A two-tiered formalization of social influence

Zoé Christoff[1] and Jens Ulrik Hansen[2]

[1] Institute for Logic, Language and Computation, University of Amsterdam
[2] Department of Philosophy, Lund University

Abstract. We propose a new dynamic hybrid logic to reason about social networks and their dynamics building on the work of "Logic in the Community" by Seligman, Liu and Girard. Our framework distinguishes between the purely private sphere of agents, namely their mental states, and the public sphere of their observable behavior, i.e., what they seem to believe. We then show how such a distinction allows our framework to model many social phenomena, by presenting the case of pluralistic ignorance as an example and discussing some of its dynamic properties.

In recent years, information dynamics and belief formation in groups of interacting agents have been widely studied within the field of *logic* [1]. The topic has also been extensively studied for agents situated in a network structure within the field of *social network analysis* [2, 3]. However, there has been very little information flow between these two fields. A recent exception is the work on influence in a community and peer pressure effects by Patrick Girard, Fenrong Liu, and Jeremy Seligman [4–8]. This paper attempts to continue building the bridge between the two fields.

A possible explanation for the lack of interaction between logic and social network analysis is their very distinct paradigmatic cases of inspiration. In social network research, the inspiration mainly comes from diffusion phenomena such as the spreading of diseases. In the logic tradition, rational agents are taken to be equipped with unlimited higher-order reasoning powers aiming for the truth. The work of Girard, Liu, and Seligman goes in both directions. On the one hand, their initial work [4–6] presents an extremely simple model of how knowledge, belief and preferences change under influence within social networks. On the other hand, their latest work [7, 8] aims at fully describing information dynamics in networks of agents with unlimited higher-order reasoning power.

Our goal is to design a framework to model real-life social phenomena and their corresponding information dynamics. As we will show, the setting of [6] cannot model situations involving a discrepancy between what the agents actually believe and what they seem to believe. Yet, we claim that the very possibility of such a discrepancy is an important feature of many social phenomena. However, we do not want to turn to much more complex frameworks such as [7] either. Therefore, we will build on the setting of [6] to design a framework which remains relatively simple but is capable of capturing more complex social phenomena.

In the next section, we briefly recall the "one-layer" framework of Seligman, Girard and Liu and we explain why it cannot model some particular social phenomena. In Section 2, we give an example of such a phenomenon, known from social psychology as *pluralistic ignorance* – a situation where all individuals of a group believe that their private attitudes differ from the ones of the rest of the group despite the fact that everyone in the group acts identically. This example illustrates the need for a more fine-grained definition of social influence which we then offer, by distinguishing what agents *privately believe* from how they *publicly behave*. In Section 3, we introduce a new general hybrid logic to reason about network dynamics taking into account these two different layers. Finally, in Section 4, we model the case of pluralistic ignorance and characterize some of its dynamic properties within this new framework.

1 The network logic of Girard, Liu, and Seligman

In [6], a hybrid logic in the original Facebook logic style of [4] is designed to model belief change induced by social influence in a community. The social network structure is represented by a set of agents and an irreflexive and symmetric relation (as in a real Facebook friendship) between them. A modal operator F quantifies over friends (or accessible agents): F reads "all of my friends" and its dual, $\langle F\rangle$, "some of my friends". Some hybrid logic machinery is also used: *nominals*, to refer to the agents, and operators $@_i$, to switch the evaluation point to the unique agent named by i. Each agent is always in one of the three following doxastic states, relatively to a given proposition φ: either she believes that φ ($B\varphi$), or she believes that $\neg\varphi$ ($B\neg\varphi$), or she is undecided about φ: ($U\varphi$ – an abbreviation of $\neg B\varphi \wedge \neg B\neg\varphi$). Sentences are interpreted indexically at an agent: if p means "I am blonde", BFp reads "I believe that all my friends are blonde" and FBp reads "each of my friends believes that s/he is blonde".

This static framework is combined with an influence operator to represent how belief repartition changes in a community, according to the following *peer pressure principle*: every agent tends to align her belief with the ones of her friends. The notions of Strong Influence and Weak Influence are defined, corresponding respectively to the belief changing operators of revision and contraction in the tradition of [9]. An agent is *strongly influenced* (SI) to believe φ when *all* of her friends (and at least one) believe that φ:

$$SI\varphi := FB\varphi \wedge \langle F\rangle B\varphi$$

An agent under strong influence with φ will come to believe φ too (assuming that revision is successful) whatever her initial attitude towards φ. An agent is already *weakly influenced* (WI) with φ when *some* of her friends believe that φ and none of her friends believe that $\neg\varphi$:

$$WI\varphi := F\neg B\neg\varphi \wedge \langle F\rangle B\varphi$$

Under weak influence, if the agent was undecided or if she already believed that φ, nothing changes; but if she believed that $\neg\varphi$, she will drop her belief and become undecided.

This simple framework makes it unproblematic to identify the stability and stabilization conditions of social-doxastic configurations, both of which can be characterized directly in the language of friendship and belief. However, this simplicity is pricey: even though this is not explicitly mentioned as such, it relies on an extremely strong assumption: *agents' belief states are influenced directly by their friends' belief states.* Thus, either all agents have direct access to their friends' beliefs (as mind-readers would), or their observed behavior always reflects their private beliefs, i.e., there is no difference between what they *seem to believe* and what they *actually believe.* This *transparency* assumption (all agents always automatically know what their friends believe) trivially rules out the modeling of situations where agents act in a way which does *not* reflect their mental states.[3]

Similar issues arise for preferences. Indeed, even if we agree that you are influenced in the very similar way described in [5], if you end up wearing a hat rather than none, it is probably not directly because all of your friends privately prefer to wear one too, but because they *act as if* they did. They could all be pretending because they all observe that everybody else is wearing a hat, and everyone could be following a trend that nobody actually likes. This is a crucial component of social science if we think for instance about real life cases where agents are enforcing a norm which they individually do not agree with. It is precisely because they do not have access to each other's preferences and beliefs that a collective behavior can result which goes against the opinions of most or even all agents, considered individually.[4] In the next section we will consider a class of similar situations called *pluralistic ignorance* and develop a two-layer notion of social influence to represent the distinction between what an agent privately believes and what beliefs she publicly expresses.

2 Pluralistic ignorance and a two-layer definition of influence

The term "pluralistic ignorance" originates in the social and behavioral sciences in the work Allport and Katz [11]. It can be roughly defined as a situation where each individual of a group believes that her private attitude towards a proposition or norm differs from the rest of the group members', even though everyone in the group acts identically. For instance, after a difficult lecture which none of the students understood, it can happen that none of them asks any question even though the teacher explicitly requested them to do so in case they did not understand the material. There are numerous examples of pluralistic ignorance in the social and psychological literature such as, in addition to this classroom

[3] Such an additional layer is also necessary for cases involving higher-order beliefs, since the complexity of such cases usually arises precisely from the fact that there might be a difference between what agent a believes that agent b believes and what agent b actually believes. However, we will not pursue the issue of higher-order beliefs any further in this paper.

[4] See for instance [10] on this issue.

example, drinking among college students, attitudes towards racial segregation, and many more.[5]

Even though different definitions have been given in the literature [16–18, 11, 19], we will follow [19] and define pluralistic ignorance as *a collective discrepancy between the agents' private attitudes and their public behavior*, namely a situation where all the individuals of a group have the same private attitude towards a proposition φ (say a belief in φ), but publicly "display" a conflicting attitude towards φ (say a belief in $\neg\varphi$).

From a dynamic perspective, pluralistic ignorance is often reported as being both a *robust* and *fragile* phenomenon. It is robust in the sense that, if nothing changes in the environment, the phenomenon might persist over a long period of time – the college students might keep obeying an unwanted drinking norm for generations. On the other hand, it is fragile in the sense that if just one agent announces her private belief, it may be enough to dissolve the phenomenon – if just one student of the classroom example starts to ask questions about the difficult lecture the rest of the students might soon follow. The two-layer definition of social influence which we develop below will allow us to explain how pluralistic ignorance may dissolve in a community by cascading effects and thus allow us to illustrate both its robustness and its fragility. Moreover, in Section 4 we will show formal results about these dynamic properties of pluralistic ignorance.

To reflect the fact that agents do *not* have access to what the others privately believe, we introduce a distinction between *private belief*, which we name "inner belief" (I_B) and *public (or observable) behavior*, which we name "expressed belief" (E_B). We define *two* undecidedness or "unbelief" notions accordingly:

$$U_{IB}\varphi := \neg I_B\varphi \wedge \neg I_B\neg\varphi \qquad (\textit{inner}\text{ unbelief})$$

$$U_{EB}\varphi := \neg E_B\varphi \wedge \neg E_B\neg\varphi \qquad (\textit{expressed}\text{ unbelief})$$

To define our new influence operator, we make the following simplifying assumption: from the subjective perspective of each agent, what matters (what influences her) is what she herself privately believes and what the others seem to believe. This reflects the fact that influence occurs (at least in good part) at the behavioral (observable, visible, displayed) level. We now redefine strong and weak influence accordingly: *2-layer strong influence* (SI^2) with respect to φ is the situation where all (and some) of my friends express the belief that φ.

$$SI^2 := FE_B\varphi \wedge \langle F\rangle E_B\varphi$$

In this case, whatever my own initial (inner and expressed) state, I end up expressing the belief that φ ($E_B\varphi$). Similarly, *2-layer weak influence* (WI^2) with

[5] An extensive study of the classroom phenomenon was done by Miller and McFarland [12]. In a study of college students, Prentice and Miller [13] found that most students believed that the average student was much more comfortable with alcohol norms than they themselves were. Fields and Schuman [14] conducted a similar study, which showed that on issues of racial and civil liberties most people perceived others to be more conservative than they actually were. O'Gorman and Garry [15] found a similar tendency among whites to overestimate other whites' support for racial segregation.

respect to φ is the situation where some of my friends express the belief that φ and none of them expresses the belief that $\neg\varphi$.

$$WI^2 := \langle F\rangle E_B\varphi \wedge F\neg E_B\neg\varphi$$

As a result, I will express the belief that φ ($E_B\varphi$) if I was initially privately undecided about φ (U_{IB}) or if I already privately believed that φ ($I_B\varphi$), and I will act as if I was indifferent ($U_{EB}\varphi$) if I initially privately believed $\neg\varphi$ ($I_B\varphi$).

According to these definitions, my reaction depends on asymmetrical information: what *I privately believe* and what *the others seem to believe*. This reflects the fundamental asymmetry between the first and third person perspectives which is needed to model pluralistic ignorance. It is symmetrical in that everybody reacts in the same way and in that everybody interprets the behavior of others in the same way; but it is asymmetric in that people don't have access to others' mental states and have a "privileged" access to their own.

	Inner state	$\langle F\rangle E_B\varphi$	$\langle F\rangle E_B\neg\varphi$	$\langle F\rangle E_U\varphi$	Type 1	Type 2	Type 3
1	$I_B\varphi$				$\rightsquigarrow E_B\varphi$	$\rightsquigarrow E_U\varphi$	$\rightsquigarrow E_B\varphi$
2	$I_B\neg\varphi$	1	1	1	$\rightsquigarrow E_B\neg\varphi$	$\rightsquigarrow E_U\varphi$	$\rightsquigarrow E_B\neg\varphi$
3	$I_U\varphi$				$\rightsquigarrow E_U\varphi$	$\rightsquigarrow E_U\varphi$	$\rightsquigarrow E_U\varphi$
4	$I_B\varphi$				$\rightsquigarrow E_B\varphi$	$\rightsquigarrow E_U\varphi$	$\rightsquigarrow E_B\varphi$
5	$I_B\neg\varphi$	1	1	0	$\rightsquigarrow E_B\neg\varphi$	$\rightsquigarrow E_U\varphi$	$\rightsquigarrow E_B\neg\varphi$
6	$I_U\varphi$				$\rightsquigarrow E_U\varphi$	$\rightsquigarrow E_U\varphi$	$\rightsquigarrow E_U\varphi$
7	$I_B\varphi$				$\rightsquigarrow E_B\varphi$	$\rightsquigarrow E_B\varphi$	$\rightsquigarrow E_B\varphi$
8	$I_B\neg\varphi$	1	0	1	$\rightsquigarrow E_U\varphi$	$\rightsquigarrow E_U\varphi$	$\rightsquigarrow E_U\varphi$
9	$I_U\varphi$				$\rightsquigarrow E_U\varphi$	$\rightsquigarrow E_U\varphi$	$\rightsquigarrow E_U\varphi$
10	$I_B\varphi$						
11	$I_B\neg\varphi$	1	0	0	$\rightsquigarrow E_B\varphi$	$\rightsquigarrow E_B\varphi$	$\rightsquigarrow E_B\varphi$
12	$I_U\varphi$						
13	$I_B\varphi$				$\rightsquigarrow E_U\varphi$	$\rightsquigarrow E_U\varphi$	$\rightsquigarrow E_U\varphi$
14	$I_B\neg\varphi$	0	1	1	$\rightsquigarrow E_B\neg\varphi$	$\rightsquigarrow E_B\neg\varphi$	$\rightsquigarrow E_B\neg\varphi$
15	$I_U\varphi$				$\rightsquigarrow E_B\neg\varphi$	$\rightsquigarrow E_B\neg\varphi$	$\rightsquigarrow E_B\neg\varphi$
16	$I_B\varphi$						
17	$I_B\neg\varphi$	0	1	0	$\rightsquigarrow E_B\neg\varphi$	$\rightsquigarrow E_B\neg\varphi$	$\rightsquigarrow E_B\neg\varphi$
18	$I_U\varphi$						
19	$I_B\varphi$				$\rightsquigarrow E_B\varphi$	$\rightsquigarrow E_B\varphi$	$\rightsquigarrow E_U\varphi$
20	$I_B\neg\varphi$	0	0	1	$\rightsquigarrow E_B\neg\varphi$	$\rightsquigarrow E_B\neg\varphi$	$\rightsquigarrow E_U\varphi$
21	$I_U\varphi$				$\rightsquigarrow E_U\varphi$	$\rightsquigarrow E_U\varphi$	$\rightsquigarrow E_U\varphi$
22	$I_B\varphi$				$\rightsquigarrow E_B\varphi$	$\rightsquigarrow E_B\varphi$	$\rightsquigarrow E_B\varphi$
23	$I_B\neg\varphi$	0	0	0	$\rightsquigarrow E_B\neg\varphi$	$\rightsquigarrow E_B\neg\varphi$	$\rightsquigarrow E_B\neg\varphi$
24	$I_U\varphi$				$\rightsquigarrow E_U\varphi$	$\rightsquigarrow E_U\varphi$	$\rightsquigarrow E_U\varphi$

Fig. 1. Influence on three different types of agents

Figure 1 lists the 24 possible situations of an individual among her friends, from her perspective, and describes her (observable) reaction. Her private attitude appears in the first column, the possible repartition of her friends' behaviors (expressed belief states) in columns 2,3,4 (in a truth table format – 1 for "true" and 0 for "false"), and her resulting behavior in one of the last three columns (depending of which type of agents we are considering). It is easy to see that our strong influence (rows 10 to 12 and 16 to 18 of the table) is still similar to the one from Seligman et al. but defined on the level of "expressed belief" instead of what was simply called "belief". However, weak influence (when not strong, rows 7 to 9 and 15 to 18) now results in a different state depending on the initial *private* belief state of the agent herself (see for instance rows 7 and 8).

There are two possible cases in which I have friends (unlike in rows 22 to 24) but I am neither strongly nor weakly influenced: whenever all of my friends express undecidedness (rows 19 to 21) and whenever some of them express the belief that φ while some express the belief that $\neg\varphi$ (rows 1 to 6). In the setting of [6], nothing happens, i.e, the agent continues to believe whatever she did before. In our setting, we have to make a choice as to what the agent expresses. The simplest one is to assume that in both these cases, agents express their true private belief (act sincerely). This corresponds to agent of type 1 in the table. However, some agents might be more inclined to follow the others, and in different ways. Types 2 and 3 in the table are examples of other possible types of agents which still comply with our definition of two-layer strong and weak influence. If I am a type 2 agent, I will be sincere (i.e., my expressed belief state will correspond to my inner belief state) whenever I face no opposition. For instance, if I privately believe that φ, I will express this belief if none of my friends expresses a belief in $\neg\varphi$. And if I am a type 3 agent, I will be sincere whenever some of my friends express support for my private belief state, I will for instance express my inner belief in φ if some of my friends express a belief in φ too. Type 1 agents are thus simply the ones that are sincere in both cases: when they get some support and when they face no opposition.

We will see in section 4 how the dynamic properties of social phenomena like pluralistic ignorance depend on the type of agents involved but let us first introduce the formal framework we will use to represent changes of the (multi-layered) state of agents in a social network.

3 A Hybrid Network Logic

In this section, we introduce a hybrid logic to reason about networks and their dynamics, which will allow us to model cases like pluralistic ignorance. We start with a static logic and then move on to give the full dynamics.

In section 2 we introduced two characteristics of each agent, namely her inner (private) belief state and her expressed belief state. Each of the two could be of three kinds. For instance, the inner belief state could be inner belief, inner non-belief, or inner undecidedness. We will generalize this idea by assuming that each agent has n different characteristics, each of which is taken from a finite set of possible values. More formally, we assume a finite set of variables/characteristics $\{V_1, V_2, ..., V_n\}$, where each variable V_l takes a value from a finite set R_l, for each $l \in \{1, ..., n\}$.

The atomic propositions of our language will then be of the form

$$V_l = r,$$

for an $l \in \{1, \ldots, n\}$ and an $r \in R_l$. We will refer to these as characteristic propositions and we will refer to the set of all characteristic propositions as PROP. If the proposition $V_l = r$ is true of an agent, we will read it as the agent possessing the particular characteristic r of type V_l.[6]

In addition to characteristic propositions, we will assume a countable infinite set of nominals (NOM) used as names for agents in possible networks, just as nominals are used to refer to possible states in traditional hybrid logic [20]. The syntax of our static language is then be given by:

$$\varphi \ ::= \ p \ | \ i \ | \ \neg\varphi \ | \ \varphi \wedge \varphi \ | \ F\varphi \ | \ G\varphi \ | \ @_i\varphi \ ,$$

where $p \in \mathsf{PROP}$ and $i \in \mathsf{NOM}$. We will use the standard abbreviations for $\vee$, $\rightarrow$, and $\leftrightarrow$ and denote the dual operator of F by $\langle F \rangle$ and the dual of G by $\langle G \rangle$. The intuitive meaning of the F and $@_i$ operators were already discussed in Section 1. The G-operator is the global modality quantifying over all agents in the network and $G\varphi$ is read as "all agents (satisfy) φ".

We now move on to define the semantics of our language. A (network) model is a tuple $\mathcal{M} = (A, \sim, g, \nu)$, where A is a non-empty set of agents, $\sim$ is a binary relation on A representing the network structure[7], $g : \mathsf{NOM} \rightarrow A$ is a function assigning an agent to each nominal, and $\nu : A \rightarrow \mathcal{V}$ is a valuation assigning characteristics to all agents in the network. Here $\mathcal{V}$ denotes the set of all *assignments* $s : \{1, ..., n\} \rightarrow R_1 \times ... \times R_n$. Hence, an assignment assigns a value in R_l to each variable V_l and given an agent $a \in A$, $\nu(a)$ is an assignment assigning characteristics to a for all variables $V_1, ..., V_n$.

Given a $\mathcal{M} = (A, \sim, g, \nu)$, an $a \in A$ and a formula φ, we define the truth of φ at a in $\mathcal{M}$ inductively by:

$$\begin{array}{lll}
\mathcal{M}, a \models V_l = r & \text{iff} & \nu(a)(l) = r \\
\mathcal{M}, a \models i & \text{iff} & g(i) = a \\
\mathcal{M}, a \models \neg\varphi & \text{iff} & \text{it is not the case that } \mathcal{M}, a \models \varphi \\
\mathcal{M}, a \models \varphi \wedge \psi & \text{iff} & \mathcal{M}, a \models \varphi \text{ and } \mathcal{M}, a \models \psi \\
\mathcal{M}, a \models G\varphi & \text{iff} & \text{for all } b \in A; \mathcal{M}, b \models \varphi \\
\mathcal{M}, a \models F\varphi & \text{iff} & \text{for all } b \in A; a \sim b \text{ implies } \mathcal{M}, b \models \varphi \\
\mathcal{M}, a \models @_i\varphi & \text{iff} & \mathcal{M}, g(i) \models \varphi
\end{array}$$

Satisfiability, validity etc. are as usual. To obtain the full dynamic language, we add dynamic modalities, which, as in standard Dynamic Epistemic Logic [21,

[6] Characteristic propositions are obviously a generalization of classical propositional variables.

[7] If we are talking about undirected networks, we will assume that $\sim$ is symmetric.

22], come from event models. On the syntactic level, given an event model $\mathcal{E}$ and a formula φ, we will add the construct $[\mathcal{E}]\varphi$ to our language. Event models are defined by simultaneous induction with the syntax of the language: An *event model* is a pair $\mathcal{E} = (\Phi, \mathsf{post})$ consisting of a finite set Φ of pairwise inconsistent formulas of our language and a post-condition function $\mathsf{post} : \Phi \to \mathcal{V}$. The set Φ will be referred to as "preconditions", and given a precondition $\varphi \in \Phi$, we will call $\mathsf{post}(\varphi) \in \mathcal{V}$ the post-condition of φ. The intuition behind this is that if an agent satisfy a $\varphi \in \Phi$ (in which case, φ is necessarily unique), then after the event $\mathcal{E}$, a will have the characteristics specified by $\mathsf{post}(\varphi)$.

As in standard Dynamics Epistemic Logic, the semantics of formulas involving event models requires a definition of product update of models with event models. Given a model $\mathcal{M} = (A, \sim, g, \nu)$ and an event model $\mathcal{E} = (\Phi, \mathsf{post})$, the product update is $\mathcal{M} \otimes \mathcal{E} = (A, \sim, g, \nu')$, where ν' is defined by:

$$\nu'(a) = \begin{cases} \mathsf{post}(\varphi) & \text{if there is a } \varphi \in \Phi \text{ such that } \mathcal{M}, a \models \varphi \\ \nu(a) & \text{otherwise} \end{cases} \tag{1}$$

Then, the semantics of a formula of the form $[\mathcal{E}]\varphi$ is given by:

$$\mathcal{M}, a \models [\mathcal{E}]\varphi \quad \text{iff} \quad \mathcal{M} \otimes \mathcal{E}, a \models \varphi$$

This way, we obtain the semantics of our full dynamic logic and satisfiability, validity etc. are extended to this in the obvious way.[8]

Given a model $\mathcal{M} = (A, \sim, g, \nu)$ and an event model $\mathcal{E} = (\Phi, \mathsf{post})$, let

$$\mathcal{M} \otimes^k \mathcal{E} := \big(...((\mathcal{M} \underbrace{\otimes \mathcal{E}) \otimes \mathcal{E}) \otimes ...\big) \otimes \mathcal{E}}_{k \text{ times}} \quad \text{for every } k \in \mathbb{N}_0.$$

An interesting question is whether the network stabilizes, that is if successive updates by $\mathcal{E}$ will result in a network model that does not change under update by $\mathcal{E}$, i.e. a fixed-point of $\mathcal{E}$. Let us formally define this.

Definition 1. *A network model $\mathcal{M} = (A, \sim, g, \nu)$ is said to be* stable *under the dynamics of an event model $\mathcal{E} = (\Phi, \mathsf{post})$ if $\mathcal{M} = \mathcal{M} \otimes \mathcal{E}$. $\mathcal{M}$ is said to* stabilize *under the dynamics of $\mathcal{E}$ if there is a $k \in \mathbb{N}$ such that $\mathcal{M} \otimes^k \mathcal{E}$ is stable.*

We can express in our language that a network is stable. Given a model $\mathcal{M} = (A, \sim, g, \nu)$, the assignment $\nu(a)$ completely describes the characteristics of a, thus the complete characteristics of a is expressed by:

$$\varphi_{\nu(a)} := \bigwedge_{l=1}^{n} V_l = \nu(a)(l).$$

[8] A sound and complete Hilbert-style proof system for the logic can be obtained from the authors.

Moreover, note that the set of all possible assignments $\mathcal{V}$ is finite. Thus, we can quantify over it in our language and express that a network model is stable by[9]:

$$\varphi_{stable} := \bigwedge_{s \in \mathcal{V}} \big(\varphi_s \to [\mathcal{E}]\varphi_s\big). \tag{2}$$

Then, it is easy to see that:

Lemma 1. *A network model $\mathcal{M}$ is stable if, and only if,*

$$\mathcal{M} \models \varphi_{stable}.$$

4 Pluralistic ignorance revisited

We will use the logic of the previous section to model pluralistic ignorance. We assume that everyone in a "group" is connected to everyone else through some finite number of steps (a "community" in the sense of [4]). In other words, we will work with connected network models, i.e, networks containing a unique community. Moreover, we will assume that the $\sim$ relation is symmetric in the rest of the section.

To begin with, we consider two variables V_I and V_E as corresponding respectively to inner belief and expressed belief. Moreover, we assume that $R_I = R_E = \{B\varphi, B\neg\varphi, UP\varphi\}$, such that $V_I = B\varphi$ corresponds to an inner belief in φ, for instance. Note that "$B\varphi$" is a value assigned to a variable, and as such, the "φ" here is NOT a formula of our formal language – φ will only occur as part of a value for a variable. However, we write $I_B\varphi$ as a short hand notation for $V_I = B\varphi$ etc..

Pluralistic ignorance, in the sense that everybody inner believes φ but expresses a belief in $\neg\varphi$, can be formalized by:

$$PI\varphi := G(I_B\varphi \wedge E_B\neg\varphi) \tag{3}$$

If $PI\varphi$ is true in a network model $\mathcal{M}$ we will say that $\mathcal{M}$ *is in a state of pluralistic ignorance.*

To investigate how social influence affects pluralistic ignorance we need to define an event model that captures the two-layer influence described in Section 2. This is fairly straightforward given the table of Figure 1. For now we assume that all agents are of type 1 mentioned in Section 2. We will return to considering other types of agents later on. For each of the 24 rows, the conjunction of the

[9] Another way of expressing that a network model is stable would be to follow the line of [6]. If $V_L = r$ is true of some agent and the network is stable, this means that none of the preconditions $\varphi \in \Phi$ of $\mathcal{E}$ for which $\mathsf{post}(\varphi)$ would change the value of V_l can be satisfied at the agent. Then, for every full characteristic we can write the conjunction of the negation of all preconditions that would change this characteristic. Finally, we can take the disjunction over all possible full characteristics and thereby obtain a formula for a network being stable.

first four columns will be a precondition. These 24 preconditions will clearly be pairwise inconsistent. For instance, the fourth row gives the precondition formula

$$I_B\varphi \wedge \langle F\rangle E_B\varphi \wedge \langle F\rangle E_B\neg\varphi \wedge \neg\langle F\rangle E_U\varphi.$$

The corresponding post-condition will be the assignment assigning $B\varphi$ to V_I and $B\varphi$ to V_E as specified by the first and the fifth column of the table. The resulting event model will be denoted $\mathcal{I}$.

As claimed in Section 2, pluralistic ignorance constitutes a "robust" state, or "equilibrium", in the sense that if a network is in a state of pluralistic ignorance it will stay in this state. We now formalize this in the following lemma:

Proposition 1. *A connected network model in a state of pluralistic ignorance is stable and the condition for being stable reduces to*

$$PI\varphi \rightarrow [\mathcal{I}]PI\varphi. \tag{4}$$

Proof. If a network model $\mathcal{M}$ satisfies $PI\varphi$, then clearly every agent satisfies the assignment that assigns $B\varphi$ to V_I and $B\neg\varphi$ to V_E and thus the truth of (2) reduces to the truth of (4). Now, an inspection of row 16 in the table of Figure 1 shows that all agents will keep expressing a belief in $\neg\varphi$ and keep their inner belief in φ after an update with $\mathcal{I}$. Thus, $PI\varphi$ will remain true after the update, i.e. $[\mathcal{I}]PI\varphi$ is true and the network is stable. □

The "fragility" component of pluralistic ignorance is a little more complex. If just one agent announces her private belief this may "dissolve" the phenomenon or it may not, depending on the structure of the network. We take pluralistic ignorance (in the form of (3)) to be *dissolved* when it is true that $G(I_B\varphi \wedge E_B\varphi)$. Assume that the network model $\mathcal{M}$ is in a state of pluralistic ignorance, i.e. $\mathcal{M}$ satisfies $PI\varphi$. Now assume that some agent (maybe by mistake) suddenly expresses her true inner belief in φ. Let us refer to this agent by the nominal i. Then the following is now satisfied in $\mathcal{M}$

$$UPI\varphi := @_i(I_B\varphi \wedge E_B\varphi) \wedge G\big(\neg i \rightarrow (I_B\varphi \wedge E_B\neg\varphi)\big).$$

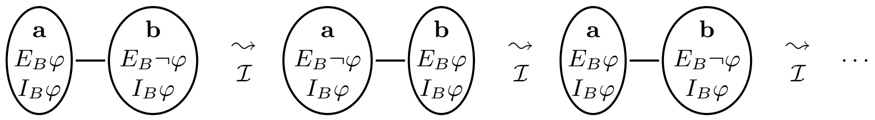
A model satisfying $UPI\varphi$ (where i might be replaced by another nominal) will be said to be in a state of *unstable pluralistic ignorance*.[10] How $\mathcal{M}$ will evolve under the influence event $\mathcal{I}$ depends on several factors. First, consider the case where i will keep expressing her true belief.[11] Then, if $\mathcal{M}$ is connected (and finite) it is easy to show that after a finite number of updates by the influence event $\mathcal{I}$, $\mathcal{M}$ will end up in a stable state where everyone expresses their true

[10] The reader should not be confused by the name of "unstable pluralistic ignorance", which does not refer to a particular case of pluralistic ignorance, but to a state of "almost" pluralistic ignorance, a state which minimally differs from it at the observable level, by one agent expressing her private beliefs, when the others do not.

[11] Formally, we have to make a small change to $\mathcal{I}$ to make sure that i will not change her expressed belief.

beliefs: By inspecting row 4 in the table of Figure 1, it follows that after one update by $\mathcal{I}$ all of i's friends will express a belief in φ and that after another update with $\mathcal{I}$ the friends of friends of i will also express their true belief. In this way, a cascade effect will spread the change throughout the network and result in a stable state where everyone expresses the same true belief.

Now, if i is only made to express her true belief for a *single* round, things get more complicated as she will, in the next round already, revert to expressing a belief in $\neg\varphi$ by the influence event $\mathcal{I}$ (as all of i's friends originally expressed a belief in $\neg\varphi$). For this reason, the network might keep "fluctuating" and never stabilize. Here is an example of the later case, where i refers to agent a[12]:



The above example shows that an unstable state of pluralistic ignorance will not necessarily stabilize, and hence not necessarily result in a state where pluralistic ignorance is dissolved. Below, we give a characterization of the ones which do result in such a state, given our assumption that all agents are of type 1.

Proposition 2. *Let $\mathcal{M} = (A, \sim, g, \nu)$ be a finite, connected, symmetric network model in a state of unstable pluralistic ignorance. Then the following are equivalent:*

(i) After a finite number of updates by the influence event $\mathcal{I}$, $\mathcal{M}$ will end up in a stable state where pluralistic ignorance is dissolved, i.e. there is a $k \in \mathbb{N}$ such that $\mathcal{M} \otimes^k \mathcal{I} \models G(I_B\varphi \wedge E_B\varphi)$ and $\mathcal{M} \otimes^k \mathcal{I} = \mathcal{M} \otimes^{k+1} \mathcal{I}$.

(ii) There is an agent that expresses her true belief in φ for two rounds in a row, i.e. there is an $a \in A$ and a $k \in \mathbb{N}$ such that $\mathcal{M} \otimes^k \mathcal{I}, a \models E_B\varphi$ and $\mathcal{M} \otimes^{k+1} \mathcal{I}, a \models E_B\varphi$.

(iii) There are two agents that are friends and both express their true beliefs in φ in the same round, i.e. there are $a, b \in A$ and a $k \in \mathbb{N}$ such that $a \sim b$, $\mathcal{M} \otimes^k \mathcal{I}, a \models E_B\varphi$, and $\mathcal{M} \otimes^k \mathcal{I}, b \models E_B\varphi$.

(iv) There are two agents that are friends and have paths of the same length to the agent named by i, i.e. there are agents $a, b \in A$ and a $k \in \mathbb{N}$ such that $a \sim b$, $\mathcal{M}, a \models \langle F\rangle^k i$, and $\mathcal{M}, b \models \langle F\rangle^k i$.

(v) There is a cycle in $\mathcal{M}$ of odd length starting at the agent named by i, i.e. there is a $k \in \mathbb{N}$ such that $\mathcal{M} \models @_i\langle F\rangle^{2k-1} i$.

(vi) There is a cycle in $\mathcal{M}$ of odd length, i.e. there is a $k \in \mathbb{N}$ and $a_1, a_2, ..., a_{2k-1} \in A$ such that $a_1 \sim a_2, a_2 \sim a_3, ..., a_{2k-2} \sim a_{2k-1}, a_{2k-1} \sim a_1$.

The proof of this proposition is a little lengthy and will be omitted here, however, it can be obtained from the authors.

[12] Here we regain the same fluctuation case that was given in [6], except that it now occurs, as wanted, at the level of expressed belief instead of "belief".

By this proposition (and its proof) we can also come up with an upper bound of the number of update-steps needed for a network model in an unstable pluralistic ignorance state to dissolve, if it stabilizes. If a network model $\mathcal{M} = (A, \sim, g, \nu)$ stabilizes it follows from (iv) that there are $a, b \in A$ and a $k \in \mathbb{N}$ such that $a \sim b$ and a and b both have a path of length k to $g(i)$. Choose the smallest such k. For all $c \in A$, let $m(c)$ be the length of the shortest path to either a or b. Then, by inspecting the proof it is not hard to see that $\mathcal{M}$ stabilizes in a state where pluralistic ignorance is dissolved in at most $k + max_{c \in A}\{m(c)\}$ steps.

As mentioned in section 2, the type of agents might also influence whether unstable pluralistic ignorance will dissolve. In the above we have focused on what happens when agents are of type 1. If one wants all agents to be of another type, then one can simply change the definition of $\mathcal{I}$. First, note that agents in a state of pluralistic ignorance will always be strongly influenced and since all the three different kinds of agents react the same to strong influence, Proposition 1 remains true for all types.

Now, let us consider a network of type 3 agents (expressing their inner belief whenever they have some support for it). The lines 1, 4, 7, and 10 of Figure 1 will stay unchanged. Thus, Proposition 2 will remain true for this type of agents. The only case left to consider is therefore whether Proposition 2 holds for type 2 agents (expressing their inner belief whenever they face no opposition). We leave this as an open problem.

Another interesting case would be networks with *mixed* types of agents. Our framework can be used to model this as well. We simply add another variable V_T to keep track of the agents' types, i.e. we take $R_T = \{1, 2, 3\}$. Now, we can modify the definition of $\mathcal{I}$ such that in the lines where the agent's type affects what they will do we split each line into three new lines distinguished by the extra preconditions of the form $V_T = k$. Then we change the corresponding post-conditions accordingly. In this way, a new event model $\mathcal{I}'$ can be defined, resulting in an influence dynamics that also depends on the agents types. We will leave the details of this for future research.

Even though we have shown that pluralistic ignorance is stable, there is a sense in which the phenomenon will not continue forever. The discrepancy between one's inner beliefs and one's expressed beliefs is a conflict which might have negative consequences for the agents and as such they may very well try to resolve it. This is a well studied issue in the social and psychological literature on pluralistic ignorance. It is usually assumed [13] that the agents have three different ways in which they can act to resolve this conflict: They can either *internalize the perceived view of their peers*, i.e. change their private beliefs, *attempt to change the perceived view of their peers*, or *alienate themselves form their peers*. In our setting, the first option simply corresponds to the agents changing their inner beliefs in φ to inner beliefs in $\neg\varphi$. the only way they can try and change the opinion of others is by their expressed belief. Thus, the most natural interpretation of the second option would be that the agents will start expressing their true beliefs in φ. Finally, one interpretation of the action of alienating oneself from one's peers would be to remove friendship links to all agents that express a belief in $\neg\varphi$.

Different agents might choose different reactions to a conflict between their inner and expressed beliefs. Therefore, it would be natural to add a new variable V_A to keep track of what action an agent will chose in case of such a conflict. Moreover, it would be natural to assume that agents only try to eliminate this conflict after experiencing it for some time, i.e., for a given number of rounds. We could also capture this by adding another variable that acts as a "counter" of rounds. These new variables can then be included in the preconditions of the influence event $\mathcal{I}$. For the first two options it is obvious what the new post-conditions should be, but for the third option we need an extension of our notion of event model such that it can also change the links in a network model. We believe this can be done, but we leave the details for future research.

5 Conclusion

We developed a hybrid logic to describe networks dynamics. Obtained as a formalization and extension of the simple framework of [6] with added dynamic modalities and event models, this new setting allows for agents to have multiple (changing) characteristics.

We extended the notion of social influence from [6] to a two-layered version, distinguishing between what agents actually (privately) believe and what they express to their friends. We argued that this distinction is a component of many social phenomena, and discussed the case of pluralistic ignorance – a phenomenon widely discussed in social psychology and behavioral economics.

We then formalized pluralistic ignorance and some of its dynamic properties in our new framework. Finally, we obtained a characterization result of the network configurations for which pluralistic ignorance will dissolve into a stable state where everybody agrees into expressing what they truly believe.

Acknowledgments. We would like to thank Johan van Benthem and Fenrong Liu for suggestions and comments during the elaboration of this paper. The research of Zoé Christoff leading to these results has received funding from the European Research Council under the European Communitys Seventh Framework Programme (FP7/2007-2013)/ERC Grant agreement no. 283963. Jens Ulrik Hansen is sponsored by the Swedish Research Council (VR) through the project "Collective Competence in Deliberative Groups: On the Epistemological Foundation of Democracy".

References

1. van Benthem, J.: Logical Dynamics of Information and Interaction. Cambridge University Press, The Netherlands (2011)
2. Easley, D., Kleinberg, J.: Networks, Crowds, and Markets: Reasoning About a Highly Connected World. Cambridge University Press, New York, USA (2010)

3. Jackson, M.O.: Social and Economic Networks. Princeton University Press (2010)
4. Seligman, J., Liu, F., Girard, P.: Logic in the community. In Banerjee, M., Seth, A., eds.: Logic and Its Applications. Volume 6521 of Lecture Notes in Computer Science. Springer Berlin Heidelberg (2011) 178–188
5. Zhen, L., Seligman, J.: A logical model of the dynamics of peer pressure. Electronic Notes in Theoretical Computer Science **278**(0) (2011) 275 – 288
6. Seligman, J., Girard, P., Liu, F.: Logical dynamics of belief change in the community, under submission. (2013)
7. Girard, P., Seligman, J., Liu, F.: General Dynamic Dynamic Logic. (forthcoming)
8. Seligman, J., Girard, P.: Facebook and the epistemic logic of friendship Extended Abstract
9. Alchourrn, C.E., Gärdenfors, P., Makinson, D.: On the logic of theory change: Partial meet contraction and revision functions. Journal of Symbolic Logic **50**(2) (1985) 510–530
10. Schelling, T.C.: Micromotives and Macrobehavior. W. W. Norton and Company (1978)
11. O'Gorman, H.J.: The discovery of pluralistic ignorance: An ironic lesson. Journal of the History of the Behavioral Sciences **22**(October) (1986) 333–347
12. Miller, D.T., McFarland, C.: Pluralistic ignorance; when similarity is interpreted as dissimilarity. Journal of Personality and Social Psychology **53** (1987) 298–305
13. Prentice, D.A., Miller, D.T.: Pluralistic ignorance and alcohol use on campus: Some consequences of misperceiving the social norm. Journal of Personality and Social Psychology **64**(2) (1993) 243–256
14. Fields, J.M., Schuman, H.: Public beliefs about the beliefs of the public. The Public Opinion Quarterly **40**(4) (1976) 427–448
15. O'Gorman, H.J., Garry, S.L.: Pluralistic ignorance – a replication and extension. The Public Opinion Quarterly **40**(4) (1976) 449–458
16. Krech, D., Crutchfield, R.S.: Theories and Problems of Social Psychology. New York: McGraw-Hill (1948)
17. Halbesleben, J.R.B., Buckley, M.R.: Pluralistic ignorance: historical development and organizational applications. Management Decision **42**(1) (2004) 126–138
18. Miller, D.T., McFarland, C.: When social comparison goes awry: The case of pluralistic ignorance. In Suls, J., Wills, T., eds.: Social comparison: Contemporary theory and research. Erlbaum, Hillsdale, NJ (1991) 287–313
19. Bjerring, J.C., Hansen, J.U., Pedersen, N.J.L.L.: On the rationality of pluralistic ignorance. Synthese (to appear)
20. Areces, C., ten Cate, B.: Hybrid logics. In Blackburn, P., van Benthem, J., Wolter, F., eds.: Handbook of Modal Logic. Elsevier, Amsterdam (2007) 821–868
21. Baltag, A., Moss, L.S., Solecki, S.: The logic of public announcements, common knowledge and private suspicious. Technical Report SEN-R9922, CWI, Amsterdam (1999)
22. van Ditmarsch, H., van der Hoek, W., Kooi, B.: Dynamic Epistemic Logic. Syntese Library volume 337. Springer, The Netherlands (2008)
23. Blackburn, P., ten Cate, B.: Pure extensions, proof rules, and hybrid axiomatics. Studia Logica **84**(2) (2006) 277–322
24. Hansen, J.U.: A hybrid public announcement logic with distributed knowledge. Electronic Notes in Theoretical Computer Science **273** (2011) 33–50 Post-Proceedings of the International Workshop on Hybrid Logic and Applications (HyLo 2010).
25. Blackburn, P.: Internalizing labelled deduction. Journal of Logic and Computation **10**(1) (2000) 137–168

Appendix A: A proof system for the logic of Section 3

In this appendix we will give a sound and complete Hilbert-style proof system for the logic of Section 3 that follows that of [23, 20]with some modifications and additions of reduction axioms as in standard Dynamic Epistemic Logic. The proof system is shown in figure 2.

The soundness of the static part of the language is straightforward and we will return to the soundness of the dynamic part later on. We will show completeness in the style of [23, 20, 24]. We will start by showing completeness with respect to the static part of the logic, i.e. we leave the operator $[\mathcal{E}]$ of the syntax and omit the axioms Red.Ax.Prop - Red.Ax.$\Box$ for now. This static logic will be referred to as *Hybrid Network Logic* and be denoted it by $\mathbf{K}_{\mathcal{HN}}$. We will return to completeness of the full dynamic logic in the next subsection. Moreover, we will leave out any talk about the global modality until after our first completeness result, which will give us an easy way to add the global modality G without any effort, just as in [23].

Axioms:	
All substitution instances of propositional tautologies	
$\bigwedge_{l=1}^{n}\big(\bigvee_{r\in R_l} V_l\!=\!r\big)$	Char.Prop.1
$\bigwedge_{l=1}^{n}\bigwedge_{r\in R_l}\big(V_l\!=\!r \to \bigwedge_{s\in R_l\setminus\{r\}}\neg V_l\!=\!s\big)$	Char.Prop.2
$X(\varphi\to\psi)\to(X\varphi\to X\psi)^1$	K_X
$@_i(\varphi\to\psi)\to(@_i\varphi\to @_i\psi)$	$\mathrm{K}_@$
$@_i\varphi\leftrightarrow\neg @_i\neg\varphi$	$\mathrm{Selfdual}_@$
$@_i i$	$\mathrm{Ref}_@$
$@_i @_j\varphi\leftrightarrow @_j\varphi$	Agree
$i\to(\varphi\leftrightarrow @_i\varphi)$	Introduction
$\langle X\rangle @_i\varphi\to @_i\varphi^1$	Back
$(@_i\langle X\rangle j\wedge @_j\varphi)\to @_i\langle X\rangle\varphi^1$	Bridge
$\langle G\rangle i$	GM
$[\mathcal{E}]V_l\!=\!r\leftrightarrow\big(\bigvee_{\varphi\in\Phi,\mathsf{post}(\varphi)(l)=r}\varphi\big)\vee\big(\neg(\bigvee_{\varphi\in\Phi}\varphi)\wedge V_l\!=\!r\big)$	Red.Ax.Prop.
$[\mathcal{E}]i\leftrightarrow i$	Red.Ax.Nom.
$[\mathcal{E}](\varphi\wedge\psi)\leftrightarrow[\mathcal{E}]\varphi\wedge[\mathcal{E}]\psi$	Red.Ax.$\wedge$
$[\mathcal{E}]\neg\varphi\leftrightarrow\neg[\mathcal{E}]\varphi$	Red.Ax.$\neg$
$[\mathcal{E}]@_i\varphi\leftrightarrow @_i[\mathcal{E}]\varphi$	Red.Ax.@
$[\mathcal{E}]X\varphi\leftrightarrow X[\mathcal{E}]\varphi$	Red.Ax.$\Box$
Rules:	
From φ and $\varphi\to\psi$, infer ψ	Modus ponens
From φ, infer $X\varphi^1$	Necessitation of X
From φ, infer $@_i\varphi$	Necessitation of @
From $@_i\varphi$, where i does not occur in φ, infer φ	Name
From $(@_i\langle X\rangle j\wedge @_j\varphi)\to\psi$, where $i\neq j$ and j does not occur in φ or ψ, infer $@_i\langle X\rangle\varphi\to\psi^1$	Paste

[1] Here X denotes either F or G.

Fig. 2. The Hilbert-style proof system

In the following, we use standard terminology for Hilbert-style proof systems: A proof of φ is a finite sequence of formulas ending with φ such that every formula in the sequence is either an axiom or follows from previous formulas in the sequence using one of the proof rules. We denote this by $\vdash \varphi$. For a set of formulas Γ, $\Gamma \vdash \varphi$ holds if there are $\psi_1, ..., \psi_n \in \Gamma$ such that $\vdash \psi_1 \wedge ... \wedge \psi_n \to \varphi$. Given a set of formulas Σ, let $\mathbf{K}_{\mathcal{HN}} + \Sigma$ denote the logic obtained by adding all the formulas in Σ as axioms. That φ is provable in the logic $\mathbf{K}_{\mathcal{HN}} + \Sigma$ will be denoted by $\vdash_{\mathbf{K}_{\mathcal{HN}}+\Sigma} \varphi$. A set of formulas Γ is said to be $\mathbf{K}_{\mathcal{HN}} + \Sigma$-inconsistent if $\Gamma \vdash_{\mathbf{K}_{\mathcal{HN}}+\Sigma} \bot$, and $\mathbf{K}_{\mathcal{HN}} + \Sigma$-consistent otherwise. A formula φ is pure if it does not contain any characteristic propositions. A set of formulas Σ is called substitution-closed if it is closed under uniform substitution of nominals by nominals.

It is now straightforward to prove the following Lindenbaum lemma in the standard way:

Lemma 2 (Lindenbaum lemma). *Let Σ be a set of pure formulas. Every $\mathbf{K}_{\mathcal{HN}} + \Sigma$-consistent set of formulas Γ can be extended to a maximal $\mathbf{K}_{\mathcal{HN}} + \Sigma$-consistent set Γ^+ (in a new language obtained by adding countable many new nominals), such that*

(1) Γ^+ contains a nominal.
(2) For all $@_i\langle F\rangle\varphi \in \Gamma^+$ there is a nominal j, such that $@_i\langle F\rangle j \in \Gamma^+$ and $@_j\varphi \in \Gamma^+$.

A canonical model in Henkin-style manner can now be constructed:

Definition 2. *Let Γ be a maximal consistent set of -formulas. Define an equivalence relation $\equiv$ on NOM by $i \equiv j$ iff $@_i j \in \Gamma$ (and denote the equivalence class of i by $|i|$). Then the canonical model $\mathfrak{M}_\Gamma = \langle A, \sim, g, \nu\rangle$ is defined by*

$$\begin{aligned} A &= \{|i| \mid i \in \mathsf{NOM}\}; \\ |i| \sim |j| &\quad \textit{iff} \quad @_i\langle F\rangle j \in \Gamma; \\ g(i) &= |i| \quad \textit{for all } i \in \mathsf{NOM}; \\ \nu(|i|)(l) = r &\quad \textit{iff} \quad @_i V_l = r \in \Gamma \quad \textit{for all } a \in A \textit{ and all } l \in \{1, ..., n\}. \end{aligned}$$

To ensure that this model is well-defined, we need the following lemma:

Lemma 3. *The following are derivable in our logic:*

i) $@_i j \to @_j i$
ii) $(@_i j \wedge @_j k) \to @_i k$
iii) $@_i j \to (@_i\varphi \leftrightarrow @_j\varphi)$

The proof of this lemma is also straightforward. The canonical model is well-defined for the following reasons: First of all, the relation $\equiv$ is an equivalence relation by the (Ref$_@$) axiom and *i*) and *ii*) of lemma 3, and thus A and g are well-defined. That $\sim$ is well-defined follows from *iii*) of lemma 3 and the Bridge axiom. Finally, consider ν. By *ii*) of lemma 3, $@_i V_l = r \in \Gamma$ iff $@_j V_l = r \in \Gamma$ whenever $i \equiv j$. Moreover, by axiom Char.Prop.1 there is always an $r \in R_l$ such that $@_i V_l = r \in \Gamma$ and by Char.Prop.2 this r is unique. Hence, ν is well-defined. An essential truth lemma can now be established:

Lemma 4 (Truth lemma). *Let Γ be a maximal consistent set of $\mathcal{HN}$-formulas that satisfy item* (2) *of the Lindenbaum lemma. Then for all $i \in \mathsf{NOM}$ and all formulas φ*

$$\mathfrak{M}_\Gamma, |i| \models \varphi \quad \textit{iff} \quad @_i\varphi \in \Gamma. \tag{5}$$

Proof. The proof goes by induction on φ. When φ is a $j \in \mathsf{NOM}$ or on the form $V_l = r$ for a $l \in \{1, ..., n\}$ and a $r \in R_l$, (5) follows directly from the definition of the canonical model.

The induction step. In the case φ is on the form $\psi \wedge \chi$, note that $@_i\psi, @_i\chi \in \Gamma$ if and only if $@_i(\psi \wedge \chi) \in \Gamma$. In the case φ is on the form $\neg\psi$, the thing to note is that $\neg @_i\psi \in \Gamma \Leftrightarrow @_i\neg\psi \in \Gamma$.

Assume now that φ has the form $@_j\psi$. First note that if $@_i@_j\psi \in \Gamma$ then $@_j\psi \in \Gamma$ by Agree. Then by induction it follows that $\mathcal{M}_\Gamma, |j| \models \psi$, which again implies that $\mathcal{M}_\Gamma, |i| \models @_j\psi$. If $\mathcal{M}_\Gamma, |i| \models @_j\psi$ then there is a $k \in \mathsf{NOM}$ such that $\mathcal{M}_\Gamma, |k| \models \psi$ and $g(j) = |k|$. By the induction hypothesis this implies that $@_k\psi \in \Gamma$. Since $g(j) = |k|$ (i.e. $@_kj \in \Gamma$) and by *iii*) of Lemma 3 we have that $@_j\psi \in \Gamma$. But then by the Agree axiom $@_i@_j\psi \in \Gamma$ follows.

The case φ is of the form $\langle F\rangle\psi$. If $\mathcal{M}_\Gamma, |i| \models \langle F\rangle\psi$, then there is a $j \in \mathsf{NOM}$ such that $|i| \sim |j|$ and $\mathcal{M}_\Gamma, |j| \models \psi$. By definition of $\sim$, $@_i\langle F\rangle j \in \Gamma$ and by the induction hypothesis $@_j\psi \in \Gamma$. But then by the bridge axiom it follows that $@_i\langle F\rangle\psi \in \Gamma$. Now assume that $@_i\langle F\rangle\psi \in \Gamma$. Then since Γ satisfies item (2) of the Lindenbaum lemma it follows that there is a nominal j such that $@_i\langle F\rangle j \in \Gamma$ and $@_j\psi \in \Gamma$. Now by the definition of $\sim$ and g and the induction hypothesis it follows that $\mathcal{M}_\Gamma, |i| \models \langle F\rangle\psi$.This concludes the proof. □

We say that a frame $\mathcal{F}$ *validates* a set of formulas Σ, if $\mathcal{M}, \models \Sigma$ for all models $\mathcal{M}$ based on $\mathcal{F}$. With this notion we state the following Frame lemma:

Lemma 5 (Frame lemma). *Let Σ be a substitution-closed set of pure $\mathcal{HN}$-formulas and let Γ be a $\mathbf{K}_{\mathcal{HN}} + \Sigma$ maximal consistent set of $\mathcal{HN}$-formulas satisfying item* (1) *and* (2) *of the Lindenbaum lemma. Then the underlying frame of $\mathfrak{M}_\Gamma$ validates all the formulas in Σ.*

Proof. See Lemma 7.1 of [25].

Finally, we can now state the completeness theorem.

Proposition 3 (Completeness of $\mathbf{K}_{\mathcal{HN}}$). *Let Σ be a substitution-closed set of pure formulas. Every set of formulas that is $\mathbf{K}_{\mathcal{HN}} + \Sigma$-consistent is satisfiable in a model whose underlying frame validates all the formulas in Σ.*

Proof. Assume that Γ is $\mathbf{K}_{\mathcal{HN}} + \Sigma$-consistent. Then it can be extended to a maximal $\mathbf{K}_{\mathcal{HN}} + \Sigma$-consistent set Γ^+ by the Lindenbaum lemma. Since there is a nominal $i \in \Gamma^+$ by item (1) of the Lindenbaum lemma it is easy to see that for all $\varphi \in \Gamma$, $@_i\varphi \in \Gamma^+$ by the Introduction axiom. But then by the truth lemma it follows that $\mathfrak{M}_{\Gamma^+}, |i| \models \varphi$ for all $\varphi \in \Gamma$. By the frame lemma the underlying frame of $\mathfrak{M}_{\Gamma^+}$ validates all the formulas in Σ and the proof is done. □

It is now easy to see that we can add the global modality G straightforwardly adding all substitution instances of the axiom (GM). Since (GM) defines the class of frames where the corresponding relation is the global relation (this is easy to check) this completeness result ensures that G will indeed be the global modality.

Completeness of the full dynamic logic

We now move on to show completeness of the full dynamic logic, that is, we include the operator $[\mathcal{E}]$ in the syntax again and add the axioms Red.Ax.Prop - Red.Ax.$\Box$. The resulting logic will be referred to as *Hybrid Dynamics Network Logic* and will be denoted by $\mathbf{K}_{\mathcal{HDN}}$. The way we will show completeness is the usual way in Dynamic Epistemic Logic, namely by providing a truth-preserving translation from $\mathbf{K}_{\mathcal{HDN}}$ into $\mathbf{K}_{\mathcal{HN}}$.

In this way of proving completeness we need the soundness of the reduction axioms Red.Ax.Prop - Red.Ax.F. This is ensured by the following lemma:

Lemma 6. *For all models $\mathcal{M} = \langle A, \sim, g, \nu\rangle$ and all $a \in A$, the following hold:*

$$\mathcal{M}, a \models [\mathcal{E}]V_l = r \quad \textit{iff} \quad \mathcal{M}, a \models \Big(\bigvee_{\varphi\in\Phi, \mathsf{post}(\varphi)(l)=r} \varphi\Big) \vee \Big(\neg(\bigvee_{\varphi\in\Phi}\varphi) \wedge V_l = r\Big) \tag{6}$$

$$\mathcal{M}, a \models [\mathcal{E}]i \quad \textit{iff} \quad \mathcal{M}, a \models i \tag{7}$$

$$\mathcal{M}, a \models [\mathcal{E}](\varphi \wedge \psi) \quad \textit{iff} \quad \mathcal{M}, a \models [\mathcal{E}]\varphi \wedge [\mathcal{E}]\psi \tag{8}$$

$$\mathcal{M}, a \models [\mathcal{E}]\neg\varphi \quad \textit{iff} \quad \mathcal{M}, a \models \neg[\mathcal{E}]\varphi \tag{9}$$

$$\mathcal{M}, a \models [\mathcal{E}]@_i\varphi \quad \textit{iff} \quad \mathcal{M}, a \models @_i[\mathcal{E}]\varphi \tag{10}$$

$$\mathcal{M}, a \models [\mathcal{E}]F\varphi \quad \textit{iff} \quad \mathcal{M}, a \models F[\mathcal{E}]\varphi \tag{11}$$

$$\mathcal{M}, a \models [\mathcal{E}]G\varphi \quad \textit{iff} \quad \mathcal{M}, a \models G[\mathcal{E}]\varphi \tag{12}$$

Proof. Proof of (6). Let $\mathcal{M} \otimes \mathcal{E}$ be $\langle A, \sim, g, \nu'\rangle$, where ν' is defined as in (1). Then we have the following equivalences:

$$\begin{aligned} \mathcal{M}, a \models [\mathcal{E}]V_l = r \quad &\text{iff} \quad \mathcal{M} \otimes \mathcal{E}, a \models V_l = r \\ &\text{iff} \quad \nu'(a)(l) = r. \end{aligned}$$

Note that, $\nu'(a)(l) = r$ is the case if, and only if, either there is a $\varphi \in \Phi$ such that $\mathcal{M}, a \models \varphi$ and $\mathsf{post}(\varphi)(l) = r$, or there is no such φ, but $\nu(a)(l) = r$. Now, he first disjunct of this disjunction is equivalent to $\mathcal{M}, a \models \big(\bigvee_{\varphi\in\Phi,\mathsf{post}(\varphi)(l)=r}\varphi\big)$ while the second disjunct it equivalent to $\mathcal{M}, a \models \big(\neg(\bigvee_{\varphi\in\Phi}\varphi) \wedge V_l = r\big)$. Hence,

$$\nu'(a)(l) = r \quad \text{iff} \quad \mathcal{M}, a \models \big(\bigvee_{\varphi\in\Phi,\mathsf{post}(\varphi)(l)=r}\varphi\big) \vee \big(\neg(\bigvee_{\varphi\in\Phi}\varphi) \wedge V_l = r\big),$$

and (6) has been proven.

Proof of (7). Note that $\mathcal{M}, a \models [\mathcal{E}]i$ if, and only if, $\mathcal{M} \otimes \mathcal{E}, a \models i$. But since the function g assigning agents to nominals does not change under the update with $\mathcal{E}$, $\mathcal{M} \otimes \mathcal{E}, a \models i$ if, and only if, $\mathcal{M}, a \models i$

Proof of (8)-(12). The proofs $(8)-(12$ are very similar, so we only do the case the case of (11). We have the following equivalences:

$$\begin{aligned}\mathcal{M}, a \models [\mathcal{E}]F\varphi \quad & \text{iff} \quad \mathcal{M} \otimes \mathcal{E}, a \models F\varphi \\ & \text{iff} \quad \mathcal{M} \otimes \mathcal{E}, b \models \varphi, \quad \text{for all } b \in A \text{ with } a \sim b \\ & \text{iff} \quad \mathcal{M}, b \models [\mathcal{E}]\varphi, \quad \text{for all } b \in A \text{ with } a \sim b \\ & \text{iff} \quad \mathcal{M}, a \models F[\mathcal{E}]\varphi.\end{aligned}$$

□

With this lemma, it is now easy to define a translation t from the language with the $[\mathcal{E}]$-operator into the language with it, such that every formula becomes provable equivalent to its translation as well as semantically equivalent. Then, the completeness result of Proposition 3 can easily be extended to the full hybrid dynamic network logic $\mathbf{K}_{\mathcal{HDN}}$.

Appendix B: A proof of Proposition 2

In this Appendix we will give the proof of Proposition 2 of Section 4. As a first step towards the proof is the following lemma:

Lemma 7. *The following is a validity of our logic for all $j \in \mathsf{NOM}$ and all $k \in \mathbb{N}$:*

$$\big(@_i(I_B\varphi \wedge E_B\varphi) \wedge GI_B\varphi \wedge @_i\langle F\rangle^k j\big) \to [\mathcal{I}]^k @_j E_B\varphi. \tag{13}$$

Intuitively, this lemma says that if i for one round expresses her true belief in φ, if everyone else truly believes φ as well, and if there is a path from i to j of length n, then after exactly n updates with $\mathcal{I}$, j will express a belief in φ.

Proof. The proof goes by induction on $k \in \mathbb{N}$. By inspecting the lines 1, 4, 7 and 10 of table of Figure 1 the validity for $k = 1$ easily follows. Now assume that (13) is true for a $k \in \mathbb{N}$ and all $j \in \mathsf{NOM}$. Assume furthermore that the antecedent is true for $k+1$ in a model $\mathcal{M} = (A, \sim, g, \nu)$. This means that there is a path of length $k+1$ from $g(i)$ to $g(j)$, in particular there is an agent a such that there is path of length k from $g(i)$ to a and a path of length 1 from a to $g(j)$. We can assume without loss of generality that there is a nominal $h \in \mathsf{NOM}$, different from i and j, such that $g(h) = a$. But then by the assumption that (13) is true for k we obtain that $[\mathcal{I}]^k @_h E_B\varphi$ is true. But then by inspecting the lines 1, 4, 7, and 10 of table of Figure 1 it follows that $[\mathcal{I}]^{k+1} @_j E_B\varphi$ is true, as well. This completes the proof. □

Another useful lemma is the following:

Lemma 8. *Let $\mathcal{M} = (A, \sim, g, \nu)$ be a finite, connected, symmetric network model in a state of unstable pluralistic ignorance. Then for all $k \in \mathbb{N}_0$,*

$$\mathcal{M} \otimes^k \mathcal{I} \models G(E_B\varphi \vee E_B\neg\varphi).$$

Thus, when a network starts out in an unstable state of pluralistic ignorance and evolves under the influence event $\mathcal{I}$, no one will ever express undecidedness.

Proof. Assume that $\mathcal{M} = (A, \sim, g, \nu)$ is a finite, connected, symmetric network model in a state of unstable pluralistic ignorance. First note that since $\mathcal{M}$ is in a state of unstable pluralistic ignorance all agents satisfy $I_B\varphi$ and as the influence event $\mathcal{I}$ does not change any agent's inner belief, $I_B\varphi$ will remain true of all agents in all models of the form $\mathcal{M} \otimes^k \mathcal{I}$.

The proof goes on induction on $k \in \mathbb{N}_0$. The induction follows trivially from the fact that $\mathcal{M} \models @_i(I_B\varphi \wedge E_B\varphi) \wedge G\big(\neg i \rightarrow (I_B\varphi \wedge E_B\neg\varphi)\big)$. Now assume that $\mathcal{M} \otimes^k \mathcal{I} \models G(E_B\varphi \vee E_B\neg\varphi)$ and consider an agent $a \in A$. Note that all a's friends either expressed a belief in φ or a belief in $\neg\varphi$ in $\mathcal{M} \otimes^k \mathcal{I}$. But then, by inspecting the lines 4, 10, and 16 of the table of Figure 1, a must either express a belief in φ or a belief in $\neg\varphi$ in $\mathcal{M} \otimes^{k+1} \mathcal{I}$. This completes the induction proof. □

We can now prove the main proposition.

Proof of Proposition 2. $(i) \Rightarrow (ii)$. This is straightforward.

$(ii) \Rightarrow (iii)$. Assume that there are $a \in A$ and $k \in \mathbb{N}$ such that $\mathcal{M} \otimes^k \mathcal{I}, a \models EB\varphi$ and $\mathcal{M} \otimes^{k+1} \mathcal{I}, a \models EB\varphi$. By Lemma 8 and an inspection of the lines 4 and 10 of the table of Figure 1, it follows that there is a $b \in A$ such that $a \sim b$ and $\mathcal{M} \otimes^k \mathcal{I}, b \models E_B\varphi$. Hence, (iii) follows.

$(iii) \Rightarrow (iv)$. Let $a, b \in A$ and $k \in \mathbb{N}$ be such that $a \sim b$, $\mathcal{M} \otimes^k \mathcal{I}, a \models E_B\varphi$, and $\mathcal{M} \otimes^k \mathcal{I}, b \models E_B\varphi$. As previously, it follows by Lemma 8 and an inspection of the lines 4 and 10 of the table of Figure 1, that there are $a_{k-1}, b_{k-1} \in A$ such that $a_{k-1} \sim a$, $b_{k-1} \sim b$, $\mathcal{M} \otimes^{k-1} \mathcal{I}, a_{k-1} \models E_B\varphi$, and $\mathcal{M} \otimes^{k-1} \mathcal{I}, b_{k-1} \models E_B\varphi$. But then, it follows by Lemma 8 and an inspection of the lines 4 and 10 of the table of Figure 1, that there are $a_{k-2}, b_{k-2} \in A$ such that $a_{k-2} \sim a_{k-1}$, $b_{k-b} \sim b_{k-1}$, $\mathcal{M} \otimes^{k-2} \mathcal{I}, a_{k-2} \models E_B\varphi$, and $\mathcal{M} \otimes^{k-2} \mathcal{I}, b_{k-2} \models E_B\varphi$. Continuing this way, we obtain $a_0, a_1, a_2, ..., a_{k-1} \in A$ and $b_0, b_1, b_2, ..., b_{k-1} \in A$ such that $a_0 \sim a_1 \sim ... \sim a_{k-1} \sim a$, $b_0 \sim b_1 \sim ... \sim b_{k-1} \sim b$, $\mathcal{M} \otimes^0 \mathcal{I}, a_0 \models E_B\varphi$, and $\mathcal{M} \otimes^0 \mathcal{I}, b_0 \models E_B\varphi$. Now since $\mathcal{M} \otimes^0 \mathcal{I} = \mathcal{M}$, and $g(i)$ is the only agent in $\mathcal{M}$ that satisfy $E_B\varphi$, it follows that $a_0 = b_0 = g(i)$. Thus, $\mathcal{M}, a \models \langle F \rangle^k i$ and $\mathcal{M}, b \models \langle F \rangle^k i$ and (iv) follows.

$(iv) \Rightarrow (v)$. Assume that there are $a, b \in A$ and $k \in \mathbb{N}$ such that $a \sim b$, $\mathcal{M}, a \models \langle F \rangle^k i$, and $\mathcal{M}, b \models \langle F \rangle^k i$. From $\mathcal{M}, a \models \langle F \rangle^k i$, it follows that there are $a_1, a_2, ..., a_{k-1} \in A$ such that $g(i) \sim a_1 \sim a_2 \sim ... \sim a_{k-1} \sim a$. Similar there are $b_1, b_2, ..., b_{k-1} \in A$ such that $g(i) \sim b_1 \sim b_2 \sim ... \sim b_{k-1} \sim b$. But then

$$g(i) \sim a_1 \sim ... \sim a_{k-1} \sim a \sim b \sim b_{k-1} \sim ... \sim b_1 \sim g(i)$$

is a path of length $2(k+1) - 1$. Hence, $\mathcal{M}, g(i) \models \langle F \rangle^{2(k+1)-1} i$ and (v) follows.

$(v) \Rightarrow (vi)$. This is trivial.

$(vi) \Rightarrow (iv)$. Assume that there is an odd cycle in $\mathcal{M}$. Since $\mathcal{M}$ is connected, there is a path from $g(i)$ to the cycle. But then is is not hard to find agents a and b of the cycle such that they both have same path length to $g(i)$ and are friends. Now, (iv) easily follows.

$(iv) \Rightarrow (i)$. Assume that there are $a, b \in A$ and a $k \in \mathbb{N}$ such that $a \sim b$, $\mathcal{M}, a \models \langle F \rangle^k i$, and $\mathcal{M}, b \models \langle F \rangle^k i$. Then by Lemma 7 it is not hard to see that after k updates with $\mathcal{I}$ both a and b will express beliefs in φ, i.e. $\mathcal{M} \otimes^k \mathcal{I}, a \models E_B \varphi$ and $\mathcal{M} \otimes^k \mathcal{I}, b \models E_B \varphi$. Then, since $a \sim b$, an inspection of the lines 1, 4, 7, and 10 of table of Figure 1 shows that a and b will keep expressing their belief in φ, i.e. $\mathcal{M} \otimes^{k+m} \mathcal{I}, a \models E_B \varphi$ and $\mathcal{M} \otimes^{k+m} \mathcal{I}, b \models E_B \varphi$ for all $m \in \mathbb{N}$. For the same reasons, for all $c \in A$ such that either $c \sim a$ or $c \sim b$, $\mathcal{M} \otimes^{k+1+m} \mathcal{I}, c \models E_B \varphi$ for all $m \in \mathbb{N}$. Similarly, for all $d \in A$ such that there is a $c \in A$ such that $d \sim c \sim a$ or $d \sim c \sim b$, we have that $\mathcal{M} \otimes^{k+2+m} \mathcal{I}, d \models E_B \varphi$ for all $m \in \mathbb{N}$. Generally for all $d \in A$ such that there are $c_1, ..., c_l \in A$ such that $d \sim c_1 \sim c_2 \sim ... \sim c_l$, $c_l = a$ or c_l, we have that $\mathcal{M} \otimes^{k+l+m} \mathcal{I}, d \models E_B \varphi$ for all $m \in \mathbb{N}$. Since $\mathcal{M}$ is finite and connected, there will be a stage l such that all agents in A have a path of length less than l to a or b and thus they all express belief in φ and will continue doing this. This completes the proof of Proposition 2. $\square$