. Using the latter expression in (17) and recalling (18) and (19), we get

$$\begin{aligned}\widehat{\boldsymbol{\gamma}}(p,s) &= \widehat{\boldsymbol{\gamma}}_0(p,s)+\widehat{\boldsymbol{\gamma}}_1(p,s) \\ &\quad +\zeta_1(p)\int_{q}\zeta_2(q)\mathbf{C}\rho_0(q)\bar{\alpha}(q)\mathbf{X}^{-1}(s,q)\widehat{\boldsymbol{\gamma}}(q,s)\,\mathrm{d}q \\ &= \zeta_1(p)\left(\widehat{\mathbf{b}}_0(s)+\widehat{\mathbf{b}}_1(s)\right) \\ &\quad +\zeta_1(p)\int_{q}\rho_0(q)\bar{\alpha}(q)\zeta_2(q)\mathbf{C}\mathbf{X}^{-1}(s,q)\zeta_1(q)\,\mathrm{d}q\widehat{\mathbf{b}}(s) \end{aligned} \tag{21}$$

under the assumption that $\mathbf{X}(s,q)=s\mathbf{I}_N+\boldsymbol{\Xi}(q)$ is invertible[1]. In (21)

$$\widehat{\mathbf{b}}_0(s) = \int_{q}\zeta_2(q)\mathbf{C}\rho_0(q)\mathbf{X}^{-1}(s,q)\int_{\mathbf{y}}\mathbf{y}\rho_0(\mathbf{y}|q)\,\mathrm{d}^N\mathbf{y}\,\mathrm{d}q \tag{22}$$

$$\widehat{\mathbf{b}}_1(s) = \int_{q}\zeta_2(q)\mathbf{C}\rho_0(q)\alpha(q)s^{-1}\mathbf{X}^{-1}(s,q)\mathbf{C}\mathbf{u}(q)\,\mathrm{d}q\,. \tag{23}$$

Note that we can also write

$$\begin{aligned}\widehat{\mathbf{b}}(s) &= \left(\int_{q}\rho_0(q)\bar{\alpha}(q)\zeta_2(q)\mathbf{C}\mathbf{X}^{-1}(s,q)\zeta_1(q)\,\mathrm{d}q\right)\widehat{\mathbf{b}}(s) \\ &\quad +\widehat{\mathbf{b}}_0(s)+\widehat{\mathbf{b}}_1(s) \end{aligned} \tag{24}$$

from which, under the assumption that matrix

$$\mathbf{T}(s) = \mathbf{I}_N - \int_{q}\rho_0(q)\bar{\alpha}(q)\zeta_2(q)\mathbf{C}\mathbf{X}^{-1}(s,q)\zeta_1(q)\,\mathrm{d}q$$

is non singular, we obtain:

$$\widehat{\mathbf{b}}(s) = \mathbf{T}(s)^{-1}\left(\widehat{\mathbf{b}}_0(s)+\widehat{\mathbf{b}}_1(s)\right)\,.$$

We now have an explicit solution (in the transformed domain) for $\mathbf{b}(t)$. Indeed, for $\mathrm{Re}(s)$ sufficiently large, matrix $\mathbf{T}(s)$ is non singular. Once $\mathbf{b}(t)$ is obtained, we can compute $\boldsymbol{\gamma}(p,t)$ through (20), then $\mathbf{m}(p,\mathbf{y},t)$ in (13), and, finally, our opinion density $\rho(p,\mathbf{x},t)$ through (11).

### 3.2 Discrete personality distribution

Suppose that the personality distribution is discrete with $M$ probability masses; then we write the opinion distribution at $t=0$ as

$$\rho_0(p) = \sum_{i=1}^{M} r_i\delta(p-p_i). \tag{25}$$

Consequently, the fixed-point equation (17) becomes

$$\begin{aligned}\widehat{\boldsymbol{\gamma}}(p_i,s) &= \sum_{k=1}^{M}\mathbf{Z}(p_i,p_k)r_k\bar{\alpha}(p_k)\mathbf{X}^{-1}(s,p_k)\widehat{\boldsymbol{\gamma}}(p_k,s) \\ &\quad +\widehat{\boldsymbol{\gamma}}_0(p_i,s)+\widehat{\boldsymbol{\gamma}}_1(p_i,s)\,. \end{aligned} \tag{26}$$

1. Note that invertibility is granted for $\mathrm{Re}(s)>\sup_q\|\boldsymbol{\Xi}(q)\|$

In (26), we defined

$$\widehat{\gamma}_0(p_i,s)=\sum_{k=1}^{M}\mathbf{Z}(p_i,p_k)r_k\mathbf{X}^{-1}(s,p_k)\mathbf{x}_0(p_k) \tag{27}$$

where $\mathbf{x}_{0,k}$ is the average opinion at time $t=0$ corresponding to personality $p_k$ and has been obtained by observing that the term $\rho_0(q)\int_{\mathbf{y}}\mathbf{y}\rho_0(\mathbf{y}|q)\,\mathrm{d}^N\mathbf{y}$ in (18) is equal to $\mathbf{m}(q,\mathbf{x},0)=\mathbf{x}_0(q)$. Also, we defined:

$$\widehat{\gamma}_1(p_i,s)=\sum_{k=1}^{M}\frac{\mathbf{Z}(p_i,p_k)r_k\alpha(p_k)}{s}\mathbf{X}^{-1}(s,p_k)\mathbf{C}\mathbf{u}(p_k)\,. \tag{28}$$

Next, recalling the notation defined in Section 1.2 we define $\underline{\widehat{\gamma}}(s)=\mathrm{cat}\{\widehat{\gamma}(p_i,s)\}$, $\underline{\widehat{\gamma}}_0(s)=\mathrm{cat}\{\widehat{\gamma}_0(p_i,s)\}$, $\underline{\widehat{\gamma}}_1(s)=\mathrm{cat}\{\widehat{\gamma}_1(p_i,s)\}$, $\underline{\mathbf{x}}_0=\mathrm{cat}\{\mathbf{x}_0(q_i)\}$, $\underline{\mathbf{u}}=\mathrm{cat}\{\mathbf{u}(q_i)\}$, $\underline{\Xi}=\mathrm{diag}\,(\Xi(p_1),\ldots,\Xi(p_M))$, $\mathbf{P}_0=\mathrm{diag}\,(r_1,\ldots,r_M)\otimes\mathbf{I}_N$, $\mathbf{P}_1=\mathrm{diag}\,(r_1\alpha(p_1),\ldots,r_M\alpha(p_M))\otimes\mathbf{I}_N$, $\mathbf{P}_2=\mathrm{diag}\,(r_1\bar{\alpha}(p_1),\ldots,r_M\bar{\alpha}(p_M))\otimes\mathbf{I}_N$, and

$$\underline{\mathbf{Z}}=\{\mathbf{Z}(p_i,p_j)\}=\{\zeta(p_i,p_j)\}\otimes\mathbf{C}\,. \tag{29}$$

Then, we can write (26) as

$$\underline{\widehat{\gamma}}(s)=\underline{\widehat{\gamma}}_0(s)+\underline{\widehat{\gamma}}_1(s)+\underline{\mathbf{Z}}\left(s\mathbf{I}_{MN}+\underline{\Xi}\right)^{-1}\mathbf{P}_2\underline{\widehat{\gamma}}(s) \tag{30}$$

while (27) and (28) become

$$\underline{\widehat{\gamma}}_0(s)=\underline{\mathbf{Z}}\left(s\mathbf{I}_{MN}+\underline{\Xi}\right)^{-1}\mathbf{P}_0\underline{\mathbf{x}}_0 \tag{31}$$

and

$$\underline{\widehat{\gamma}}_1(s)=\underline{\mathbf{Z}}\left(s\mathbf{I}_{MN}+\underline{\Xi}\right)^{-1}s^{-1}\mathbf{P}_1\left(\mathbf{I}_M\otimes\mathbf{C}\right)\underline{\mathbf{u}}. \tag{32}$$

Let $\underline{\mathbf{X}}(s)=s\mathbf{I}_{MN}+\underline{\Xi}$. Then we can solve explicitly the fixed-point equation of (30) (in the transform domain) as

$$\begin{aligned}\underline{\widehat{\gamma}}(s) &= \left(\mathbf{I}_{MN}-\underline{\mathbf{Z}}\underline{\mathbf{X}}(s)^{-1}\mathbf{P}_2\right)^{-1}\left(\underline{\widehat{\gamma}}_0(s)+\underline{\widehat{\gamma}}_1(s)\right)\\ &= \left(\mathbf{I}_{MN}-\underline{\mathbf{Z}}\underline{\mathbf{X}}(s)^{-1}\mathbf{P}_2\right)^{-1}\underline{\mathbf{Z}}\underline{\mathbf{X}}(s)^{-1}\\ &\qquad\cdot\left(\mathbf{P}_0\underline{\mathbf{x}}_0+s^{-1}\mathbf{P}_1\left(\mathbf{I}_M\otimes\mathbf{C}\right)\underline{\mathbf{u}}\right)\\ &= \left(\underline{\mathbf{X}}(s)\underline{\mathbf{Z}}^{-1}-\mathbf{P}_2\right)^{-1}\left(\mathbf{P}_0\underline{\mathbf{x}}_0+s^{-1}\mathbf{P}_1\left(\mathbf{I}_M\otimes\mathbf{C}\right)\underline{\mathbf{u}}\right)\\ &= \underline{\mathbf{Z}}\left(\underline{\mathbf{X}}(s)-\mathbf{P}_2\underline{\mathbf{Z}}\right)^{-1}\left(\mathbf{P}_0\underline{\mathbf{x}}_0+s^{-1}\mathbf{P}_1\left(\mathbf{I}_M\otimes\mathbf{C}\right)\underline{\mathbf{u}}\right)\,.\end{aligned} \tag{33}$$

Now, let

$$\mathbf{\Psi}=\underline{\Xi}-\mathbf{P}_2\underline{\mathbf{Z}}\,. \tag{34}$$

Provided that $\mathbf{\Psi}$ is invertible, the above solution can be inverse-transformed to get the solution in the time domain as

$$\underline{\gamma}(t)=\underline{\mathbf{Z}}\left(\mathrm{e}^{-\mathbf{\Psi}t}\mathbf{P}_0\underline{\mathbf{x}}_0+\mathbf{\Psi}^{-1}\left(\mathbf{I}_{MN}-\mathrm{e}^{-\mathbf{\Psi}t}\right)\mathbf{P}_1\left(\mathbf{I}_M\otimes\mathbf{C}\right)\underline{\mathbf{u}}\right). \tag{35}$$

Then $\rho(p,\mathbf{x},t)$ can be obtained as described in the previous section.

Finally, the following theorem complements the above result:

***Theorem 1.*** i) For any finite $t$ the Volterra equation (14) admits a unique solution which is Lipschitz-continuous. ii) The solution $\gamma(p,t)$ of the Volterra equation (14) , under any distribution $\rho_0(p)$, which is continuous at every point in $p$, is the uniform limit of solutions $\gamma_n(p,t)$, obtained by replacing distribution $\rho_0(p)$ with its discrete approximation $\rho_n(p)$ whose mesh-size is $\frac{1}{n}$.

The proof of this theorem is given following exactly the same lines of Appendix B (which proves a similar statement under steady state conditions).

## 4 Stability analysis

Let us define the system as *stable* if, for every initial condition, the opinion distribution converges to the stationary solution for $t\to\infty$. In this section, we first focus on the case where the personality distribution is discrete, then we generalize the result to the the case of continuous personality distributions.

### 4.1 Discrete personality distribution

Let us state the stability conditions as follows:

- $\underline{\Xi}$ is Hurwitz-stable (i.e., all the eigenvalues of $\underline{\Xi}$ have positive real part).
- $\mathbf{\Psi}$ is Hurwitz-stable.

Indeed, recalling the expression of the covariance in (12), its value remains limited if the first condition is met, while, looking at (13) and (35), the mean of the distribution remains limited if both the above conditions are satisfied. In particular, when the first condition is not met, there are some personalities for which the opinion distribution scatters about along some directions. We call this phenomenon *type-I instability*. Instead, if the first condition is met and the second one is not, for all personalities the opinion covariance remains limited but there are some personalities for which the mean opinion value drifts to infinity. We will refer to this case as *type-II instability*.

Below we elaborate on the conditions that are needed to ensure the system stability.

#### 4.1.1 Condition to avoid Type-I instability

Let $\lambda_k(\mathbf{A})=\lambda_k^R(\mathbf{A})+j\lambda_k^I(\mathbf{A})$ be the $k$-th eigenvalue of matrix $\mathbf{A}$. Recall that $\Xi(p_i)=\bar{\alpha}(p_i)\mathbf{\Gamma}(p_i)+\alpha(p_i)\mathbf{C}$, thus the necessary and sufficient condition to avoid type-I instability is given by

$$\min_k\lambda_k^R(\Xi(p_i))>0,\quad\forall i\,.$$

Since $\mathbf{\Gamma}(p_i)=\mathbf{C}\sum_h\zeta(p_i,p_h)\rho_0(p_h)$, we can write

$$\Xi(p_i)=\left[\bar{\alpha}(p_i)\sum_h\zeta(p_i,p_h)\rho_0(p_h)+\alpha(p_i)\right]\mathbf{C}\,.$$

It follows that the stability condition becomes

$$\left[\bar{\alpha}(p_i)\sum_h\zeta(p_i,p_h)\rho_0(p_h)+\alpha(p_i)\right]\min_k\lambda_k^R(\mathbf{C})>0\,.$$

Assuming that $\bar{\alpha}(p_i)\sum_h\zeta(p_i,p_h)\rho_0(p_h)+\alpha(p_i)>0$ for every $i$, then stability is ensured when

$$\min_k\lambda_k^R(\mathbf{C})>0$$

The above expression highlights that if the opinion dynamics along every topic are stable, introducing a Hurwitz stable matrix $\mathbf{C}$ preserves stability.

#### 4.1.2 Condition to avoid Type-II instability

The necessary and sufficient condition to avoid type-II instability is given by

$$\min_k\lambda_k^R(\mathbf{\Psi}(p_i))>0,\quad\forall i\,.$$

Let us focus on the case where $\alpha(p) \equiv \alpha$. By recalling (34), we write

$$\begin{aligned} \mathbf{\Psi} &= \underline{\mathbf{\Xi}} - \mathbf{P}_2\underline{\mathbf{Z}} \\ &= \bar{\alpha}\underline{\mathbf{\Gamma}} + \alpha\mathbf{I}_M \otimes \mathbf{C} - \bar{\alpha}\mathbf{P}_0\underline{\mathbf{Z}} \\ &= \alpha\mathbf{I}_M \otimes \mathbf{C} + \bar{\alpha}\left(\underline{\mathbf{\Gamma}} - \mathbf{P}_0\underline{\mathbf{Z}}\right) \end{aligned} \tag{36}$$

where $\underline{\mathbf{\Gamma}} = \operatorname{diag}\left(\mathbf{\Gamma}(p_1), \ldots, \mathbf{\Gamma}(p_M)\right)$, with $\mathbf{\Gamma}(p_i)$ obtained by discretizing (7). Then we define $\underline{\mathbf{\Gamma}} - \mathbf{P}_0\underline{\mathbf{Z}} \triangleq \mathbf{\Theta} \otimes \mathbf{C}$, where $\mathbf{\Theta}$ is the $M \times M$ matrix whose $(i,j)$-th entry is given by

$$\theta_{ij} = \begin{cases} \sum_{k \neq i} \zeta(p_i, p_k) r_k, & i = j \\ -\zeta(p_i, p_j) r_i, & i \neq j\,. \end{cases}$$

As a consequence, considering that $\lambda_{(j-1)N+i}(\mathbf{\Theta} \otimes \mathbf{C}) = \lambda_i(\mathbf{\Theta})\lambda_j(\mathbf{C})$ and using (36), the condition for stability reads as follows:

$$\Re\{\lambda_i(\mathbf{\Theta})\lambda_j(\mathbf{C})\} > -\frac{\alpha}{1-\alpha}\Re\{\lambda_j(\mathbf{C})\} \quad \forall i, j\,.$$

**Remark.** In the scalar case ($N = 1$), the following propositions hold.

***Proposition 1.*** If $\zeta(p_i, p_j) \geq 0$ for all $i, j$, $\mathbf{\Psi}$ is Hurwitz-stable.

**Proof:** If $\zeta(p_i, p_j) \geq 0$ for all $i, j$, $\mathbf{\Psi}$ is a matrix whose off-diagonal elements are nonpositive and we can apply Theorem 1 of [36]. Precisely the Hurwitz stability of $\mathbf{\Psi}$ (condition J29 in Theorem 1 of [36]) is implied by the fact that the row sums of $\mathbf{\Psi}$ are all positive (condition K35 applied with diagonal matrix $\mathbf{D} = \mathbf{I}$). Since the row sums of $\mathbf{\Psi}$ are all equal to $\alpha$, the proposition follows. ■

***Proposition 2.*** $\mathbf{\Psi}$ is Hurwitz-stable if $\min_{i,j} \zeta(p_i, p_j) < 0$ and, for every $i$

$$\sum_{i \neq j} \left(\zeta(p_i, p_j) r_j - |\zeta(p_i, p_j)| r_i\right) > -\frac{\alpha}{1-\alpha}\,. \tag{37}$$

**Proof:** If $\min_{i,j} \zeta(p_i, p_j) < 0$, the Hurwitz stability of $\mathbf{\Psi}$ (condition J29 in Theorem 1 of [36]) is implied by condition N39 applied with diagonal matrix $\mathbf{D} = \mathbf{I}$, which is expressed in (37). ■

Note that the above propositions still hold in the case of continuous personalities. We now present an example on the stability conditions for a simple case as described below.

**Example 1:** Consider $N = 1$ and that there are $M$ personalities $p_i = \frac{(2i-1)}{M} - 1$, $i = 1, \ldots, M$ ($M$ even), with $r_i = \frac{1}{M}$ and

$$\zeta(p_i, p_j) = \begin{cases} \zeta_1, & p_i p_j > 0 \\ -\zeta_2, & p_i p_j < 0 \end{cases}$$

with $\zeta_1, \zeta_2$ being arbitrary positive values. Thus, $\mathbf{\Xi}_i$ is a scalar equal to $\alpha + \bar{\alpha}\frac{\zeta_1 - \zeta_2}{2}$ for all $i$. Moreover,

$$\underline{\mathbf{Z}} = \begin{bmatrix} \zeta_1 & -\zeta_2 \\ -\zeta_2 & \zeta_1 \end{bmatrix} \otimes \mathbf{1}_{M/2}$$

where $\mathbf{1}_n$ is a size-$n$ square matrix with all entries equal to 1. Thus,

$$\lambda_i(\underline{\mathbf{Z}}) = \begin{cases} 0, & 1 \leq i \leq M-2 \\ M\frac{\zeta_1 - \zeta_2}{2}, & i = M-1 \\ M\frac{\zeta_1 + \zeta_2}{2}, & i = M\,. \end{cases}$$

Since $\mathbf{\Psi} = \left(\alpha + \bar{\alpha}\frac{\zeta_1 - \zeta_2}{2}\right)\mathbf{I}_M - \frac{\bar{\alpha}}{M}\underline{\mathbf{Z}}$, it follows that

$$\lambda_i(\mathbf{\Psi}) = \begin{cases} \alpha + \bar{\alpha}\frac{\zeta_1 - \zeta_2}{2}, & 1 \leq i \leq M-2 \\ \alpha, & i = M-1 \\ \alpha - \bar{\alpha}\zeta_2, & i = M\,. \end{cases}$$

We then have type-I instability if $\zeta_2 \geq \zeta_1 + 2\frac{\alpha}{1-\alpha}$ and type-II instability if $\frac{\alpha}{1-\alpha} \leq \zeta_2 < \zeta_1 + 2\frac{\alpha}{1-\alpha}$. Conversely, the system is stable iff $\zeta_2 < \frac{\alpha}{1-\alpha}$. It is easy to see that this is equivalent to condition (37), which in this case is necessary and sufficient. ■

### 4.2 Continuous personality distribution

We consider the case of a continuous personality distribution as the limit case of a family of discrete distributions with increasingly small discretization steps, $\Delta p$. Then similarly to what done in the previous section, we assume that the following inequality holds:

$$\begin{aligned} &\lim_{\Delta p \to 0}\left(\sum_{j \neq i} \zeta(p_i, p_j)\rho_0(p_j)\Delta p - \sum_{j \neq i} |\zeta(p_i, p_j)|\rho_0(p_i)\Delta p\right) \\ &= \int \zeta(p, q)\rho_0(q)\,\mathrm{d}q - \int |\zeta(p, q)|\,\mathrm{d}q\rho_0(p) \\ &> -\frac{\alpha(p)}{1-\alpha(p)}\,. \end{aligned} \tag{38}$$

We now prove that the above is the stability condition for the continuous case, provided that some technical conditions, specified below, are met. Suppose that (38) is true for our continuous-personality system. Then, for $\Delta p$ sufficiently small, (37) is also satisfied for the discretized system, so that the corresponding $\mathbf{\Psi}(\Delta p)$ is Hurwitz-stable. main-field

Next we assume that for $\Delta p$ sufficiently small:

- $\mathbf{\Psi}(\Delta p)$ is uniformly Hurwitz-stable, i.e., the minimum real part of its eigenvalues is bounded away from zero by an amount $\epsilon$ and
- $\mathbf{\Psi}^{-1}(\Delta p)$ has a uniformly bounded norm, i.e., $\|\mathbf{\Psi}^{-1}(\Delta p)\| < K, \forall \Delta p$ sufficiently small and $K < \infty$ [2].

It follows that, for $\Delta p$ sufficiently small, the fixed-point solution in (35) is uniformly bounded from above. As $\Delta p \to 0$, the fixed-point solution for the continuous-personality system is uniformly

2. This last condition is implied by the previous one when $\mathbf{\Psi}(\Delta p)$ is diagonalizable.

bounded from above (by Theorem 1 and norm continuity) for every finite $t$. Hence, the system is stable.

## 5 STEADY-STATE ANALYSIS

Under previous stability/ergodicity conditions, it is interesting to analyse the limiting solution for $t \to \infty$, which can be obtained as solution of the associated steady state FP equation. In the most general case, we can rewrite (5) disregarding the dependence on $t$, as the following multi-dimensional FP equation:

$$\sum_{n=1}^{N} \frac{\partial}{\partial x_n} \left(\mu_n(p,\mathbf{x})\rho(p,\mathbf{x})\right) = \sum_{m,n=1}^{N} \frac{D_{mn}}{2} \frac{\partial^2 \rho(p,\mathbf{x})}{\partial x_m \partial x_n} \tag{39}$$

where $\rho(p,\mathbf{x}) = \lim_{t\to\infty} \rho(p,\mathbf{x},t)$ and, according to the second line in (6), $\boldsymbol{\mu}$ is given by

$$\begin{aligned} \boldsymbol{\mu}(p,\mathbf{x}) &= \bar{\alpha}(p)\left[\boldsymbol{\gamma}(p) - \boldsymbol{\Gamma}(p)\mathbf{x}\right] + \alpha(p)\mathbf{C}[\mathbf{u}(p) - \mathbf{x}] \\ &= \bar{\alpha}(p)\mathbf{C}(\boldsymbol{\beta}(p) - \eta(p)\mathbf{x}) + \alpha(p)\mathbf{C}[\mathbf{u}(p) - \mathbf{x}] \\ &= -w(p)\mathbf{C}(\mathbf{x} - \mathbf{f}(p)) \,. \end{aligned} \tag{40}$$

In the above equation, we used the expressions of $\boldsymbol{\gamma}(p)$ and $\boldsymbol{\Gamma}(p)$ in (8) and (7), respectively, and defined $\boldsymbol{\gamma}(p) = \mathbf{C}\boldsymbol{\beta}(p)$ and $\boldsymbol{\Gamma} = \eta(p)\mathbf{C}$ where $\eta(p) = \int_q \zeta(p,q)\rho_0(q)\,\mathrm{d}q$ and $\boldsymbol{\beta}(p) = \int_{\mathbf{y}}\int_q \zeta(p,q)\mathbf{y}\rho(q,\mathbf{y})\,\mathrm{d}^N\mathbf{y}\,\mathrm{d}q$. Moreover, we defined $w(p) = \bar{\alpha}(p)\eta(p) + \alpha(p)$ and $\mathbf{f}(p) = \frac{\bar{\alpha}(p)\boldsymbol{\beta}(p) + \alpha(p)\mathbf{u}(p)}{w(p)}$.

Interestingly, if vector $\boldsymbol{\beta}(p)$ is given, then $\boldsymbol{\mu}(p,\mathbf{x})$ does not depend on $\rho(p,\mathbf{x})$ any longer, rather it depends only on the personality $p$, the opinion vector $\mathbf{x}$, and the initial distribution $\rho_0(p)$. Furthermore, observe that $\boldsymbol{\phi}(p) = \frac{\beta(p)}{\eta(p)}$ can be expressed as the fixed point of the multidimensional Fredholm equation:

$$\begin{aligned} \boldsymbol{\phi}(p) &= \underbrace{\int_q \frac{\zeta(p,q)\alpha(q)\mathbf{u}(q)\rho_0(q)}{\eta(p)w(q)}\,\mathrm{d}q}_{\mathbf{h}(q)} \\ &\quad + \int_q \underbrace{\frac{\zeta(p,q)\bar{\alpha}(q)\eta(q)}{\eta(p)w(q)}}_{\Phi(p,q)} \boldsymbol{\phi}(q)\rho_0(q)\,\mathrm{d}q \\ &= \mathbf{h}(p) + \int_q \Phi(p,q)\boldsymbol{\phi}(q)\rho_0(q)\,\mathrm{d}q \end{aligned} \tag{41}$$

(See Appendix B in the Supplemental Material for details.)

From (40), we observe that if $-w(p)\mathbf{C}$ is stable for every $p$, the stationary solution of the FP equation under steady-state conditions is given by [33]:

$$\rho(p,\mathbf{x}) = \frac{\rho_0(p)}{\sqrt{2\pi|\mathbf{W}(p)|}} \mathrm{e}^{-1/2(\mathbf{x}-\mathbf{f}(p))^{\mathsf{T}}\mathbf{W}(p)^{-1}(\mathbf{x}-\mathbf{f}(p))} \tag{42}$$

where $\mathbf{W}(p)$ is the unique solution of the Lyapunov equation $\mathbf{A}(p)\mathbf{W}(p) + \mathbf{W}(p)\mathbf{A}(p)^{\mathsf{T}} = -\mathbf{D}$ and can be expressed as:

$$\mathbf{W}(p) = -\int_0^\infty \mathrm{e}^{\mathbf{A}(p)\tau}\mathbf{D}\mathrm{e}^{\mathbf{A}(p)^{\mathsf{T}}\tau}\,\mathrm{d}\tau \,.$$

A more explicit expression of $\rho(p,\mathbf{x})$ can be obtained when $\mathbf{C}$ is symmetric. Let us write $\boldsymbol{\mu}(p,\mathbf{x})$ as

$$\boldsymbol{\mu}(p,\mathbf{x}) = -\nabla_{\mathbf{x}} \left[ \underbrace{\frac{w(p)}{2}(\mathbf{x} - \mathbf{f}(p))^{\mathsf{T}}\mathbf{C}(\mathbf{x} - \mathbf{f}(p))}_{V(p,\mathbf{x})} \right] \tag{43}$$

where $V(p,\mathbf{x})$ is a potential. If $\mathbf{D}$ is symmetric, it can be written as $\mathbf{D} = \mathbf{Q}\boldsymbol{\Sigma}\mathbf{Q}^{\mathsf{T}}$ where $\mathbf{Q}$ is an orthogonal matrix and $\boldsymbol{\Sigma}$ is diagonal. The multi-dimensional FP equation in (39) can be rewritten as

$$-\nabla_{\mathbf{x}}^{\mathsf{T}}\left[\nabla_{\mathbf{x}}V(p,\mathbf{x})\rho(p,\mathbf{x})\right] = \frac{1}{2}\mathsf{Tr}\left\{\mathbf{D}\mathbf{H}_{\mathbf{x}}(\rho(p,\mathbf{x}))\right\} \tag{44}$$

where $\mathbf{H}_{\mathbf{x}}(\rho(p,\mathbf{x}))$ is the Hessian matrix of $\rho(p,\mathbf{x})$ with respect to the variable $\mathbf{x}$. By defining $\mathbf{L} = \mathbf{Q}\boldsymbol{\Sigma}^{-1/2}$, we then have $\mathbf{L}^{\mathsf{T}}\mathbf{D}\mathbf{L} = \mathbf{I}$. Now consider a new system of coordinates, $\mathbf{y}$, such that $\mathbf{x} = \mathbf{L}\mathbf{y}$. Then, for any twice differentiable function $f(\mathbf{x})$, we have $\nabla_{\mathbf{x}}f(\mathbf{x}) = \mathbf{L}\nabla_{\mathbf{y}}f(\mathbf{L}\mathbf{y}) = \mathbf{L}\nabla_{\mathbf{y}}f_1(\mathbf{y})$ and $\mathbf{H}_{\mathbf{x}}(f(\mathbf{x})) = \mathbf{L}\mathbf{H}_{\mathbf{y}}(f(\mathbf{L}\mathbf{y}))\mathbf{L}^{\mathsf{T}} = \mathbf{L}\mathbf{H}_{\mathbf{y}}(f_1(\mathbf{y}))\mathbf{L}^{\mathsf{T}}$. By replacing these expressions in (44), we obtain

$$-\nabla_{\mathbf{y}}^{\mathsf{T}}\mathbf{L}^{\mathsf{T}}\left[\mathbf{L}\nabla_{\mathbf{y}}V_1(p,\mathbf{y})\rho_1(p,\mathbf{y})\right] = \frac{\mathsf{Tr}\left\{\mathbf{D}\mathbf{L}\mathbf{H}_{\mathbf{y}}(\rho_1(p,\mathbf{y}))\mathbf{L}^{\mathsf{T}}\right\}}{2}$$

which, after some algebra, reduces to

$$\begin{aligned} -\nabla_{\mathbf{y}}^{\mathsf{T}}\left[\boldsymbol{\Sigma}^{-1}\nabla_{\mathbf{y}}V_1(p,\mathbf{y})\rho_1(p,\mathbf{y})\right] &= \frac{1}{2}\mathsf{Tr}\left\{\mathbf{H}_{\mathbf{y}}(\rho_1(p,\mathbf{y}))\right\} \\ &= \frac{1}{2}\nabla_{\mathbf{y}}^2\rho_1(p,\mathbf{y}) \,. \end{aligned} \tag{45}$$

In the above equation, $V_1(p,\mathbf{y})$ has the same structure as $V(p,\mathbf{x})$ in (43) and, thus, (45) is a FP equation in standard form whose solution is given by the following distribution:

$$\rho_1(p,\mathbf{y}) = \frac{\mathrm{e}^{-2V_1(p,\mathbf{y})}}{\int\int \mathrm{e}^{-2V_1(q,\mathbf{z})}\mathrm{d}q\mathrm{d}\mathbf{z}} \,. \tag{46}$$

Therefore the steady state solution of the FP equation can be expressed as a Gibbs distribution (46) associated with a potential $V_1(p,\mathbf{y})$.

**Remark 1:** Note that the same expression can be obtained from the transient solution for $t \to \infty$ under certain conditions. Specifically, consider (14)-(16) and their Laplace transforms (17)-(19); using the final-value theorem, we get

$$\lim_{t\to\infty} \boldsymbol{\gamma}(p,t) = \lim_{s\to 0} s\widehat{\boldsymbol{\gamma}}(p,s) \,. \tag{47}$$

Assuming that both $\boldsymbol{\Psi}$ and $\boldsymbol{\Xi}(p)$ are Hurwitz-stable for all $p$, then $\lim_{t\to\infty}\boldsymbol{\gamma}_0(p,t) = \lim_{s\to 0} s\widehat{\boldsymbol{\gamma}}_0(p,s) = 0$ and

$$\begin{aligned} \lim_{t\to\infty} \boldsymbol{\gamma}_1(p,t) &= \lim_{s\to 0} s\widehat{\boldsymbol{\gamma}}_1(p,s) \\ &= \int_q \mathbf{Z}(p,q)\rho_0(q)\alpha(q)\boldsymbol{\Xi}^{-1}(q)\mathbf{C}\mathbf{u}(q)\,\mathrm{d}q \,. \end{aligned} \tag{48}$$

So doing, the stationary solution satisfies the integral equation

$$\begin{aligned} \lim_{t\to\infty} \boldsymbol{\gamma}(p,t) &= \lim_{t\to\infty} \boldsymbol{\gamma}_1(p,t) \\ &\quad + \int_q \mathbf{Z}(p,q)\rho_0(q)\bar{\alpha}(q)\boldsymbol{\Xi}^{-1}(q) \lim_{t\to\infty} \boldsymbol{\gamma}(p,t)\,\mathrm{d}q \,. \end{aligned} \tag{49}$$

Note that the same observations made for (17) hold also for the stationary solution (49).

**Remark 2:** When $t \to \infty$, the expression of the average in

(13) becomes:

$$\begin{aligned}
&\mathbf{m}(p,\cdot,\infty)\\
&= \lim_{t\to\infty} \mathrm{e}^{-\boldsymbol{\Xi}(p)t}\mathbf{y} + \alpha(p)\boldsymbol{\Xi}^{-1}(p)\left(\mathbf{I}_N - \mathrm{e}^{-\boldsymbol{\Xi}(p)t}\right)\mathbf{C}\mathbf{u}(p)\\
&\qquad + \bar{\alpha}(p)\int_0^t \mathrm{e}^{-\boldsymbol{\Xi}(p)(t-\tau)}\boldsymbol{\gamma}(p,\tau)\,\mathrm{d}\tau\\
&= \alpha(p)\boldsymbol{\Xi}^{-1}(p)\mathbf{C}\mathbf{u}(p)\\
&\qquad + \lim_{t\to\infty}\bar{\alpha}(p)\int_0^\infty u(t-\tau)\mathrm{e}^{-\boldsymbol{\Xi}(p)(t-\tau)}\boldsymbol{\gamma}(p,\tau)\,\mathrm{d}\tau\\
&\overset{(a)}{=} \alpha(p)\boldsymbol{\Xi}^{-1}(p)\mathbf{C}\mathbf{u}(p)\\
&\qquad + \bar{\alpha}(p)\lim_{s\to 0} s\mathcal{L}\left\{u(t)\mathrm{e}^{-\boldsymbol{\Xi}(p)t}\right\}\mathcal{L}\left\{\boldsymbol{\gamma}(p,t)\right\}\\
&= \alpha(p)\boldsymbol{\Xi}^{-1}(p)\mathbf{C}\mathbf{u}(p) + \bar{\alpha}(p)\boldsymbol{\Xi}^{-1}(p)\lim_{s\to 0} s\widehat{\boldsymbol{\gamma}}(p,s)\\
&= \boldsymbol{\Xi}^{-1}(p)\left[\alpha(p)\mathbf{C}\mathbf{u}(p) + \bar{\alpha}(p)\boldsymbol{\gamma}(p,\infty)\right]
\end{aligned} \tag{50}$$

where in $(a)$ we applied the final value theorem and we noted that the integral can be written as the convolution of two functions, thus its Laplace transform is the product of the transforms of the aforementioned functions.

**Remark 3:** When $\rho_0(p)$ is discrete, letting $t \to \infty$ in (35), we obtain an expression of the steady state distribution as: $\underline{\boldsymbol{\gamma}}(\infty) = \underline{\mathbf{Z}}\boldsymbol{\Psi}^{-1}\mathbf{P}_1\mathbf{C}\underline{\mathbf{u}}$.

At last, observe that the following result holds:

***Theorem 2.*** i) The Fredholm equation defining $\phi(p)$ (given explicitly in (41)) admits a unique solution which is Lipschitz-continuous. ii) The solution $\phi(p)$ of the Fredholm equation (41), under any distribution $\rho_0(p)$, which is continuous at every point in $p$, is the uniform limit of solutions $\phi_n(p)$, obtained by replacing distribution $\rho_0(p)$ with its discrete approximation $\rho_n(p)$ whose mesh-size is $\frac{1}{n}$.

The proof is provided in Appendix B in the Supplemental Material.

## 6 Summary

For the sake of clarity, here we summarize the main steps and analytical tools that we used in our derivations.

- We first adopted the mean-field approach to model the opinion dynamic evolution through a continuous distribution function whose expression can be obtained by solving a FP equation;
- Then, by taking the Fourier transform of the FP equation and using the method of characteristics, we rewrote it as a system of first-order partial-differential equations;
- Such a system was solved and the final solution was obtained by the Fourier inverse transform;
- The conditions ensuring the system stability were derived for the matrices characterizing the system, by exploiting the definition of Hurwitz stability;
- Finally, we carried out the steady state analysis starting from the FP equation and letting $t \to \infty$. We obtained an expression for $\boldsymbol{\mu}(p,\mathbf{x})$ which is a function of quantities that can be computed by solving a multidimensional Fredholm equation.

## 7 Numerical results

In this section, we show some numerical examples, which shed light on the impact of the model parameters on the stationary state of opinions as well as on their dynamics. In the following, we will always consider a uniform distribution of personalities in the range $[-1,1]$ and a two-dimensional opinion space, i.e., $\mathbf{x} = [x_1, x_2]^\mathsf{T}$.

### 7.1 Sensitivity of the stationary distribution on the model parameters

In this first subsection, we evaluate the impact of the noise variance, the prejudice, and the coupling matrix $\mathbf{C}$, on the stationary distribution.

We first show the effect of noise variance $\sigma_n^2$ and the prejudice $\mathbf{u}$. To this purpose, we consider an asymmetric coupling matrix

$$\mathbf{C} = \begin{bmatrix} 1 & \rho \\ \epsilon & 1 \end{bmatrix} \tag{51}$$

where $\rho = 0.3$ and $\epsilon$ is a very small positive number (namely, $10^{-10}$ while obtaining the results)[3]. Such model is well suited for a case where the reciprocal influence of the opinions on two subjects is unidirectional (e.g., from subject 2 to subject 1): a possible example could be the appreciation of the government action (subject 1), and the opinion on the right level of taxation (subject 2), where subject 2 is more likely to affect subject 1 than vice versa. We assume a constant level of stubbornness $\alpha(p) = 0.01$, while the strength of opinion interaction is given by

$$\zeta(p,q) = \frac{1}{1+|p-q|^2} \tag{52}$$

i.e., interactions are stronger between agents with similar personalities, a model which we will refer to as *proximity* model. Fig. 1 shows the contour lines of the stationary opinion distribution for different values of $\sigma_n^2$, for two different prejudice scenarios. In the top row, the prejudice is given by

$$\mathbf{u}(p) = \begin{cases} [-1,0]^\mathsf{T}, & p < 0 \\ [1,0]^\mathsf{T}, & p \geq 0 \end{cases} \tag{53}$$

while in the bottom it satisfies

$$\mathbf{u}(p) = \begin{cases} [0,-1]^\mathsf{T}, & p < 0 \\ [0,1]^\mathsf{T}, & p \geq 0\,. \end{cases} \tag{54}$$

In both rows, we set $\sigma_n^2 = 10^{-3}$ for the left plot, $\sigma_n^2 = 5\times 10^{-3}$ for the center plot, and $\sigma_n^2 = 10^{-2}$ for the right plot. Observe that the stationary distribution features two peaks, corresponding to the two different prejudice points, with the same height and width. The width increases with increasing noise variance, for the highest noise variance, the peaks start to merge.

Next, in Fig. 2 we investigate the effect of topic correlation, as expressed by the coupling matrix $\mathbf{C}$. We use the same interaction strength as per (52), and the prejudice given in (54). Also, as before, $\alpha(p) = 0.01$ and $\sigma_n^2 = 10^{-3}$. Finally, we consider the coupling matrix as in (51), with $\rho \in \{-10,-5,0,5,10\}$. For $\rho = 0$ there is no interaction between the two opinion components; changing $\rho$ does not have any effect on the mean of the stationary distribution for each personality ((13)), leading to peaks whose locations are essentially invariant with respect to $\rho$. Notice that the distribution of $x_2$ is independent of $\rho$. Moreover, due to the symmetric scenario, in all cases the stationary distribution shows a reflectional symmetry around the origin. Finally, changing $\rho$ into $-\rho$ has, in the considered scenario, the effect of reflecting the stationary distribution around the $x_2$-axis, or, in other words, of

3. We have used $\epsilon$ as an approximation to 0, since setting $\epsilon = 0$ would yield a non-diagonalizable matrix $\mathbf{C}$.

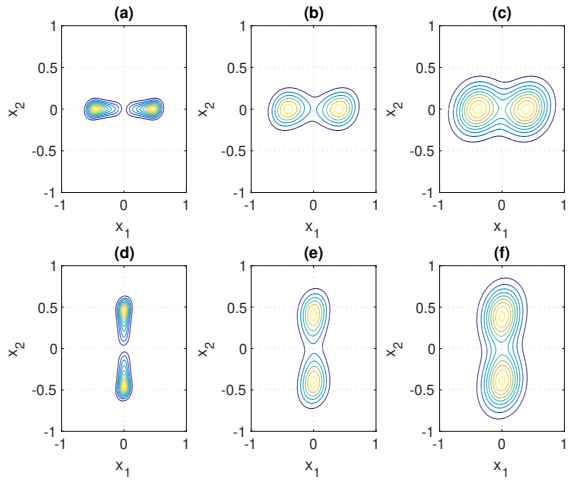


Fig. 1. Contour lines of stationary opinion distribution for $\alpha = 0.01$ and $\zeta(p,q)$ as in (52). Plots (a), (b), (c): prejudice as in (53). Plots (d), (e), (f): prejudice as in (54). Plots (a), (d): $\sigma_n^2 = 10^{-3}$. Plots (b), (e): $\sigma_n^2 = 5 \times 10^{-3}$; plots (c), (f): $\sigma_n^2 = 10^{-2}$.

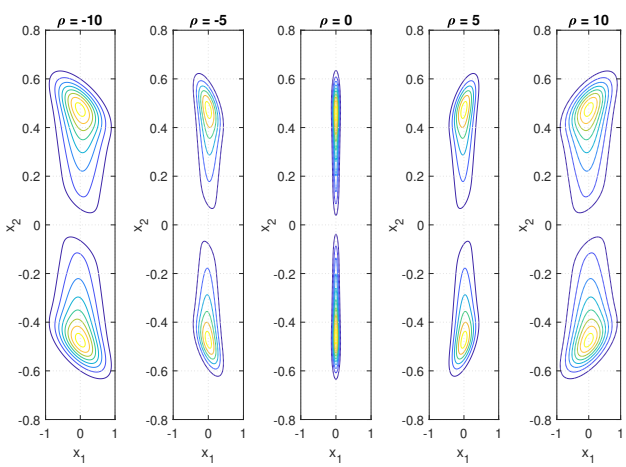


Fig. 2. Contour lines of stationary opinion distribution for $\alpha = 0.01$, $\sigma_n^2 = 10^{-3}$, and $\zeta(p,q)$ as in (52). Coupling matrix given by (51). From left to right: $\rho = -10, -5, 0, 5, 10$.

changing $x_1$ into $-x_1$. From the system point of view, the effect of different correlation values implies a larger share of agents that have strong positive opinions on both subjects (positive correlation value) or one strong positive and one strong negative opinion (negative correlation value). Going back to the previous practical example, agents which are in favor of a low level of taxation will judge more positively or more negatively the government action, depending on the correlation coefficient sign.

### 7.2 Community-based scenario, the influence of Z and the insurgence of instability

We now assess the impact of the interaction strength matrix $\mathbf{Z}$ on the stability of opinions. We consider a different scenario, in which there are $M$ personalities ($M$ even), all with the same stubbornness level $\alpha = 0.01$, organized in two communities and with an interaction similar to that considered in Example 1, but with $N = 2$, i.e.,

$$\underline{\mathbf{Z}} = \left( \begin{bmatrix} \zeta_1 & -\zeta_2 \\ -\zeta_2 & \zeta_1 \end{bmatrix} \otimes \mathbf{1}_{M/2} \right) \otimes \mathbf{C} \tag{55}$$

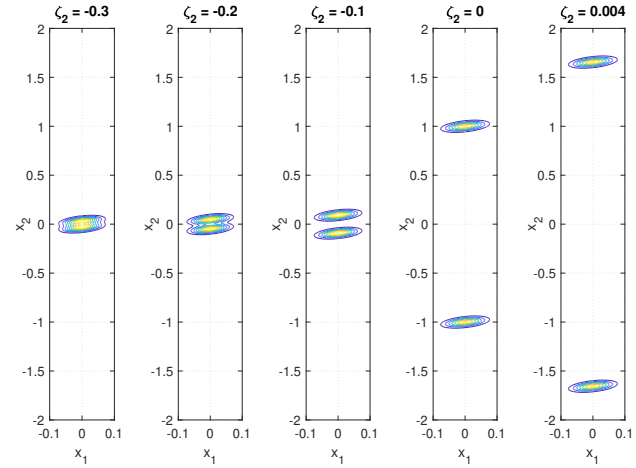


Fig. 3. Stationary opinion distribution for $\alpha = 0.01$, $\sigma_n^2 = 10^{-3}$, $\mathbf{Z}$ as in (55) ($\zeta_1 = 1$), for several values of $\zeta_2$ in the stability region. Coupling matrix given by (51) with $\rho = 1$. From left to right: $\zeta_2 = -0.3, -0.2, -0.1, 0, 0.004$.

with $\zeta_1 = 1$ and $\zeta_2$ used as a parameter. We consider $\mathbf{C}$ as given by (51) with $\rho = 1$, the prejudice as given in (54), and $\sigma_n^2 = 10^{-3}$. Also, $\rho_0(p)$ is given by (25) with $r_i = 1/M$ for all $i$. As a possible practical example of such a scenario, we can think of two religious sects, say Bogumils and Cathars, which have generally different views on two topics, represented by the two opinions $x_1$ and $x_2$.

It is not difficult to show that the stability region boundaries are the same as for Example 1, since $\mathbf{C}$ is Hurwitz-stable. Furthermore, we can derive the asymptotic expression of the mean in (50) by exploiting Remark 3 in Section 5 and writing from (7) and (9) $\underline{\mathbf{\Xi}} = \beta \mathbf{I}_M \otimes \mathbf{C}$, where $\beta = \bar{\alpha}\frac{1-\zeta_2}{2} + \alpha$. If $\beta > 0$, through some algebra and by applying the properties of the Kronecker product, we obtain

$$\mathbf{m}(p, \cdot, \infty) = \frac{\alpha}{\alpha - \overline{\alpha}\zeta_2}\mathbf{u}(p). \tag{56}$$

Thus, if $\zeta_2 < \alpha/\overline{\alpha} \simeq 0.0101$, the system is stable.

Fig. 3 shows the stationary distribution for increasing value of $\zeta_2$. As it can be seen, Bogumils and Cathars merge because of the attractive forces for $\zeta_2 = -0.2$ or lower, while, for $-0.1 < \zeta_2 < 0$, notwithstanding their reciprocal attraction, they remain separated because of the effect of the prejudice. For $0 < \zeta_2 < \alpha/\overline{\alpha}$, the two communities repel each other, but this repulsion is not strong enough to win the effect of prejudice, so stability is preserved while the means grow larger and larger for $\zeta_2 \uparrow \alpha/\overline{\alpha}$.

For $\zeta_2 > \alpha/\overline{\alpha}$, the system is not stable anymore, and (56) does not hold. In particular, when $\zeta_2 < \zeta_1 + 2\alpha/\overline{\alpha} = 1.0202$, the intra-community attraction and the inter-community repulsion have such a relative strength that the system experiences type-II instability: the communities are preserved but their respective means tend to diverge. In our example, the two religious sects are so enemy of each other, to radicalize their views while retaining a strong identity within themselves, giving rise to religious fanaticism. We now look at the time evolution of opinions in the case where, for all personalities, opinions start deterministically from the origin. Fig. 4 shows the opinions at time instants $t = 5, 10, 15, 20$, for $\zeta_2 = 0.1$. Notice that, because of the correlation, the two communities diverge along the bisector of the I-III quadrants.

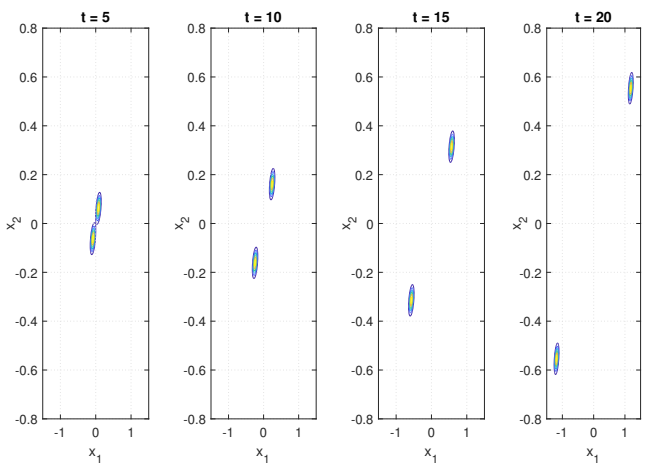


Fig. 4. Contour lines of opinion distribution for $\alpha = 0.01$, $\sigma_n^2 = 10^{-3}$, $\mathbf{Z}$ as in (55) ($\zeta_1 = 1$, $\zeta_2 = 0.1$). Coupling matrix given by (51) with $\rho = 1$. From left to right: $t = 5, 10, 15, 20$.

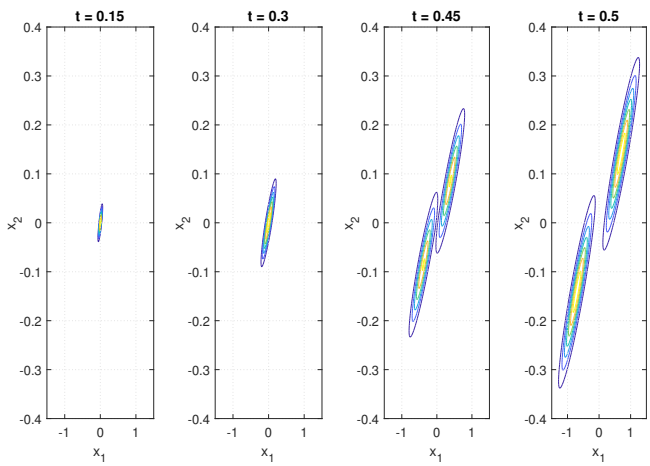


Fig. 5. Contour lines of opinion distribution for $\alpha = 0.01$, $\sigma_n^2 = 10^{-3}$, $\mathbf{Z}$ as in (55) ($\zeta_1 = 1$, $\zeta_2 = 10$). Coupling matrix given by (51) with $\rho = 1$. From left to right: $t = 0.15, 0.3, 0.45, 0.5$.

Finally, for $\zeta_2 > \zeta_1 + 2\alpha/\overline{\alpha}$, the inter-community repulsion prevails on intra-community attraction, the system experiences type-I instability, and the variance within the two communities also grows to infinity. In other words, Bogumils and Cathars dissolve themselves into heterogeneous crowds not having definite views on the two topics of interest. Fig. 4 shows the opinions at time instants $t = 0.15, 0.3, 0.45, 0.5$ for $\zeta_2 = 10$, again when the opinions all start deterministically from the origin. The dynamic is similar, but faster than in the previous case, and the communities expand, so that, for $t$ sufficiently large, their boundaries disappear.

#### 7.2.1 *Finite-network behavior*

In order to assess the validity of the mean-field approach, in Fig. 6 we show the opinion distribution behavior for a finite set of $U = 500$ agents, by solving numerically (3). In particular, we consider the same model as in Figs. 3-4 for $\zeta_2 \in \{-0.1, 0, 0.004, 0.1\}$. The first half of the agents are Bogumils, while the others are Cathars. For the first three values of $\zeta_2$, which correspond to a stable system, we show the stationary opinion distribution. For $\zeta_2 = 0.1$, we show the distribution at time $t = 20$. As it can be seen, the results in Figs. 3-4 match those in Fig. 6, indicating that

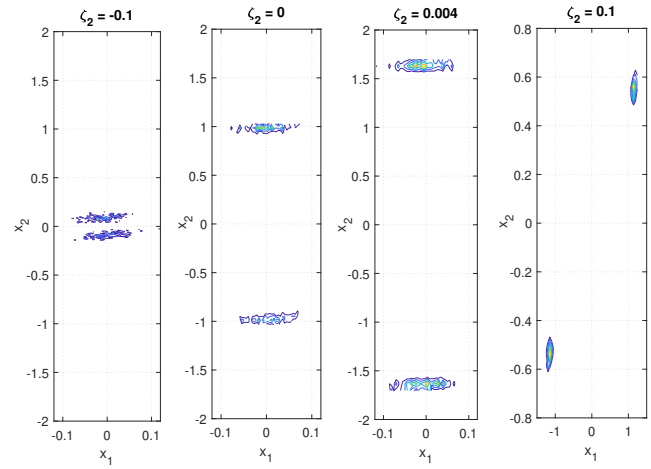


Fig. 6. Contour lines of simulated opinion distribution for $U = 500$, $\alpha = 0.01$, $\sigma_n^2 = 10^{-3}$, $\mathbf{Z}$ as in (55) ($\zeta_1 = 1$), where the first half of users belong to community 1, while the second half to community 2. Coupling matrix given by (51) with $\rho = 1$. From left to right: $\zeta_2 = -0.1, 0, 0.004, 0.1$. For the rightmost graph, $t = 20$.

the asymptotic analysis well represents the opinion dynamics of relatively small populations.

### 7.3 Rotational effects

We now consider the case in which the dynamics is stable in both dimensions, but the effect of the coupling matrix yields instability (of Type I). Consider again the two-community scenario given by (55), but with $\zeta_1 = 1$ and $\zeta_2 = -0.1$. Note that, in the scalar case, this would be a stable scenario. However, let us now consider a coupling matrix given by

$$\mathbf{C} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \tag{57}$$

which is easily seen to have eigenvalues equal to $\pm j$. This is quite an artificial scenario, since this coupling matrix has rather an ad-hoc shape, to which it is difficult to associate any real-world situation. All the other parameters are the same as in the previous example, except for the prejudice, which is given by

$$\mathbf{u}(p) = \begin{cases} [0, -10], & p < 0 \\ [0, 10], & p \geq 0\,. \end{cases} \tag{58}$$

It is easy to see that, for the $i$-th discrete personality, the covariance matrix in (12) is given by:

$$\mathbf{\Sigma}\left(p_i, t\right) = \sigma_n^2 t \mathbf{I}_2 \tag{59}$$

independently from $i$. Note that such covariance matrix does not reach any finite limit for $t \to \infty$, hence the system is unstable. Fig. 7 shows the temporal evolution of the opinion distribution, for $t = 1, 10, 20, 30, 40, 50, 60, 70, 80$. The peaks corresponding to the two communities widen along time, as expected, while their means undergo a rotation around the origin, at some instants (such as at $t = 60$) making the peaks temporarily merge.

## 8 Conclusions

Several analytical models representing the dynamics of social opinions have been proposed in the literature. By drawing on this body of work, we developed a model that accounts for the individual endogenous process of opinion evolution as well as

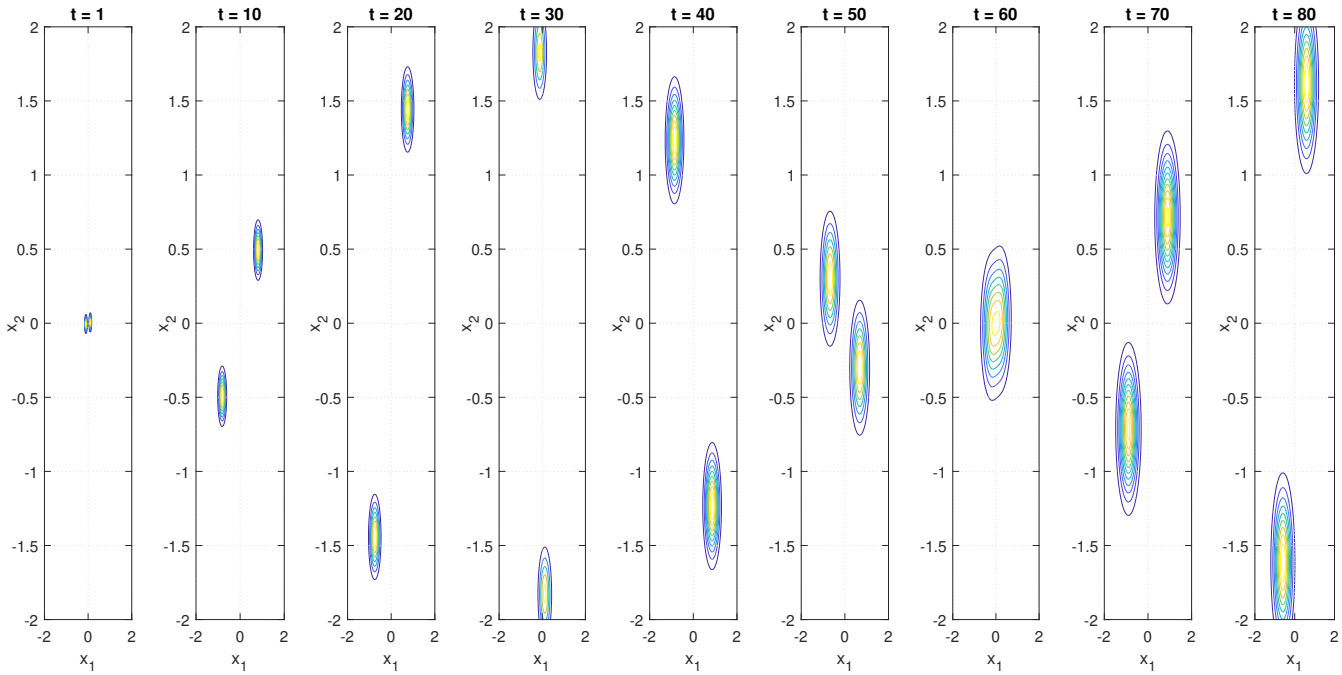


Fig. 7. Contour lines of opinion distribution for $\alpha = 0.01$, $\sigma_n^2 = 10^{-3}$, $\mathbf{Z}$ as in (55) ($\zeta_1 = 1$, $\zeta_2 = -0.1$). Coupling matrix given by (57). From left to right: $t = 1, 10, 20, 30, 40, 50, 60, 70, 80$.

for the possible adversarial behavior of individuals. Importantly, our model also represents the interdependence between opinions on different, yet correlated, subjects. Under asymptotic conditions on the size of the individuals' population, we obtained the time evolution of the opinions distribution as the solution of a multidimensional Fokker-Planck equation. We then discussed the stability conditions and derived the steady-state solution. Our numerical results match the stability conditions we obtained and show interesting phenomena in collective belief dynamics.

## REFERENCES


[1] R. Colbaugh, K. Glass, and P. Ormerod, *Predictability and Prediction for an Experimental Cultural Market*. Springer Berlin Heidelberg, 2010, pp. 79–86.

[2] S. Asur and B. A. Huberman, "Predicting the future with social media," in *IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, Washington, DC, USA, 2010, pp. 492–499.

[3] G. Shi, M. Johansson, and K. H. Johansson, "How agreement and disagreement evolve over random dynamic networks," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 6, pp. 1061–1071, 2013.

[4] F. Baccelli, A. Chatterjee, and S. Vishwanath, "Pairwise stochastic bounded confidence opinion dynamics: Heavy tails and stability," *IEEE Transactions on Automatic Control*, 2017.

[5] B. Pang and L. Lee, "Opinion mining and sentiment analysis," *Found. Trends Inf. Retr.*, vol. 2, no. 1-2, pp. 1–135, Jan. 2008.

[6] L. Dall'Asta, A. Baronchelli, A. Barrat, and V. Loreto, "Nonequilibrium dynamics of language games on complex networks," *PHYSICAL REVIEW E*, vol. 74, pp. 1–13, 2006.

[7] H.-X. Yang and B.-H. Wang, "Disassortative mixing accelerates consensus in the naming game," *Journal of Statistical Mechanics: Theory and Experiment*, Jan. 2015.

[8] T. M. Liggett, "Stochastic models of interacting systems," *Ann. Probab.*, vol. 25, no. 1, pp. 1–29, 01 1997. [Online]. Available: https://doi.org/10.1214/aop/1024404276

[9] G. J. Baxter, S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, "Bootstrap percolation on complex networks," *Phys. Rev. E*, vol. 82, p. 011103, Jul 2010. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevE.82.011103

[10] D. Kempe, J. Kleinberg, and É. Tardos, "Maximizing the spread of influence through a social network," in *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2003, pp. 137–146.

[11] N. E. Friedkin and E. C. Johnsen, "Social positions in influence networks," *Social Networks*, vol. 19, no. 3, pp. 209–222, 1997.

[12] C. Ravazzi, P. Frasca, R. Tempo, and H. Ishii, "Ergodic randomized algorithms and dynamics over networks," *IEEE Transactions on Control of Network Systems*, vol. 2, no. 1, pp. 78–87, March 2015.

[13] G. Como and F. Fagnani, "Scaling limits for continuous opinion dynamics systems," vol. 21, no. 4. Institute of Mathematical Statistics, 2011, pp. 1537–1567. [Online]. Available: http://www.jstor.org/stable/23033379

[14] J. Garnier, G. Papanicolaou, and T.-W. Yang, "Consensus convergence with stochastic effects," *Vietnam Journal of Mathematics*, vol. 1, no. 1, pp. 1–25, March 2016. [Online]. Available: arXiv:1508.07313

[15] V. D. Blondel, J. M. Hendrickx, and J. N. Tsitsiklis, "On krause's multi-agent consensus model with state-dependent connectivity," *IEEE Transactions on Automatic Control*, vol. 54, no. 11, pp. 2586–2597, Nov 2009.

[16] G. Weisbuch, G. Deffuant, F. Amblard, and J.-P. Nadal, "Meet, discuss, and segregate!" *Complexity*, vol. 7, no. 3, pp. 55–63, 2002. [Online]. Available: http://dx.doi.org/10.1002/cplx.10031

[17] A. Bhattacharyya, M. Braverman, B. Chazelle, and H. L. Nguyen, "On the convergence of the hegselmann-krause system," in *4th Conference on Innovations in Theoretical Computer Science*, ser. ITCS '13. New York, NY, USA: ACM, 2013, pp. 61–66. [Online]. Available: http://doi.acm.org/10.1145/2422436.2422446

[18] A. Nordio, A. Tarable, C. F. Chiasserini, and E. Leonardi, "Belief dynamics in social networks: A fluid-based analysis," *IEEE Transactions on Network Science and Engineering*, vol. 5, no. 4, pp. 276–287, 2018.

[19] G. Deffuant, D. Neau, F. Amblard, and G. Weisbuch, "Mixing beliefs among interacting agents," *Advances in Complex Systems*, vol. 03, no. 01n04, pp. 87–98, 2000. [Online]. Available: http://www.worldscientific.com/doi/abs/10.1142/S0219525900000078

[20] R. Hegselmann and U. Krause, "Opinion dynamics and bounded confidence: models, analysis and simulation," *Journal of Artificial Societies and Social Simulation*, vol. 5, no. 3, pp. 87–98, June 2002. [Online]. Available: http://jasss.soc.surrey.ac.uk/5/3/2.html

[21] M. H. DeGroot, "Reaching a consensus," *Journal of the American Statistical Association*, vol. 69, no. 345, pp. 118–121, 1974.

[22] D. Acemoglu, G. Como, F. Fagnani, and A. Ozdaglar, "Opinion fluctua-

tions and persistent disagreement in social networks," in *2011 50th IEEE Conference on Decision and Control and European Control Conference*, Dec 2011, pp. 2347–2352.

[23] A. Nedic and B. Touri, "Multi-dimensional hegselmann-krause dynamics," *Urbana*, vol. 51, p. 61801, 2012.

[24] S. Fortunato, V. Latora, A. Pluchino, and A. Rapisarda, "Vector opinion dynamics in a bounded confidence consensus model," *International Journal of Modern Physics C*, vol. 16, no. 10, pp. 1535–1551, 2005.

[25] L. Li, A. Scaglione, A. Swami, and Q. Zhao, "Consensus, polarization and clustering of opinions in social networks," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 6, pp. 1072–1083, 2013.

[26] N. E. Friedkin, A. V. Proskurnikov, R. Tempo, and S. E. Parsegov, "Network science on belief system dynamics under logic constraints," *Science*, vol. 354, no. 6310, pp. 321–326, 2016.

[27] S. E. Parsegov, A. V. Proskurnikov, R. Tempo, and N. E. Friedkin, "Novel multidimensional models of opinion dynamics in social networks," *IEEE Transactions on Automatic Control*, vol. 62, no. 5, pp. 2270–2285, May 2017.

[28] L. Festinger, *A Theory of Cognitive Dissonance*. Stanford Univ. Press, 1957.

[29] B. Gawronski and F. Strack, Eds., *Cognitive Consistency: A Fundamental Principle in Social Cognition*. New York: Gulford Press, 2012.

[30] A. Baronchelli, "The Emergence of Consensus," *ArXiv e-prints*, Apr. 2017.

[31] J. Gärtner, "On the McKean-Vlasov limit for interacting diffusions," *Mathematische Nachrichten*, vol. 137, no. 1, pp. 197–248, 1988.

[32] D. A. Dawson, "Critical dynamics and fluctuations for a mean-field model of cooperative behavior," *Journal of Statistical Physics*, vol. 31, no. 1, pp. 29–85, 1983.

[33] H. Risken, "Fokker-planck equation," in *The FokkerPlanck Equation: Method of Solution and Applications*. Springer, 1996, pp. 63–95.

[34] T. Burton, *Volterra Integral and Differential Equations*, 2nd ed. Elsevier, 2005, vol. 202.

[35] N. Kolmogorov and S. V. Fomin, *Elements of tyhe Theory of Functions and Functional Analysis*. Martino Publishing, 2012.

[36] R. Plemmons, "M-matrix characterizations.inonsingular m-matrices," *Linear Algebra and its Applications*, vol. 18, no. 2, pp. 175 – 188, 1977. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0024379577900738

[37] S. Sarra, "The method of characteristics with applications to conservation laws," *Journal of Online Mathematics and its Applications*, 2003.

[38] A. N. Kolmogorov and S. V. Fomin, *Elements of the Theory of Functions and Functional Analysis*. Martino Publishing, 2012.

## APPENDIX A
## DERIVATION OF (11)

By using (6) and by neglecting dependencies on the system variables unless strictly necessary, we can rewrite the FP equation in (5) as:

$$\frac{\partial}{\partial t}\rho(\mathbf{x}) = \sum_{n=1}^{N}\frac{\partial}{\partial x_n}\left[(\mathbf{e}_n^{\mathsf{T}}\boldsymbol{\Xi}\mathbf{x}-\phi_n\rho(\mathbf{x})\right]+\frac{1}{2}\sum_{m,n=1}^{N}D_{mn}\frac{\partial^2}{\partial x_m\partial x_n}\rho(\mathbf{x}) \tag{60}$$

where $\mathbf{e}_n$ is the vector with all zero components but the $n$-th one equal to 1. Then, by taking the Fourier transform of (60) with respect to $\mathbf{x}$, we obtain

$$\frac{\partial}{\partial t}\widehat{\rho}(\boldsymbol{\nu})+\boldsymbol{\nu}^{\mathsf{T}}\boldsymbol{\Xi}\nabla_{\boldsymbol{\nu}}\widehat{\rho}(\boldsymbol{\nu}) = -\left[j2\pi\boldsymbol{\nu}^{\mathsf{T}}\boldsymbol{\phi}+2\pi^2\boldsymbol{\nu}^{\mathsf{T}}\mathbf{D}\boldsymbol{\nu}\right]\widehat{\rho}(\boldsymbol{\nu}) \tag{61}$$

where $\boldsymbol{\nu}$ is the variable in the transformed domain, while $\hat{f}(\boldsymbol{\nu})$ is the Fourier transform of the generic function $f(\mathbf{x})$. The operator $\nabla_{\boldsymbol{\nu}}$ is the gradient with respect to the vector $\boldsymbol{\nu}$. Next we transform the above first-order partial-derivative equation into a system of first-order ordinary differential equations by using the method of characteristics [37]. To this end we introduce an auxiliary parameter $u$ and consider $t=t(u)$, $\boldsymbol{\nu}=\boldsymbol{\nu}(u)$ and $\widehat{\rho}(p,\boldsymbol{\nu}(u),t(u))=\widehat{\rho}(u)$, with initial conditions $t(0)=0$, $\boldsymbol{\nu}(0)=\boldsymbol{\nu}_0$ and $\widehat{\rho}(p,\boldsymbol{\nu}(0),t(0))=\widehat{\rho}_0(p,\boldsymbol{\nu}_0)=\widehat{\rho}(0)$. According to the rule of total derivative, $\frac{\mathrm{d}\widehat{\rho}}{\mathrm{d}u}$ can be written as

$$\frac{\mathrm{d}\widehat{\rho}}{\mathrm{d}u}=\frac{\partial\widehat{\rho}}{\partial t}\frac{\mathrm{d}t}{\mathrm{d}u}+\frac{\mathrm{d}\boldsymbol{\nu}^{\mathsf{T}}}{\mathrm{d}u}\nabla_{\boldsymbol{\nu}}\widehat{\rho}\,.$$

If we set $\frac{\mathrm{d}t}{\mathrm{d}u}=1$ and $\frac{\mathrm{d}\boldsymbol{\nu}}{\mathrm{d}u}=\boldsymbol{\Xi}^{\mathsf{T}}\boldsymbol{\nu}$, the above total derivative corresponds to the left hand side of (61). With such setting we can reduce (61) to the following system of differential equations:

$$\begin{cases}\frac{\mathrm{d}t}{\mathrm{d}u} = 1\\ \frac{\mathrm{d}\boldsymbol{\nu}}{\mathrm{d}u} = \boldsymbol{\Xi}^{\mathsf{T}}\boldsymbol{\nu}\\ \frac{\mathrm{d}\widehat{\rho}}{\mathrm{d}u} = -\left[j2\pi\boldsymbol{\nu}^{\mathsf{T}}\boldsymbol{\phi}+2\pi^2\boldsymbol{\nu}^{\mathsf{T}}\mathbf{D}\boldsymbol{\nu}\right]\widehat{\rho}\,.\end{cases} \tag{62}$$

The first two equations are easily solved as $t(u)=u$, and $\boldsymbol{\nu}(u)=\mathrm{e}^{\boldsymbol{\Xi}^{\mathsf{T}}u}\boldsymbol{\nu}_0$. By substituting these solutions in the third equation, the latter can be rearranged as

$$\frac{\mathrm{d}\widehat{\rho}(u)}{\widehat{\rho}(u)}=-\left[j2\pi\boldsymbol{\nu}_0^{\mathsf{T}}\mathrm{e}^{\boldsymbol{\Xi}u}\boldsymbol{\phi}+2\pi^2\boldsymbol{\nu}_0^{\mathsf{T}}\mathrm{e}^{\boldsymbol{\Xi}u}\mathbf{D}\mathrm{e}^{\boldsymbol{\Xi}^{\mathsf{T}}u}\boldsymbol{\nu}_0\right]\mathrm{d}u \tag{63}$$

We now recall that $\boldsymbol{\phi}=\boldsymbol{\phi}(t)$ is a function of $t$. Since $t=u$, by integrating with respect to $u$ the above equation, we get

$$\log\frac{\widehat{\rho}(u)}{\widehat{\rho}(0)}=-j2\pi\boldsymbol{\nu}_0^{\mathsf{T}}\int_0^u\mathrm{e}^{\boldsymbol{\Xi}w}\boldsymbol{\phi}(w)\,\mathrm{d}w-2\pi^2\boldsymbol{\nu}_0^{\mathsf{T}}\left(\int_0^u\mathrm{e}^{\boldsymbol{\Xi}w}\mathbf{D}\mathrm{e}^{\boldsymbol{\Xi}^{\mathsf{T}}w}\,\mathrm{d}w\right)\boldsymbol{\nu}_0 \tag{64}$$

The solution for $\hat{\rho}(u)$ is given by

$$\widehat{\rho}(u) = \widehat{\rho}(0)\exp\left\{-j2\pi\boldsymbol{\nu}_0^{\mathsf{T}}\int_0^u\mathrm{e}^{\boldsymbol{\Xi}w}\boldsymbol{\phi}(w)\,\mathrm{d}w-2\pi^2\boldsymbol{\nu}_0^{\mathsf{T}}\left(\int_0^u\mathrm{e}^{\boldsymbol{\Xi}w}\mathbf{D}\mathrm{e}^{\boldsymbol{\Xi}^{\mathsf{T}}w}\,\mathrm{d}w\right)\boldsymbol{\nu}_0\right\} \tag{65}$$

By substituting $u=t$ and $\boldsymbol{\nu}_0=\mathrm{e}^{-\boldsymbol{\Xi}^{\mathsf{T}}t}\boldsymbol{\nu}(t)$ into (65), and after reintroducing all dependencies we finally obtain

$$\begin{aligned}\widehat{\rho}(p,\boldsymbol{\nu},t) &= \widehat{\rho}_0(p,\boldsymbol{\nu}_0)\exp\left\{-j2\pi\boldsymbol{\nu}^{\mathsf{T}}\underbrace{\mathrm{e}^{-\boldsymbol{\Xi}(p)t}\left[\int_0^t\mathrm{e}^{\boldsymbol{\Xi}(p)\tau}\boldsymbol{\phi}(p,\tau)\,\mathrm{d}\tau\right]}_{\widetilde{\mathbf{m}}(p,t)}-2\pi^2\boldsymbol{\nu}^{\mathsf{T}}\underbrace{\left(\int_0^t\mathrm{e}^{-\boldsymbol{\Xi}(p)\tau}\mathbf{D}\mathrm{e}^{-\boldsymbol{\Xi}(p)^{\mathsf{T}}\tau}\,\mathrm{d}\tau\right)}_{\boldsymbol{\Sigma}(p,t)}\boldsymbol{\nu}\right\}\\ &= \widehat{\rho}_0\left(p,\boldsymbol{\nu}\mathrm{e}^{-\boldsymbol{\Xi}(p)t}\right)\exp\left\{-j2\pi\boldsymbol{\nu}^{\mathsf{T}}\widetilde{\mathbf{m}}(p,t)-2\pi^2\boldsymbol{\nu}^{\mathsf{T}}\boldsymbol{\Sigma}(p,t)\boldsymbol{\nu}\right\}\end{aligned} \tag{66}$$

where we used the fact that

$$\begin{aligned}\boldsymbol{\nu}_0^{\mathsf{T}}\left(\int_0^t\mathrm{e}^{\boldsymbol{\Xi}w}\mathbf{D}\mathrm{e}^{\boldsymbol{\Xi}^{\mathsf{T}}w}\,\mathrm{d}w\right)\boldsymbol{\nu}_0 &= \boldsymbol{\nu}^{\mathsf{T}}\left(\int_0^t\mathrm{e}^{-\boldsymbol{\Xi}(p)t}\mathrm{e}^{\boldsymbol{\Xi}w}\mathbf{D}\mathrm{e}^{\boldsymbol{\Xi}^{\mathsf{T}}w}\mathrm{e}^{-\boldsymbol{\Xi}(p)t}\,\mathrm{d}w\right)\boldsymbol{\nu}\\ &= \boldsymbol{\nu}^{\mathsf{T}}\left(\int_0^t\mathrm{e}^{\boldsymbol{\Xi}(w-t)}\mathbf{D}\mathrm{e}^{\boldsymbol{\Xi}^{\mathsf{T}}(w-t)}\,\mathrm{d}w\right)\boldsymbol{\nu}\\ &\overset{(a)}{=} \boldsymbol{\nu}^{\mathsf{T}}\left(\int_0^t\mathrm{e}^{-\boldsymbol{\Xi}\tau}\mathbf{D}\mathrm{e}^{-\boldsymbol{\Xi}^{\mathsf{T}}\tau}\,\mathrm{d}\tau\right)\boldsymbol{\nu}\end{aligned} \tag{67}$$

and in $(a)$ we defined $\tau=t-w$. Now, by taking the inverse Fourier transform w.r.t. $\boldsymbol{\nu}$ of (66), we get

$$\rho(p,\mathbf{x},t)=\mathrm{e}^{\boldsymbol{\Xi}(p)t}\rho_0\left(p,\mathbf{x}\mathrm{e}^{\boldsymbol{\Xi}(p)t}\right)\star\mathcal{G}\left(\mathbf{x},\widetilde{\mathbf{m}}(p,t),\boldsymbol{\Sigma}(p,t)\right) \tag{68}$$

where the symbol $\star$ represents the convolution operator and $\mathcal{G}\left(\mathbf{x},\widetilde{\mathbf{m}}(p,t),\boldsymbol{\Sigma}(p,t)\right)$ is the pdf of the multivariate Gaussian distribution with mean $\widetilde{\mathbf{m}}(p,t)$ and covariance $\boldsymbol{\Sigma}(p,t)$. Finally, after a suitable change of variable, and by recalling that $\rho_0(p,\mathbf{y})=\rho_0(\mathbf{y}|p)\rho_0(p)$,

we rewrite the convolution product in (68) as

$$\rho(p,\mathbf{x},t) \quad = \quad \rho_0(p)\int_{\mathbf{y}} \rho_0(\mathbf{y}|p)\,\mathcal{G}\left(\mathbf{x},\mathbf{m}(p,\mathbf{y},t),\mathbf{\Sigma}(p,t)\right)\,\mathrm{d}^N\mathbf{y} \tag{69}$$

where $\mathbf{m}(p,\mathbf{y},t) \triangleq \mathrm{e}^{-\mathbf{\Xi}(p)t}\mathbf{y} + \widetilde{\mathbf{m}}(p,t) = \mathrm{e}^{-\mathbf{\Xi}(p)t}\left[\mathbf{y} + \int_0^t \mathrm{e}^{\mathbf{\Xi}(p)\tau}\boldsymbol{\phi}(p,\tau)\,\mathrm{d}\tau\right]$.

# APPENDIX B
# PROOF OF THEOREM 2

## B.1 Preliminaries

Under steady state conditions, using (42) and the definition of $\mathbf{f}(p)$, we can rewrite $\boldsymbol{\beta}(p)$ as

$$\begin{aligned}\boldsymbol{\beta}(p) \quad &= \quad \int_q \zeta(p,q)\rho_0(q)\int_{\mathbf{x}} \frac{\mathbf{x}}{\sqrt{2\pi|\mathbf{W}(q)|}}\mathrm{e}^{-\frac{1}{2}(\mathbf{x}-\mathbf{f}(q))^{\mathsf{T}}\mathbf{W}(q)^{-1}(\mathbf{x}-\mathbf{f}(q))}\,\mathrm{d}^N\mathbf{x}\,\mathrm{d}q \\ &= \quad \int_q \zeta(p,q)\rho_0(q)\mathbf{f}(q)\,\mathrm{d}q \end{aligned} \tag{70}$$

By changing variable and considering $\boldsymbol{\phi}(p) = \frac{\boldsymbol{\beta}(p)}{\eta(p)}$, we rewrite the above expression as

$$\begin{aligned}\boldsymbol{\phi}(p) \quad &= \quad \underbrace{\int_q \frac{\zeta(p,q)\alpha(q)\mathbf{u}(q)\rho_0(q)}{\eta(p)w(q)}\,\mathrm{d}q}_{\mathbf{h}(q)} + \int_q \underbrace{\frac{\zeta(p,q)\bar{\alpha}(q)\eta(q)}{\eta(p)w(q)}}_{\Phi(p,q)}\boldsymbol{\phi}(q)\rho_0(q)\,\mathrm{d}q \\ &= \quad \mathbf{h}(p) + \int_q \Phi(p,q)\boldsymbol{\phi}(q)\rho_0(q)\,\mathrm{d}q \end{aligned} \tag{71}$$

which is the Fredholm equation of the second type [38]. In the following, we will consider a per-element solution of the above equation.

In order to solve the Fredholm equation, let us denote by $\mathcal{A}[\phi](p)$ the following operator

$$\mathcal{A}[\phi](p) = \int_p \Phi(p,q)\phi(q)\rho_0(q)\,\mathrm{d}q \tag{72}$$

operating on a Banach space of Lipschits continuous functions $\phi(q)$, where $\phi(q)$ is the generic component of $\boldsymbol{\phi}(p)$. We have that $\phi(p)$ is equipped with the following norm:

$$\begin{aligned}\|\phi(p)\|_{\mathrm{Lip}} \quad &= \quad c_0\sup_{\mathcal{P}}|\phi(p)| + c_1\sup_{p,q\in\mathcal{P},p\neq q}\frac{|\phi(p)-\phi(q)|}{|p-q|} \\ &= \quad c_0\|\phi(p)\|_\infty + c_1\|\phi(p)\|_{\mathrm{L}} \end{aligned} \tag{73}$$

with $c_0, c_1 > 0$ and $c_0 + c_1 = 1$. Then we can rewrite (71) as

$$\phi(p) = h(p) + \mathcal{A}[\phi](p)$$

where $h(p)$ is the generic component of $\mathbf{h}(p)$. From the above expression, we get

$$\phi(p) = (\mathcal{I}-\mathcal{A})^{-1}[h](p)$$

whenever $(\mathcal{I}-\mathcal{A})^{-1}[\cdot]$ exists and is continuous over the aforementioned Banach space.

Now let $R(p)$, $R:\mathcal{P}\to[0,1]$ be a continuous initial cumulative distribution with pdf $\rho_0(p)$, and let $R_n(p)$ be the stepwise approximation of $R(p)$ with meshsize equal to $1/n$. Then, using (72), we can write

$$\mathcal{A}_n[\phi] = \int_{\mathcal{P}} \Phi(p,q)\phi(q)\,\mathrm{d}R_n(q), \quad \mathcal{A}[\phi] = \int_{\mathcal{P}} \Phi(p,q)\phi(q)\,\mathrm{d}R(q),$$

and

$$h_n(p) = \int_{\mathcal{P}} \frac{\zeta(p,q)\alpha(q)u(q)}{\eta(p)w(q)}\,\mathrm{d}R_n(q), \quad h(p) = \int_{\mathcal{P}} \frac{\zeta(p,q)\alpha(q)u(q)}{\eta(p)w(q)}\,\mathrm{d}R(q),$$

Given the above definitions, we introduce $\phi_n(p)$ as

$$\phi_n(p) = \left(\mathcal{I}-\mathcal{A}_n\right)^{-1}[h_n](p)\,.$$

## B.2 Main Theorem

We can now prove Theorem 2 whose statement is reported again below for completeness:

***Theorem 3.*** i) The Fredholm equation (71) admits a unique solution which is Lipschitz-continuous. ii) The solution $\phi(p)$ of the Fredholm equation (71), under any distribution $\rho_0(p)$, which is continuous at every point in $p$, is the uniform limit of solutions $\phi_n(p)$, obtained by replacing distribution $\rho_0(p)$ with its discrete approximation $\rho_n(p)$ whose mesh-size is $\frac{1}{n}$.

*Proof:* In order to prove the thesis we proceed as follows:

- i) descends from the properties of operator $\mathcal{A}[\cdot]$ over the Banach space of Lipschitz-continuous functions equipped with norm $\|\cdot\|_{\rm Lip}$, which satisfies $\|\mathcal{A}[\cdot]\|_{\rm Lip} < 1$, since we can derive the existence and the continuity of the operator $(\mathcal{I}-\mathcal{A})^{-1}[\cdot]$ with respect to norm $\|\cdot\|_{\rm Lip}$ (see Lemma 1 below).
- ii) descends from Lemma 2 which shows that operators $\mathcal{A}_n[\cdot]$ converge to the operator $\mathcal{A}[\cdot]$, and from Lemma 3 which shows that convergence holds also for the inverse operators $(\mathcal{I}-\mathcal{A}_n)^{-1}[\cdot]$. Then, using Lemma 4, we show that $\phi_n(p) = (\mathcal{I}-\mathcal{A}_n)^{-1}[h_n](p)$ converges to $\phi(p) = (\mathcal{I}-\mathcal{A})^{-1}[h](p)$.

□

***Lemma 1.*** Given the operator $\mathcal{A}[\cdot]$ defined above we have that $\|\mathcal{A}[\phi](p)\|_{\rm Lip} < 1$ for opportunely chosen $c_0$ and $c_1$ under the assumption that $\Phi(p,q)$ is $C^1(\mathcal{P}^2)$. Furthermore the operator $(\mathcal{I}-\mathcal{A})^{-1}[\cdot]$ exists continuous with respect to norm $\|\cdot\|_{\rm Lip}$.

*Proof:* We have that

$$\begin{aligned} |\mathcal{A}[\phi](p)| &\leq \int_q |\Phi(p,q)\phi(q)\rho_0(q)\,\mathrm{d}q| \leq \int_q |\Phi(p,q)\rho_0(q)||\phi(q)|\,\mathrm{d}q \\ &\leq \|\phi(p)\|_\infty \int_q |\Phi(p,q)\rho_0(q)|\,\mathrm{d}q \end{aligned} \tag{74}$$

Therefore, by assuming that $\alpha(p)$ and $\zeta(p,p')$ are regular in their argmmuments, we have

$$\begin{aligned} \|\mathcal{A}[\phi](p)\|_\infty &= \sup_p \|\mathcal{A}[\phi](p)\| \\ &\leq \|\phi(p)\|_\infty \sup_p \int_q \Phi(p,q)\rho_0(q)\,\mathrm{d}q \\ &\leq \|\phi(p)\|_\infty \frac{1}{\eta(p)} \sup_q \frac{\bar{\alpha}(q)\eta(q)}{\alpha(q)+\bar{\alpha}(q)\eta(q)} \underbrace{\int_q \zeta(p,q)\rho_0(q)\,\mathrm{d}q}_{\eta(p)} \\ &= \|\phi(p)\|_\infty \sup_q \frac{\bar{\alpha}(q)\eta(q)}{\alpha(q)+\bar{\alpha}(q)\eta(q)} \\ &= \kappa\|\phi(p)\|_\infty \end{aligned} \tag{75}$$

where $\kappa < 1$. Similarly, using the Lagrange theorem, we have

$$\sup_{p,q\in\mathcal{P},p\neq q} \frac{|y(p)-y(q)|}{|p-q|} = \sup_p \left|\frac{\mathrm{d}y(p)}{\mathrm{d}p}\right|$$

for any continuous and differentiable function $y(p)$. Therefore,

$$\begin{aligned} \left|\frac{\mathrm{d}\mathcal{A}[\phi](p)}{\mathrm{d}p}\right| &= \left|\int_q \frac{\partial\Phi(p,q)}{\partial p}\phi(q)\rho_0(q)\,\mathrm{d}q\right| \leq \\ &\leq \int_q \left|\frac{\partial\Phi(p,q)}{\partial p}\phi(q)\rho_0(q)\right|\,\mathrm{d}q \\ &\leq \left|\sup_{p,q\in\mathcal{P}} \frac{\partial\Phi(p,q)}{\partial p}\rho_0(q)\right| \|\phi(q)\|_\infty \\ &= \|\Psi(p,q)\|_L \|\phi(q)\|_\infty \end{aligned} \tag{76}$$

Note that, to obtain the last expression, we defined $\|\Psi(p,q)\|_{\rm L} = \sup_{p,q\in\mathcal{P}} \left|\frac{\partial\Phi(p,q)}{\partial p}\rho_0(q)\right|$. Now combining (75) and (76), we have:

$$\|\mathcal{A}[\phi](p)\|_{\rm Lip} \leq c_0\kappa\|\phi(p)\|_\infty + c_1\|\Psi(p,q)\|_{\rm L}\|\phi(p)\|_\infty$$

Dividing both sides by $\|\phi(p)\|_{\rm Lip}$ we get

$$\frac{\|\mathcal{A}[\phi](p)\|_{\rm Lip}}{\|\phi(p)\|_{\rm Lip}} \leq c_0\frac{\kappa\|\phi(p)\|_\infty}{\|\phi(p)\|_{\rm Lip}} + c_1\|\Psi(p,q)\|_{\rm L}\frac{\|\phi(p)\|_\infty}{\|\phi(p)\|_{\rm Lip}}.$$

Now, since by construction $\frac{\|\phi(p)\|_\infty}{\|\phi(p)\|_{\rm Lip}} \leq \frac{1}{c_0}$, we have

$$\|\mathcal{A}[\phi](p)\|_{\rm Lip} = \sup\frac{\|\mathcal{A}[\phi](p)\|_{\rm Lip}}{\|\phi(p)\|_{\rm Lip}} \leq \kappa + \frac{c_1}{c_0}\|\Psi(p,q)\|_{\rm L}$$

which can be made smaller than 1 by opportunely setting $c_0$ and $c_1$, i.e. by setting $\frac{c_1}{c_0} < \frac{1-\kappa}{\|\Psi(p,q)\|_{\rm L}}$.

For a generic linear operator $\mathcal{A}[\cdot]$ defined over a Banach space such that $\|\mathcal{A}[\cdot]\| < 1$, the associated operator $(\mathcal{I}-\mathcal{A})^{-1}[\cdot]$ exists continuous and can be written as [38, Th. 8 p. 102 and Th. 1 p. 111]

$$(\mathcal{I}-\mathcal{A})^{-1}[\cdot] = \sum_{k=0}^{\infty}(\mathcal{A})^k[\cdot]\,. \tag{77}$$

□

We now consider the sequence of pertubated operators $\mathcal{A}_n[\cdot]$. The general result below applies.

***Lemma 2.*** Given a sequence of operators $\mathcal{A}_n[\cdot]$ and a sequence of functions $h_n(p)$ as defined above, they converge to, respectively, the above expressions of $\mathcal{A}[\cdot]$ and $h(p)$ in Lipschitz norm (i.e., $\|\mathcal{A}_n[\cdot]-\mathcal{A}[\cdot]\|_{\mathrm{Lip}} \to 0$ and $\|h_n(p)-h(p)\|_{\mathrm{Lip}} \to 0$).

*Proof:* To simplify the notation, without loss of generality we assume $\mathcal{P}=[0,1]$ (we recall that $\mathcal{P}$ is assumed to be compact). Then,

$$\mathcal{A}_n[\phi(p)] = \int_{\mathcal{P}} \Phi(p,q)\phi(q)\,\mathrm{d}R_n(q) \quad = \quad \sum_{m=0}^{n-1}\Phi\left(p,\frac{m}{n}\right)\phi\left(\frac{m}{n}\right)\int_{\frac{m}{n}}^{\frac{m+1}{n}}\mathrm{d}R(q)$$

Furthermore,

$$|\mathcal{A}[\phi(p)]-\mathcal{A}_n[\phi(p)]| = \sum_{m=0}^{n-1}\int_{\frac{m}{n}}^{\frac{m+1}{n}}\left[\Phi(p,q)\phi(q)-\Phi\left(p,\frac{m}{n}\right)\phi\left(\frac{m}{n}\right)\right]\mathrm{d}R(q)$$

We therefore obtain:

$$\begin{aligned}
|\mathcal{A}[\phi(p)]-\mathcal{A}_n[\phi(p)]| &\leq \sum_{m=0}^{n-1}\int_{\frac{m}{n}}^{\frac{m+1}{n}}|\Phi(p,q)|\left|\phi(q)-\phi\left(\frac{m}{n}\right)\right|\mathrm{d}R(q)\\
&+ \sum_{m=0}^{n-1}\int_{\frac{m}{n}}^{\frac{m+1}{n}}\left|\Phi(p,q)-\Phi\left(p,\frac{m}{n}\right)\right|\left|\phi\left(\frac{m}{n}\right)\right|\mathrm{d}R(q)\\
&\leq \|\Phi(p,q)\|_\infty\frac{\|\phi(p)\|_L}{n}\int_0^1\mathrm{d}R(q)\\
&+ \frac{\|\Phi(p,q)\|_{\mathrm{L}}}{n}\|\phi(p)\|_\infty\int_0^1\mathrm{d}R(q)\\
&= \frac{1}{n}\left(\|\Phi(p,q)\|_\infty\|\phi(p)\|_L+\|\Phi(p,q)\|_L\|\phi(p)\|_\infty\right)
\end{aligned}$$

It follows that:

$$\|\mathcal{A}[\phi(p)]-\mathcal{A}_n[\phi(p)]\|_\infty \leq \frac{1}{n}\left(\|\Phi(p,q)\|_\infty\|\phi(p)\|_L+\|\Phi(p,q)\|_L\|\phi(p)\|_\infty\right)$$

Similarly, since $\Phi(p,p')$ is assumed to be differentiable with continuous derivative with respect to $p$ and $q$, both $\mathcal{A}[\phi(p)]$ and $\mathcal{A}_n[\phi(p)]$ are differentiable at every point with continuous derivative and:

$$\left|\frac{\mathrm{d}\mathcal{A}[\phi(p)]}{\mathrm{d}p}-\frac{\mathrm{d}\mathcal{A}_n[\phi(p)]}{\mathrm{d}p}\right| \quad = \quad \left|\sum_{m=0}^{n-1}\int_{\frac{m}{n}}^{\frac{m+1}{n}}\left[\frac{\partial\Phi(p,q)}{\partial p}\phi(q)-\frac{\partial\Phi\left(p,\frac{m}{n}\right)}{\partial p}\phi\left(\frac{m}{n}\right)\right]\mathrm{d}R(q)\right|$$

Proceeding as before, we get:

$$\|\mathcal{A}[\phi(p)]-\mathcal{A}_n[\phi(p)]\|_L \quad \leq \quad \frac{1}{n}\left(\left\|\frac{\partial\Phi(p,q)}{\partial p}\right\|_\infty\|\phi(p)\|_L+\left\|\frac{\partial\Phi(p,q)}{\partial p}\right\|_L\|\phi(p)\|_\infty\right)$$

and

$$\begin{aligned}
\|\mathcal{A}[\phi(p)]-\mathcal{A}_n[\phi(p)]\|_{\mathrm{Lip}} \quad \leq \quad & \frac{c_0}{n}\left(\|\Phi(p,q)\|_\infty\|\phi(p)\|_L+\Phi(p,q)\|_L\|\phi(p)\|_\infty\right)\\
&+\frac{c_1}{n}\left(\left\|\frac{\partial\Phi(p,q)}{\partial p}\right\|_\infty\|\phi(p)\|_L+\left\|\frac{\partial\Phi(p,q)}{\partial p}\right\|_L\|\phi(p)\|_\infty\right)
\end{aligned}$$

Since in the right hand side of the above expressions none of the norms depend on $n$, it easy to see that $\|\mathcal{A}[\phi(p)]-\mathcal{A}_n[\phi(p)]\|_{\mathrm{Lip}} \to 0$ as $n\to\infty$. With similar arguments, we can prove that $\|h(p)-h_n(p)\|_{\mathrm{Lip}} \to 0$ as $n\to\infty$. □

***Lemma 3.*** Given a Banach space with norm $\|\cdot\|$, and a sequence of linear operators $\mathcal{A}_n[\cdot]\to\mathcal{A}[\cdot]$ in norm, with $\|\mathcal{A}[\cdot]\|<1$, we have that the continuous operators $(\mathcal{I}[\cdot]-\mathcal{A}_n[\cdot])^{-1}\to(\mathcal{I}[\cdot]-\mathcal{A}[\cdot])^{-1}$ in norm.

*Proof:* Given that $\|\mathcal{A}[\cdot]\|<1$ by the continuity of norm $\|\mathcal{A}_n[\cdot]\|\to\|\mathcal{A}[\cdot]\|<1$, for $n$ sufficiently large we can assume $\|\mathcal{A}_n[\cdot]\|<1$. For any of such $n$, we define $\mathcal{B}_n[\cdot]=(\mathcal{I}[\cdot]-\mathcal{A}_n[\cdot])^{-1}-(\mathcal{I}[\cdot]-\mathcal{A}[\cdot])^{-1}$. By (77), we can write

$$\mathcal{B}_n[\cdot] = \sum_{k=0}^{\infty}(\mathcal{A}_n[\cdot])^k-\sum_k(\mathcal{A}[\cdot])^k\,.$$

Since both series on the right hand side of the above expression converge, we can write:

$$\mathcal{B}_n[\cdot] = \sum_{k=0}^{\infty} (\mathcal{A}_n[\cdot])^k - (\mathcal{A}[\cdot])^k$$

Now, we have:

$$\|\mathcal{B}_n[\cdot]\| = \left\| \sum_{k=0}^{\infty} (\mathcal{A}_n[\cdot])^k - (\mathcal{A}[\cdot])^k \right\| \leq \sum_{k=0}^{\infty} \|(\mathcal{A}_n[\cdot])^k - (\mathcal{A}[\cdot])^k\|$$

where the last inequality follows by the sub-additivity and continuity of the norm.

Denoted with $c_k = \|(\mathcal{A}_n[\cdot])^k - (\mathcal{A}[\cdot])^k\| = \|[(\mathcal{A}_n[\cdot] - \mathcal{A}[\cdot]) + \mathcal{A}[\cdot]]^k - (\mathcal{A}[\cdot])^k\|$, and considering that the operator algebra is, in general, non-commutative, we have:

$$\begin{aligned}
c_k &= \left\| \sum_{x_i \in \{0,1\}, i=1,\ldots,k} \prod_i (\mathcal{A}[\cdot])^{x_i} (\mathcal{A}_n[\cdot] - \mathcal{A}[\cdot])^{1-x_i} - (\mathcal{A}[\cdot])^k \right\| \\
&= \left\| \sum_{\substack{x_i \in \{0,1\}, i=1,\ldots,k \\ \{x_1, x_2 \cdots x_k\} \neq \{1,1,\cdots,1\}}} \prod_i (\mathcal{A}[\cdot])^{x_i} (\mathcal{A}_n[\cdot] - \mathcal{A}[\cdot])^{1-x_i} \right\| \\
&\leq \sum_{\substack{x_i \in \{0,1\}, i=1,\ldots,k \\ \{x_1, x_2 \cdots x_k\} \neq \{1,1,\cdots,1\}}} \prod_i \|(\mathcal{A}[\cdot])^{x_i}\| \|(\mathcal{A}_n[\cdot] - \mathcal{A}[\cdot])^{1-x_i}\| \\
&= \sum_{i=1}^{k} \binom{k}{i} \|\mathcal{A}_n[\cdot] - \mathcal{A}[\cdot]\|^i \|\mathcal{A}[\cdot]\|^{k-i} \\
&= \sum_{i=1}^{k} \binom{k}{i} \|\mathcal{A}[\cdot]\|^i \|\mathcal{A}_n[\cdot] - \mathcal{A}[\cdot]\|^{k-i} \\
&= (\|\mathcal{A}[\cdot]\| + \|\mathcal{A}_n[\cdot] - \mathcal{A}[\cdot]\|)^k - \|\mathcal{A}[\cdot]\|^k .
\end{aligned}$$

Therefore, by monotonicity of positive series, we have:

$$\|\mathcal{B}_n[\cdot]\| \leq \sum_{k=0}^{\infty} c_k \leq \sum_{k=0}^{\infty} (\|\mathcal{A}[\cdot]\| + \|\mathcal{A}_n[\cdot] - \mathcal{A}[\cdot]\|)^k - \|\mathcal{A}[\cdot]\|^k \tag{78}$$

Since $\|\mathcal{A}_n[\cdot] - \mathcal{A}[\cdot]\|_\infty \to 0$ and $\|\mathcal{A}[\cdot]\|_\infty < 1$, we can assume $n$ sufficiently large so that $\|\mathcal{A}[\cdot]\| + \|\mathcal{A}_n[\cdot] - \mathcal{A}[\cdot]\| < 1$, thus:

$$\begin{aligned}
\|\mathcal{B}_n[\cdot]\| &\leq \sum_{k=0}^{\infty} (\|\mathcal{A}[\cdot]\| + \|\mathcal{A}_n[\cdot] - \mathcal{A}[\cdot]\|)^k - \|\mathcal{A}[\cdot]\|^k \\
&= \left| \sum_{k=0}^{\infty} (\|\mathcal{A}[\cdot]\| + \|\mathcal{A}_n[\cdot] - \mathcal{A}[\cdot]\|)^k - \sum_{k=0}^{\infty} \|\mathcal{A}_n[\cdot]\|^k \right| \\
&= \left| \frac{1}{1 - (\|\mathcal{A}[\cdot]\| + \|\mathcal{A}_n[\cdot] - \mathcal{A}[\cdot]\|)} - \frac{1}{1 - \|\mathcal{A}[\cdot]\|} \right|
\end{aligned}$$

The thesis follows immediately since $\|\mathcal{A}_n[\cdot] - \mathcal{A}[\cdot]\| \to 0$, hence $\|\mathcal{B}_n[\cdot]\| \to 0$.

□

***Lemma 4.*** Given the Banach space $\mathrm{Lip}(\mathcal{P})$, a sequence of linear and continuous operators $\mathcal{C}_n[\cdot]$: $\mathrm{Lip}[\mathcal{P}] \to \mathrm{Lip}[\mathcal{P}]$ converging in $\|\cdot\|_{\mathrm{Lip}}$-norm to the continuous operator $\mathcal{C}$, and a sequence of functions $y_n \in \mathrm{Lip}(\mathcal{P})$ converging to $y$ in $\|\cdot\|_{\mathrm{Lip}}$-norm, then $\mathcal{C}_n[y_n]$ converges in $\|\cdot\|_{\mathrm{Lip}}$-norm to $\mathcal{C}[y]$.

*Proof:*

To simplify the notation, we denote $\|\cdot\|_{\mathrm{Lip}}$ simply with $\|\cdot\|$:

$$\begin{aligned}
\|\mathcal{C}_n[y_n] - \mathcal{C}[y]\| &= \|\mathcal{C}_n[y_n] - \mathcal{C}_n[y_0] + \mathcal{C}_n[y_0] - \mathcal{C}[y]\| \\
&\leq \|\mathcal{C}_n[y_n - y_0]\| + \|\mathcal{C}_n[y] - \mathcal{C}[y]\|
\end{aligned}$$

Now, on the one hand, $\|\mathcal{C}_n[y_n - y]\| \leq \|\mathcal{C}_n[\|y_n - y\|]\| \to 0$ given that $\|\mathcal{C}_n[\cdot]\|$ is bounded since it is continuous, and $\|y_n - y\| \to 0$. On the other hand, $\|\mathcal{C}_n[y] - \mathcal{C}[y]\| \to 0$ since $\|\mathcal{C}_n[\cdot] - \mathcal{C}[\cdot]\| \to 0$, while $y$ is bounded in norm. The assertion follows immediately.

□