Data-Informed Modeling of the Formation, Persistence, and Evolution of Social Norms and Conventions

Data-informed modeling of the formation, persistence, and evolution of social norms and conventions

Mengbin Ye* and Lorenzo Zino

Abstract Social norms and conventions are commonly accepted and adopted behaviors and practices within a social group that guide interactions —e.g., how to spell a word or how to greet people— and are central to a group's culture and identity. Understanding the key mechanisms that govern the formation, persistence, and evolution of social norms and conventions in social communities is a problem of paramount importance for a broad range of real-world applications, spanning from preparedness for future emergencies to promotion of sustainable practices. In the past decades, mathematical modeling has emerged as a powerful tool to reproduce and study the complex dynamics of norm and convention change, gaining insights into their mechanisms, and ultimately deriving tools to predict their evolution. The first goal of this chapter is to introduce some of the main mathematical approaches for modeling social norms and conventions, including population models and agent-based models relying on the theories of dynamical systems, evolutionary dynamics, and game theory. The second goal of the chapter is to illustrate how quantitative observations and empirical data can be incorporated into these mathematical models in a systematic manner, establishing a data-based approach to mathematical modeling of formation, persistence, and evolution of social norms and conventions. Finally, current challenges and future opportunities in this growing field of research are discussed.

Mengbin Ye
Centre for Optimisation and Decision Science, Curtin University, Bentley 6149 WA, Australia, e-mail: mengbin.ye@curtin.edu.au

Lorenzo Zino
Department of Electronics and Telecommunications, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy, e-mail: lorenzo.zino@polito.it

* Corresponding author

Keywords

agent-based models; cascade; complex networks; data science; dynamical systems; evolutionary dynamics; psychology experiments; game theory

Introduction

Mathematical modeling has increasingly emerged as a powerful tool to describe, study, and predict complex social dynamics. Such models are applied to a broad range of applications; several examples are mentioned here to provide an impression, while the interested reader may consider review papers for broader coverage (Castellano et al 2009; Jusup et al 2022). Different application contexts include, for instance, studying social influence and opinion formation (French Jr 1956; DeGroot 1974; Anderson and Ye 2019; Noorazar 2020), understanding the diffusion of social innovations (Bass 1969; Montanari and Saberi 2010; Peyton Young 2011; Zino et al 2022) and the spread of misinformation (Del Vicario et al 2016; Franceschi and Pareschi 2022), anticipating crimes and violence (D'Orsogna and Perc 2015; Succar and Porfiri 2024), unveiling complex political processes (Leonard et al 2021; Fontan and Altafini 2021), and predicting the behavioral response of people to an epidemic outbreak (Cinelli et al 2020; Ye et al 2021b).

The use of mathematical models to study social phenomena can be traced back to pioneering efforts developed in the 1940s, which were not met with ready acceptance by social scientists at large. For a survey of such pioneering works, see Rapoport (1963); Abelson (1967). Mathematical models started to become widely used and accepted in the social science community from the end of the 1960s, thanks to an array of seminal works including the diffusion model proposed by Frank Bass (Bass 1969), the opinion dynamics model proposed by Morris H. DeGroot (DeGroot 1974), and the linear threshold model due to Mark Granovetter (Granovetter 1978). These works paved the way for a flourishing growth of the field of mathematical modeling of social systems over the past several decades, witnessing the development of more refined mathematical models tailored to reproduce specific phenomena of interest (Edling 2002; Castellano et al 2009; Jusup et al 2022). In particular, a crucial advancement can be pinpointed to the design and refinement of agent-based models (Bonabeau 2002), in which a population of heterogeneous agents interact following simple agent-based rules on complex network structures (Vega-Redondo 2007; Easley and Kleinberg 2010). This modeling paradigm allows one to encapsulate within the mathematical model key features of human behavior, such as the tendency to be influenced by their peers in a non-additive manner and the non-homogeneous and often time-varying pattern of social interactions Holme and Saramäki (2012); Guilbeault et al (2018).

In order to move beyond using such models to explain observed phenomena and towards direct application, a crucial challenge is the integration of empirical observation and data into the design and validation of the models, and in their calibration to real-world scenarios. In the past, models were typically developed by first proposing mechanisms that were grounded or inspired by established social psychology theories and general observations drawn from empirical evidence. Then, the models were validated in controlled experiments and fitted to field data. This is the case, for instance, of classical opinion dynamics models, which are based on the mathematical mechanism of linear averaging. The seminal DeGroot model of opinion formation (DeGroot 1974) was built on the social psychological theories of social influence and social power (French Jr 1956). It has been validated experimentally subsequently (Becker et al 2017; Chandrasekhar et al 2020) and fitted to online social media data (Kozitsin 2021). Similarly, one of its most successful extensions, the Friedkin–Johnsen model, was proposed based on the social psychological theory that people are unwilling to depart from existing prejudices (Friedkin and Johnsen 1990), and has since then been validated experimentally (Friedkin and Bullo 2017) and fitted to the discussion dynamics of the Paris Climate Accords (Bernardo et al 2021).

Recently, the unprecedented availability of large datasets (e.g. online social media data) and the tremendous advancements in data science have dramatically changed the picture around modeling social systems, leading to the birth of the novel emergent field of computational social science (Lazer et al 2009). Besides the development of model-free data-driven approaches, even in the context of model-based approaches such a paradigm shift has increased the centrality of data integration into mathematical models. In this context, data are not just used a posteriori to calibrate existing models as noted above, but also to inform the design and refinement of new and improved models. This ultimately yields a data-based approach to mathematical modeling of social systems.

The previous paragraphs have illustrated how the field of mathematical modeling of social systems is extremely broad and continuously evolving, with the continuous development of new models, methodologies, and applications. This chapter focuses on the specific application to the formation, persistence, and evolution of social norms and conventions. In other words, the main focus of the rest of this chapter is on mathematical models tailored to capture and reproduce the collective adoption, maintenance, and replacement of a behavior, idea, or action by a social community (Peyton Young 2015). Shedding light onto this phenomenon and developing mathematical models that are able to predict it and, ultimately, control said phenomenon have important societal implications in the context of social change towards a sustainable economy (Selin 2021; Hoffmann et al 2024) and preparedness to future emergencies and disasters (Van Bavel et al 2020). The aim of this chapter is to introduce the reader to data-driven mathematical modeling of social norms and conventions, and as such, the chapter will elaborate on several illustrative examples rather than provide a comprehensive account of the literature.

The rest of the chapter is organized as follows. First, the problem of modeling social norms and conventions is presented, with a brief discussion of these concepts from a psychological and sociological perspective, illustrated by means of real-world examples and highlighting their key features. Second, the main mathematical approaches used to model these social phenomena are discussed. The chapter starts presenting population models, which are the first and simplest class of models, and highlighting their inherent limitations in capturing key features of social systems, such as interpersonal interactions, social influence, and heterogeneity. Then, different classes of agent-based models that have been developed to overcome the limitations of population models are presented: cascading models, evolutionary dynamic models, and game-theoretic decision-making models. Third, the use and integration of real-world data to design, inform, validate, and calibrate these mathematical models is extensively discussed, illustrated by different examples for different classes of agent-based models. Fourth, the chapter is concluded by providing a summary of the main advances in the modeling of formation and evolution of social norms and conventions and on the integration of real-word data, and discussing current and future challenges and opportunities in this promising research field.

Background on norms and conventions

In order to develop appropriate mathematical models of social norms and conventions, one first needs to understand what norms and conventions are from a theoretical psychological and sociological perspective. For this reason, this section is devoted to a brief discussion of relevant concepts, with explanations by means of classical examples from the related literature concerning the formation, persistence, and evolution of norms and conventions.

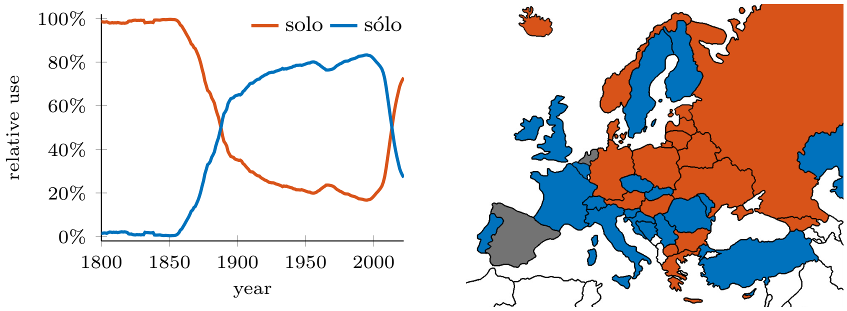
Social norms and social conventions are fundamental aspects of our societies and their cultures, and help to provide a framework for interactions between people (Lewis 2002; Bicchieri 2014; Gelfand and Jackson 2016). Despite conceptual and functional differences between norms and conventions which are theoretically debated upon in the social psychology literature (Southwood and Eriksson 2011), they have some important common traits. Thus, from a modeling perspective, one can consider norms and conventions together. Throughout this chapter, the term "norm" and "convention" will be used interchangeably. A common feature of both is that their value mostly depends on their widespread adoption and acceptance (Bicchieri 2005, 2014; Lewis 2002). For instance, when walking on a crowded sidewalk, it matters little whether one walks on the left-side or right-side of the sidewalk, provided that others do likewise (Lewis 2002). When greeting others, there is not much difference between the use of handshakes versus bowing, but it would be embarrassing to not coordinate on the same gesture (Marmor 2009). When collaboratively writing a document in English, especially for instance a scientific manuscript featuring authors from different countries, one could use British or American spelling, as far as the document is consistent and all the authors use the same spelling (Lieberman et al 2007). These are all examples of social conventions, where individuals benefit from coordinating with peers to adopt the same behavior among several equivalent alternatives. Norms share many similarities, but one key difference is that the the normative behavior (whatever is accepted as the standard in the group) is often enforced by potential or actual punishment for deviants. For instance, it is now a generally accepted norm within warfare that medical personnel and wounded soldiers are considered noncombatants (Finnemore and Sikkink 1998), and is now part of the 1949 Geneva Convention. The processes of formation, persistence, and evolution of social norms are fundamental for the well functioning of our societies (Lewis 2002; Marmor 2009). The understanding of these key societal processes is critical for public authorities who, for many different reasons, may be interested in changing the conventions or social norms currently adopted by a population in a systematic manner, e.g., to promote a persistent change towards more sustainable behaviors to face the ongoing climate change crisis (Selin 2021) or being able to promptly adapt conventions and norms to respond to societal emergencies, as happened during the COVID-19 health crisis (Van Bavel et al 2020).

Real-world examples

Here, some examples of real-world social norms and conventions are presented and discussed. These examples illustrate some of the key features of norms and conventions and show how they form, persist in time, and evolve. Readers interested in further examples can enjoy the excellent survey by Peyton Young (2015).

Footbinding in rural China

Footbinding practice among Chinese women is a classical example of the evolution of a social norm, extensively studied in the sociology literature (Mackie 1996; Brown and Satterthwaite-Phillips 2018). This practice involved the painful application of a binding cloth over a girl's feet (often from a young age) for a long period, in order to change their shape and reduce their size. Those that did not undergo footbinding were seen as less desirable in terms of marriage, illustrating the disadvantages in deviating from the norm. It is recorded that footbinding was the normative practice in rural China for many centuries, during the Ming Dynasty (1368-1644), and that efforts to abolish such a practice started in 1665 by the Manchu government (Mackie 1996). However, footbinding persisted as a social norm for several centuries: at the beginning of the 20th century, the vast majority of women in Northern rural counties underwent footbinding (Mackie 1996). It was only in the 1910s that women in Northern and Central China started abandoning footbinding, with a complete change of the social norm occurring over a 30-year span. This famous example illustrates some key aspects of social norms, such as the value associated with their widespread adoption, their persistence, and their sudden change.



(a) (b)

Fig. 1: Examples of formation, persistence, and evolution of social norms and conventions. In panel (a), relative use of "sólo" (blue) and "solo" (orange) in written Spanish from Google Ngram data (Michel et al 2011). In panel (b), map of Europe illustrating the hand used for wedding ring in different countries: blue for vast majority of left hand, orange for vast majority of right hand, gray for regional/cultural differences within the country.

Spelling of "sólo" vs "solo" in Spanish

Other relevant examples of conventions come from the field of linguistic. In fact, the spelling of a specific word or even the meaning of a word do not have intrinsic value. Rather, their value is clearly associated with agreement on a specific spelling or interpretation within a community. An example often used in this literature is the spelling of the adverb "only" in Spanish. In fact, two spellings ("sólo" and "solo") have been used over the past two centuries (Amato et al 2018). At the beginning of the 19th century, the only spelling used was "solo." Then, according to Google Ngram data (Michel et al 2011), the use of "sólo" quickly became dominant at the end of the 19th century and remained preferred for almost a century, until the 2010s, when a swift change of convention reverted to the spelling "solo", potentially as it is more convenient on US keyboards. This evolution is illustrated in Fig. 1a.

Hand on which to wear a wedding ring

Another interesting example of convention change is the use of the wedding ring, which can be traced back to the Roman Empire or even earlier, to the Ancient Egyptians. While it is common in all Western countries to wear a ring (typically made of gold) on the fourth finger (called "ring finger" for this reason) to indicate that its wearer is married, there is not global agreement on which hand to wear it (Monger 2004). These differences are illustrated in Fig. 1b. Interestingly, the figure shows a pattern with high local conformity, but global diversity. In fact, in Europe, the vast majority of the population within a given country wears the wedding ring on the same hand. However, there is diversity between different countries, due to different cultural factors. There are two exceptions where there is no agreement within the same country, but even in these cases, there is strong local conformity. In the Netherlands, Protestants wear their wedding ring on the right hand, while Catholics on the left; in Spain, wedding rings are typically worn on the right, except for Catalonia and adjacent regions.

Key features of norms and conventions

The examples discussed in the previous section have highlighted some key universal features of social norms and conventions that one needs to keep in mind in order to develop mathematical modeling frameworks that can reproduce and predict their formation, persistence, and evolution.

First, as already discussed in the above, a key feature of social norms and conventions is the fact that their value depends on their widespread adoption and acceptance. Thus, norms and conventions are often enforced by an individual's tendency and desire to act in conformity with societal expectations (Marques and Paez 1994; Vega-Redondo 2007). In fact, there is no clear advantage in using the spelling "sólo" instead of "solo", besides conforming with others. Moreover, it is important to notice that some social norms can be characterized by *global diversity but strong local conformity*, as in the example of the use of the wedding ring, where the hand on which the ring is worn can be different in different communities/countries, but are locally consistent within the same community.

Second, in many of the examples discussed in the above, it is possible to observe the persistence of a social norm or convention over a long period. In fact, footbinding persisted as the norm in rural China for centuries, even when the government tried to ban it (Mackie 1996). However, a norm or convention cannot persist indefinitely for all time, for obvious reasons. From a modeling point of view, this can be associated with the mathematical concept of meta-stability. In fact, many real-world examples feature a replacement of the status quo norm or convention by an alternative (e.g., footbinding replaced by non-footbinding, or the spelling "solo" replaced by "sólo"). Interestingly, replacement often occurs over a time window that is substantially shorter when compared to the phase in which the status quo persisted, so that the fraction of adopters of the alternative convention typically follows an S-shaped curve, following the terminology of the existing literature (Rogers 2003) (see, e.g., Fig. 1a). Finally, it is important to also notice that this replacement itself will not remain as the new permanent convention, but rather, it becomes the status quo for a period before also likely being replaced by another alternative (or even by the original one, as in the spelling example), in a continuous process of social innovation that is key for the functioning and progression of our societies.

The third important feature, which is really at the core of this chapter, is that the collective adoption and evolution of social norms and conventions is typically the emergent behavior of a population, whereby individuals engage in repeated interactions over time, through which one individual can observe the behavior of others (e.g., by exchanging messages or e-mails, one becomes aware of how others spell a word; or by meeting, one can observe how others greet) and is influenced by such observations (Peyton Young 2015). The effect of these interactions is often complex and non-linear. It is known from many empirical studies that the pattern of social interactions is inherently complex, characterized by spatial and temporal heterogeneity, and can be conveniently captured by complex network structures (Boccaletti et al 2006). This observation explains the success of agent-based models within this field. In fact, in order to capture this important feature of social norms, one needs to have a mathematical description of the dynamics at the granularity of an individual. Building on this individual-level description, a mathematical model that describes how the individual behavior is affected by the interpersonal interactions which occur on a complex network is derived. In agent-based models, these interactions result in an emergent behavior at the population level that can be highly organized, counter-intuitive, and both robust and fragile to system shocks.

Mathematical modeling of norms and conventions

This section illustrates and discusses some of the main mathematical approaches used to model the formation and evolution of social norms and conventions. The development of these approaches began from the empirical observations of real-world instances of the formation and evolution of norms and conventions. From these observations, the mathematical models that are developed aim to encapsulate some of the key features of social norms that have been highlighted and extensively discussed in the previous section of this chapter. The models considered aim to balance (arriving at different points on the spectrum) between being sufficiently complex to capture the desired phenomena and sufficiently simple so as to be tractable (both analytically and numerically). The models also frame the problem in a similar manner: initially, there is widespread adoption within the population of one norm or convention, termed the *status quo*, and a new norm or convention is introduced into the population, termed the *innovation* or *alternative* (here, innovation refers to the fact that it is novel to the population, and not necessarily an entirely new norm).

This section is divided into two main parts. The first part focuses on discussing population models, where the formation and evolution of social conventions is described at the granularity of the entire population. In other words, the proposed models keep track of the temporal evolution of the population fraction adopting the innovation. The first and most famous model belonging to this class is the Bass model (Bass 1969), which is extensively presented and discussed in the sequel. Population models, however, have inherent limitations as they struggle at capturing the complex and heterogeneous pattern of interpersonal interactions and social influence that drive social change, as explained in detail below. For this reason, the second and larger part of this section is devoted to agent-based models. These reproduce the phenomenon of formation and evolution of social conventions as the emergent behavior of a population of individuals (agents), whose dynamics are described at the granularity of the individual (Bonabeau 2002). In particular, this section presents and discusses four classes of agent-based models that have been extensively used in the context of social norms and conventions: cascading models, evolutionary dynamic models, and game-theoretic decision-making models.

In many fields, one can often rely on established laws or principles within said field to derive the equations to describe the dynamics of the system of interest. For instance, the dynamical equations for electromagnetic circuits can be derived from Maxwell's equations, while the equations of motions of mechanical systems can be derived from Newton's laws of motion. Even in mathematical epidemiology, models of infectious disease spread are typically derived following mass conservation principles. In contrast, while scientists have established theories and observations (qualitative and quantitative) to describe human behavior, there are no laws or principles in the same sense as in electromagnetism or mechanics. Thus, many models of social dynamics involve formulating terms within the equations to conceptually represent social and behavioral mechanisms of relevance. For the models presented below, the details of the derivations will be provided for some of them to give context. For others, explicit derivations are omitted due to their complexity (giving instead references to articles or books that contain a complete mathematical derivation), but descriptions of the intuition underpinning the specific equations are provided.

Population models

Population models are the first class of mathematical models developed to study the adoption and evolution of conventions and norms. In this model class, the focus is in studying and reproducing the social phenomenon of interest with a mathematical description of the system at the population level. In other words, by defining a variable $z(t) \in [0,1]$ that measures the fraction of adopters of a specific norm or convention at time t. For consistency, in this chapter all models will be presented as discrete-time dynamics (i.e, $t \in \mathbb{N}_{\geq 0}$). However, one should keep in mind that there are continuous-time formulations for most of the models presented in the following. The population model establishes a mathematical rule that determines how $z(t)$ evolves in time, ultimately yielding a dynamical system that is analytically tractable due to its low dimensionality. Different mathematical approaches can used to define the evolution of the variable $z(t)$, including ordinary differential equations, recursive equations, and stochastic processes.

One of the first and most successful models developed within the framework of population model is the famous Bass diffusion model proposed by Frank Bass in Bass (1969). This model is grounded in the marketing and management science theory of innovation diffusion developed by Everett M. Rogers from the early 1960s (Rogers 2003), and was initially proposed within a marketing framework to study how new products are adopted by a population. Since then, it has been generalized and applied to many scenarios involving a population adopting an innovation, including a new social norm or convention; see, e.g., Mahajan et al (1990); Horvat et al (2020).

This model consists of a recursive equation that describes how $z(t)$ evolves in time. Specifically, the temporal evolution of the fraction of adopters is determined by the following nonlinear autonomous recursive equation:

$$z(t+1) = z(t) + p\big(1 - z(t)\big) + qz(t)\big(1 - z(t)\big), \tag{1}$$

where $p, q \in \mathbb{R}$ are two constant parameters (typically non-negative in healthy markets). In plain words, Eq. (1) establishes that the fraction of adopters of the innovation grows according to two distinct mechanisms. Besides current adopters captured by the first term, the second term on the right hand side of Eq. (1), $p(1 - z(t))$, accounts for new adoptions via innovation: a constant fraction of those who are not yet adopting the innovation will decide to adopt it. Hence, such a term is proportional to the fraction of non-adopters (i.e, $1 - z(t)$), and multiplied by the parameter p that represents the strength of the innovation process. The third term, $qz(t)(1 - z(t))$, instead, accounts for imitation: non-adopters who interact with adopters are then convinced to adopt at a given rate q. For this reason, the second term is non-linear, given by the product between adopters ($z(t)$) and non-adopters ($1 - z(t)$). Observe that the solution of the recursive equation in Eq. (1) is a logistic equation, which is characterized by a S-shaped curve, as illustrated in Fig. 2. It is worth noticing that the original formulation of the model was in continuous-time as a differential equation, while the formulation in Eq. (1) is instead based on Satoh (2001).

In his seminal paper, Frank Bass fitted his model by calibrating the parameters p and q via statistical regression, using real-world marketing data on sales of different innovative products in the decades before the publication of Bass (1969), such as home freezers and black-and-white televisions. Through such a fitting, he obtained respectable approximations of the empirically observed S-shaped adoption curves (Rogers 2003). Subsequent studies have extended the validation using different datasets from different application fields (Sultan et al 1990).

Despite its ability to reproduce real-world adoption curves, the Bass model is subject to some critical limitations. The fact that it provides a description of the phenomenon only at the population level prevents the model from encapsulating some of the features mentioned above, including the key role of interpersonal interactions, social influence, and heterogeneity across the population. These shortcomings limit the possibility to adopt the Bass model (and, more in general, population models) to study how the complex network of social interactions and the inherent complexity of the behavioral factors influencing individual-level decision-making mechanisms can affect the adoption curve. In particular, the shortcomings become especially noticeable when, taking on the perspective of a policymaker or authority, one wishes to apply individual-level interventions (e.g. monetary incentives) to facilitate widespread adoption of the innovation. Since the granularity is at the population level, such interventions cannot be easily incorporated. These limitations have thus led to the development of a new class of mathematical models that describe the social dynamics at the granularity of the individual.

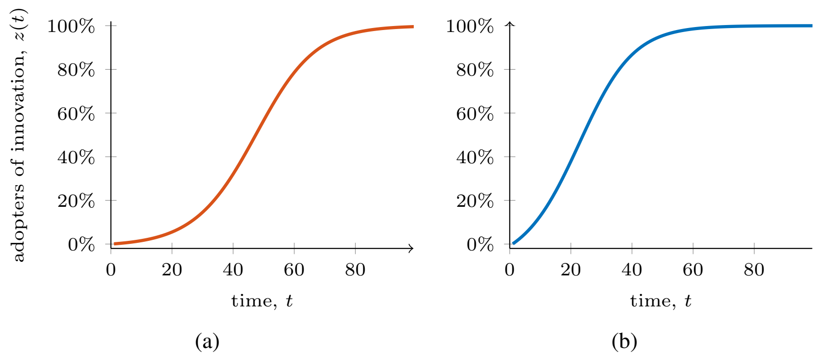


Fig. 2: Two exemplary trajectories of the adoption curve obtained with the discrete-time Bass model in Eq. (1) with (a) $p = 0.001$ and $q = 0.01$; and (b) $p = 0.001$ and $q = 0.1$. Both plots reproduce an S-shaped curve, with different initial slope and velocity to reach the inflection point.

Agent-based models

In the past few decades, agent-based models have emerged as a powerful approach to go beyond the limitations of population models, being able to reproduce the adoption process at the level of an individual. Such an approach allows one to study how individual-level mechanisms affect the population-level behavior, providing a novel array of tools to study the role of individual social and psychological factors in the formation and evolution of social norms and conventions. Importantly, such models are also better suited to studying the effects of intervention policies.

Agent-based models can differ in many aspects concerning the rationale and the mathematical implementation of the mechanisms that govern each individual's behavior. However, they all share key common traits. First, all agent-based models consider a population of n agents (individuals), which are denoted by integer indices, i.e., the set $\mathcal{V} := \{1,\dots,n\}$. Each individual $i \in \mathcal{V}$ is characterized by a state $x_i(t)$, which represents the choice of individual i with respect to the norm or convention considered in the model. This chapter considers the simplest scenario, involving the diffusion of a single innovation (alternative norm or convention) in a population, and thus the model posits that $x_i(t)$ is a binary variable, defined as

$$x_i(t) = \begin{cases} 1 \text{ if } i \text{ adopts the innovation at time } t, \\ 0 \text{ if } i \text{ does not adopt the innovation at time } t. \end{cases} \tag{2}$$

Further nuances can be introduced, e.g. by allowing $x_i(t)$ to take a value from a discrete set (if multiple options are possible), or assuming it to be continuous. Given the formulation in Eq. (2), the adoption curve can be observed as an emergent behavior of the dynamics of the agents. In fact, the total fraction of adopters is equal

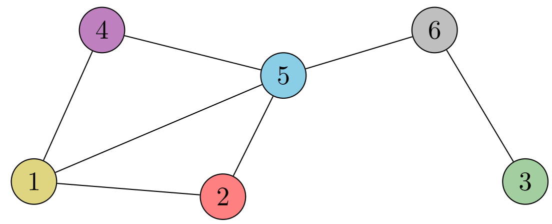


Fig. 3: Example of a network structure, with node set $\mathcal{V} = \{1,2,3,4,5,6\}$ and edge set $\mathcal{E} = \{(1,2),(1,4),(1,5),(2,5),(3,6),(4,5),(5,6)\}$. Note that node 5 (in blue) has three neighbors ($|\mathcal{N}_5| = 3$), namely $\mathcal{N}_5 = \{2,4,6\}$. Hence, its state $x_5(t)$ evolves as a function of its own current state, and the states of its three neighbors.

to

$$z(t) = \frac{1}{n}\sum_{i\in\mathcal{V}} x_i(t). \tag{3}$$

A second common feature of agent-based models is the fact that, as reflected in the real world, individuals interact with their peers on a complex network structure, exchanging information on their state and mutually influencing one another. Real-world social networks have been extensively investigated utilizing large datasets from online social networks (Fu et al 2008) and face-to-face experiments with wearable proximity sensors (Cattuto et al 2010). These studies have allowed researchers to derive a clear picture of their structure and key characteristics, including heterogeneity (Newman 2018). Based on these empirical observations, agent-based models typically represent the interacting individuals as nodes of a network $\mathcal{G} = (\mathcal{V},\mathcal{E})$, where the edge set $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is such that an edge $(i, j) \in \mathcal{E}$ if and only if i is influenced by j. If not differently stated, it is typically assumed that edges are bidirectional, i.e., the presence of edge (i, j) means that i is influenced by j and j is influenced by i. Hence, the set of neighbors of i, denoted by $\mathcal{N}_i(t) := \{j : (i, j) \in \mathcal{E}\}$ represents all agents who (directly) influence individual i.

Figure 3 provides a simple illustration of a network structure. Different agent-based models can assume different network structures, which can be time-invariant, or evolve in time, even in a state-dependent manner (Holme and Saramäki 2012). For the sake of simplicity, this chapter will focus on models on time-invariant networks, while briefly mentioning some extensions to time-varying frameworks. It should also be noted that the edges connecting agents are abstract representations of social influence; edges exist if one agent can influence the state of another agent in some way, and edges can represent a range of interactions such as friendships, communication in online social media platforms, or agents observing each other in the physical world.

Finally, a third common trait of agent-based models is the presence of agent dynamics, which determines how the agent state evolves in time to become $x_i(t+1)$, as a function of its current state $x_i(t)$ and the state of neighboring nodes of the network, i.e., $x_j(t)$ with $j \in \mathcal{N}_i$. The main differences between different agent-based models typically lie in the agent dynamics, i.e., on how agents revise their state on the basis of some internal dynamics and due to social influence from their peers on the network.

A first, intuitive, modeling approach can be to embed the Bass model described in the previous section onto a network structure. This is done by assuming that the imitation mechanism present in the Bass model is regulated by a network, whereby an individual tends to imitate only those adopters among their neighbors. This approach has many similarities with epidemic models, whereby contagion (by a disease or by an innovation) is driven by spontaneous mechanisms (spontaneous recovery or adoption of the innovation) and pairwise interactions (transmission of the disease via contagion or imitation), and has been extensively adopted in the literature to derive social diffusion models (Rizzo and Porfiri 2016; Bertotti et al 2016; Fibich 2016; Fagnani and Zino 2017).

However, these imitation-driven models rely on the simplifying assumption that social dynamics evolve similarly to epidemics, according to a so-called simple contagion mechanism, whereby each interaction with an adopter (infected individual) yields a certain rate or probability of adopting (being infected), and the effects of each interaction are independent. On the contrary, it is known that many social processes follow a complex contagion mechanism (Centola 2018). Here, multiple interactions with different neighbors are required to influence an individual, and interactions have, in general, a nonlinear impact on the dynamics. The rest of this section presents three important classes of agent-based models that differ in the agent dynamics but are all characterized by the presence of a complex contagion mechanism: namely, models based on threshold phenomena, evolutionary dynamics, and game theory.

Linear threshold models

One of the first and most successful approaches proposed to capture a complex contagion mechanism within a mathematical modeling framework is the linear threshold model, proposed by Mark Granovetter Granovetter (1978). This model was primarily proposed to reproduce the diffusion of an innovation (e.g., a novel convention) within a network of individuals. Hence, individuals have a binary state representing whether they have adopted the novel convention or not, as in Eq. (2), and each individual is associated with a parameter $\theta_i \in [0,1]$ that represents the resistance of individual i to change. In its simplest implementation, the linear threshold model is a discrete-time deterministic process. At each time-step, individual i decides to adopt the innovation if and only if more than a fraction θ_i of their neighbors have already adopted the innovation, i.e.,

$$x_i(t+1) = \begin{cases} 1 \text{ if } \frac{1}{d_i} \sum_{j \in \mathcal{N}_i} x_j(t) \geq \theta_i, \\ 0 \text{ otherwise}, \end{cases} \tag{4}$$

where $d_i := |\mathcal{N}_i|$ is the number of neighbors (degree) of i on the network.

In this model, it is typically assumed that one or multiple seeds are present and act as innovators (Rogers 2003), while all others agents are adopting the status quo. This can be easily modeled by setting the threshold of a seed node to $\theta_i = 0$, which means that they adopt the innovation, regardless of what their neighbors do. Then, the entire population is initialized to adopt the status quo ($x_i(t) = 0$, for all $i \in \mathcal{V}$), and it can be easily proved that the state is monotonically increasing (i.e., if an individual switches to adopt the innovation, then they never revert to the status quo). In this setting, initial seeds may trigger a cascade, where at each time-step one or multiple agents may decide to adopt the innovation, in a recursive manner, resulting in the innovation spreading through the network.

Researchers have studied the model in Eq. (4) on different network structures, in order to understand how social interactions facilitate or hinder triggering diffusion cascades and affect the size of the cascade. For instance, one can refer to Watts (2002) for an extensive study on random networks and to Rossi et al (2019) for a study of the impact of the degree distribution of the network. Moreover, several extensions have been proposed including, e.g., time-varying thresholds (Arditti et al 2024). One popular optimization problem studied using the linear threshold model is known as the influence maximization problem (Kempe et al 2003). In its simplest form, a policymaker has the objective to maximize $\lim_{t\to\infty} z(t)$, i.e., the number of individuals who end up adopting the innovation, by selecting the network locations of a fixed number of seed nodes. Readers may refer to Li et al (2018) for an introductory survey.

The linear threshold model and its extensions have been widely used to study innovation diffusion processes, largely thanks to their simplicity and intuitive implementation. However, they have some limitations. First, they typically do not allow individuals to drop the innovation and switch back to the status quo, which has been observed in many real-world examples. Second, in order to create a cascade leading to the diffusion of the innovation, Eq. (4) typically requires heterogeneous thresholds θ_i, which translates into a model with potentially many parameters, and thus difficult to be calibrated without the risk of overfitting. Third, despite being able to capture some features of complex contagion through the threshold mechanism, it is difficult to easily incorporate other behavioral mechanisms (besides imitation/social pressure) into the model, meaning key mechanisms that affect real-world decisions are omitted.

Evolutionary dynamics models

Another approach to model social norms and conventions is by means of evolutionary dynamics. Evolutionary dynamics were originally proposed in the context of biological and ecological systems, to study the competition between different species and the evolution of genetic traits in the same species. In plain words, evolutionary dynamics consider a network of nodes (often representing geographical locations), which can be occupied by individuals of different species, who can interact with

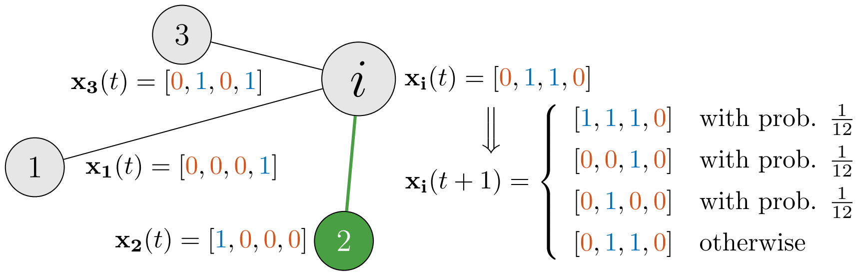


Fig. 4: Example of an iteration of the evolutionary dynamics in Axelrod (1997). Individual i is chosen and has three neighbors: 1 and 2 have a single entry of the state in common with i (first and last, respectively), while 3 has two entries in common (first and second). Assuming that neighbor 2 is chosen at random (in green), then, i imitates 2 by copying one of the three entries in which they differ (selected at random), with probability equal to the fraction of entries that they have in common (i.e., $\frac{1}{4}$), ultimately yielding the update rule written on the right of the figure.

those in neighboring nodes and replace them according to a fitness mechanism. For a simple and intuitive formulation of evolutionary dynamics, one can refer to Lieberman et al (2005). In the context of social dynamics, different species can represent different conventions, and a person will imitate the conventions of others after interacting with them. Here, fitness-based imitation is the key novelty with respect to simpler models (e.g., the Bass model, where imitation is simply due to an interaction), and it is able to encapsulate some aspects of the complex contagion process.

A seminal implementation of evolutionary dynamics is the model of the dissemination of culture, proposed in Axelrod (1997). In this model, each individual is characterized by a multi-dimensional state variable $\mathbf{x_i}(t) = [x_i^1(t), \dots, x_i^m(t)]$, where each of the m entries represents a cultural trait, which can be assumed to be binary $x_i^k(t) \in \{0,1\}$, for $k = 1,2,\dots,m$. Each trait represents an alternative convention on related topics (e.g., different spelling conventions for a word). At each time step, an individual $i \in \mathscr{V}$ is picked uniformly at random and this individual interacts with one of their neighbors $j \in \mathscr{N}_i$, selected at random.

When i interacts with j, i may imitate j, depending on how much the two individuals are similar. In particular, with probability proportional to the number of entries of the state on which the two individuals agree, i updates a random entry ℓ of their state $\mathbf{x_i}(t)$ on which i and j disagree (if any), by copying the corresponding entry of j, i.e., $x_i^\ell(t+1) = x_j^\ell(t)$. In other words, the more two individuals are culturally similar, the more likely one imitates the other after an interaction on a trait where there is no similarity. All other cultural traits remain unchanged. An explanatory iteration of this update rule is illustrated in Fig. 4.

In its original formulation in Axelrod (1997), the model was proposed and studied on regular lattices. Interestingly, numerical simulations for the original model often yielded the emergence of local consensus on the same convention, but global disagreement. In fact, multiple stable regions characterized by different conventions between different regions were often observed, capturing an important feature of social norms and conventions. Then, building on this seminal work, many efforts have been made to extend the use of evolutionary dynamics in the context of social norms and convention. This body of literature includes establishing analytical results for some specific implementations (Lanchier and Moisson 2015; Pedraza et al 2021), generalizing the model and its results to complex networks (Klemm et al 2003; Guerra et al 2010), proposing different implementations, e.g., tailored to the emergence of innovation in linguistic conventions (Baronchelli et al 2006), and designing control strategies to promote diffusion of innovation over status quo (Zino et al 2023a).

Game-theoretic models

The models described above capture some features of complex contagion, but often overlook an important aspect: in social interactions, individuals act strategically in a rational manner. That is, the adoption or otherwise of a certain innovation is the result of a decision-making process, through which an individual tends to achieve some notion of a payoff or reward. This is the underlying rationale beyond the development of mathematical models grounded in evolutionary game theory.

In a game-theoretic framework, each individual is seen as a player, who has two possible actions (0 and 1), corresponding to the state of the individual x_i, as in Eq. (2). Each player $i \in \mathcal{V}$ who has an interaction with a player j receives a payoff that depends on the state of individual i (x_i) and the state of individual j (x_j). This payoff can be encoded in a payoff matrix:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \tag{5}$$

with $a, b, c, d \in \mathbb{R}$ being constants. Namely, player i receives a payoff equal to a (or b) for selecting action $x_i = 0$ against an opponent j who plays $x_j = 0$ (or $x_j = 1$); and a payoff equal to c (or d) for selecting action $x_i = 1$ against an opponent who plays $x_j = 0$ (or $x_j = 1$) (von Neumann et al 1944). For network games, the total payoff that individual i receives for selecting action 0 (denoted by $u_i(0, \mathbf{x})$) or 1 (denoted by $u_i(1, \mathbf{x})$) is simply given by the average of all payoffs that i would receive from each of their neighbors (Jackson and Zenou 2015), i.e.,

$$u_i(0,\mathbf{x}) = \frac{1}{d_i}\Big(a|\{j \in \mathcal{N}_i : x_j = 0\}| + b|\{j \in \mathcal{N}_i : x_j = 1\}|\Big)$$
$$= \frac{1}{d_i}\sum_{j\in\mathcal{N}_i}\Big(a(1-x_j) + bx_j\Big), \tag{6a}$$
$$u_i(1,\mathbf{x}) = \frac{1}{d_i}\Big(c|\{j \in \mathcal{N}_i : x_j = 0\}| + d|\{j \in \mathcal{N}_i : x_j = 1\}|\Big)$$
$$= \frac{1}{d_i}\sum_{j\in\mathcal{N}_i}\Big(c(1-x_j) + dx_j\Big). \tag{6b}$$

Note that the definition $u_i(\cdot,\mathbf{x})$ highlights that the payoff is dependent on the action (state) of others in the network. Clearly, this approach can be extended, e.g., by associating weights to each edge of the network and transforming Eq. (6) into a weighted sum, so that each peer provides a different contribution to the payoff of i.

In the context of social norms and convention, coordination games are often used to capture the human tendency to conform. In fact, in the most basic implementation of coordination games, the payoff matrix A is the diagonal matrix

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1+\alpha \end{bmatrix}, \tag{7}$$

where the parameter $\alpha \in (-1,\infty)$ captures the relative advantage (if positive) or disadvantage (if negative) of the innovation with respect to the status quo. In other words, an individual receives a unit reward for coordinating with a peer on the status quo, and a reward $1+\alpha$ for coordinating on the innovation, as illustrated in Fig. 5. For the coordination game with payoff matrix in Eq. (7), the payoff functions in Eq. (6) reduce to the following expressions:

$$u_i(0,\mathbf{x}) = \frac{1}{d_i}\sum_{j\in\mathcal{N}_i}(1-x_j), \tag{8a}$$
$$u_i(1,\mathbf{x}) = \frac{1}{d_i}(1+\alpha)\sum_{j\in\mathcal{N}_i}x_j. \tag{8b}$$

Individuals revise their action (i.e., decide whether to adopt the status quo, 0, or the innovation, 1) with the aim of maximizing their payoff. In its simplest implementation, one can assume that individuals are fully rational and always decide to choose the action that gives the maximal payoff, given the action of the others, i.e., to adopt a best-response update rule:

$$x_i(t+1) = \text{argmax}_{s\in\{0,1\}} u_i(s,\mathbf{x}(t)). \tag{9}$$

In the case of a tie-break situation, one can presuppose a tie-breaker rule, e.g. to stick with the current action $x_i(t)$. The update rule in Eq. (9), ultimately induces a threshold-like dynamics as in Eq. (4), where an individual's threshold is determined by the parameters in the payoff matrix. For the coordination game in Eq. (7), notice that from Eq. (9), it follows that $x_i(t+1) = 1$ if and only if $\text{argmax}_{s\in\{0,1\}} u_i(s,\mathbf{x}(t)) =$

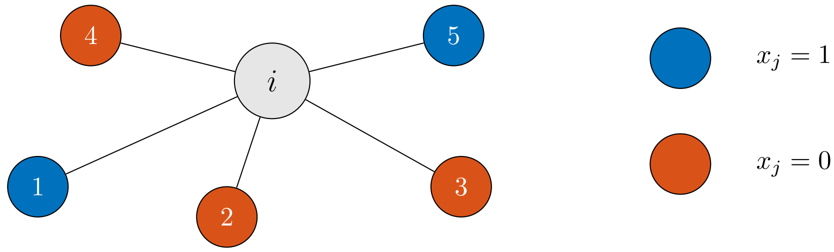


Fig. 5: Example of a network coordination game. Individual i has five neighbors: three of them are currently adopting the status quo (in orange), and two the innovation (blue). Hence, $u_i(0,\mathbf{x}) = 3$ and $u_i(1,\mathbf{x}) = 2+2\alpha$. Clearly, i would get a larger payoff for adopting the innovation if and only if $\alpha > \frac{1}{2}$.

1, i.e., if $u_i(1,\mathbf{x}) > u_i(0,\mathbf{x})$. By substituting the expressions from Eq. (8) into this condition, one obtains that $u_i(1,\mathbf{x}) > u_i(0,\mathbf{x})$ if and only if

$$\frac{1}{d_i}\sum_{j\in\mathcal{N}_i} x_j > \frac{1}{2+\alpha}. \tag{10}$$

Evidently, this yields Eq. (4) with threshold $\theta_i = \frac{1}{2+\alpha}$, as in Morris (2000), shaped by the relative advantage parameter α of the coordination game. However, note that in the linear threshold model, Eq. (4) does not permit agent i to switch from $x_i(t) = 1$ to $x_i(t+1) = 0$, unlike Eq. (9). Thus, the linear threshold model is more suited for some scenarios where one cannot revert to the status quo, such as installing solar panels on a home, whereas the game-theoretic model is relevant for norms where one can switch back and forth, e.g. the spelling of a word.

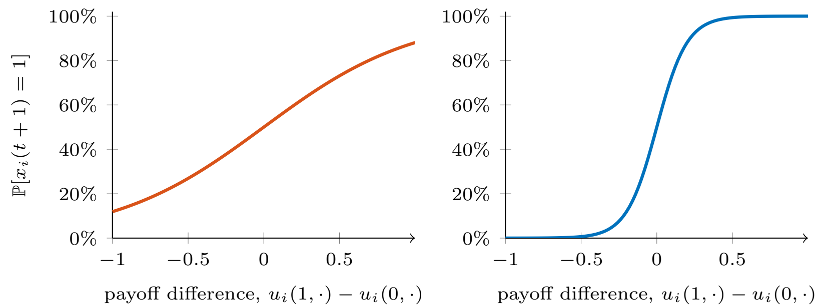
For the sake of completeness, it is worth noticing that the payoff matrix of the coordination game in Eq. (7) is quite general. In particular, for the best-response update rule in Eq. (9), the payoff matrix in Eq. (7) is representative of all payoff matrices with $a > c$ and $d > b$. In fact, from Eq. (9), one observes that $x_i(t+1) = 1$ if and only if $u_i(1,\mathbf{x}) > u_i(0,\mathbf{x})$. Using the general expression of the payoff function in Eq. (6), one obtains

$$\frac{1}{d_i}\sum_{j\in\mathcal{N}_i}\Big(c(1-x_j)+dx_j\Big) > \frac{1}{d_i}\sum_{j\in\mathcal{N}_i}\Big(a(1-x_j)+bx_j\Big), \tag{11}$$

which can be re-written as

$$c-a+(a-b-c+d)\frac{1}{d_i}\sum_{j\in\mathcal{N}_i} x_j > 0. \tag{12}$$

This in turn reduces to the following condition:



(a) small β_i (b) large β_i

Fig. 6: Probability of adopting the innovation $x_i(t+1)=1$ according to the log-linear dynamics in Eq. (14) as function of the payoff difference $u_i(1,\cdot)-u_i(0,\cdot)$, for (a) small and (b) large values of the rationality parameter β_i.

$$\frac{1}{d_i}\sum_{j\in\mathcal{N}_i} x_j > \frac{a-c}{a-c+d-b} = \frac{1}{1+\frac{d-b}{a-c}}. \tag{13}$$

Hence, any game with payoff matrix as given in Eq. (5), with $a>c$ and $d>b$, can be reduced to an equivalent coordination game of the form in Eq. (8), by setting the relative advantage equal to $\alpha=\frac{d-b}{a-c}-1$. For more details on this derivation, one can refer to Zino et al (2022).

More complex dynamics have been proposed for the state update, which allows to account for bounded rationality (Mäs and Nax 2016). In particular, substantial attention has been devoted to the study of a noisy version of the best-response update rule, where it is assumed that individuals revise their state in a probabilistic fashion, according to a log-linear learning rule (Blume 1995), also known as the Gibbs model, i.e.,

$$\mathbb{P}[x_i(t+1)=s] = \frac{\exp\{\beta_i u_i(s,\mathbf{x}(t))\}}{\exp\{\beta_i u_i(1,\mathbf{x}(t))\}+\exp\{\beta_i u_i(0,\mathbf{x}(t))\}}, \tag{14}$$

where $\beta_i \geq 0$ is a non-negative parameter that captures the rationality of individual i. In particular, when $\beta_i=0$, individual i will pick either 0 or 1 uniformly at random; whereas, in the limit $\beta_i\to\infty$, Eq. (14) reduces to Eq. (9), and i always selects the action that gives the maximal payoff. For intermediate values of the rationality parameter $\beta_i\in(0,\infty)$, Eq. (14) posits that individual i adopts the best-response action with a probability greater than $1/2$ but less than 1, depending on the difference between the payoff for adopting the best response action and the payoff for adopting the other action, as illustrated in Fig. 6. In particular, the larger β_i is, the larger the probability that i adopts the action with larger payoff. From another perspective, given a fixed β_i, the greater the difference between the payoff of the two actions, the greater the probability of selecting the best response action.

The role of the rationality parameter β_i, the relative advantage α, and the network structure have been extensively investigated for Eq. (14) in several papers, including Montanari and Saberi (2010); Peyton Young (2011), where the presence of clusters and cohesive sets has been highlighted as a key factor to favor the adoption of innovation. Alternatively, the diffusion of innovation can be tied to the presence of committed individuals who consistently choose the innovation ($x_i(t) = 1$ for all $t \geq 0$), including the number and location in the network (Ye et al 2021a; Gao et al 2023); more details are presented below. Similar to Axelrod's model presented above, it is also possible to observe local conformity (local in the sense of the network neighborhood) and global diversity, whereby different clusters in the network adopt 0 or 1 (Lambiotte et al 2007).

One of the main advantages of this game-theoretic formulation for not only modeling social norms and conventions but social dynamics more generally, is its flexibility. For instance, one can change the structure of the payoff matrix in Eq. (5) to capture other types of two-player, two-strategy games besides the coordination game, such as anti-coordination and prisoner's dilemma games, to model other types of social dynamics; see (Riehl et al 2018, Appendix C). One can also add terms to the function in Eq. (6) to account for further behavioral mechanisms besides conformity, such as the role of emerging trends (Zino et al 2022), existing beliefs or preferences (Peyton Young 2015), the presence of anti-conformist individuals (Ramazi et al 2016; Vanelli et al 2020), and social influence (Zino et al 2020; Aghbolagh et al 2023). Moreover, further sources of complexity can be incorporated into game-theoretic models by considering update rules in which agents have partial information on the game structure, such as mechanisms based on imitation (Como et al 2021; Govaert et al 2021), and by extending the set of possible actions to more than two (Dal Forno et al 2012).

Data integration

After having presented some of the main mathematical approaches used to model the formation and evolution of social norms and conventions, this section illustrates how data science has changed the picture around modeling of social dynamics and how different types of data have been integrated towards refining existing approaches and developing novel models validated by data. This section showcases three examples illustrating three different approaches to data-based mathematical modeling. The first example concerns the empirical validation of a threshold model, and a description of how data can be used to inform the thresholds for different types of individuals. The second example considers an evolutionary dynamics model tailored to a linguistic conventions problem, and focuses on the design and conduct of an experiment to validate model-based predictions on tipping points in social conventions. The third example takes a step further in the integration of data in the modeling phase, illustrating how experiments can be designed in order to understand the key terms that should be added to a coordination game in order to faithfully model

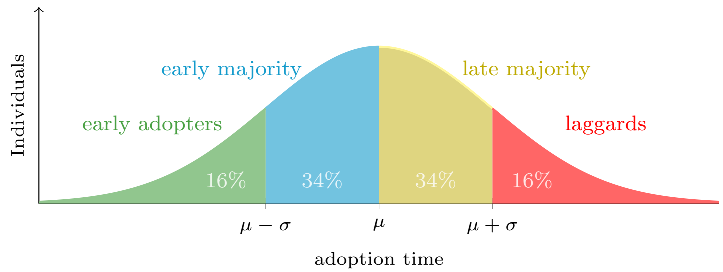


Fig. 7: Categories of individuals in the classical diffusion innovation theory (Rogers 2003), where μ and σ are mean and standard deviation, respectively.

the formation of a consensus on a convention both at the individual level and at the population level.

Data-based determination of thresholds for innovation diffusion

In his seminal paper Valente (1996), Thomas W. Valente used available data on the diffusion of innovation to validate a threshold-like model, demonstrating how population-level models can fail in capturing important features of the adoption process. To this aim, he built on the categories of early adopters, early majority, late majority, and laggards, extensively used in the theory of diffusion of innovation (Ryan and Gross 1943; Rogers 2003). In the classical innovation diffusion theory it is assumed that the distribution of the time to adoption across the population is a Gaussian distribution, which ultimately yields the famous S-shaped adoption curve. Under this assumption, individuals involved in the adoption process can be divided into four categories:

1. early adopters ($\approx 16\%$), who adopt the innovation earlier than one standard deviation before the average time of adoption;
2. early majority ($\approx 34\%$), who adopt the innovation after the early adopters, but before the average time of adoption;
3. late majority ($\approx 34\%$), who adopt the innovation after the early majority, but one standard deviation before the average;
4. laggards ($\approx 16\%$), who adopt the innovation at least one standard deviation after the average time of adoption;

as illustrated in Fig. 7. For the sake of completeness, note that in some more recent works (Rogers 2003), a fifth category of “innovators” is sometimes used to refer to early adopters who adopt the innovation earlier than one standard deviation before the average time of adoption.

However, these categories refer to the timing at which an individual adopts the innovation with respect to the whole population. In his work, Valente proposed to focus on the proportion of adopters among the neighbors of a selected individual at the time said individual adopts the innovation, which is term exposure at adoption, which is nothing but an empirical estimate of the threshold in a linear threshold model. Similar to time of adoption, individuals can be divided into four categories depending on their exposure at adoption, namely, individuals with

1. very low threshold ($\approx 16\%$), who have exposure at adoption one standard deviation smaller than the average;
2. low threshold ($\approx 34\%$), who have exposure at adoption greater than those with very low threshold, but smaller than the average;
3. high threshold ($\approx 34\%$), who have exposure at adoption greater than the average but smaller than one standard deviation greater than the average;
4. very high threshold ($\approx 16\%$), who have exposure at adoption one standard deviation greater than the average.

In Valente (1996), the author considered three datasets involving the adoption of an innovation, for which not only population-level data and individual-level data were available, but also data on social interactions between individuals is present, in order to reconstruct the underlying social network. Specifically, the three datasets concerned i) the prescription of a new drug (tetracycline) by physicians in Illinois in the 1950s, 2) the adoption of hybrid corn in Brazilian villages in 1966, and 3) family planning methods for married women in Korea in 1973. From the analysis of all three datasets, Valente found a highly statistically significant correlation between the two categories mentioned above, used to distinguish individuals. That is, early adopters are more likely to have a very low threshold of exposure at adoption, early majority individuals have a low threshold, et cetera. However, on average, only slightly more than half of the individuals were observed to be in the very same category in both classifications. For instance, an individual may belong to the laggard category either because they have a very large threshold, or because they had a low threshold but due to the way in which the innovation spread through the network, they did not receive any exposure until much later. In conclusion, classical categories based on population-level adoption may fail in capturing important aspects of individual-level adoption processes, especially when nontrivial network structures govern the interaction processes.

The results of this analysis confirm that population models (e.g., the Bass model and its generalizations), despite being able to reproduce the observed S-shaped adoption curves, may neglect important aspects related to how the structure of the social network actually shapes the adoption process. This limits their practical prediction ability, providing strong motivation for the use of agent-based models, which are instead able to describe the process at the individual-level, and thus incorporate the impact of the social network.

Experimental determination of tipping points in social conventions

A problem of paramount importance in the context of social norms and conventions is to understand whether a sufficiently large committed minority can change the societal norm. Committed minority are individuals who promote the innovation actively and consistently when the innovation is still new. Empirical evidence (Kuran 1995), supported by mathematical models (Xie et al 2011), suggested that there is a critical mass that determines a tipping point: a committed minority smaller than this critical mass cannot overturn an existing status quo, while a committed minority greater than this critical mass is sufficient to trigger social change and the collective adoption of a novel social convention. Several empirical and model-based studies have focused on determining this critical mass, with results spanning from 3.5% to 40% (Chenoweth and Stephan 2011; Grey 2006).

In Centola et al (2018), the authors aimed to design an experiment in order to empirically validate the results obtained via mathematical models. Specifically, they considered a problem related to linguistics, where groups of people have to achieve a consensus on the name of one or multiple objects. For this classical linguistic problem, a mathematical model inspired by evolutionary dynamics has been proposed termed the Naming Game model (Baronchelli et al 2006). Briefly, each individual i is characterized by a state $x_i(t)$ that is called the inventory, consisting of word-object pairs. Initially, all inventories are empty. At each time step $t \in \mathbb{N}_{\geq 0}$, the following procedure is iterated:

1) A pair of (neighbors) individuals i, j are picked at random and one of them plays as speaker (say i) and the other as listener;
2) The speaker selects randomly an object and retrieves a word from the inventory $x_i(t)$ associated with the object, or, if the inventory is empty, invents a new word;

3a) If the listener has the word-object pair named by the speaker in the inventory $x_j(t)$, both individuals maintain in their inventories at time $t+1$ only the winning word (i.e, the one word presented by the speaker), deleting all other words associated with the same object;

3b) If the listener does not have the word presented by the speaker in the inventory, the listener updates the inventory at time $t+1$ by adding the word presented by the speaker.

4) The time-step is incremented to $t+1$ and the process resumes from item 1).

Centola et al (2018) proposed the following experimental setting to validate this model. They recruited 194 participants to play an online repeated game, divided into 10 different groups. In the first stage, in each group, participants were randomly coupled and they were shown a picture of a face and are asked to insert the name of that face. If the players entered the same name, they were given a monetary reward, otherwise they were penalized; this procedure was iterated, until the whole population reached a consensus on the name of that face. Then, in a second stage of the experiment, in each of the 10 groups, a different fraction of computer bots were introduced (from 15% to 35%), who acted as a committed minority, supporting a novel name instead of the established status quo. Then, the same procedure of random matches (including also bots) and monetary rewards/punishments were iterated.

The results of this experiment confirmed the model-based predictions from Baronchelli et al (2006). In fact, calibrating the parameters of the model (namely, number of objects, population size, and number of iterations) to the experimental setting, one obtained that the critical mass was approximately 25%. In Centola et al (2018), all groups with less than 25% bots showed only a very small minority of participants would eventually switch to the novel name. On the contrary, groups with at least 25% bots showed 72% uptake of the new name. Interestingly, the experiments provide empirical support to the use of agent-based models to derive not only qualitative observations concerning formation and evolution of social norms and conventions, but also in determining precise quantitative predictions.

Key factors in decision-making via online experiments

In mathematical models built on coordination games, as described above, individuals tend to make decisions on which action to take among two (or more) conventions by maximizing a payoff function that depends solely on how many of their neighbors adopt a certain convention, as per Eqs. (6–7). However, while social coordination and the tendency to conform with others are key factors in human decision-making concerning social norms and conventions (Marques and Paez 1994), social scientists have observed and theorized the presence of other behavioral factors that play a role in individual-level decision-making. The social and behavioral science community agree on the presence of two important factors, namely the presence of inertia (sometimes referred to as status-quo bias) and the sensitivity to trends. Concerning the first aspect, individuals often prefer to stick to their current decisions and have some inherent resistance to change even if there are clear benefits to switching (Samuelson and Zeckhauser 1988). The second aspect, instead, is related to the fact that people tend to follow the trends observed in the population by being more attracted behaviors or products that are becoming more popular, even if it is currently only adopted by a minority (Sparkman and Walton 2017). These two factors are well known in the social psychology literature, and many researchers have empirically validated their presence in individual-level decision making in different contexts. However, they are not present in the mechanism of coordination game models, potentially limiting the ability of such models to faithfully reproduce real-world phenomena.

In Ye et al (2021a), the authors built on these social psychology observations and theories to design an experimental paradigm to investigate whether inertia and sensitivity to trends have downstream consequences on the emergent behavior of the population concerning the formation and evolution of conventions, and thus to determine whether it is important to incorporate these factors in a game-theoretic model. In the experiments, participants were enrolled online and divided into small groups (each group comprised of 12 agents: 8–10 human participants and 2–4 com-

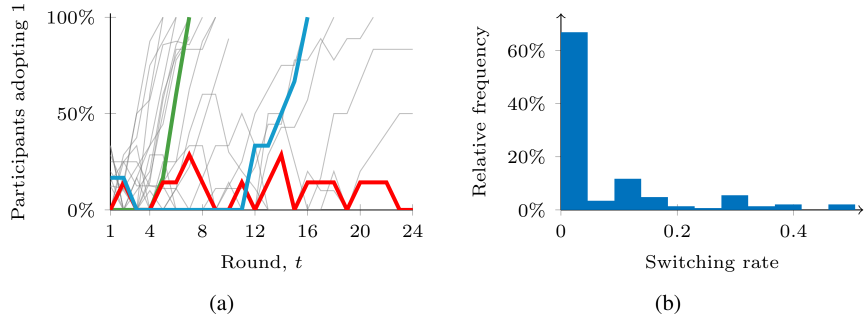


Fig. 8: Results of the experiment to determine key factors in decision-making concerning the adoption of conventions from Ye et al (2021a). Panel (a) illustrates different adoption curves for the innovation, illustrating scenarios of fast (green) and delayed (blue) adoption, and no adoption (red). Panel (b) shows the switching rate, which is used in Ye et al (2021a) to calibrate the model.

puter bots). In each group, participants made repeated decisions, choosing between two options (say, 0 or 1), with the final goal of reaching a consensus. At each round, each participant was able to see the proportion of the rest of the group that chose each option in the previous round. Computer bots were added to initially enforce a majority on a status quo (say 0), and then flipped to try to promote the adoption of the innovation (say 1) by selecting the other option. The results of the experiment are summarized in Fig. 8. From a statistical analysis of the experimental results conducted based on individual-level data, the authors concluded that coordination, inertia, and sensitivity to trends were all present in the individual-level decision making mechanism. Specifically, the state of an individual i at time $t+1$ was significantly influenced by i) the state of others at time t, ii) their own current state $x_i(t)$, and iii) whether the number of adopters of the alternative had increased or decreased over the previous iteration of the game. For more details on the experimental paradigm, see Mlakar et al (2024).

Building on this experimental evidence, the authors in Ye et al (2021a) conjectured that individuals revise their state according to the log-linear learning rule in Eq. (14), with payoff functions that extend those of a coordination game in Eq. (8) to account for inertia and sensitivity to trends. In particular, it was proposed that:

$$u_i(1,\mathbf{x}(t)) = \frac{b_i}{d_i}\sum_{j\in\mathcal{N}_i} x_j(t) + k_i x_i(t) + r_i \hat{x}_i(t)\,, \tag{15a}$$

$$u_i(0,\mathbf{x}(t)) = \frac{b_i}{d_i}\sum_{j\in\mathcal{N}_i} \big(1-x_j(t)\big) + k_i\big(1-x_i(t)\big) + r_i\big(1-\hat{x}_i(t)\big)\,, \tag{15b}$$

where

$$\hat{x}_i(t) = \frac{1}{2}\Big(1 + \frac{1}{n-1}\sum_{j\in\mathcal{V}\setminus\{v\}}\big(x_j(t) - x_j(t-1)\big)\Big) \tag{16}$$

captures the trend, being $\hat{x}_i(t) > \frac{1}{2}$ if the fraction of adopters of action 1 has increased in the previous time-step and $\hat{x}_i(t) < \frac{1}{2}$ otherwise. The payoff in Eq. (15) expands Eq. (8) by adding two terms. In fact, the first term captures coordination with no relative advantage for the innovation, and coincides with the standard term from Eq. (8) with $\alpha = 0$. The second term ($k_i x_i$ and $k_i(1-x_i)$, respectively) increases the payoff for keeping the current action. The third term ($k_i x_i$ and $k_i(1-x_i)$, respectively) increases the payoff for selecting the action whose support has increased in the previous time-step.

The three parameters b_i, k_i, r_i are non-negative scalar constants, weighting the contribution of the three terms, and can be assumed to sastify $b_i + k_i + r_i = 1$ without loss of generality. In Ye et al (2021a), these parameters were calibrated by fitting the individual-level experimental data illustrated in Fig. 8b, assuming that there are two classes of individuals, and with individuals of the same class having the same parameters to avoid overfitting. In other words, by exploring a grid on the parameter space, one could determined the configuration that minimized the discrepancy between the number of times individuals change their action (from 0 to 1 or vice versa) in the simulations and the same quantity recorded from experimental data. This calibration demonstrated that, in the absence of the two additional terms (i.e, by enforcing $k_i = 0$ and/or $r_i = 0$), a model calibrated with individual-level data fails in capturing the range of population-level emergent behavior observed in the experiment (see Fig. 8a). On the contrary, when all three terms are included into Eq. (15), the model is able to faithfully reproduce the experimental data at both the individual- and population-level.

In summary, this work describes how experiments can be designed in order to improve existing mathematical models, going beyond a simple calibration of a model to actually informing the very mathematical formulation of the model dynamics and, ultimately, validating its improvement with respect to existing models. This procedure allows to adopt existing modeling paradigms and refine them using experimental data to ultimately derive improved models, which can be used to explore the social phenomena of interest beyond the experimental limitations. An example of this can be found, e.g., in Gao et al (2023), where the experimentally-validated model from Ye et al (2021a) is used to explore the impact of the network structure on the evolution of conventions.

Summary, conclusions, and vision

This chapter has provided an overview of the mathematical modeling of social systems, applied directly to the formation, persistence, and evolution of social norms and conventions. After reviewing the basic notions and features of norms and conventions, several different frameworks were presented. Then, the integration of data into these models was discussed, using examples from the literature. These data ranged from those obtained in field surveys to controlled psychology experiments. In this last section, the key challenges of data-driven mathematical models of social systems are discussed, and a future vision is proposed based on potential opportunities. While the following commentary is centered around modeling of social norms and conventions, it largely extends to broader problems, such as modeling of opinion and belief formation, dis/misinformation propagation, and collective action. Addressing these challenges will not only require the advancement of mathematical modeling and analysis techniques, but crucially, close collaboration with colleagues in the theoretical and experimental social and behavioral sciences.

Despite several decades of sustained development involving efforts of the scientific community spanning multiple disciplines, when it comes to direct applications, mathematical modeling of social systems still lags behind other fields such as ecology and epidemiology. In fact, while the COVID-19 pandemic created global disruption and tragically resulted in millions of deaths (World Health Organization 2024), it was also a crowning moment for the real-world application of mathematical models of epidemic spreading (Vespignani et al 2020; Estrada 2020). A range of different models, from agent-based models to population models, were used to predict COVID-19 spread, and crucially, inform government and help plan public policy and medical interventions (Giordano et al 2020; Della Rossa et al 2020; Zhao and Chen 2020; Casella 2021; Parino et al 2021; Truszkowska et al 2021). A similar systematic and widespread adoption of mathematical models in the context of social systems is still far from our daily life, due to the presence of several challenges that need to be addressed. In the following, two challenges that are especially associated with modeling social systems are discussed.

Challenges

One challenge, perhaps unique to modeling of human behavior (as opposed to e.g., modeling of infectious diseases or predator-prey relations in ecology) is that people are by nature highly context-dependent. For instance, the decision-making processes for collective action and protest movements (whether to participate and with what tactics) share some commonalities with norms and conventions (which norm to adopt), such as the fact that peer pressure and internal preferences can play a key role. However, there are also fundamental differences; a person's willingness to protest is closely tied to how strongly they (internally) identify with the protest movement and group, as well as interactions with the authority (Louis et al 2022, 2020), while what matters most for norms is to coordination with others and conform to the majority (Lewis 2002; Bicchieri 2005). Thus, one may require a different framework and starting point for modeling protest movements (Thomas et al 2024) as opposed to norms and conventions. Separately, the fact the people learn through social interactions and adapt over time is well-known (Fay et al 2010), and thus may also require modeling (Acemoglu et al 2011). These considerations can create a difficult balance between a mathematician's natural instinct to develop models and methods that can generalize across multiple application domains and contexts, and the established view in social and behavioral science that people are highly context-dependent. Finding this balance, and being able to identify the level of generalizability of various models, will contribute greatly towards the real-world applicability of mathematical models of social systems. It will require concerted efforts in collaboration with other disciplines, and close scrutiny of how various decision-making and behavioral processes are modeled, and to what degree there is alignment with theoretical foundations and empirical evidence from the behavioral sciences.

Another limit to the use of real-world data to inform mathematical models of social dynamics is the limited possibility to access high-quality data of individuals' behaviors. In fact, for many other domains of applications of mathematical models (e.g., physical systems or ecological systems), many tools have been developed for obtaining accurate and reliable estimations of the quantities involved in the dynamics, e.g., physical quantities or amount of individuals for each species in a geographical region (Tredennick et al 2021). For social dynamics, while it is indeed true that the ubiquitous nature of social media platforms and the development of data science techniques have allowed to process huge amounts of data, high quality data are often difficult to obtain. In fact, data is often available at the population level (e.g., the adoption curve), but rarely available at the individual-level (e.g., the pattern of adoption in a social network). Moreover, many datasets are owned by private companies, making it difficult for researchers to access. In addition, data collected via experiments are strongly dependent on the specific experimental setting, so are difficult to re-use in other contexts or generalize. Similar issues related to the data quality were present in epidemic modeling before 2020. However, the global call for standardized and high-quality epidemiological data for COVID-19 has enabled giant leaps in the field of mathematical epidemiology (Gardner et al 2021). A key challenge for the future research in mathematical models of social systems is the development of common guidelines to create repositories of high-quality data of social dynamics, which will be used to improve and inform mathematical models.

Opportunities

So far, there has been much focus on using experimental data to inform, refine, and validate models. One obvious opportunity is to explore the converse, viz. using calibrated mathematical models to explore and design different interventions, which can then be tested using real-world experiments. Identifying ways to create "social tipping points" that lead to mass change in norms and conventions has become of increasing interest, seen as a way to address pressing societal issues ranging from the climate crisis to pandemic response (Otto et al 2020; Bak-Coleman et al 2021; Mlakar et al 2024). There exist many possible levers for interventions, and it is impractical and costly to empirically explore all of these (e.g. with experiments or field tests) to identify the most effective and relevant intervention, especially since interventions can be context-dependent (see above for a separate discussion on this issue). Moreover, many interventions at the individual level are based on social and cognitive psychology methods, and the effect size can be small even if statistically significant; it is difficult to use experiments to establish whether these small effects at the individual level can lead to population level tipping points. Having validated and calibrated mathematical models would allow one to systematically and rapidly (at least relative to planning and running a field experiment) examine an array of interventions, including those that vary in time (Walsh et al 2023; Zino et al 2023b). In the modeling framework, one can identify interventions most effective at generating tipping points, including an "optimal" intervention by associating each intervention with a cost. These insights could then inform which interventions to test in an empirical setting, and how. Employing mathematical models in this manner is slowly becoming accepted in the social and behavioral sciences community, and such an approach is a major opportunity for mathematicians to contribute in the future.

Another opportunity for future developments in the mathematical modeling of social conventions and norms lies in the explosive growth of machine learning techniques and artificial intelligence. In particular, developments in machine learning have paved the way for the use of mathematical paradigms such as that of universal differential equations (Rubel 1981), which augment model-based approaches with machine-learnable structures within the general context of physics-informed machine learning (Karniadakis et al 2021). Briefly, universal differential equations consists of parameterizable systems of equations that do not dictate the precise evolution of a dynamical system, but rather determine certain conditions that any possible evolution must fulfill (e.g., enforcing the emergence of a S-shaped curve in social diffusion). The key aspect is that these systems of equations can be augmented and trained with deep learning techniques, as proposed in (Rackauckas et al 2020; Koch et al 2023). Recently, some research groups have started using these machine learning techniques in the context of developing models of human behavior with applications in epidemiology (Kuwahara and Bauch 2024). These new methods may pave the way for a promising direction of future exploration, toward a systematic methodology that merges theoretically-informed and data-driven approaches to develop effective mathematical models for social dynamics.

Finally, it is worth noticing that social systems are increasingly integrated within complex cyber-physical systems (Annaswamy and Yildiz 2021). Hence, also the formation, persistence, and evolution of social norms and conventions are phenomena that are deeply intertwined with physical and technical dynamics. For instance, a paradigm shift towards increasing the use of public transportation is needed to reduce the carbon footprint of our cities. However, human decisions concerning the use of public transportation are clearly dependent on the efficiency of the traffic network, which in turn is affected by the number and type of vehicles present in the system and, in turn, human decisions, thus completing the feedback loop (Hull 2008). In a different application field, the virality of new trends, memes and products on a social media platform are increasingly guided by recommender systems, which are algorithms that filter the information and content an individual is presented with (Jannach et al 2010). In turn, the current decisions of individuals to engage with the content (watching, liking, sharing) is considered by the recommender system in terms of future content to present to individuals. Cyber-physical-human systems have emerged as a novel and promising paradigm that integrates a model of human behavior with a model of a cyber-physical system (Annaswamy and Yildiz 2021), offering a unified approach to represent and study the complex socio-technical systems that are becoming pervasive in our daily life. Given their recent emergence, the exploration of data-based modeling approaches for human behavior within the context of cyber-physical-human systems, building on the techniques discussed in this chapter, is still a research direction mostly unexplored and highly promising.

Cross-References

The following chapters could potentially be cross-referenced, depending on their precise content and clarification on how the cross-referencing works.

- Data-Driven Approach to Analysis of SIR-Type Models: The Principle of Parsimony Applied to Epidemics Modeling in the Age of COVID-19
- MONOPOLI: A Customizable Model for Forecasting COVID-19 Around the World Using Alternative Nonpharmaceutical Intervention Policy Scenarios, Human Movement Data, and Regional Demographics
- Markov Decision Processes and Best Responses of Epidemic Models
- Retrospective Modeling of the Impact of Vaccination and Non-pharmaceutical Interventions on the COVID-19 Pandemic: A Case Study of Mexico
- Quantitative Approaches in the Life Sciences
- Preparing for the Next Pandemic: Learning Lessons From the Recent Past
- Potential Games in the Economics and Management of Pollution

References

Abelson RP (1967) Mathematical Models in Social Psychology, Elsevier, p 1–54. DOI 10.1016/s0065-2601(08)60341-x

Acemoglu D, Dahleh MA, Lobel I, Ozdaglar A (2011) Bayesian learning in social networks. The Review of Economic Studies 78(4):1201–1236, DOI 10.1093/restud/rdr004

Aghbolagh HD, Ye M, Zino L, Chen Z, Cao M (2023) Coevolutionary dynamics of actions and opinions in social networks. IEEE Transactions on Automatic Control 68(12):7708–7723, DOI 10.1109/tac.2023.3290771

Amato R, Lacasa L, Díaz-Guilera A, Baronchelli A (2018) The dynamics of norm change in the cultural evolution of language. Proceedings of the National Academy of Sciences 115(33):8260–8265, DOI 10.1073/pnas.1721059115

Anderson BDO, Ye M (2019) Recent advances in the modelling and analysis of opinion dynamics on influence networks. Int J Autom Comput 16(2):129–149, DOI 10.1007/s11633-019-1169-8

Annaswamy AM, Yildiz Y (2021) Cyber-physical-human systems. In: Encyclopedia of Systems and Control, Springer, DOI 10.1007/978-3-030-44184-5_100113

Arditti L, Como G, Fagnani F, Vanelli M (2024) Robust coordination of linear threshold dynamics on directed weighted networks. IEEE Transactions on Automatic Control 69(10):6515–6529, DOI 10.1109/tac.2024.3371882

Axelrod R (1997) The dissemination of culture: A model with local convergence and global polarization. The Journal of Conflict Resolution 41(2):203–226, DOI 10.1177/0022002797041002001

Bak-Coleman JB, Alfano M, Barfuss W, Bergstrom CT, Centeno MA, Couzin ID, Donges JF, Galesic M, Gersick AS, Jacquet J, Kao AB, Moran RE, Romanczuk P, Rubenstein DI, Tombak KJ, Van Bavel JJ, Weber EU (2021) Stewardship of global collective behavior. Proceedings of the National Academy of Sciences 118(27), DOI 10.1073/pnas.2025764118

Baronchelli A, Felici M, Loreto V, Caglioti E, Steels L (2006) Sharp transition towards shared vocabularies in multi-agent systems. Journal of Statistical Mechanics: Theory and Experiment 2006(06):P06,014–P06,014, DOI 10.1088/1742-5468/2006/06/p06014

Bass FM (1969) A new product growth for model consumer durables. Management Science 15(5):215–227, DOI 10.1287/mnsc.15.5.215

Becker J, Brackbill D, Centola D (2017) Network dynamics of social influence in the wisdom of crowds. Proceedings of the National Academy of Sciences 114(26), DOI 10.1073/pnas.1615978114

Bernardo C, Wang L, Vasca F, Hong Y, Shi G, Altafini C (2021) Achieving consensus in multilateral international negotiations: The case study of the 2015 Paris Agreement on climate change. Science Advances 7(51), DOI 10.1126/sciadv.abg8068

Bertotti M, Brunner J, Modanese G (2016) The bass diffusion model on networks with correlations and inhomogeneous advertising. Chaos, Solitons & Fractals 90:55–63, DOI 10.1016/j.chaos.2016.02.039

Bicchieri C (2005) The Grammar of Society: The Nature and Dynamics of Social Norms, 1st edn. Cambridge University Press, Cambridge, UK

Bicchieri C (2014) Philosophy of social science: A new introduction, Oxford University Press Oxford, chap Norms, Conventions, and the Power of Expectations, pp 208–229

Blume L (1995) The statistical mechanics of best-response strategy revision. Games and Economic Behavior 11(2):111–145, DOI 10.1006/game.1995.1046

Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang D (2006) Complex networks: Structure and dynamics. Physics Reports 424(4–5):175–308, DOI 10.1016/j.physrep.2005.10.009

Bonabeau E (2002) Agent-based modeling: Methods and techniques for simulating human systems. Proceedings of the National Academy of Sciences 99(suppl 3):7280–7287, DOI 10.1073/pnas.082080899

Brown MJ, Satterthwaite-Phillips D (2018) Economic correlates of footbinding: Implications for the importance of Chinese daughters' labor. PloS One 13(9):e0201,337, DOI 10.1371/journal.pone.0201337

Casella F (2021) Can the COVID-19 epidemic be controlled on the basis of daily test reports? IEEE Control Systems Letters 5(3):1079–1084, DOI 10.1109/LCSYS.2020.3009912

Castellano C, Fortunato S, Loreto V (2009) Statistical physics of social dynamics. Reviews of Modern Physics 81(2):591–646, DOI 10.1103/revmodphys.81.591

Cattuto C, Van den Broeck W, Barrat A, Colizza V, Pinton J, Vespignani A (2010) Dynamics of person-to-person interactions from distributed RFID sensor networks. PLOS ONE 5(7):e11,596, DOI 10.1371/journal.pone.0011596

Centola D (2018) How Behavior Spreads: The Science of Complex Contagions . Princeton university Press, Princeton NJ, US

Centola D, Becker J, Brackbill D, Baronchelli A (2018) Experimental evidence for tipping points in social convention. Science 360(6393):1116–1119, DOI 10.1126/science.aas8827

Chandrasekhar AG, Larreguy H, Xandri JP (2020) Testing models of social learning on networks: Evidence from two experiments. Econometrica 88(1):1–32, DOI 10.3982/ecta14407

Chenoweth E, Stephan MJ (2011) Why Civil Resistance Works: The Strategic Logic of Nonviolent Conflict. Columbia University Press

Cinelli M, Quattrociocchi W, Galeazzi A, Valensise CM, Brugnoli E, Schmidt AL, Zola P, Zollo F, Scala A (2020) The covid-19 social media infodemic. Scientific Reports 10(1), DOI 10.1038/s41598-020-73510-5

Como G, Fagnani F, Zino L (2021) Imitation dynamics in population games on community networks. IEEE Transactions on Control of Network Systems 8(1):65–76, DOI 10.1109/tcns.2020.3032873

Dal Forno A, Gardini L, Merlone U (2012) Ternary choices in repeated games and border collision bifurcations. Chaos, Solitons & Fractals 45(3):294–305, DOI 10.1016/j.chaos.2011.12.003

DeGroot MH (1974) Reaching a Consensus. Journal of the American Statistical Association 69(345):118–121, DOI 10.2307/2285509

Del Vicario M, Bessi A, Zollo F, Petroni F, Scala A, Caldarelli G, Stanley HE, Quattrociocchi W (2016) The spreading of misinformation online. Proceedings of the National Academy of Sciences 113(3):554–559, DOI 10.1073/pnas.1517441113

Della Rossa F, Salzano D, Di Meglio A, De Lellis F, Coraggio M, Calabrese C, Guarino A, Cardona-Rivera R, De Lellis P, Liuzza D, Lo Iudice F, Russo G, di Bernardo M (2020) A network model of italy shows that intermittent regional strategies can alleviate the covid-19 epidemic. Nature Communications 11(1), DOI 10.1038/s41467-020-18827-5

D'Orsogna MR, Perc M (2015) Statistical physics of crime: A review. Physics of Life Reviews 12:1–21, DOI 10.1016/j.plrev.2014.11.001

Easley DA, Kleinberg JM (2010) Networks, Crowds, and Markets - Reasoning About a Highly Connected World. Cambridge University Press

Edling CR (2002) Mathematics in sociology. Annual Review of Sociology 28(1):197–220, DOI 10.1146/annurev.soc.28.110601.140942

Estrada E (2020) COVID-19 and SARS-CoV-2. Modeling the present, looking at the future. Physics Reports 869:1–51, DOI 10.1016/j.physrep.2020.07.005

Fagnani F, Zino L (2017) Diffusion of innovation in large scale graphs. IEEE Transactions on Network Science and Engineering 4(2):100–111, DOI 10.1109/tnse.2017.2678202

Fay N, Garrod S, Roberts L, Swoboda N (2010) The interactive evolution of human communication systems. Cognitive science 34(3):351–386, DOI 10.1111/j.1551-6709.2009.01090.x

Fibich G (2016) Bass-SIR model for diffusion of new products in social networks. Physical Review E 94(3), DOI 10.1103/physreve.94.032305

Finnemore M, Sikkink K (1998) International norm dynamics and political change. International Organization 52(4):887–917, DOI 10.1162/002081898550789

Fontan A, Altafini C (2021) A signed network perspective on the government formation process in parliamentary democracies. Scientific Reports 11(1), DOI 10.1038/s41598-021-84147-3

Franceschi J, Pareschi L (2022) Spreading of fake news, competence and learning: kinetic modelling and numerical approximation. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 380(2224), DOI 10.1098/rsta.2021.0159

French Jr JRP (1956) A Formal Theory of Social Power. Psychological Review 63(3):181–194, DOI 10.1037/h0046123

Friedkin NE, Bullo F (2017) How truth wins in opinion dynamics along issue sequences. Proceedings of the National Academy of Sciences 114(43):11,380–11,385, DOI 10.1073/pnas.1710603114

Friedkin NE, Johnsen EC (1990) Social influence and opinions. The Journal of Mathematical Sociology 15(3-4):193–206, DOI 10.1080/0022250X.1990.9990069

Fu F, Liu L, Wang L (2008) Empirical analysis of online social networks in the age of web 2.0. Physica A: Statistical Mechanics and its Applications 387(2–3):675–684, DOI 10.1016/j.physa.2007.10.006

Gao T, Zino L, Ye M (2023) Effect of network structure and committed minority placement in promoting social diffusion. IEEE Transactions on Computational Social Systems pp 1–14, DOI 10.1109/TCSS.2023.3303568

Gardner L, Ratcliff J, Dong E, Katz A (2021) A need for open public data standards and sharing in light of COVID-19. The Lancet Infectious Diseases 21(4):e80, DOI 10.1016/s1473-3099(20)30635-6

Gelfand MJ, Jackson JC (2016) From one mind to many: the emerging science of cultural norms. Current Opinion in Psychology 8:175–181, DOI 10.1016/j.copsyc.2015.11.002

Giordano G, Blanchini F, Bruno R, Colaneri P, Di Filippo A, Di Matteo A, Colaneri M (2020) Modelling the covid-19 epidemic and implementation of population-wide interventions in italy. Nature Medicine 26(6):855–860, DOI 10.1038/s41591-020-0883-7

Govaert A, Ramazi P, Cao M (2021) Rationality, imitation, and rational imitation in spatial public goods games. IEEE Transactions on Control of Network Systems 8(3):1324–1335, DOI 10.1109/tcns.2021.3065655

Granovetter M (1978) Threshold Models of Collective Behavior. American Journal of Sociology 83(6):1420–1443, DOI 10.1086/226707

Grey S (2006) Numbers and beyond: The relevance of critical mass in gender research. Politics & Gender 2(04), DOI 10.1017/s1743923x06221147

Guerra B, Poncela J, Gómez-Gardeñes J, Latora V, Moreno Y (2010) Dynamical organization towards consensus in the axelrod model on complex networks. Physical Review E 81(5), DOI 10.1103/physreve.81.056105

Guilbeault D, Becker J, Centola D (2018) Complex Contagions: A Decade in Review, Springer, pp 3–25. DOI 10.1007/978-3-319-77332-2_1

Hoffmann T, Ye M, Zino L, Cao M, Rauws W, Bolderdijk JW (2024) Overcoming inaction: An agent-based modelling study of social interventions that promote systematic pro-environmental change. Journal of Environmental Psychology 94:102,221, DOI 10.1016/j.jenvp.2023.102221

Holme P, Saramäki J (2012) Temporal networks. Physics Reports 519(3):97–125, DOI 10.1016/j.physrep.2012.03.001

Horvat A, Fogliano V, Luning PA (2020) Modifying the bass diffusion model to study adoption of radical new foods–the case of edible insects in the netherlands. PLOS ONE 15(6):e0234,538, DOI 10.1371/journal.pone.0234538

Hull A (2008) Policy integration: What will it take to achieve more sustainable transport solutions in cities? Transport Policy 15(2):94–103, DOI 10.1016/j.tranpol.2007.10.004

Jackson MO, Zenou Y (2015) Games on Networks. In: Handbook of Game Theory with Economic Applications, vol 4, Elsevier, chap 3, pp 95–163

Jannach D, Zanker M, Felfernig A, Friedrich G (2010) Recommender Systems: An Introduction. Cambridge University Press, DOI 10.1017/CBO9780511763113

Jusup M, Holme P, Kanazawa K, Takayasu M, Romić I, Wang Z, Geček S, Lipić T, Podobnik B, Wang L, Luo W, Klanjšček T, Fan J, Boccaletti S, Perc M (2022) Social physics. Physics Reports 948:1–148, DOI 10.1016/j.physrep.2021.10.005

Karniadakis GE, Kevrekidis IG, Lu L, Perdikaris P, Wang S, Yang L (2021) Physics-informed machine learning. Nature Reviews Physics 3(6):422–440, DOI 10.1038/s42254-021-00314-5

Kempe D, Kleinberg J, Tardos E (2003) Maximizing the spread of influence through a social network. In: Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp 137–146, DOI 10.1145/956750.956769

Klemm K, Eguíluz VM, Toral R, San Miguel M (2003) Nonequilibrium transitions in complex networks: A model of social interaction. Physical Review E 67(2), DOI 10.1103/physreve.67.026120

Koch J, Chen Z, Tuor A, Drgona J, Vrabie D (2023) Structural inference of networked dynamical systems with universal differential equations. Chaos 33(2), DOI 10.1063/5.0109093

Kozitsin IV (2021) Opinion dynamics of online social network users: a micro-level analysis. The Journal of Mathematical Sociology 47(1):1–41, DOI 10.1080/0022250x.2021.1956917

Kuran T (1995) The inevitability of future revolutionary surprises. American Journal of Sociology 100(6):1528–1551, DOI 10.1086/230671

Kuwahara B, Bauch CT (2024) Predicting Covid-19 pandemic waves with biologically and behaviorally informed universal differential equations. Heliyon 10(4):e25,363, DOI 10.1016/j.heliyon.2024.e25363

Lambiotte R, Ausloos M, Hołyst JA (2007) Majority model on a network with communities. Physical Review E 75(3), DOI 10.1103/physreve.75.030101

Lanchier N, Moisson PH (2015) Fixation results for the two-feature axelrod model with a variable number of opinions. Journal of Theoretical Probability 29(4):1554–1580, DOI 10.1007/s10959-015-0623-y

Lazer D, Pentland A, Adamic L, Aral S, Barabási AL, Brewer D, Christakis N, Contractor N, Fowler J, Gutmann M, Jebara T, King G, Macy M, Roy D, Van Alstyne M (2009) Computational social science. Science 323(5915):721–723, DOI 10.1126/science.1167742

Leonard NE, Lipsitz K, Bizyaeva A, Franci A, Lelkes Y (2021) The nonlinear feedback dynamics of asymmetric political polarization. Proceedings of the National Academy of Sciences 118(50), DOI 10.1073/pnas.2102149118

Lewis D (2002) Convention: A Philosophical Study, 1st edn. Wiley-Blackwell, Hoboken NJ, US

Li Y, Fan J, Wang Y, Tan KL (2018) Influence maximization on social graphs: A survey. IEEE Transactions on Knowledge and Data Engineering 30(10):1852–1872, DOI 10.1109/TKDE.2018.2807843

Lieberman E, Hauert C, Nowak MA (2005) Evolutionary dynamics on graphs. Nature 433(7023):312–316, DOI 10.1038/nature03204

Lieberman E, Michel JB, Jackson J, Tang T, Nowak MA (2007) Quantifying the evolutionary dynamics of language. Nature 449(7163):713–716, DOI 10.1038/nature06137

Louis W, Thomas E, McGarty C, Lizzio-Wilson M, Amiot C, Moghaddam F (2020) The volatility of collective action: Theoretical analysis and empirical data. Political Psychology 41(S1):35–74, DOI 10.1111/pops.12671

Louis WR, Lizzio-Wilson M, Cibich M, McGarty C, Thomas EF, Amiot CE, Weber N, Rhee J, Davies G, Rach T, et al (2022) Failure leads protest movements to support more radical tactics. Social Psychological and Personality Science 13(3):675–687, DOI 10.1177/19485506211037296

Mackie G (1996) Ending footbinding and Infibulation: A Convention Account. American Sociological Review pp 999–1017, DOI 10.2307/2096305

Mahajan V, Muller E, Bass FM (1990) New Product Diffusion Models in Marketing: A Review and Directions for Research. Journal of Marketing 54(1):1–26

Marmor A (2009) Social Conventions: From Language to Law, 1st edn. Princeton University Press, Princeton NJ, US

Marques JM, Paez D (1994) The 'black sheep effect': Social catergorization, rejection of ingroup deviates, and perception of group variability. European Review of Social Psychology 5(1):37–68, DOI 10.1080/14792779543000011

Michel JB, Shen YK, Aiden AP, Veres A, Gray MK, Team TGB, Pickett JP, Hoiberg D, Clancy D, Norvig P, Orwant J, Pinker S, Nowak MA, Aiden EL (2011) Quantitative analysis of culture using millions of digitized books. Science 331(6014):176–182, DOI 10.1126/science.1199644

Mlakar Ž, Bolderdijk JW, Risselada H, Fennis BM, Ye M, Zino L, Cao M (2024) Social tipping games: Experimental paradigms for studying consumer movements. Journal of the Association for Consumer Research DOI 10.1086/731916

Monger G (2004) Marriage Customs of the World: From Henna to Honeymoons. Bloomsbury Academic

Montanari A, Saberi A (2010) The spread of innovations in social networks. Proceedings of the National Academy of Sciences 107(47):20,196–20,201, DOI 10.1073/pnas.1004098107

Morris S (2000) Contagion. The Reviews of Economic Studies 67(1):57–78, DOI 10.1111/1467-937X.00121

Mäs M, Nax HH (2016) A behavioral study of "noise" in coordination games. Journal of Economic Theory 162:195–208, DOI 10.1016/j.jet.2015.12.010

Newman M (2018) The structure of real-world networks. In: Networks, 2nd edn, Oxford University Press, DOI 10.1093/oso/9780198805090.003.0010

Noorazar H (2020) Recent advances in opinion propagation dynamics: a 2020 survey. The European Physical Journal Plus 135(6), DOI 10.1140/epjp/s13360-020-00541-2

Otto IM, Donges JF, Cremades R, Bhowmik A, Hewitt RJ, Lucht W, Rockström J, Allerberger F, McCaffrey M, Doe SS, et al (2020) Social tipping dynamics for stabilizing Earth's climate by 2050. Proceedings of the National Academy of Sciences 117(5):2354–2365, DOI 10.1073/pnas.1900577117

Parino F, Zino L, Porfiri M, Rizzo A (2021) Modelling and predicting the effect of social distancing and travel restrictions on COVID-19 spreading. Journal of the Royal Society Interface 18(175):20200,875, DOI 10.1098/rsif.2020.0875

Pedraza L, Pinto S, Pinasco JP, Balenzuela P (2021) Analytical approach to the Axelrod model based on similarity vectors. Physical Review E 103(1), DOI 10.1103/physreve.103.012307

Peyton Young H (2011) The dynamics of social innovation. Proceedings of the National Academy of Sciences 108(Supplement 4):21,285–21,291, DOI 10.1073/pnas.1100973108

Peyton Young H (2015) The Evolution of Social Norms. Annual Review of Economics 7(1):359–387, DOI 10.1146/annurev-economics-080614-115322

Rackauckas C, Ma Y, Martensen J, Warner C, Zubov K, Supekar R, Skinner D, Ramadhan A, Edelman A (2020) Universal differential equations for scientific machine learning. arXiv preprint DOI Rackauckas2020

Ramazi P, Riehl J, Cao M (2016) Networks of conforming or nonconforming individuals tend to reach satisfactory decisions. Proceedings of the National Academy of Sciences 113(46):12,985–12,990, DOI 10.1073/pnas.1610244113

Rapoport A (1963) Mathematical models of social interaction. In: Luce D (ed) Handbook of Mathematical Psychology, John Wiley & Sons., pp 2–493

Riehl J, Ramazi P, Cao M (2018) A survey on the analysis and control of evolutionary matrix games. Annu Rev Control 45:87–106

Rizzo A, Porfiri M (2016) Innovation diffusion on time-varying activity driven networks. The European Physical Journal B 89(1), DOI 10.1140/epjb/e2015-60933-3

Rogers E (2003) Diffusion of Innovations, 5th edn. Free Press, New York NY, US

Rossi WS, Como G, Fagnani F (2019) Threshold models of cascades in large-scale networks. IEEE Transactions on Network Science and Engineering 6(2):158–172, DOI 10.1109/tnse.2017.2777941

Rubel LA (1981) A universal differential equation. Bulletin of the American Mathematical Society 4(3):345–349, DOI 10.1090/s0273-0979-1981-14910-7

Ryan B, Gross NC (1943) The diffusion of hybrid seed corn in two iowa communities. Rural Sociology 8(1):15

Samuelson W, Zeckhauser R (1988) Status Quo Bias in Decision Making. J Risk Uncertain 1(1):7–59

Satoh D (2001) A discrete bass model and its parameter estimation. Journal of the Operations Research Society of Japan 44(1):1–18, DOI 10.15807/jorsj.44.1

Selin NE (2021) Lessons from a pandemic for systems-oriented sustainability research. Science Advances 7(22), DOI 10.1126/sciadv.abd8988

Southwood N, Eriksson L (2011) Norms and conventions. Philosophical Explorations 14(2):195–217, DOI 10.1080/13869795.2011.569748

Sparkman G, Walton GM (2017) Dynamic norms promote sustainable behavior, even if it is counternormative. Psychological Science 28(11):1663–1674, DOI 10.1177/0956797617719950

Succar R, Porfiri M (2024) Urban scaling of firearm violence, ownership and accessibility in the united states. Nature Cities 1(3):216–224, DOI 10.1038/s44284-024-00034-8

Sultan F, Farley JU, Lehmann DR (1990) A meta-analysis of applications of diffusion models. Journal of Marketing Research 27(1):70–77, DOI 10.1177/002224379002700107

Thomas EF, Ye M, Angus SD, Mathew TJ, Louis W, Walsh L, Ellery S, Lizzio-Wilson M, McGarty C (2024) From mobilisation to radicalisation: Probing the persistence and radicalisation of social movements using an agent-based model. Preprint Available at `https://arxiv.org/abs/2408.12795`

Tredennick AT, Hooker G, Ellner SP, Adler PB (2021) A practical guide to selecting models for exploration, inference, and prediction in ecology. Ecology 102(6), DOI 10.1002/ecy.3336

Truszkowska A, Behring B, Hasanyan J, Zino L, Butail S, Caroppo E, Jiang Z, Rizzo A, Porfiri M (2021) High-resolution agent-based modeling of COVID-19 spreading in a small town. Advanced Theory and Simulations 4(3), DOI 10.1002/adts.202000277

Valente TW (1996) Social network thresholds in the diffusion of innovations. Social Networks 18(1):69–89, DOI 10.1016/0378-8733(95)00256-1

Van Bavel JJ, et al (2020) Using social and behavioural science to support COVID-19 pandemic response. Nature Human Behaviour 4(5):460–471, DOI 10.1038/s41562-020-0884-z

Vanelli M, Arditti L, Como G, Fagnani F (2020) On games with coordinating and anti-coordinating agents. In: Proceedings of the 21st IFAC World Congress, vol 53(2), pp 10,975–10,980, DOI 10.1016/j.ifacol.2020.12.2848

Vega-Redondo F (2007) Complex Social Networks. Cambridge University Press, DOI 10.1017/cbo9780511804052

Vespignani A, Tian H, Dye C, Lloyd-Smith JO, Eggo RM, Shrestha M, Scarpino SV, Gutierrez B, Kraemer MUG, Wu J, Leung K, Leung GM (2020) Modelling covid-19. Nature Reviews Physics 2(6):279–281, DOI 10.1038/s42254-020-0178-4

von Neumann J, Morgenstern O, Rubinstein A (1944) Theory of Games and Economic Behavior. Princeton University Press

Walsh L, Ye M, Anderson BD, Sun Z (2023) Decentralised adaptive-gain control for the sis network epidemic model. In: Proceedings of the 22nd IFAC World Congress, Elsevier BV, vol 56(2), p 8506–8511, DOI 10.1016/j.ifacol.2023.10.1140

Watts DJ (2002) A simple model of global cascades on random networks. Proceedings of the National Academy of Sciences 99(9):5766–5771, DOI 10.1073/pnas.082090499

World Health Organization (2024) Who coronavirus disease (covid-19) dashboard. `https://data.who.int/dashboards/covid19/cases`

Xie J, Sreenivasan S, Korniss G, Zhang W, Lim C, Szymanski BK (2011) Social consensus through the influence of committed minorities. Physical Review E 84(1), DOI 10.1103/physreve.84.011130

Ye M, Zino L, Mlakar Ž, Bolderdijk JW, Risselada H, Fennis B, Cao M (2021a) Collective patterns of social diffusion are shaped by individual inertia and trend-seeking. Nature Communications 12:5698, DOI 10.1038/s41467-021-25953-1

Ye M, Zino L, Rizzo A, Cao M (2021b) Game-theoretic modeling of collective decision making during epidemics. Physical Review E 104(2), DOI 10.1103/physreve.104.024314

Zhao S, Chen H (2020) Modeling the epidemic dynamics and control of COVID-19 outbreak in China. Quantitative Biology 8(1):11–19, DOI 10.1007/s40484-020-0199-0

Zino L, Ye M, Cao M (2020) A two-layer model for coevolving opinion dynamics and collective decision-making in complex social systems. Chaos 30(8):083,107, DOI 10.1063/5.0004787

Zino L, Ye M, Cao M (2022) Facilitating innovation diffusion in social networks using dynamic norms. PNAS Nexus 1, DOI 10.1093/pnasnexus/pgac229

Zino L, Como G, Fagnani F (2023a) Fast spread in controlled evolutionary dynamics. IEEE Transactions on Control of Network Systems 10(3):1555–1567, DOI 10.1109/tcns.2023.3234593

Zino L, Ye M, Rizzo A, Calafiore GC (2023b) On adaptive-gain control of replicator dynamics in population games. In: Proceedings of the 62nd IEEE Conference on Decision and Control, IEEE, vol 1, p 485–490, DOI 10.1109/cdc49753.2023.10383983