Group Reasoning in Social Environments

Erman Acar
University of Mannheim
Mannheim, Germany
erman@informatik.uni-mannheim.de

Gianluigi Greco
University of Calabria
Rende, Italy
ggreco@mat.unical.it

Marco Manna
University of Calabria
Rende, Italy
manna@mat.unical.it

ABSTRACT

While modeling group decision making scenarios, the existence of a central authority is often assumed which is in charge of amalgamating the preferences of a given set of agents with the aim of computing a socially desirable outcome, for instance, maximizing the utilitarian or the egalitarian social welfare. Departing from this classical perspective and inspired by the growing body of literature on opinion formation and diffusion, a setting for group decision making is studied where agents are selfishly interested and where each of them can adopt her own decision without a central coordination, hence possibly disagreeing with the decision taken by some of the other agents. In particular, it is assumed that agents belong to a social environment and that their preferences on the available alternatives can be influenced by the number of "neighbors" agreeing/disagreeing with them. The setting is formalized and studied by modeling agents' reasoning capabilities in terms of weighted propositional logics and by focusing on Nash-stable solutions as the prototypical solution concept. In particular, a thoroughly computational complexity analysis is conducted on the problem of deciding the existence of such stable outcomes. Moreover, for the classes of environments where stability is always guaranteed, the convergence of Nash dynamics consisting of sequences of best response updates is studied, too.

1. INTRODUCTION

Understanding how global behavior emerges from local interactions among individuals is a well-established topic of research in a number of different areas, including economics, finance, epidemiology, social psychology, and political science. More recently, due to the rapid proliferation of social networking services, such as Facebook and Twitter, which created novel and highly-dynamic forms of techno-social ecosystems, computer scientists have been fascinated by the problem, too (see, e.g., [16]).

A social environment can be naturally modeled as network, whose nodes correspond to the individuals and whose edges encode their social interconnections which give rise to influence phenomena. In particular, most of the classical studies on this topic assume that neighbors communicate by propagating and diffusing "items", such as technologies or diseases, because of reasons ranging from similarity and social ties [3], to conformity [33], and to compliance [14], just to name a few. Well-known diffusion models are the cascade [27], the tipping/treshold [23], and the homophilic model [4]. Moreover, in some cases, these models are mixed together (e.g., [13]) or they are extended to accommodate the diffusion of different and competing information [11, 25, 31, 12, 29].

More recently, richer models of social environments have been proposed in the literature, which are tailored to study how "immaterial" things, in particular *opinions*, form and diffuse over the network. Many of these works [2, 18, 20, 26] build on a basic model of [17], where each individual is equipped with a real number (for example, representing a position on a political spectrum or a probability assigned to a certain belief), which is updated, at each time step, to be a weighted average of that opinion with the current opinions of the neighbors. By doing so, the diffusion processes will converge to a state of consensus where all individuals hold the same opinion. A natural extension, first proposed by [19], is to equip each individual with an innate opinion in addition to the expressed opinion. At each time step, the expressed opinion is then updated to minimize the disagreement with the innate opinion and the opinions expressed by the neighbors—see, e.g., [9, 10, 8, 5].

Enhancing the expressiveness of such models for information diffusion is an important research issue. Indeed, modeling the reasoning capabilities of the individuals as real numbers is a clear limitation from the knowledge representation viewpoint. However, proposing and analyzing formalizations where individuals are seen as thinking entities equipped with their own (in particular, logical) theories, and where mechanisms are conceived to reason about how these theories interact within the dynamics of the network is an avenue widely unexplored in the literature. The goal of the paper is to fill this gap, by proposing a semantically rich framework for studying both information diffusion and social influence phenomena based on a group decision-making setting where agents have their own view of the world modeled in terms of *weighted propositional logics* [15]. Indeed, this modeling language is convenient for our aim, since it has been shown to express all common classes of utility functions, and it also provides a convenient means to elicit user's preferences while balancing expressivity and complexity [34]. Furthermore, it practically enables to exploit SAT-Solvers to deal with instances involving tens of thousands of variables and formulas consisting of millions of symbols, today [21].

The proposed framework is reminiscent of a number of earlier approaches in social choice, where such logic-based agents have been studied in the context of computing socially desirable outcomes, for instance, maximizing the utilitarian [34] or the egalitarian social welfare [24]. In particular, from this literature we borrow the notion of *goalbase* — to associate numerical weights with goals specified in terms of propositional formulas — but we depart from the aforementioned approaches by designing a novel framework where agents are selfishly interested; so, each of them

Appears in: *Proceedings of the 16th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2017), S. Das, E. Durfee, K. Larson, M. Winikoff (eds.), May 8–12, 2017, São Paulo, Brazil.*

can adopt her own decision without a central coordination, hence possibly disagreeing with the decision taken by some of the other agents. New problems have been studied accordingly.

Details on the proposed formalization are provided in Section 3, and natural questions arising therein are addressed in the subsequent sections. In fact, we study the complexity of problems related to the existence of *Nash* stable profiles, i.e., where no agent finds convenient to adopt a different decision/opinion (see Section 4), as well as the convergence of fully-decentralized dynamics consisting of sequences of best response updates (see Section 5). In particular,

► We first study the computational question of checking whether a given environment admits a *Nash* stable profile, i.e., where there is no agent finding convenient to adopt a different decision. We show that the problem is intractable, by conducting a fine-grained analysis w.r.t. the size of the domains on which the theories of the agents are expressed (and by also considering the related problem of checking whether a profile given at hand is stable). The reductions exploited to prove these results shed lights on the expressivity of the framework, too.

► As a way to circumvent the above bad news, we then focus on a specific class of social environments where agents' reasoning capabilities are suitably bounded—we call them *linear* agents. Computational problems are reconsidered on this class. In particular, we identify a number of conditions a-priori guaranteeing the existence of stable profiles and we show that, whenever these conditions do not hold, checking whether a stable profile exists is still an intractable problem, formally **NP**-complete. Therefore, these results precisely chart the tractability frontier for reasoning about Nash stable profile.

► Finally, for all the cases where stable profiles are guaranteed to exist by the conditions mentioned above, we study the question of whether a dynamics consisting of sequences of best response moves, without any central coordination, is guaranteed to converge.

2. PRELIMINARIES

Throughout the paper, we assume that a universe $\mathcal{V}$ of variables is given and, for any structure ζ defined on $\mathcal{V}$, we denote by $\mathtt{dom}(\zeta)$ the set of all the variables occurring in ζ.

We consider the propositional language $\mathcal{L}$ consisting of all formulas built over $\mathcal{V}$ by using the Boolean connectives $\wedge$, $\vee$, and $\neg$, plus the Boolean constants $\top$ (*true*) and $\bot$ (*false*). Moreover, given two formulas φ_1 and φ_2 in $\mathcal{L}$, we use $\varphi_1 \Rightarrow \varphi_2$ as a shorthand for $\neg\varphi_1 \vee \varphi_2$, and $\varphi_1 \Leftrightarrow \varphi_2$ as a shorthand for $(\varphi_1 \Rightarrow \varphi_2) \wedge (\varphi_2 \Rightarrow \varphi_1)$. An *interpretation* I is a function assigning a Boolean value to each variable in its domain, i.e., I: $\mathtt{dom}(I) \rightarrow \{\top, \bot\}$. We often describe I extensively, i.e., as the set of literals $\{x \in \mathtt{dom}(I) \mid I(x) = \top\} \cup \{\neg x \mid x \in \mathtt{dom}(I) \wedge I(x) = \bot\}$. The restriction of I to any set $V \subseteq \mathtt{dom}(I)$ is denoted by $I|_V$. We deal with I under the usual semantics for propositional logic, by $I \models \varphi$ denoting that I satisfies a formula $\varphi \in \mathcal{L}$ with $\mathtt{dom}(I) \supseteq \mathtt{dom}(\varphi)$.

A *weighted formula* is a pair (φ, w), where $\varphi \in \mathcal{L}$ and where $w \in \mathbb{Q}$ is its weight [15]. A *goalbase* G is a finite set of weighted formulas. For any interpretation I with $\mathtt{dom}(I) \supseteq \mathtt{dom}(G)$, we define $G(I)$ as the sum of the weights over all the pairs $(\varphi, w) \in G$ such that $I \models \varphi$ holds, that is, $G(I) = \sum_{(\varphi,w)\in G \text{ s.t. } I\models\varphi} w$.

3. FORMAL FRAMEWORK

In what follows, we introduce our framework step by step with the aid of a running example. Later, we are going to trace the tractability frontier of the main problems about Nash stability and Nash dynamics.

3.1 Agents and environments

We assume that a set $[n] = \{1, ..., n\}$ of agents is given. A *social environment* is a triple $\mathcal{G} = ([n], E, \kappa)$, which we often transparently view as a graph whose nodes are the agents in $[n]$ and whose edges in $E \subseteq [n] \times [n]$ encode the fact that such agents are related by some relationship, because of reasons that might range from physical limitations and constraints to legal banishments and friendships. The environment $\mathcal{G}$ is also equipped with a function κ storing the knowledge base $\kappa(i)$ of each agent $i \in [n]$, which we model by using weighted propositional logic.

Formally, every agent $i \in [n]$ is characterized by a knowledge base $\kappa(i)$ having the form $\langle G_i^0, ..., G_i^{d_i}\rangle$, where $d_i \geq 0$ is a natural number, called the *degree* of i, and where G_i^δ is a goalbase, for each natural number $\delta \in \{0, ..., d_i\}$. To explain the role played by these goalbases, we start with the simplest kind of agent, whose degree is 0. An agent of this kind is *autonomous*, as she does not care about the other agents and acts in the environment as if she were alone.

EXAMPLE 1. Consider an agent, say 1, such that $\kappa(1) = \langle G_1^0\rangle$ with $G_1^0 = \{(hil \Leftrightarrow dem, 1), (\neg dem, 1)\}$. This goalbase models the intuition that 1 "will vote for Hillary iff she is a Democrat" and that actually "she is not a Democrat". Indeed, G_1^0 is meant to encode the *utility* of agent 1, ranging between 0 and 2. And, its maximum value is achieved over the interpretation $\{\neg dem, \neg hil\}$. That is, based on her innate opinion i.e., based solely on $\langle G_1^0\rangle$, agent 1 will not vote for Hillary. $\lhd$

Higher degrees are meant to encode agents whose reasoning mechanisms are influenced by their social relationships. In particular, we assume that each agent selects her own interpretation and that the utility of an agent i sums the value returned by G_i^0 up with a *polynomial function*, having degree d_i, in the number of the agents "agreeing" on the interpretation selected by i and whose coefficients are returned by the values of the goalbases G_i^δ, for $\delta > 0$.

In order to formalize our intuition, let $\mathtt{dom}(i)$ denote the set $\mathtt{dom}(G_i^0) \cup \cdots \cup \mathtt{dom}(G_i^{d_i})$, let $\mathtt{space}(i)$ be the set of all interpretations over $\mathtt{dom}(i)$, and let us define a *profile* Π as a function $\{i \mapsto \Pi_i\}_{i\in[n]}$ mapping each agent $i \in [n]$ to interpretation $\Pi_i \in \mathtt{space}(i)$. Moreover, define the *partners* of i with respect to Π as the set

$$\mathtt{partners}(i,\Pi) = \{j \mid (j,i) \in E \text{ and } \Pi_i|_{\mathtt{dom}(j)} = \Pi_j|_{\mathtt{dom}(i)}\},$$

that is, as the set of agents j for which there is an edge from j to i in $\mathcal{G}$ and interpretation Π_j is "compatible" with Π_i. Then, the *utility* of agent i with respect to Π is the following rational number:

$$u_i(\Pi) = G_i^0(\Pi_i) + \sum_{\delta\in\{1,...,d_i\}} G_i^\delta(\Pi_i) \cdot |\{i\} \cup \mathtt{partners}(i,\Pi)|^\delta.$$

With this definition of utility in place, we can now summarize the main working assumptions of our formal framework: for every agent, (1) any influence is restricted to neighboring agents only (i.e., there is no direct influence from neighbors of neighbors), and (2) all the neighbors have the same degree of influence. These assumptions are aimed to keep things as simple as possible; relaxing them is a part of the future work. Note that when the value of $G_i^\delta(\Pi_i)$ in the above expression is negative (resp., positive) for each $\delta > 0$, it is the case that agent i behaves as a *dissenter* (resp., *conformist*), namely she prefers an interpretation that helps to minimize (resp., maximize) the number of her partners in order to maximize her utility. (The notions of dissenters and conformists will be studied more in depth in Section 5.) In the following, we use the convention that dissenters are graphically represented as nodes labeled with the symbol '$-$', while conformists as nodes labeled with the opposite symbol '$+$'.

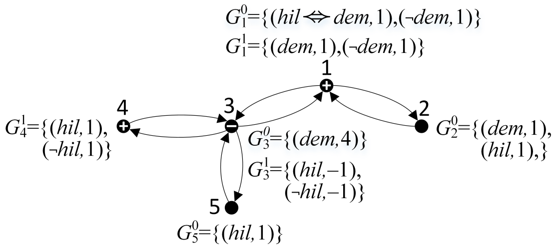

Figure 1: Knowledge bases and interactions for the social environment of Example 2; $\forall i \in [n]$, $d_i \leq 1$.

EXAMPLE 2. Let us extend Example 1, by considering the five agents whose goalbases are as reported in Figure 1. Note that the knowledge base of agent 1 is modified so that $\kappa(1) = \langle G_1^0, G_1^1 \rangle$, where $G_1^1 = \{(dem, 1), (\neg dem, 1)\}$. That is, agent 1 gets an additional utility if she agrees with some neighbor on the decision about being a Democrat.

Assume that agents interact according to the edges depicted in Figure 1, and consider the profile Π such that $\Pi_1 = \{hil, dem\}$, $\Pi_2 = \{hil, dem\}$, $\Pi_3 = \{dem, \neg hil\}$, $\Pi_4 = \{\neg hil\}$, and $\Pi_5 = \{hil\}$. Note that agent 1 — acting as a conformist — can be influenced only by agents 2 and 3 due to network structure, and that her current interpretation Π_1 is compatible with Π_2 only (according to Π, agent 2 is the only partner of agent 1). Thus, $\texttt{partners}(1, \Pi) = \{2\}$ and $u_1(\Pi) = G_1^0(\Pi_1) + G_1^1(\Pi_1) \cdot |\{1\} \cup \{2\}|^1 = 1 + 1 \cdot 2 = 3$. In particular, observe that given her social interconnections with 2 and 3, agent 1 is now inclined to support the Democratic Party and hence, to vote for Hillary—just check that agent 1 gets a lower utility with any interpretation different from Π_1. Hence, the expressed opinion of agent 1 now differs from her innate one. For the other agents we have $u_2(\Pi) = 2$, $u_3(\Pi) = 2$, $u_4(\Pi) = 2$, and $u_5(\Pi) = 1$. (We implicitly assume that $G_4^0 = \emptyset$. However, alternative equivalent definitions would be, for example, $G_4^0 = \{(hil, 0), (\neg hil, 0)\}$ and even $G_4^0 = \{(hil \vee \neg hil, 0)\}$. In any such cases, it holds that $G_4^0(\Pi_4) = 0$.) $\lhd$

3.2 Nash stability

We assume hereinafter that a social environment $\mathcal{G}$=$([n], E, \kappa)$ is given. For any agent $i \in [n]$, the *restriction* Π_{-i} of a profile Π to the agents $[n] \setminus \{i\}$ is the function $\{j \mapsto \Pi_j\}_{j \in [n] \setminus \{i\}}$. Given a profile Π, an agent $i \in [n]$, and an interpretation $I \in \texttt{space}(i)$, the mapping $\{i \mapsto I\}$ is a *best response move* for i (w.r.t. Π) if $u_i(\Pi_{-i} \cup \{i \mapsto I\}) \geqslant u_i(\Pi_{-i} \cup \{i \mapsto J\})$, for each $J \in \texttt{space}(i)$. A profile Π is *(Nash) stable* in $\mathcal{G}$ if each agent is playing one of her best response moves. That is, for each $i \in [n]$ and for each interpretation $I \in \texttt{space}(i)$, it holds that $u_i(\Pi) \geqslant u_i(\Pi_{-i} \cup \{i \mapsto I\})$.

A *(Nash) dynamics* is a sequence of profiles, where each of them is either the initial one in the sequence or has the form $\Pi_{-i} \cup \{i \mapsto I\}$, with $\{i \mapsto I\}$ being a best response move for some agent i w.r.t. the previous profile Π.

EXAMPLE 3. The profile Π discussed in Example 2 is the only stable one. In particular, if we start from any initial profile where $\{1 \mapsto \Pi_1\}$, any Nash dynamics will converge to Π. Indeed, just notice that agents 2 and 5 are autonomous and converge to Π_2 and Π_5, respectively. Moreover, given the weighted formula $(dem, 4)$ belonging to G_3^0, agent 3 always converges to an interpretation where she is a Democrat.

Now, concerning the decision about whether to vote for Hillary, note that agent 3 acts as a dissenter; hence, she converges to Π_3 (she does not vote for Hillary, while being a Democrat), because the majority of her neighbors vote for Hillary: she prefers to minimize the number of her partners. Instead, agent 4 is a conformist and she adapts Π_4 to the opinion by agent 3.

Consider instead, the modified setting where the goalbase G_2^0 is updated as $\{(\neg dem, 1), (\neg hil, 1)\}$. In this case, agent 1 will be faced with two contrasting opinions about being a Democrat, so that she will eventually come back to her innate view by selecting the interpretation $\{\neg dem, \neg hil\}$. Then, agent 3 will have similarly one neighbor voting for Hillary and one not voting for her. So, the decision of 3 depends only on 4, and vice versa. But, these two agents cannot find an agreement since 3 is a dissenter, while 4 is a conformist. Hence, no stable profile exists at all. $\lhd$

4. REASONING ABOUT SOCIAL ENVIRONMENTS

After that the framework has been formalized, it is now our goal to test its expressivity and study analytical and computational properties related to the convergence of Nash dynamics. And as promised, we would like to trace the tractability frontier of the main problems concerning Nash stability and Nash dynamics.

Motivated by Example 3, we start by studying the computational complexity of the problem $\exists$-NASH of deciding, given a social environment as input, whether it admits any stable profile at all. In addition, it is also sensible to study the problem IS-NASH of deciding where some given profile is actually stable.

A summary of our results is presented in Table 1. Note that the analysis is provided parametrically w.r.t. the value $\texttt{maxDomSize}(\mathcal{G})$ of the social environment $\mathcal{G}$, which we define as the maximum cardinality of the domain over all the agents. Indeed, a logarithmic bound on this parameter ensures that the complexity drops down one level in the polynomial hierarchy (see Table 1). Proofs for these results are elaborated in the rest of the section.

$\texttt{maxDomSize}(\mathcal{G})$	IS-NASH	$\exists$-NASH
$\mathcal{O}(\log n)$	**PTIME** → Thm. 1	**NP**-c → Thm. 2
arbitrary	**coNP**-c → Thm. 3	$\Sigma_2^{\mathbf{P}}$-c → Thm. 4

Table 1: Complexity of the problems IS-NASH and $\exists$-NASH in the general case, and in case $\texttt{maxDomSize}(\mathcal{G})$ is bounded logarithmically by n.

4.1 "Logarithmic" agents

Consider first the setting where $\texttt{maxDomSize}(\mathcal{G}) \in \mathcal{O}(\log n)$. In this case, for each $i \in [n]$, $\texttt{space}(i)$ can contain polynomially-many interpretations only; therefore, checking whether some agent has an incentive to deviate from some given profile is clearly feasible in polynomial time.

THEOREM 1. IS-NASH *is in* **PTIME** *if* $\texttt{maxDomSize}(\mathcal{G})$ *belongs to* $\mathcal{O}(\log n)$.

PROOF. Consider a social environment $\mathcal{G}$ and a profile Π. For each $i \in [n]$ and for each $I \in \texttt{space}(i)$, if $u_i(\Pi) < u_i(\Pi_{-i} \cup \{i \mapsto I\})$, then *reject*; otherwise, after this double for-loop, *accept*. (Note that, by hypothesis, $\texttt{space}(i)$ may contain only polynomially many interpretations.) $\square$

On the other hand, we shall next show that $\exists$-NASH is intractable, formally **NP**-complete. It is instructive to overview the proof, as it sheds lights on the expressiveness of our framework. Indeed, we shall next show that our agents are able to reason about forming *coalitions*, even though this concept does not appear as a first-class citizen in the framework. In particular, social environments can simulate *anonymous hedonic games*, which are games in which the preferences of each player can be completely summarized as preferences over the coalitions where she is included (see, e.g., [6]). Formally, a hedonic game H is a pair $\langle [n], \succeq \rangle$ where $\succeq$ is a function associating to each player $i \in [n]$ a *preference* order (complete and transitive binary relation) $\succeq_i$ over $\{S \subseteq [n] \mid i \in S\}$. In particular, H is *anonymous* if $\succeq_i$ is given as a function $\mu_i : [n] \to \mathbb{Q}$ such that $C_1 \succeq_i C_2$ iff $\mu_i(|C_1|) \geqslant \mu_i(|C_2|)$. That is, in anonymous games preferences of the agents depend only on the size of the coalitions where they belong to.

THEOREM 2. $\exists$-NASH *is* **NP**-*complete if* $\mathtt{maxDomSize}(\mathcal{G})$ *belongs to* $\mathcal{O}(\log n)$.

PROOF. For the hardness, consider an anonymous hedonic game H on $[n]$. A *coalition* in H is any subset of $[n]$. A *configuration* $\mathcal{C}$ is a set of coalitions that forms a partition of $[n]$. For each $i \in [n]$, the set $\mathtt{coal}(i, \mathcal{C})$ denotes the coalition of $\mathcal{C}$ containing i. Configuration $\mathcal{C}$ is *stable* if, for each $i \in [n]$ and $C \in \mathcal{C} \cup \emptyset$, $\mu_i(\mathtt{coal}(i, \mathcal{C})) \geq \mu_i(\{i\} \cup C)$. Deciding whether stable configuration exists for these games is **NP**-complete [7, 30].

Given H, observe first that μ_i, for each $i \in [n]$, can be always written (in polynomial time) in the form $\mu_i(s) = c_i^{n-1} \cdot s^{n-1} + \ldots + c_i^1 \cdot s + c_i^0$, via standard interpolation techniques (e.g., [1]). Then, we can build the social environment $\mathcal{G} = ([n], E, \kappa)$ where: $E = \{(i,j) \mid i,j \in [n] \text{ and } i \neq j\}$ and where, for each $i \in [n]$, $\kappa(i)$ is such that:

- $G_i^0 = \{(\bigwedge_{q \leqslant \lceil \log n \rceil} (x_q \vee \neg x_q), c_i^0)\}$, and
- $G_i^\delta = (\top, c_i^\delta)$, for each $\delta \in \{1, ..., n-1\}$.

Note that, over the variables x_q, we can encode the natural numbers in the set $[n]$, which we can univocally use to identify the coalition each agent belongs to. Then, it can be easily checked that H admits a stable configuration if, and only if, $\mathcal{G}$ admits a stable profile.

To conclude, note that $\exists$-NASH is in **NP**. Indeed, it can be solved by guessing a profile, and by subsequently checking in polynomial time whether it is stable (cf. Theorem 1). $\square$

4.2 Arbitrary agents

We now analyze the case of arbitrary agents. Contrasted with Theorem 1, the following shows that a complexity blow-up occurs moving from the case of "logarithmic" agents.

THEOREM 3. IS-NASH *is* **coNP**-*complete. Hardness holds even for* $n = 1$ *and* $d_1 = 0$.

PROOF. (Membership) Consider a social environment $\mathcal{G}$ and a profile Π. The complementary problem is in **NP**: (a) Guess an agent $i \in [n]$ and an interpretation $I \in \mathtt{space}(i)$; (b) Check that $u_i(\Pi) < u_i(\Pi_{-i} \cup \{i \mapsto I\})$. In particular, note that the utility function can be computed in polynomial time.

(Hardness) Consider the **coNP**-complete problem of deciding whether a Boolean formula ϕ with $\mathtt{dom}(\phi) = \{x_1, \ldots, x_m\}$ is not satisfiable. Based on ϕ, we construct the social environment $\mathcal{G} = ([1], \emptyset, \kappa)$ with $\kappa(1) = \langle G_1^0 \rangle$ and such $G_1^0 = \{(x_0 \wedge \phi, 1)\}$. Consider the profile Π with $\Pi_1 = \{\neg x_0, \neg x_1, ..., \neg x_m\}$, and note that $u_1(\Pi) = 0$ since $I \not\models x_0 \wedge \phi$. Eventually, Π is stable iff ϕ is unsatisfiable. $\square$

Similarly, $\exists$-NASH becomes more difficult; in fact complete for the second level of the polynomial hierarchy.

THEOREM 4. $\exists$-NASH *is* $\Sigma_2^{\mathbf{P}}$-*complete. Hardness holds even for* $n = 5$ *and* $d_i \leqslant 1$, *for each* $i \in [n]$.

i	G_i^0	G_i^1
1	$\{(z \Leftrightarrow \phi(\mathbf{x}, \mathbf{y}), 3/2)\}$	$\{(\top, 1)\}$
2	$\{(x_1 \wedge \cdots \wedge x_p, 0)\}$	$\emptyset$
3	$\emptyset$	$\{(z, -1), (\neg z, -1)\}$
4	$\emptyset$	$\{(z, 1), (\neg z, 1)\}$
5	$\{(z, 1)\}$	$\emptyset$

Table 2: Knowledge base κ from the proof of Theorem 4.

PROOF SKETCH. (Membership) Guess a profile Π (in **NP**), and then check whether Π is stable (in **coNP** by Theorem 3).

(Hardness) Deciding whether a quantified formula $F = \exists \mathbf{x}. \forall\, \mathbf{y}. \phi(\mathbf{x}, \mathbf{y})$ with $\mathbf{x} = x_1, ..., x_p$ and $\mathbf{y} = y_1, ..., y_q$ is satisfiable is a $\Sigma_2^{\mathbf{P}}$-complete problem. Based on F, we build the environment $\mathcal{G} = ([5], E, \kappa)$ by keeping the underlying graph of Figure 1 and where κ is now specified in Table 2. We claim that F is satisfiable $\Leftrightarrow$ $\mathcal{G}$ admits a stable profile.

($\Rightarrow$) Let $I : \mathbf{x} \to \{\top, \bot\}$ be such that, for each $J : \mathbf{y} \to \{\top, \bot\}$, $I \cup J \models \phi$. Fix any given $\bar{J} : \mathbf{y} \to \{\top, \bot\}$, construct the profile Π where $\Pi_1 = I \cup \bar{J} \cup \{z \mapsto \top\}$, $\Pi_2 = I$, $\Pi_3 = \Pi_4 = \{z \mapsto \bot\}$, and $\Pi_5 = \{z \mapsto \top\}$, which implies $u_1(\Pi) = 7/2$, $u_2(\Pi) = 0$, $u_3(\Pi) = -2$, $u_4(\Pi) = 2$, and $u_5(\Pi) = 1$. By exhaustively enumerating all possible best response moves for the agents, it can be checked that Π is stable.

($\Leftarrow$) We shall show that if F is not satisfiable, then $\mathcal{G}$ does not admit any stable profile. Accordingly, for each $I : \mathbf{x} \to \{\top, \bot\}$, let $J_I : \mathbf{y} \to \{\top, \bot\}$ be such that $I \cup J_I \not\models \phi$. By contradiction, assume that there is a stable profile Π and take an arbitrary assignment I. First, one can notice that stability implies $\Pi_1 = I \cup J_I \cup \{z \mapsto \bot\}$, $\Pi_2 = I$, and $\Pi_5 = \{z \mapsto \top\}$, leading to $u_1(\Pi) = 9/2$, $u_2(\Pi) = 0$, and $u_5(\Pi) = 1$. Eventually, this makes either agent 3 or agent 4 "unstable", which is impossible. $\square$

5. CONSTANT DOMAINS AND NASH DYNAMICS

In the light of the results of Section 4, in order to isolate tractable scenarios, we need to further constraint the reasoning capabilities of the agents. Accordingly, in this section, we study environments $\mathcal{G}$ such that $\mathtt{maxDomSize}(\mathcal{G}) = \mathcal{O}(1)$, i.e., every agent can reason about a constant number of variables only. In addition, we focus on the case of *linear* agents, that is, $d_i \leqslant 1$ holds, for each $i \in [n]$, with the question of extending our results to agents that are not linear being an interesting avenue for further research (see Section 6 for further observations on this issue).

The analysis that follows is parametric w.r.t. some salient features of the agents. Formally, consider an agent $i \in [n]$. Then, the *base* and the *gradient* of i are defined by the sets $\mathtt{base}(i) = \{G_i^0(I) \mid I \in \mathtt{space}(i)\}$ and $\mathtt{grad}(i) = \{G_i^1(I) \mid I \in \mathtt{space}(i)\}$, respectively. Agent i is said:

- *autonomous* (of gradient 0) if $d_i = 0$ or $\mathtt{grad}(i) = \{0\}$;
- *conformist* (of gradient +) if $\mathtt{grad}(i) = \{c\}$ and $c > 0$;
- *dissenter* (of gradient $-$) if $\mathtt{grad}(i) = \{d\}$ and $d < 0$;

Moreover, in a completely orthogonal way, agent i is said:

- *adaptable* (of base 1) if $|\mathtt{base}(i)| = 1$;
- *resolute* (of base ∞) if $|\mathtt{base}(i)| > 1$;

For each fixed natural number $h > 0$, let us now denote by $\mathcal{C}_h[base, grad, graph]$ the class collecting all the environments $\mathcal{G}$

base	*grad*	*graph*: `sym`	*graph*: `arb`
1	0	*ever* → Thm. 5	*ever* → Thm. 5
	${}^0/{+}$	*ever* → Thm. 5	*ever*^{nc} → Thm. 6
	${}^0/{-}$	*ever* → Thm. 5	**NP**-c → Thm. 9
	${}^0/{\pm}$	**NP**-c → Thm. 8	**NP**-c → Cor. 1
∞	0	*ever* → Thm. 5	*ever* → Thm. 5
	${}^0/{+}$	*ever* → Thm. 5	*ever*^{oc} → Thm. 7 (h=1)
			NP-c → Thm. 10 (h>1)
	${}^0/{-}$	*ever* → Thm. 5	**NP**-c → Cor. 2
	${}^0/{\pm}$	**NP**-c → Cor. 1	**NP**-c → Cor. 1

Table 3: Summary of the results for the existence of stable profiles in Section 5. For the entries marked as *ever*, a stable profile always exists and can be computed in polynomial time; here, Nash dynamics converge in polynomial time, but for the entries with superscript nc and oc—where they do not necessarily converge and where this question remains open, respectively.

such that: (1) $\texttt{maxDomSize}(\mathcal{G}) = h$; (2) each agent of $\mathcal{G}$ is adaptable if $base = 1$, and resolute if $base = \infty$; (3) each agent is autonomous if $grad = 0$; autonomous or conformist if $grad = {}^0/{+}$; autonomous or dissenter if $grad = {}^0/{-}$; autonomous, conformist or dissenter if $grad$ is ${}^0/{\pm}$; (4) $\mathcal{G}$ is symmetric (i.e., undirected) if $graph = \texttt{sym}$ and arbitrary if $graph = \texttt{arb}$. E.g., $\mathcal{C}_2[1, {}^0/{+}, \texttt{sym}]$ is the class of environments where the domain of each agent has at most 2 variables, each agent is adaptable, either autonomous or conformist, and the underlying graph is symmetric.

A summary of our results is reported in Table 3. There, we separate all the fragments, defined as discussed above, where stable profiles are always guaranteed to exist (marked with the label *ever*), from those where $\exists$-NASH is instead an intractable problem (**NP**-complete). For the former fragments, we can show that Nash dynamics always converges in polynomial time—except in two cases.

5.1 Convergence and tractability

We start by identifying the cases in which a Nash dynamics is guaranteed to converge, implying that a Nash stable profile always exists. In particular, in these cases, we show that any dynamics converges after polynomially many steps.

Given an interpretation $I = \{l_1, \dots, l_{|I|}\}$, φ_I denotes the formula $l_1 \wedge \dots \wedge l_{|I|}$. An environment $\mathcal{G} = ([n], E, \kappa)$ is in *normal form* if, for each $i \in [n]$ with $\kappa(i) = \langle G_i^0, G_i^1 \rangle$, it holds that: (1) $\texttt{base}(i) \subset \mathbb{Q}^+ \cup \{0\}$ and $G_i^0 = \{(\varphi_I, G_i^0(I)) \mid I \in \texttt{space}(i)\}$; and (2) $\texttt{grad}(i) = \{g_i\} \subset \{-1, 0, 1\}$ and $G_i^1 = \{(\top, g_i)\}$. A social environment $\mathcal{G} = ([n], E, \kappa)$ in $\mathcal{C}_h[\infty, {}^0/{+}, \texttt{sym}]$ or in $\mathcal{C}_h[\infty, {}^0/{-}, \texttt{sym}]$ is in *discrete linear form* if, for each $i \in [n]$ with $\kappa(i) = \langle G_i^0, G_i^1 \rangle$, it holds that: (1) $\texttt{base}(i) \subset \{0, \dots, n \cdot 2^{2h+1}\}$ and $G_i^0 = \{(\varphi_I, G_i^0(I)) \mid I \in \texttt{space}(i)\}$; and (2) $\texttt{grad}(i) = \{g_i\} \subset \{-2^h - 1, 0, 2^h + 1\}$ and $G_i^1 = \{(\top, g_i)\}$.

THEOREM 5. *For each fixed $h > 0$ and for each social environment $\mathcal{G}$=$([n], E, \kappa)$ in $\mathcal{C}_h[\infty, 0, \texttt{arb}]$, $\mathcal{C}_h[\infty, {}^0/{+}, \texttt{sym}]$, or $\mathcal{C}_h[\infty, {}^0/{-}, \texttt{sym}]$, any Nash dynamics converges in polynomially many steps.*

PROOF. Any environment in $\mathcal{C}_h[\infty, 0, \texttt{arb}]$ converges in at most n steps after the best move of each agent. For the remaining two cases we provide a more sophisticated argument. First, we construct from $\mathcal{G}$ an environment $\mathcal{G}' = ([n], E, \kappa')$ in discrete linear form, sharing with $\mathcal{G}$ both the underlying graph and its dynamics. Intuitively, in κ' we get rid of fractional or "too large" values, and of negative values occurring in the bases. Consider an agent i with $\texttt{grad}(i) = \{m\}$. Let $\kappa(i) = \langle G_i^0, G_i^1 \rangle$, $\texttt{space}(i) = \{I_1, ..., I_c\}$ with $c \leqslant 2^h$. For each $x \leqslant c$, compute $\alpha_x = G_i^0(I_x)$. Tuple $\tau_0(i) = (m, \alpha_1, ..., \alpha_c)$ together with $\texttt{space}(i)$ is an alternative encoding of $\kappa(i)$. Assume $I_1, ..., I_c$ are sorted in a way that $x < y$ implies $\alpha_x \leqslant \alpha_y$. We build $\kappa'(i)$ by replacing $\tau_0(i)$ with the result $\tau_3(i)$ of the following transformations (step 1 modifies $\mathcal{G}$ to be in normal form, while the remaining two steps construct $\mathcal{G}'$).

1. Let $w = \min_x \alpha_x$. Build $\tau_1(i) = (\mathsf{sgn}(m), \beta_1, ..., \beta_c)$ as follows: if $w < 0$, then $\beta'_x = \alpha_x - w$, else $\beta'_x = \alpha_x$; moreover, if $m \neq 0$, then $\beta_x = \beta'_x/|m|$, else $\beta_x = \beta'_x$;

2. Build $\tau_2(i) = (\mathsf{sgn}(m), \gamma_1, ..., \gamma_c)$ as follows: if $\beta_1 < n$, then $\gamma_1 = \beta_1$, else $\gamma_1 = n$; if $\beta_x - \beta_{x-1} < n$, then $\gamma_x = \gamma_{x-1} + \beta_x - \beta_{x-1}$, else $\gamma_x = \gamma_{x-1} + n$;

3. Let $F = f_1, ..., f_c$ be the fractional parts of $\beta_1, ..., \beta_c$, and $\mu : F \to \{0, ..., c\}$ s.t. $\mu(0) = 0$ and $f_x < f_y$ implies $\mu(f_x) < \mu(f_y)$. $\tau_3(i) = (\mathsf{sgn}(m) \cdot z, \delta_1, ..., \delta_c)$ where $z = 2^{h+1}$ and each $\delta_x = z \times \lfloor \gamma_x \rfloor + \mu(f_x)$.

Note that each δ_x is linear in $\mathcal{G}$ as it belongs to the set $\{0, ..., z \cdot n \cdot c + c\}$. Consider now a profile Π. Let $com(\Pi)$ denote the edges of E connecting agents with compatible interpretations, and $\hat{z} = z$ in the case ${}^0/{+}$ or $\hat{z} = -z$ otherwise. Also, for each $i \in [n]$, let δ_i^Π be the weight in $\tau_3(i)$ associated with interpretation Π_i. Consider now the function

$$\Phi(\Pi) = \hat{z} \cdot (|com(\Pi)| + 1) + \sum_{j \in [n]} \delta_j^\Pi.$$

Assume that agent i with gradient $m_i \neq 0$ is going to move. Let $com_i(\Pi) = \{\{i, j\} \in E \mid j \in \texttt{partners}(i, \Pi)\}$, and $E_i = \{\{x, j\} \in E \mid x = i\}$. Function $\Phi(\Pi)$ can be rewritten as

$$\delta_i^\Pi + \hat{z} \cdot |1 + com_i(\Pi)| + \hat{z} \cdot |com(\Pi) \setminus E_i| + \sum_{j \neq i} \delta_j^\Pi.$$

The term $\delta_i^\Pi + \hat{z} \cdot |1 + com_i(\Pi)|$ is exactly the utility of i in Π, while the remaining terms do not change after the move of i from Π to Π'. Hence, $\Phi(\Pi') - \Phi(\Pi) = u_i(\Pi') - u_i(\Pi)$, implying that Φ behaves as an *exact potential function*. And, according to our normalization, Φ is also discrete and its modulus is bounded by $\mathcal{O}(n^2)$. But it is not potential any more in the presence of autonomous agents. In fact, if $m_i = 0$, then $u_i(\Pi') - u_i(\Pi) = \delta_i^{\Pi'} - \delta_i^\Pi$. Therefore, Φ is an exact potential function during any sub-dynamic involving non-autonomous agents only. However, since autonomous agents may move at most once in any dynamics, the full dynamics converges in $\mathcal{O}(n^3)$ steps. $\square$

Note that from the above proof, one can observe that the running time — although polynomial — exponentially depends on h, which is assumed here to be a fixed parameter. We now consider two fragments where a stable profile always exists and can be computed in polynomial time.

THEOREM 6. *For each fixed $h > 0$ and for each social environment $\mathcal{G}$=$([n], E, \kappa)$ in $\mathcal{C}_h[1, {}^0/{+}, \texttt{arb}]$, a stable profile always exist and can be computed in polynomial time. Yet, Nash dynamics do not necessarily converge.*

PROOF. Consider any interpretation $I : \texttt{dom}(\mathcal{G}) \to \{\top, \bot\}$. Build from I the profile $\Pi_I = \{i \mapsto I|_{\texttt{dom}(i)}\}_{i \in [n]}$. Clearly, the number of edges incoming to i is exactly $|\texttt{partners}(i, \Pi)|$. Hence, since agents are adaptable, Π_I is stable.

We now provide an environment $\mathcal{G} = ([n], E, \kappa)$ and an initial profile Π which does not necessarily converge. In particular,

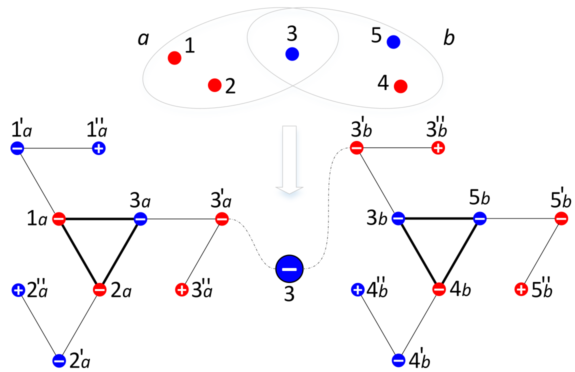

Figure 2: Gadget in the proof of Theorem 8.

$[n] = \{i, a_i, ..., d_i \mid 1 \leqslant i \leqslant 3\}$; $E = \{(1,3),(3,2),(2,1)\} \cup \{(x_i, i) \mid 1 \leqslant i \leqslant 3 \wedge\ a \leqslant x \leqslant c\} \cup \{(1,d_1),(d_1,3)\} \cup \{(3,d_3),(d_3,2)\} \cup \{(2,d_2),(d_2,1)\}$. Regarding κ, the domain of each agent is $\{p\}$, implying the possible interpretations I_1,I_2. Agents 1, 2, 3, d_1, d_2 and d_3 are adaptable conformists with base and gradient 1. The remaining agents are adaptable autonomous with base and gradient 0. Initial profile Π is s.t.: 1, a_1, b_1 ,d_1, c_2, c_3 play I_1, and the remaining agents play I_2. Since there are only two interpretations, it is sufficient to give the sequence of deviating agents which ends up with the initial profile: 3, d_3, 1, d_1, 2, d_2, 3, d_3, 1, d_1, 2, d_2. Since this sequence can repeat indefinitely, it does not necessarily converge. $\square$

THEOREM 7. *For each social environment $\mathcal{G}$=$([n], E, \kappa)$ in $\mathcal{C}_1[\infty, {}^0/{+}, \mathtt{arb}]$, a stable profile always exist and can be computed in polynomial time.*

PROOF. From the interpretation I that maps every variable of $\mathtt{dom}(\mathcal{G})$ to $\top$, we build the profile $\Pi_I = \{i \mapsto I|_{\mathtt{dom}(i)}\}_{i\in[n]}$. Since each agent i has at most one variable, say x, and since I maximizes already the number of her partners, we are sure that if i wants to change the value of x to $\bot$, then she will never reconsider this choice. Hence, we can fix $x = \bot$, remove i from the analysis, and iteratively apply the argument on the reduced environment. $\square$

5.2 Hardness for symmetric environments

We now proceed with the cases in which a stable profile might not exist. In particular, we show that the complexity results in Section 5.1 precisely identify the maximal subclass of tractable social environments. Contrasted with the **NP**-hardness results exhibited in Section 4, the results that follow (for symmetric and, later on, for arbitrary environments) are technically deeper and require more sophisticated elaborations. Indeed, it might be even as a surprise that hardness results can emerge in the settings we shall next analyze, due to the limited resources made available to the agents therein.

THEOREM 8. *For each fixed $h > 0$, problem $\exists$-NASH over $\mathcal{C}_h[1, {}^0/{\pm}, \mathtt{sym}]$ is **NP**-complete.*

PROOF. (Membership) Inherited from Theorem 2.

(Hardness) We first show the statement for $h = 1$. Deciding whether a 3-uniform hypergraph $H = (V, E)$ is 2-colorable is an **NP**-complete problem [28]. (A 2-coloring is valid if each (hyper)edge contains two distinct vertices mapped to different colors, namely no edge is monochromatic.) Based on H, we build the environment $\mathcal{G} = ([n], E', \kappa)$ and we claim that H is 2-colorable iff $\mathcal{G}$ admits a stable profile. In particular, (1) $[n] = \{u, u_e, u'_e, u''_e \mid u \in e$ and $e \in E\}$, (2) $u_e, v_e \in [n]$ implies $\{u_e, v_e\} \in E'$, (3) $u_e \in [n]$ implies $\{u_e, u'_e\} \in E'$, $\{u'_e, u''_e\} \in E'$, and $\{u'_e, u\} \in E'$, (4) for each $i \in [n]$, formula $(x, 0)$ belongs to G^0_i, and (5) for each $i \in [n]$, if i is of the form u''_e, then formula $(\top, 1)$ belongs to G^1_i, else formula $(\top, -1)$ belongs to G^1_i. Note that $\mathtt{dom}(\kappa) = \{x\}$. See Figure 2 for an example, where red nodes choose $\{x\}$ and blue ones choose $\{\neg x\}$. In particular, each closed triangle is stable iff it is not monochromatic. Moreover, each gadget of the form $\{u'_e, u''_e\}$ behaves as a “color propagator”. In fact, to preserve stability, it “propagates” to node u_e the color associated to node u.

To show that the problem remains hard for any fixed $h > 1$, we relay on Lemma 2 which requires a more sophisticated construction, which is given in the remaining part of this proof. (Note that more variables imply more interpretations, which in turn may satisfy more dissenter agents.)

For any $k \geqslant 3$, let C^k_x denote the complete graph on the vertices $\{x_1, ..., x_k\}$. Clearly, C^k_x is both k-chromatic[1] and $(k-1)$-regular[2]. Given a $k \geqslant 3$ and an edge $e = \{a, b\}$, we define graph $G^k_e = (V^k_e, E^k_e)$ from $C^k_e = (V, E)$ as follows: (1) $V^k_e = \{a, b\} \cup V$; and (2) $E^k_e = \{e\} \cup E \cup \{\{a, e_i\} \mid 1 \leqslant i \leqslant k-1\} \cup \{\{b, e_k\}\}$. See Figure 3 for an illustration of G^3_e in case $e = \{a, b\}$. Note that, whenever a and b are colored in the same way, after collapsing the vertices a and b we obtain a complete k-regular graph of order $k+1$ which cannot be k-colorable. (This fact is at the basis of Lemma 1.) More in general, G^k_e enjoys the following property.

PROPOSITION 1. *For each $k \geqslant 3$ and for each $e = \{a, b\}$, G^k_e is k-chromatic and each vertex of $\{e_1, ..., e_k\}$ has exactly k incident edges.*

Fix a $k \geqslant 3$. Consider a graph $H = (V, E)$, and the collection of graphs $\{G^k_e = (V^k_e, E^k_e) \mid e \in E\}$. We define the graph $H^k = (V^k, E^k)$ in such a way that $V^k = \bigcup_{e\in E} V^k_e$ and $E^k = \bigcup_{e\in E} E^k_e$. Of course, H^k can be constructed in polynomial time. Hereafter, for reasons that will be clear later, H^k is called k-*invalid*. Accordingly, given a graph H, kCOL-kINV is the problem of checking whether H^k is k-colorable.

PROPOSITION 2. *kCOL-kINV is **NP**-complete for any $k \geqslant 3$.*

PROOF. Let kCOL be the problem: given a graph, is it k-colorable? We reduce kCOL to kCOL-kINV, since kCOL is **NP**-complete for any given $k \geqslant 3$ [32]. More specifically, a graph H is k-colorable iff H^k is k-colorable.

($\Rightarrow$) Fix a $k \geqslant 3$. Let $H = (V, E)$, $H^k = (V^k, E^k)$, and $C = \{c_1, ..., c_k\}$ be k colors. By Proposition 1, from a valid k-coloring $\gamma : V \rightarrow C$ of H we obtain a valid k-coloring γ^k :

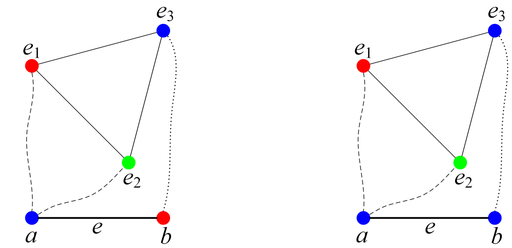

Figure 3: Graph G^3_e, connecting $e = \{a, b\}$ and C^3_e.

[1]A graph is called k-chromatic if it is k-colorable but it is not $(k-1)$-colorable.

[2]A graph is called k-regular if the number of edges incident to each vertex is exactly k.

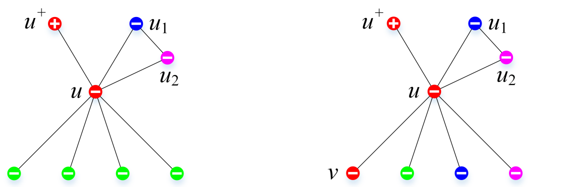

Figure 4: Gadgets in the proof of Lemma 2 for $k = 4$.

$V^k \to C$ of H^k as follows: (1) $u \in V$ implies $\gamma^k(u) = \gamma(u)$; and (2) $e = \{a, b\} \in E$ implies $\gamma^k(e_1) = \gamma(b)$, $\gamma^k(e_k) = \gamma(a)$, and γ^k restricted to the vertices $V_e^k \setminus \{a, b\}$ is any bijection to $C \setminus \{\gamma(a), \gamma(b)\}$. See Figure 3 (left-hand-side), for an example.

($\Leftarrow$) This direction is trivial since $E^k \supset E$. Hence, any k-coloring of H^k is also a k-coloring of H. $\square$

The next lemma clarifies the name k-invalid.

LEMMA 1. *For any given $k \geqslant 3$, if a graph H^k is not k-colorable then, for each (invalid) coloring, there exists a vertex having k incident edges which is colored as one of its neighbors.*

PROOF. Fix a $k \geqslant 3$. Let $H = (V, E)$, $H^k = (V^k, E^k)$, and $\gamma^k : V^k \to \{c_1, ..., c_k\}$ an invalid k-coloring of H^k. First, we observe that there is necessarily an edge $e = \{a, b\} \in V$ such that $\gamma(a) = \gamma(b)$. Otherwise, by Proposition 1, it is possible to modify γ^k to make it valid. Let us now focus on the vertices $\{e_1, ..., e_k\}$ (recall that each of them has exactly k incident edges). If for some $i \neq j$ we have that $\gamma^k(e_i) = \gamma(e_j)$, the statement is true. Conversely, since e_k is the only one in C_e^k not connected to a, it must hold that $\gamma^k(a) = \gamma^k(e_k)$. (See Figure 3, right-hand-side.) But this implies that $\gamma^k(e_k) = \gamma^k(b)$, which completes the proof. $\square$

We are now ready the complete the proof of our main result.

LEMMA 2. *For each fixed $h > 1$, problem* $\exists$-NASH *over* $\mathcal{C}_h[1, 0/\pm, \mathtt{sym}]$ *is* **NP***-complete.*

PROOF. Fix some $h > 1$. Let $k = 2^h$. We provide a reduction from kCOL-kINV to $\exists$-NASH where, for each agent i, $\mathtt{dom}(i) = \{x_1, ..., x_h\}$. Consider a k-invalid graph $H^k = (V, E)$, and let D collect all the vertices of V with k incident edges. We construct from H^k the social environment $\mathcal{G} = (V', E', \kappa)$ where each $u \in V$ is a dissenter, and where each $u \in D$ is connected both to the conformist agent u^+ and to the complete graph $C_u^{k-2} = (V_u, E_u)$ of dissenters. See Figure 4 for an example. More precisely, (1) $V' = V^- \cup V^+$, where $V^- = V \cup \bigcup_{u \in D} V_u$ and $V^+ = \{u^+ \mid u \in D\}$; (2) $E' = E \cup E^c \cup \{\{u, u^+\} \mid u \in D\} \cup \bigcup_{u \in D} E_u$, where $E^c = \{\{u, u_j\} \mid u \in D \text{ and } 1 \leqslant j \leqslant k-2\}$; (3) $i \in V'$ implies $G_i^0 = \{(\bigvee_{j \leqslant h} x_j, 0)\}$; (4) $i \in V^+$ implies $G_i^1 = \{(\top, 1)\}$, while $i \in V^-$ implies $G_i^1 = \{(\top, -1)\}$. Let $\mathcal{I} = \{I_1, ..., I_k\}$ be all the interpretations over $\{x_1, ..., x_h\}$. For notational convenience, we also use $\mathcal{I}$ as colors for H^k.

($\Rightarrow$) From a valid k-coloring $\gamma : V \to \mathcal{I}$ of H^k we construct a stable profile Π for $\mathcal{G}$. For each $u \in V$, $\Pi_u = \Pi_{u^+} = \gamma(u)$. To guarantee stability, for each $u \in D$ and for each interpretation $I \neq \Pi_u$, at least one of the other $2k - 2$ neighbors of u must be mapped to I. Let $\mathcal{J} = \{\gamma(v) \mid \{u, v\} \in E\}$. In the worst case, $|\mathcal{J}| = 1$. But since $\Pi_u \notin \mathcal{J}$, then the interpretations not already associated to the neighbors of u are at most $k - 2$. But these can be safely associated to the agents in V_u. See Figure 4 (left-hand-side).

($\Leftarrow$) We prove the contrapositive. Since $\mathcal{G}$ contains a supergraph of H^k, any profile Π of $\mathcal{G}$ necessarily maps two connected agents u and v to the same interpretation. And by Lemma 1 we can assume that the number of edges incident to u in H^k is k. Hence u has in total $2k - 1$ neighbors in $\mathcal{G}$. Assume now that Π is stable. Necessarily, $\Pi_u = \Pi_{u^+} = \Pi_v$. This imposes that, for each interpretation $I \neq \Pi_u$, at least two of the other $2k - 3$ neighbors of u must be mapped to I. But this is not possible since there are $k - 1$ interpretations and we would need at least $2k - 2$ neighbors. See Figure 4 (right-hand-side). $\square$

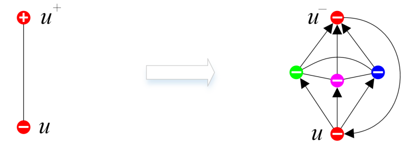

Figure 5: Gadgets of Lemma 2 and Theorem 9 resp., for $k = 4$.

This completes the proof of Theorem 8.

5.3 Hardness results for arbitrary environments

We now move to analyze the case where no restriction at all is considered on the underlying network topologies. In fact, a simple consequence of the above elaboration is stated below.

COROLLARY 1. *For each fixed $h > 0$, problem* $\exists$-NASH *over* $\mathcal{C}_h[1, 0/\pm, \mathtt{arb}]$ *is* **NP***-complete.*

The result can be further strengthened by focusing on autonomous and dissenter agents only (with gradient $0/-$).

THEOREM 9. *For each fixed $h > 0$, problem* $\exists$-NASH *over* $\mathcal{C}_h[1, 0/-, \mathtt{arb}]$ *is* **NP***-complete.*

PROOF SKETCH. The proof is an adaptation of the one of Theorem 8. Regarding the reduction from the 2-colorability of 3-uniform hypergraphs, the set $[n]$ is enriched by the set of nodes $\{u_e''' \mid u \in e \text{ and } e \in E\}$. Moreover, for each $u_e \in [n]$, the (undirected) edges $\{u_e, u_e'\}$, $\{u_e', u_e''\}$ and $\{u_e', u\}$ are replaced by the (directed) edges (u_e, u_e'), (u_e', u_e'''), (u_e''', u_e''), (u_e'', u_e'), and (u, u_e'). (The edges in the triangles remain undirected.) Finally, for each $i \in [n]$, formula $(x, 0)$ belongs to G_i^0 and formula $(\top, -1)$ belongs to G_i^1. See Figure 6 for an example. The gadget on the left hand side is extracted from Figure 2, while the one on the right is the new one. Both gadgets admit a stable profile iff node u and u_e have the same color.

Regarding the reduction from kCOL-kINV given in the proof of Lemma 2, we need similar tricks. Consider some $k = 2^h$. For each node $u \in D$, each agent u^+ is replaced by an agent u^-, and $k - 1$ extra agents, say $u_1, ..., u_{k-1}$, are added. All new agents are adaptable dissenters. Agents $u_1, ..., u_{k-1}$ form a symmetric clique. Moreover, for each $i \in \{1, ..., k - 1\}$ we add the extra directed edges (u, u_i) and (u_i, u^-). See Figure 5 for an example when $k = 4$. The extra agents, as dissenters, try to get different colors, which force u^- to have the same color with u. $\square$

COROLLARY 2. *For each fixed $h > 0$, problem* $\exists$-NASH *over* $\mathcal{C}_h[\infty, 0/-, \mathtt{arb}]$ *is* **NP***-complete.*

We conclude our analysis with a reduction from the satisfiability of 3-CNF Boolean formulas, which proves the following result. In particular, we construct an environment where, for each $i \in [n]$, $\mathtt{dom}(i) = \{x, y\}$ and agent i is a conformist. Moreover, we exploit a gadget, inspired by the one given in [31], that is stable if and only if the given formula is satisfiable.

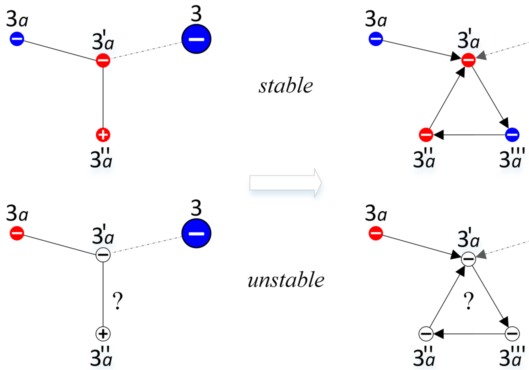

Figure 6: Gadget in Theorem 8 vs. gadget in Theorem 9.

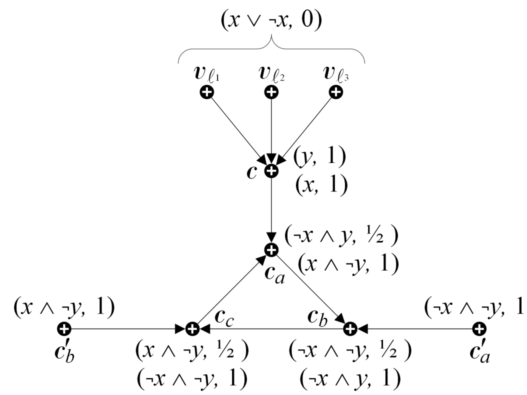

Figure 8: Second gadget in the proof of Theorem 10.

THEOREM 10. *For each fixed $h > 1$, problem* $\exists$-NASH *over* $\mathcal{C}_h[\infty, {}^0/{+}, \texttt{arb}]$ *is* **NP**-*complete.*

PROOF SKETCH. We reduce 3SAT to $\exists$-NASH where, for each $i \in [n]$, $\texttt{dom}(i) = \{x, y\}$ and agent i is a conformist. Given a 3-CNF formula $\phi = c_1 \wedge ... \wedge c_m$ with $\texttt{dom}(\phi) = \{z_1, \ldots, z_k\}$, we construct the social environment $\mathcal{G} = (V, E, \kappa)$ as follows. First, for each $i \in V$, $G_i^1 = (\top, 1)$. For each variable $z \in \texttt{dom}(\phi)$, we add to $\mathcal{G}$ the gadget given in Figure 7, where the weighted formulas refer to the goalbases of degree 0.

This gadget is stable iff the interpretations of v_z and $v_{\neg z}$ are incompatible. For each clause $c = \ell_1 \vee \ell_2 \vee \ell_3$ we add to $\mathcal{G}$ the gadget given in Figure 8 (still weighted formulas refer to the goalbases of degree 0), inspired by the one given in [31].

Thus, ϕ is satisfiable iff $\mathcal{G}$ admits a stable profile. In particular, a profile Π of $\mathcal{G}$ is stable only if each of the triangles (of vertices c_a, c_b and c_c) given in Figure 8 is stable, which can happen only if each Π_c maps both x and y to $\top$. But this means that necessarily, for some ℓ of c, the value of x in Π_{v_ℓ} is true (this intuitively means that clause c is satisfied) and $u_c(\Pi) \geqslant 4$. Otherwise if, for each ℓ of c, the value of x in Π_{v_ℓ} is false, then $u_c(\Pi) = 3$ and Π_c could be changed to map x to $\perp$ and y to $\top$ to have $u_c(\Pi) = 5$. Finally, we observe that one can add harmless variables to show that the statement holds for each $h > 2$. $\square$

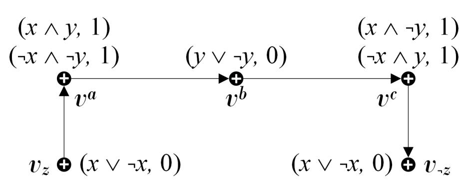

Figure 7: First gadget in the proof of Theorem 10.

6. DISCUSSION AND CONCLUSION

We have proposed and studied a setting for analyzing influence phenomena over social networks, where the reasoning capabilities of the agents are modeled via weighted propositional logic. Moreover, it shares the perspective of the work by [5], by allowing to deal with expressed opinions differing from innate ones (see Example 2). In fact, the setting studied in that work (adopting *generalized discrete preferences*) is very close to our class $\mathcal{C}_h[\infty, {}^0/{+}, \texttt{sym}]$. Both of them admit convergent Nash dynamics only, and are able to reason about internal beliefs and social constraints. However, that approach uses monotone non-decreasing real functions by assuming exactly one belief for agent, while we offer a more logic-based approach that admits multiple beliefs, even though our utility functions are more specific. Another recent study sharing some features with our formalization is the one by [22]. There, propositional logic is used to model the opinions of the agents, but there is no difference between expressed opinions and innate ones. Indeed, agents are not associated with utility functions and the diffusion process does not follow Nash dynamics. In addition, convergence processes are studied w.r.t. the topologies of the underlying graphs rather than w.r.t. the knowledge bases of the agents.

Our results include a thorough complexity analysis precisely isolating those scenarios for which Nash equilibria can be easily computed (and Nash dynamics are guaranteed to converge). Avenues of further research naturally include the study of well-known game-theoretic concepts in our setting, such as the price of anarchy and stability. Moreover, our results are mainly focused on linear agents but it would be interesting to conduct a similar analysis for agents of degree 2. Indeed, concave functions are powerful functions which have been already recognized as an important class [34]. In particular, they can capture autonomous, conformist and dissenter agents, and may also represent inflation phenomena: the utility of an agent normally behaving as a conformist can decrease if she gets too many partners (hence, exceeding the value where the parabola has its maximum).

Finally, the paper has now precisely charted the tractability frontier for reasoning on social environments. Hence, it would be relevant to focus on the maximal tractable subclass identified in Section 5.1 by pairing our results with machine learning techniques –aimed to discover the attitudes of linear agents based on their logged interactions. This way, one can assess the ability of these classes to predict evolutions in (portions of) real-world social environments, such as Facebook or Twitter.

Acknowledgments

The work was supported by project "Ba2Know (Business Analytics to Know) Service Innovation - LAB", No. PON03PE_00001_1 funded by the Italian Ministry of University and Research (MIUR), and by project "Smarter Solutions in the Big Data World (S2BDW)", funded by the Italian Ministry for Economic Development (MISE) within the programme PON "Imprese e competitivitá" 2014-2020. G. Greco's work was partially supported by a Kurt Gödel Research Fellowship, awarded by the Kurt Gödel Society and founded by the John Templeton Foundation.

REFERENCES

[1] M. Abramowitz. *Handbook of Mathematical Functions, With Formulas, Graphs, and Mathematical Tables*. Dover Publications, Incorporated, 1974.

[2] D. Acemoglu and A. Ozdaglar. Opinion dynamics and learning in social networks. *Dynamic Games and Applications*, 1(1):3–49, 2011.

[3] A. Anagnostopoulos, R. Kumar, and M. Mahdian. Influence and correlation in social networks. In *Proc. of KDD 2008*, pages 7–15. ACM, 2008.

[4] S. Aral, L. Muchnik, and A. Sundararajan. Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proceedings of the National Academy of Sciences*, 106(51):21544–21549, 2009.

[5] V. Auletta, I. Caragiannis, D. Ferraioli, C. Galdi, and G. Persiano. Generalized discrete preference games. In S. Kambhampati, editor, *Proc. of IJCAI 2016*, pages 53–59. IJCAI/AAAI Press, 2016.

[6] H. Aziz and R. Savani. *Handbook of Computational Social Choice*, chapter Hedonic games. Cambridge University Press, 2015.

[7] C. Ballester. Np-completeness in hedonic games. *Games and Economic Behavior*, 49(1):1 – 30, 2004.

[8] K. Bhawalkar, S. Gollapudi, and K. Munagala. Coevolutionary opinion formation games. In *Proc. of STOC 2013*, pages 41–50. ACM, 2013.

[9] D. Bindel, J. Kleinberg, and S. Oren. How bad is forming your own opinion? In *Proceedings of the 2011 IEEE 52Nd Annual Symposium on Foundations of Computer Science*, FOCS '11, pages 57–66. IEEE Computer Society, 2011.

[10] D. Bindel, J. Kleinberg, and S. Oren. How bad is forming your own opinion? *Games and Economic Behavior*, 92:248–265, 2015.

[11] A. Borodin, Y. Filmus, and J. Oren. Threshold models for competitive influence in social networks. In *Internet and Network Economics*, pages 539–550. Springer, 2010.

[12] T. Carnes, C. Nagarajan, S. M. Wild, and A. van Zuylen. Maximizing influence in a competitive social network: A follower's perspective. In *Proceedings of the Ninth International Conference on Electronic Commerce*, pages 351–360. ACM, 2007.

[13] M. Cha, A. Mislove, and K. P. Gummadi. A measurement-driven analysis of information propagation in the flickr social network. In *Proceedings of the 18th International Conference on World Wide Web*, pages 721–730. ACM, 2009.

[14] R. B. Cialdini and N. J. Goldstein. Social influence: Compliance and conformity. *Annu. Rev. Psychol.*, 55:591–621, 2004.

[15] S. Coste-Marquis, J. Lang, P. Liberatore, and P. Marquis. Expressive power and succinctness of propositional languages for preference representation. In D. Dubois, C. A. Welty, and M. Williams, editors, *Proc. of KR 2004*, pages 203–212. AAAI Press, 2004.

[16] E. David and K. Jon. *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge University Press, 2010.

[17] M. H. DeGroot. Reaching a consensus. *J. American Statistical Association*, 69:118–121, 1974.

[18] P. DeMarzo, D. Vayanos, and J. Zwiebel. Persuasion bias, social influence, and unidimensional opinions. *The Quarterly Journal of Economics*, 118(3):909–968, 2003.

[19] N. E. Friedkin and E. C. Johnsen. Social influence and opinions. *Journal of Mathematical Sociology*, 15:193–206, 1990.

[20] B. Golub and M. O. Jackson. Naïve learning in social networks and the wisdom of crowds. *American Economic Journal: Microeconomics*, 2(1):112–49, February 2010.

[21] C. P. Gomes, H. Kautz, A. Sabharwal, and B. Selman. Chapter 2 satisfiability solvers. In V. L. Frank van Harmelen and B. Porter, editors, *Handbook of Knowledge Representation*, volume 3 of *Foundations of Artificial Intelligence*, pages 89 – 134. Elsevier, 2008.

[22] U. Grandi, E. Lorini, and L. Perrussel. Propositional opinion diffusion. In *Proc. of AAMAS 2015*, pages 989–997. International Foundation for Autonomous Agents and Multiagent Systems, 2015.

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

[24] G. Greco and J. Lang. Group decision making via weighted propositional logic: Complexity and islands of tractability. In Q. Yang and M. Wooldridge, editors, *Proc. of IJCAI 2015*, pages 3008–3014. AAAI Press, 2015.

[25] X. He, G. Song, W. Chen, and Q. Jiang. Influence blocking maximization in social networks under the competitive linear threshold model. In *SDM*, pages 463–474. SIAM, 2012.

[26] M. O. Jackson. *Social and Economic Networks*. Princeton University Press, Princeton, NJ, USA, 2008.

[27] D. Kempe, J. Kleinberg, and E. Tardos. Maximizing the spread of influence through a social network. In *Proc. of KDD 2003*, pages 137–146. ACM, 2003.

[28] L. Lovász. Coverings and colorings of hypergraphs. In *Proceedings of the 4th Southeastern Conference on Combinatorics, Graph Theory and Computing*, Utilitas Mathematica, pages 3–12, 1973.

[29] N. Pathak, A. Banerjee, and J. Srivastava. A generalized linear threshold model for multiple cascades. In *Data Mining (ICDM), 2010 IEEE 10th International Conference on*, pages 965–970, 2010.

[30] D. Peters. Complexity of hedonic games with dichotomous preferences. In D. Schuurmans and M. P. Wellman, editors, *Proc. of AAAI 2016*, pages 579–585. AAAI Press, 2016.

[31] S. Simon and K. Apt. Choosing products in social networks. In *Internet and Network Economics*, volume 7695, pages 100–113. Springer, 2012.

[32] L. Stockmeyer. Planar 3-colorability is polynomial complete. *SIGACT News*, 5(3):19–25, July 1973.

[33] J. Tang, S. Wu, and J. Sun. Confluence: Conformity influence in large social networks. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 347–355. ACM, 2013.

[34] J. Uckelman, Y. Chevaleyre, U. Endriss, and J. Lang. Representing utility functions via weighted goals. *Math. Log. Q.*, 55(4):341–361, 2009.