

# The Sound of Silence in Social Networks \*

Jesús Aranda<sup>1</sup>, Juan Francisco Díaz<sup>1</sup>, David Gaona<sup>1</sup>, and Frank Valencia<sup>2,3</sup>

<sup>1</sup> Universidad del Valle, Colombia

<sup>2</sup> CNRS-LIX, École Polytechnique de Paris, France

<sup>3</sup> Pontificia Universidad Javeriana Cali, Colombia

**Abstract.** We generalize the classic multi-agent DeGroot framework for opinion dynamics by incorporating the Spiral of Silence theory from political science, which posits that individuals may withhold their opinions when they perceive them to be in the minority. As in the original DeGroot model, the social network is represented as a weighted directed graph encoding how agents influence one another. However, agents holding minority opinions become *silent*, meaning they do not express their views.

We introduce two families of models. In *Silence Opinion Memoryless* (SOM<sup>-</sup>) models, agents update their opinions by averaging those of their *non-silent* neighbors. In *Silence Opinion Memory-based* (SOM<sup>+</sup>) models, agents average the opinions of *all* neighbors, but for silent ones, only the most recently expressed opinion is used. We show that SOM<sup>-</sup> models guarantee consensus on clique graphs but, unlike the classic DeGroot model, not on all strongly connected aperiodic graphs. For SOM<sup>+</sup> models, even cliques may fail to reach consensus, illustrating that even minimal memory can significantly affect opinion dynamics. Finally, we validate our models through large-scale simulations on small-world networks with over *two million* agents. The results support the Spiral of Silence theory and reveal inherent limitations to consensus in more realistic settings.

## 1 Introduction

Social networks play a significant role in *opinion formation* often having strong impact on polarization and the ability to reach *consensus*. Broadly, the dynamics of opinion formation in social networks involve users expressing their opinions, being exposed to the opinions of others, and potentially adapting their own views based on these interactions. Modeling these dynamics enables us to glean insights into how opinions form and spread within social networks.

The DeGroot framework [14] is one of the most prominent formalisms for opinion formation and consensus-building in social networks. In the models of this framework, a social network is represented as a weighted directed graph, where edges denote the degree to which individuals (i.e., *agents*) influence one another. Each agent holds an opinion, expressed as a value in  $[0, 1]$ , indicating their level of agreement with an underlying proposition (e.g., “*AI is a threat to humanity*”). Agents repeatedly update their

---

\* This work has been partially supported by the SGR project PROMUEVA (BPIN 2021000100160) under the supervision of Colombian Ministry of Science, Technology and Innovation (Minciencias) and by the CNRS project TOBIAS under the MITI interdisciplinary program.

opinions by taking the weighted average of their opinion differences with those who influence them (i.e., their *contacts*). There is empirical evidence validating the opinion formation through averaging of the model in controlled sociological experiments [9].

*Consensus*, i.e., convergence to a common opinion, is a central property in models of social learning and opinion formation [17]. In fact, difficulties in achieving consensus are a sign of a polarized society. A fundamental result for (classic) DeGroot models shows that agents converge to consensus if the influence graph is strongly connected and aperiodic. The DeGroot framework continues to be a focus of research for constructing frameworks for understanding opinion formation dynamics in social networks (e.g., [17,4,3,2,22,2,13,11,31,30,15,10]).

Nevertheless, the classic DeGroot formalism makes an assumption that could be overly constraining within social network contexts. It assumes that *all agents express their opinions at each time unit*. This assumption, which renders models tractable, may hold in some controlled scenarios as participants are often encouraged to express their views freely and consistently. However, in many real-world situations, some individuals may choose not to express their opinions due to personal choice or social pressure.

Indeed, the *Spiral of Silence* [23] is a well-established social theory that describes how individuals may be unwilling to express their opinions when they perceive themselves to be in the minority. This reluctance can lead to the *reinforcement of dominant views* within a social network. The theory asserts that individuals have a natural tendency to avoid social isolation and seek acceptance within their social groups. When people believe their opinions are unpopular or likely to be met with disapproval, they may opt to remain silent.

The relevance of the Spiral of Silence has been validated in social media environments. Large-scale empirical studies show that *social media platforms do not offer alternative spaces for minority opinion expression*, but instead reinforce silencing behaviors [18]. Meta-analyses confirm that the link between perceived opinion support and willingness to speak out remains strong in digital contexts [21]. Recent work finds that users are particularly sensitive to perceived disagreement among their online contacts, leading to self-censorship that can *spill over into offline behavior* [18,26]. During the COVID-19 pandemic, individuals frequently conformed publicly to dominant narratives while privately holding dissenting views, further illustrating the theory's enduring relevance [6]. These findings suggest that the dynamics of opinion suppression described by The Spiral of Silence not only persist but may be exacerbated in digital social networks.

In this paper, we generalize DeGroot models into a framework where agents may choose to remain silent at a given time following the Spiral of Silence. We consider two possibilities, leading to the two families of models described below.

*Memoryless framework SOM<sup>-</sup>*. In these models, silent agents are excluded from the opinion updates of the agents they would otherwise influence. Additionally, agents become silent at a given time if their views do not align with the majority of their non-silent contacts. This framework is called the *silence opinion memoryless* (SOM<sup>-</sup>) models, as the previous opinions of silent agents are not retained. This corresponds to a social scenario in which opinions (e.g., expressed in posts) are removed once they have been accessed.

Notice that ignoring silent agents at a given time unit amounts to removing certain edges from the influence graph at that time. Thus, a fundamental distinction from DeGroot models is that  $\text{SOM}^-$  models exhibit *dynamic influence* as edges may disappear and reappear during opinion evolution.

*Memory-based framework  $\text{SOM}^+$ .* In these models, agents choose to be silent if their opinion does not align with the most recent public opinions of *the majority of* their contacts. Unlike in  $\text{SOM}^+$  models, silent agents *are not excluded* from the opinion updates of the agents they influence: when their current opinion is unknown, their most recent public opinion is taken into account in the update. This framework is called the *silence opinion memory-based* ( $\text{SOM}^+$ ) models, as the most recent opinion of each agent is retained.

A property that distinguishes  $\text{SOM}^+$  models from classic DeGroot models (and  $\text{SOM}^-$  models) is that the latter are *Markovian* processes: The next state depends on the current state but not the past states. Thus,  $\text{SOM}^+$  models are *history-dependent* but with very limited memory; only most recent public opinions are remembered. We will show that this minimal notion of memory has an impact on consensus.

*Contributions.* We make the following theoretical and experimental contributions:

1. We generalize the DeGroot model to incorporate key aspects of the Spiral of Silence theory. To the best of our knowledge, this is the first extension of the DeGroot framework to do so.
2. We show that in  $\text{SOM}^-$ , convergence to consensus is guaranteed in clique graphs (i.e., fully connected graphs) with more than two agents. This result highlights that consensus remains possible under Spiral of Silence dynamics, even when silent agents are excluded from updates.
3. We prove that, unlike in the classical DeGroot model,  $\text{SOM}^-$  does not guarantee consensus in all strongly connected aperiodic graphs.
4. We demonstrate that  $\text{SOM}^+$  fails to guarantee consensus even in clique graphs. This negative result underscores that the limited memory in  $\text{SOM}^+$  can significantly alter the dynamics, making consensus harder to achieve under Spiral of Silence assumptions.
5. We validate our models through examples and large-scale simulations involving over *two million agents* on randomly generated networks with small-world properties typical of real social systems. These simulations support core claims of the Spiral of Silence theory, particularly the *reinforcement of dominant views in social networks*. The simulation code is available at: [https://github.com/](https://github.com/DavidGaona/belief_evolution_simulator.git)  
[DavidGaona/belief\\_evolution\\_simulator.git](https://github.com/DavidGaona/belief_evolution_simulator.git)

All in all, this paper highlights the impact of silence dynamics and memory on opinion formation and highlights the limitations of consensus in more nuanced models.

The paper is organized as follows: The new silence opinion models are introduced in Section 2. The study of consensus for these models is presented in Section 3. Our large-scale simulations and case studies emerging from our models and highlighting the spiral of silence effect are presented in Section 4. The concluding remarks are given in Section 5. For the sake of space, all proofs have been moved to the Appendix.

## 2 Opinion Models

In the DeGroot framework [14], each agent updates their opinion by taking the weighted average of the opinions of those who influence them. The models of this framework, however, do not account for the social phenomenon known as the Spiral of Silence [23], where some agents may choose to become or remain silent if their opinion does not align with the majority. As a result, their *current* opinion may not influence their contacts.

In this section, we generalize the DeGroot framework to take into account the Spiral of Silence. If an agent  $j$  decides to be silent, there are at least two natural options when updating the opinions of the agents having  $j$  as a contact: (1) agent  $j$  is simply ignored in the update since their current opinion is unknown (or not public), or (2) the most recent opinion when  $j$  was not silent is taken into account in the update. The former corresponds to a scenario where, for privacy purposes, opinions (messages) are removed once they have been accessed. The latter represents a typical scenario in social networks where previous opinions are kept and thus continue to influence others despite the agent's current silence.

The above options lead us to the two generalizations of DeGroot models studied in this paper: the *silence opinion memoryless* ( $\text{SOM}^-$ ) models, where previous opinions are forgotten, and the *silence opinion memory-based (or history-dependent)* ( $\text{SOM}^+$ ) models, where previous opinions are remembered. In both, agents become silent by a majority rule for each case. Below we introduce the elements of the models.

### 2.1 The Influence Graph

In social learning models, a *community/society* is typically represented as a directed weighted graph with edges between individuals (agents) representing the direction and strength of the influence that one carries over the other. This graph is referred to as the *Influence Graph*.

**Definition 1 (Influence Graph).** An ( $n$ -agent) influence graph is a weighted directed graph  $G = (A, E, I)$ , where  $A = \{1, \dots, n\}$ ,  $E \subseteq A \times A$ , and  $I : A \times A \rightarrow [0, 1]$  a weight function such that  $I(i, j) = 0$  iff  $(i, j) \notin E$  and for each  $i \in A$ ,  $\sum_{j \in N_i \cup \{i\}} I(j, i) = 1$  where  $N_i = \{j \in A \setminus \{i\} : (j, i) \in E\}$ .

The vertices in  $A$  represent  $n$  agents of a given community or network. The set of edges  $E \subseteq A \times A$  represents the (direct) influence relation between these agents; i.e.,  $(i, j) \in E$  means that agent  $i$  (*directly*) influences agent  $j$ . The value  $I(i, j)$ , for simplicity written  $I_{ij}$ , denotes the strength of the influence: 0 means no influence, and a higher value means stronger influence. The normalization condition ensures that the total influence on each agent sums to 1. The set  $N_i$  represents the *neighbors* of agent  $i$ .

We recall some notions from graph theory [16]. A sequence in  $E$  of the form  $(i, i_1)(i_1, i_2) \dots (i_{m-1}, j)$  is a path (of length  $m$ ) from  $i$  to  $j$ . The graph  $G$  is said to be *strongly connected* if for every pair  $(i, j)$  of distinct nodes in  $A$ , there is a path from  $i$  to  $j$ . A graph  $G$  is a *clique* if for every pair  $(i, j)$  of distinct nodes in  $A$ ,  $(i, j) \in E$ . A cycle is a path  $(i, i_1)(i_1, i_2) \dots (i_{m-1}, i)$  with all  $i, i_1, \dots, i_{m-1}$  being distinct. Finally,  $G$  is *aperiodic* if the greatest common divisor of the lengths of its cycles is one.

## 2.2 Silence Opinion Models

To incorporate the Spiral of Silence into the DeGroot framework, we will model the evolution of agents' opinions alongside their decisions to remain silent about a given underlying *statement* or *proposition*. Such a proposition could include controversial statements like, for example, “*AI poses a threat to humanity*” or “*pineapple belongs on pizza*”. Thus, the state of the agents (system) with respect to the proposition involves both the *state of opinion* and the *state of silence*.

The *state of opinion* of all agents is represented as a vector in  $[0, 1]^n$ . If  $\mathbf{B}$  is a state of opinion, then  $\mathbf{B}_i$  denotes the opinion of agent  $i$  with respect to a given proposition. If  $\mathbf{B}_i = 0$  ( $\mathbf{B}_i = 1$ ) agent  $i$  completely disagrees (agrees) with the proposition. The higher the value, the stronger the agreement.

The *state of silence* is represented as a vector in  $\{0, 1\}^n$ . If  $\mathbf{S}$  is a state of silence,  $S_i = 1$  ( $S_i = 0$ ) means that agent  $i$  is *speaking* (is *not speaking*; i.e., agent  $i$  is silent).

At each time  $t \in \mathbb{N}$ , every agent  $i \in A$  updates their opinion and their silence state. We shall use  $\mathbf{B}^t$  and  $\mathbf{S}^t$  to denote the state of opinion and silence at time  $t$ . We can now define a general Silence DeGroot opinion model as follows.

**Definition 2 (SO Models).** A Silence Opinion (SO) model is a tuple  $(G, \mathbf{B}^0, \mathbf{S}^0, \mu_G)$  where  $G = (A, E, I)$  is an  $n$ -agent influence graph,  $\mathbf{B}^0$  the initial state of opinion,  $\mathbf{S}^0$  the initial state of silence,  $\mu_G : [0, 1]^n \times \{0, 1\}^n \times \mathbb{N} \rightarrow [0, 1]^n \times \{0, 1\}^n$  the state-transition function, called (state) update function. For every  $t \in \mathbb{N}$ , the state of the system at time  $t + 1$  is  $(\mathbf{B}^{t+1}, \mathbf{S}^{t+1}) = \mu_G(\mathbf{B}^t, \mathbf{S}^t, t)$ .

The update functions can be used to express any deterministic and discrete transition from one state to the next, possibly taking into account the influence graph, the current and even previous states. These functions are typically expressed by means of equations between states. In what follows, we will define particular update functions that take the spiral of silence into account.

## 2.3 Spiral of Silence Models

To build intuition we recall that the opinion update from the DeGroot model states that each agent adopts the weighted average of the opinions of the agents that directly influence them. This update can be equivalently expressed in terms of opinion differences with their neighbors, as shown in the rightmost formula of the following equation:

$$\mathbf{B}_i^{t+1} = \sum_{j \in N_i \cup \{i\}} I_{ji} \cdot \mathbf{B}_j^t = \mathbf{B}_i^t + \sum_{j \in N_i} I_{ji} \cdot (\mathbf{B}_j^t - \mathbf{B}_i^t) \quad (1)$$

for each  $i \in A$ ,  $t \in \mathbb{N}$ . Thus, in the DeGroot model each agent updates their opinion with the weighted average of the opinion differences with their neighbors.

We now generalize the above DeGroot update (Eq. 1) as opinion update functions that depend not only on the current state of opinion but also on the state of silence and, possibly, on previous states.

**Memoryless Update** Our first update corresponds to the first option mentioned at the beginning of Sec. 2: The opinions of silent neighbors are ignored in the update. This can be realized by modifying Eq. 1 as shown in the following opinion update equation:

$$\mathbf{B}_i^{t+1} = \mathbf{B}_i^t + \sum_{j \in N_i} I_{ji} \cdot \mathbf{S}_j^t \cdot (\mathbf{B}_j^t - \mathbf{B}_i^t) \quad (2)$$

We now define the corresponding silence update function following the Spiral of Silence Theory. First, we need some notation. Let  $x, y, \tau \in [0, 1]$ . The  $\tau$ -proximity relation  $x \sim_\tau y$  holds true iff  $|x - y| \leq \tau$ , i.e., if  $x$  and  $y$  are within a tolerance radius  $\tau$ . Also, let  $N_i^t = \{j \in N_i : \mathbf{S}_j^t = 1\}$  be the sets of non-silent neighbors of  $i$  at time  $t$ . The silence update function is given as follows:

$$\mathbf{S}_i^{t+1} = \begin{cases} 1 & \text{if } \mathcal{M}_i \cdot |N_i^t| \leq |\{j \in N_i^t \mid \mathbf{B}_i^t \sim_{\tau_i} \mathbf{B}_j^t\}| \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

For each agent  $i$ , the constants  $\tau_i, \mathcal{M}_i \in [0, 1]$  represent the agent's *tolerance radius* and *majority threshold* i.e., the minimum proportion of neighbors required for a group to be regarded as a majority by agent  $i$ .<sup>4</sup>

Intuitively, an agent  $i$  considers the opinion of  $j$  to be close enough to theirs if it is within their tolerance radius  $\tau_i$ . Agent  $i$  decides to speak iff at least a fraction  $\mathcal{M}_i$ , their majority threshold, of their *non-silent* contacts have opinions close enough to theirs.

We can now define the memoryless models for the spiral of silence.

**Definition 3 (SOM<sup>-</sup>).** Let  $M = (G, \mathbf{B}^0, \mathbf{S}^0, \mu_G)$  be an SO model with  $G = (A, E, I)$ . Then  $M$  is said to be a SO Memoryless (SOM<sup>-</sup>) model if for each  $i \in A$  and  $t \in \mathbb{N}$ ,  $\mu_G(\mathbf{B}^t, \mathbf{S}^t, t) = (\mathbf{B}^{t+1}, \mathbf{S}^{t+1})$  where  $\mathbf{B}_i^{t+1}$  and  $\mathbf{S}_i^{t+1}$  are determined by Eq. 2 and Eq. 3, respectively.

Clearly, we can recover the DeGroot update (Eq. 1) by setting each tolerance radius constant  $\tau_i$  in Eq. 3 to 1 (or by setting each majority threshold  $\mathcal{M}_i$  to 0) and the initial state of silence  $\mathbf{S}^0$  to the unit vector  $\mathbf{1}_n = (1, 1, \dots, 1)$  of size  $n$ .

*Remark 1.* The dynamic nature of the influence graph in SOM<sup>-</sup> models sets them apart from the static influence in the DeGroot model. Silencing an agent  $j$  at a given time amounts to removing all edges  $(j, i) \in E$  from the graph at that moment. This allows for more complex opinion formation behaviors.

Furthermore, we could have normalized the sum in Eq. 2 by dividing it by  $\sum_{j \in N_i^t} I_{ji}$  when this divisor is not equal to zero.<sup>5</sup> While this would not impact our technical results in Section 3, it would amplify the influence of the non-silent neighbors of  $i$  at time  $t$ , which may seem unnatural. Instead, notice that from Eq. 2 we get:

$$\mathbf{B}_i^{t+1} = \mathbf{B}_i^t + \sum_{j \in N_i} I_{ji} \cdot \mathbf{S}_j^t \cdot (\mathbf{B}_j^t - \mathbf{B}_i^t) = (1 - \sum_{j \in N_i} I_{ji} \cdot \mathbf{S}_j^t) \cdot \mathbf{B}_i^t + \sum_{j \in N_i} I_{ji} \cdot \mathbf{S}_j^t \cdot \mathbf{B}_j^t. \quad (4)$$

---

<sup>4</sup> Although the constraint  $\mathcal{M}_i \geq 0.5$  may seem natural, we neither require nor assume it in our technical results or simulations.

<sup>5</sup> Notice we are using  $N_i^t$  as index set in the summation rather than  $N_i$ .

Thus, the influence  $I_{ji}$  of a silent agent  $j$  at time  $t$  may be seen as increasing the weight of agent  $i$ 's opinion at that time. This can be interpreted as agent  $i$  increasing confidence in their own opinion in the absence of external influence from agent  $j$ .

**Memory-based Update.** We now introduce the models corresponding to second option in the beginning of Sec. 2: If  $j$  is silent at time  $t$ , the opinion update takes into account the opinion they had the last time unit  $u$  (where  $u \leq t$ ) when they were not silent. For this to be well-defined, we assume that initially all agents are not silent; i.e.,  $\mathbf{S}^0 = \mathbf{1}_n$ .

Let  $\bar{t}_j = \max\{u \leq t \mid \mathbf{S}_j^u = 1\}$ . The *public state of opinion* at time  $t$  is a state of opinion  $\text{pubB}^t$  such that  $\text{pubB}_j^t = \mathbf{B}_j^{\bar{t}_j}$  for each  $j \in A$ . The following opinion update equation captures the above intuition:

$$\mathbf{B}_i^{t+1} = \mathbf{B}_i^t + \sum_{j \in N_i} I_{ji} \cdot (\text{pubB}_j^t - \mathbf{B}_i^t) \quad (5)$$

The corresponding silence update tells us that an agent  $i$  becomes or remains silent at time  $t+1$  precisely when the public opinion of the majority of *all* their neighbors are not close enough to their own. More precisely:

$$\mathbf{S}_i^{t+1} = \begin{cases} 1 & \text{if } \mathcal{M}_i \cdot |N_i| \leq |\{j \in N_i \mid \mathbf{B}_i^t \sim_{\tau_i} \text{pubB}_j^t\}| \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

where  $\tau_i, \mathcal{M}_i \in [0, 1]$  are the *tolerance radius* and the *majority threshold* constants, respectively, for agent  $i$ . The memory-based models are defined thus:

**Definition 4 (SOM<sup>+</sup>).** Let  $M = (G, \mathbf{B}^0, \mathbf{S}^0, \mu_G)$  be an SO model where  $G = (A, E, I)$  is an  $n$ -agent influence graph. Then  $M$  is said to be an SO memory-based (SOM<sup>+</sup>) model if  $\mathbf{S}^0 = \mathbf{1}_n$  and for each  $i \in A$  and  $t \in \mathbb{N}$ ,  $\mu_G(\mathbf{B}^t, \mathbf{S}^t, t) = (\mathbf{B}^{t+1}, \mathbf{S}^{t+1})$  where  $\mathbf{B}_i^{t+1}$  and  $\mathbf{S}_i^{t+1}$  are determined by Eq. 5 and Eq. 6, resp.

The DeGroot update (Eq. 1) is a particular case of the SOM<sup>+</sup> opinion update (Eq. 5): We only need to set each tolerance radius constant  $\tau_i$  in Eq. 6 to 1 since  $\mathbf{S}^0$  is already required to be the unit vector of ones  $\mathbf{1}_n$  in SOM<sup>+</sup> models.

*Remark 2.* The main difference between SOM<sup>+</sup> and the DeGroot and SOM<sup>-</sup> models is that the latter two are *Markovian* processes. The next state depends on the current state but not past states. In fact, much of the tractability of DeGroot models derives from its connection to Markov chains. Nevertheless, the next state in SOM<sup>+</sup> models does not depend on the entire state history but just on the most recent public opinions. In the next sections, we will see the impact of this limited amount of memory on opinion evolution.

### 3 Results on Consensus

*Consensus* is a central problem in social learning models. Often, an inability to reach a consensus is a sign of polarization. In the DeGroot framework, consensus represents convergence to the same opinion value over time.

**Definition 5 (Consensus).** Let  $(G, \mathbf{B}^0, \mathbf{S}^0, \mu_G)$  be an SO model with  $G = (A, E, I)$ . We say that the agents in  $A$  converge to consensus if there exists a value  $v \in [0, 1]$  such that for all  $i \in A$ ,  $\lim_{t \rightarrow \infty} \mathbf{B}_i^t = v$ .

Conversely, we refer to the lack of (convergence to) consensus as *dissensus*, which occurs when agents fail to converge to a single opinion value.

In this section, we explore consensus for both types of models on two different graph topologies: *clique* and *strongly connected* graphs. We show that consensus can only be guaranteed in arbitrary  $SOM^-$  models on clique graphs. In all other cases, consensus cannot be guaranteed due to the existence of *perpetual silence* in the case of  $SOM^-$  models and public opinions in the case for  $SOM^+$  models.

Due to space limitations, in this section we provide only sketches of the proofs for the lemmas and theorems; the complete proofs can be found in the Appendix.

### 3.1 SOM<sup>-</sup> Properties

A key property of  $SOM^-$  models is that if all agents become silent at time  $t$ , they will all speak up at the very next round  $t + 1$ . The following lemma formalizes this property:

**Lemma 1.** Let  $(G, \mathbf{B}^0, \mathbf{S}^0, \mu_G)$  be an  $SOM^-$  model with  $G = (A, E, I)$ . For any  $t \in \mathbb{N}$ , if  $\mathbf{S}_i^t = 0$  for all  $i \in A$  then for all  $i \in A$ ,  $\mathbf{S}_i^{t+1} = 1$ .

From Lem. 1, in  $SOM^-$  models we cannot have all agents silent forever.

**Corollary 1.** Let  $(G, \mathbf{B}^0, \mathbf{S}^0, \mu_G)$  be an  $SOM^-$  model with  $G = (A, E, I)$ . For every  $t \in \mathbb{N}$ , there exist  $i \in A$  such that  $\mathbf{S}_i^t = 1$  or  $\mathbf{S}_i^{t+1} = 1$ .

We will now prove that the sequences of maximum and minimum opinion values,  $\{\max(\mathbf{B}^t)\}_{t \in \mathbb{N}}$  and  $\{\min(\mathbf{B}^t)\}_{t \in \mathbb{N}}$ , are (bounded) monotonically non-increasing and non-decreasing, respectively, so they must converge to some opinion value, say  $U$  and  $L$  with  $L \leq U$ . In Section 3.2, we will prove that, under certain conditions,  $U = L$ , which implies that the model converges to consensus.

First, we show that the opinion values in a state are bounded by the extreme opinions in the previous state.

**Lemma 2 (Opinion Bounds).** Let  $(G, \mathbf{B}^0, \mathbf{S}^0, \mu_G)$  be an  $SOM^-$  model with  $G = (A, E, I)$ . For any  $t \in \mathbb{N}$ ,  $\min(\mathbf{B}^t) \leq \mathbf{B}_i^{t+1} \leq \max(\mathbf{B}^t)$  for all  $i \in A$ .

Notice that monotonicity does not necessarily hold for the opinion values of agents. Nevertheless, it follows from Lem. 2 that  $\max(\mathbf{B}^t)$  is monotonically non-increasing and  $\min(\mathbf{B}^t)$  is monotonically non-decreasing with respect to  $t$ .

**Corollary 2 (Monotonicity of Extremes).** Let  $(G, \mathbf{B}^0, \mathbf{S}^0, \mu_G)$  be an  $SOM^-$  model with  $G = (A, E, I)$ . For all  $t \in \mathbb{N}$ ,  $\max(\mathbf{B}^{t+1}) \leq \max(\mathbf{B}^t)$  and  $\min(\mathbf{B}^{t+1}) \geq \min(\mathbf{B}^t)$ .

The monotonicity (and boundedness) of extremes and the Monotonic Convergence Theorem [27], lead us to the existence of limits for the opinion values of extreme agents.

**Theorem 1 (Limits of Extremes).** Let  $(G, \mathbf{B}^0, \mathbf{S}^0, \mu_G)$  be an  $SOM^-$  model with  $G = (A, E, I)$ . There must exist  $U, L \in [0, 1]$  such that  $\lim_{t \rightarrow \infty} \{\max(\mathbf{B}^t)\} = U$  and  $\lim_{t \rightarrow \infty} \{\min(\mathbf{B}^t)\} = L$ .

Henceforth,  $U$  and  $L$  refer to the limit values from Th. 1. Notice that if the limits for the opinion values of extreme agents are the same (i.e.,  $U = L$ ) and by the squeeze theorem [27], we can conclude that all the agents converge to consensus.

### 3.2 Consensus in $SOM^-$ Cliques

In this section, we show that consensus is guaranteed for  $SOM^-$  models whose influence graphs are cliques with at least three agents.

We consider the *minimum influence* of the underlying (clique) graph  $G = (A, E, I)$ , defined as the constant  $I_{\min} = \min_{(i,j) \in E} I(i, j)$ <sup>6</sup> and the difference between the maximum and minimum opinion values at time  $t$ , defined as  $R_t = \max(\mathbf{B}^t) - \min(\mathbf{B}^t)$ , notice that  $R_t \in [0, 1]$  at any time  $t$ .

The proof strategy is based on the following key observations: (a) From any time  $t$  onwards, there will always be non-silent agents (Cor. 1). (b) The maximum and minimum opinion values, which are monotonically non-increasing and non-decreasing respectively, must converge to some value, say  $U$  and  $L$  respectively, with  $L \leq U$  (Th. 1). (c) As the graph is a clique, the non-silent agents will influence, through the update function, all other graph agents infinitely often (Cor. 1). In each such update, the span between the maximum and minimum opinions is multiplied by at most the constant  $1 - I_{\min} < 1$  (Lem. 3). Since this contraction recurs infinitely often, the span tends to zero, which forces all opinions to converge to a common value (Th. 2).

To prove consensus (i.e.,  $U = L$ ), we now demonstrate how the difference between the maximum and minimum opinion values is reduced over different time units.

**Lemma 3 (Convergent Extremes Differences).** Let  $(G, \mathbf{B}^0, \mathbf{S}^0, \mu_G)$  be an  $SOM^-$  model with an  $n$ -agent influence graph  $G = (A, E, I)$  where  $G$  is a clique with  $n \geq 3$ . For all  $m \in \mathbb{N}$ , there exists  $t \in \mathbb{N}$  such that  $R_t \leq R_0 \cdot (1 - I_{\min})^m$ .

As the difference between the maximum and minimum opinions reduces over time, approaching zero, we can now state our consensus result for cliques with at least three agents.

**Theorem 2 (Consensus in  $SOM^-$  Cliques).** Let  $(G, \mathbf{B}^0, \mathbf{S}^0, \mu_G)$  be an  $SOM^-$  model with an  $n$ -agent influence graph  $G = (A, E, I)$  where  $G$  is a clique. If  $n \geq 3$ , then the agents in  $A$  converge to consensus.

*Remark 3.* Notice that for 2-agent cliques ( $n = 2$ ) consensus is not guaranteed; let us consider a clique with two agents: Agent 1 and Agent 2, where the opinions are  $\mathbf{B}_1^0 = 1$  and  $\mathbf{B}_2^0 = 0$ , the influences are  $I(1, 2) = I(2, 1) = 1$ , the tolerance radii are  $\tau_1 = \tau_2 = 1$  and the majority thresholds are  $\mathcal{M}_1 = \mathcal{M}_2 = 0.5$ . In this case, the opinion evolution of agent 1, starting with opinion 1, and agent 2, starting with opinion 0, always alternates between the values 1 and 0. Figure 1 illustrates the opinion evolution of this clique.

---

<sup>6</sup> Notice that  $I_{\min} > 0$  since we are assuming that the underlying graph is a clique.

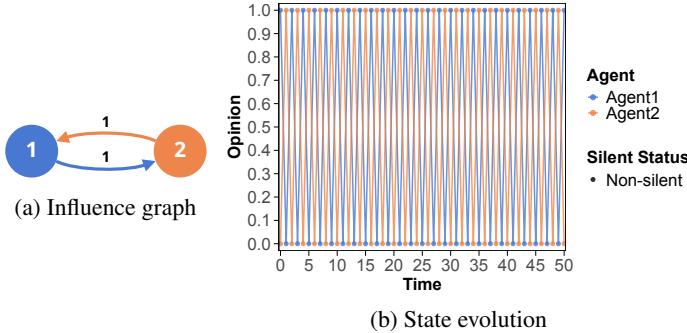


Fig. 1: 2-agent clique opinion evolution;  $\mathbf{B}^0 = (1, 0)$ ,  $\tau_i = 1$ ,  $\mathcal{M}_i = 0.5$ ,  $i \in \{1, 2\}$ .

### 3.3 Dissensus in SOM<sup>-</sup> Models

In strongly connected aperiodic graphs, the SOM<sup>-</sup> models differs from DeGroot models by no longer guaranteeing consensus. Agents can enter a state of perpetual silence, effectively disrupting opinion propagation as if severing connections from the graph. This phenomenon is particularly critical when silent agents form bridges between connected components. Their opinions, influenced by opposing connected components, may remain below their major threshold indefinitely. As a result, they prevent opinion exchange between components, obstructing the possibility of achieving consensus. The following example illustrates this scenario by showing the agents state evolution and influence graph.

*Remark 4.* For visual clarity, self-influences are omitted. As a result, the visible incoming influences for each agent may not sum to one. Nevertheless, self-influences are implicitly present to ensure conformity with Definition 1. Readers can infer an agent's self-influence by subtracting the sum of its visible incoming influences from one.

In the example shown in Fig. 2, Agent 3 serves as the sole bridge connecting two distinct graph components: {Agents 1, 2} and {Agents 4, 5}. Due to the tolerance radius and majority threshold parameters, Agent 3 becomes effectively isolated due to its central position. Specifically, Agent 3's belief (0.5) lies outside the tolerance radius of agents in both components, causing it to have no neighbors with sufficiently similar opinions. Consequently, Agent 3 becomes perpetually silent after the first round.

Meanwhile, the opposing influences from both components on Agent 3's belief update equation cancel each other out the “pull” from the left component (Agents 1, 2) is balanced by an equal and opposite “pull” from the right component (Agents 4, 5). This results in Agent 3's belief remaining fixed while it stays silent. In contrast, agents within each component maintain beliefs within each other's tolerance radii (after the first round), ensuring they continue to speak and influence one another. As a result, each component converges to its own local consensus, while the perpetually silent Agent 3 can no longer transmit influence between them. The graph thus becomes effectively disconnected, preventing global consensus despite the original strong connectivity.

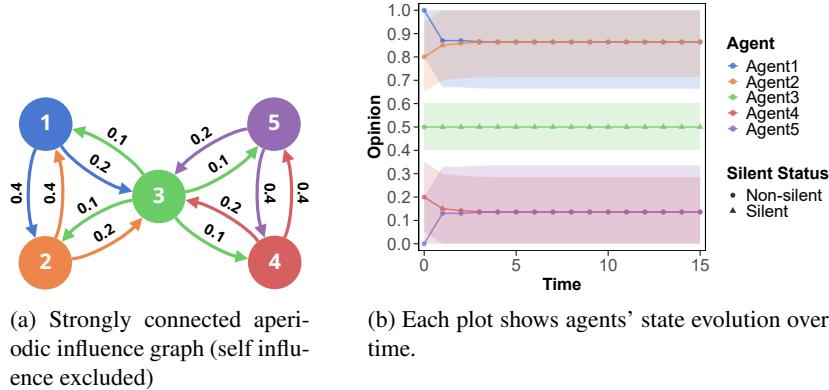


Fig. 2: Examples of Dissensus in  $\text{SOM}^-$  Models. Triangles represent silent agents, circles non-silent ones. Colored areas indicate opinion values within each agent's tolerance radius. Initial state vector:  $\mathbf{B}^0 = (1.0, 0.8, 0.5, 0.2, 0.0)$ ; tolerance radii  $\tau = (0.2, 0.15, 0.1, 0.15, 0.2)$ ; majority thresholds equal to 0.5 for each agent.

### 3.4 Dissensus in $\text{SOM}^+$ Models

We now prove through an insightful counterexample that, unlike in the *memoryless*  $\text{SOM}^-$  models, consensus is not guaranteed in  $\text{SOM}^+$  models for clique graphs. This negative result shows that incorporating a minimal yet natural memory mechanism (where only the most recent public opinions are considered) can actually hinder consensus.

While  $\text{SOM}^+$  models share similarities with  $\text{SOM}^-$  ones regarding perpetual silence, it introduces a unique phenomenon where the entire graph can enter and indefinitely remain in a silent state. This distinction stems from how silent agents influence the opinion update in  $\text{SOM}^+$ .

Recall from Section 2.3 that  $\text{pubB}^t$  is the public state of opinion at time  $t$  and represents the most recent public opinion of each agent. Henceforth, we will refer to  $\text{pubB}_i^t$  as the *public opinion of agent  $i$* . In  $\text{SOM}^+$ , speaking agents influence the opinions of their neighbors as in  $\text{SOM}^-$ . Nevertheless, unlike in  $\text{SOM}^-$ , silent agents influence their neighbors with their public opinion.

Hence, if agent  $i$ 's opinion converges to a value where most recent public opinions of more than a fraction  $\mathcal{M}_i$  of its neighbors fall outside its tolerance range, the agent will remain perpetually silent. This leads to scenarios where agents withdraw from discourse, while other agents continue to be influenced by these outdated public opinions.

Unlike  $\text{SOM}^-$ , where opinions of silent agents are disregarded,  $\text{SOM}^+$  allows for the persistence of unchanged public opinions indefinitely. This can result in dissensus (i.e., lack of convergence to a consensus) due to the formation of public opinions that no longer reflect the current opinions of silent agents.

**Counter-example to Consensus.** Consider the  $\text{SOM}^+$  model  $M = (G, \mathbf{B}^0, \mathbf{S}^0, \mu_G)$  with  $G$  as the Clique in Fig. 3a and  $\mathbf{B}^0 = (1.0, 0.9, 0.1, 0.0)$ , tolerance radius  $\tau_i = 0.1$ ,

majority threshold  $\mathcal{M}_i = 0.5$  for each agent  $i$  in  $G$ . All agents initially have more than half of their neighbors' opinions outside their tolerance radius. Consequently, all agents become silent at  $t = 1$ . At this point, each agent's updated opinion places all other agents' opinions outside its tolerance radius. The agents' opinions then converge to distinct values, determined by the initial opinions  $\mathbf{B}^0$ , which perpetuates the condition for silence. This state persists indefinitely, as the convergence values maintain the silence condition for all agents. This results in dissensus as illustrated in Fig.3b .

The above counter-example demonstrates how the SOM<sup>+</sup> models can lead to complete silence and opinion divergence in a clique, preventing the possibility of consensus.

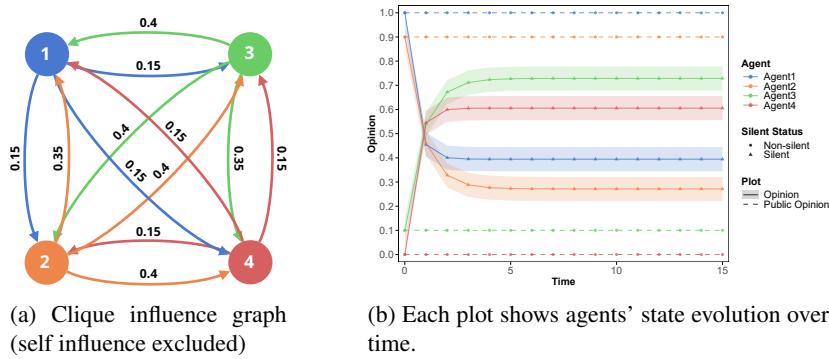


Fig. 3: Counter-Example to Consensus in an SOM<sup>+</sup> Model. Triangles represent silent agents, circles non-silent ones. Colored areas indicate opinion values within each agent's tolerance radius. Initial state vector:  $\mathbf{B}^0 = (1.0, 0.9, 0.1, 0.0)$ , tolerance radius and majority threshold equal to 0.1 and 0.5 for each agent.

## 4 Experimenting with the Spiral of Silence

This section demonstrates how SOM<sup>-</sup> and SOM<sup>+</sup> models can capture key dynamics of the spiral of silence through agent-based simulations. We first showcase an SOM<sup>-</sup> model that reinforces dominant opinions, enabling vocal minorities to disproportionately influence public discourse despite numerical inferiority. Next, we illustrate an SOM<sup>+</sup> model revealing an apparent dichotomy: agents privately converge to *consensus* while maintaining *divergent* public stances, mirroring social environments where outdated expressions mask underlying agreement [18].

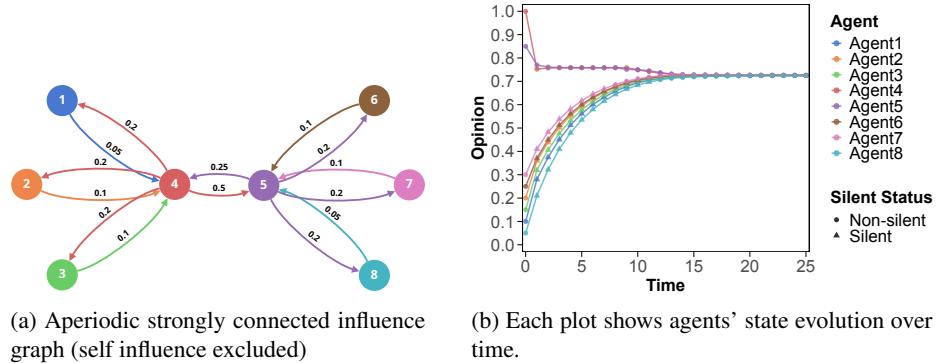
To validate scalability, we simulate networks ranging from 4 to over 2.1 million agents ( $2^2$  to  $2^{21}$ ) with varying connectivity, leveraging a modified version of Ross et al.'s algorithm [24] to generate networks with power-law ("rich-get-richer") degree distributions and small-world properties. Small-world properties are characterized by short average path lengths between nodes while maintaining high local clustering. This reflects *real social networks'* structure where information spreads efficiently despite

limited local connections, making our simulations more realistic for studying opinion dynamics. Indeed, studies have shown social networks exhibit 6 degrees of separation [28], that is, any one person is on average 6 connections away from any other.

#### 4.1 Reinforcement of Dominant Views

Figure 4 illustrates how SOM<sup>-</sup> models capture the essence of the Spiral of Silence theory. In this strongly connected graph, a vocal minority effectively dominates the discourse, causing the silent majority to converge toward the perceived majority opinion.

The network has two groups: Group 1 (Agents 1, 2, 3, 6, 7, and 8) with opinions in the lower half of the spectrum, and Group 2 (Agents 4 and 5) with opposing views near the higher end. Despite Group 1 being the actual majority, the network topology and differing tolerance radii  $\tau$  lead to Group 2 dominating the opinion dynamics. Group 1 agents lack intra-group connections but are all linked to Group 2 agents. Also, Group 2's significantly larger tolerance radii allow them to remain non-silent more frequently.



(a) Aperiodic strongly connected influence graph (self influence excluded) (b) Each plot shows agents' state evolution over time.

Fig. 4: Silent Majority vs. Vocal Minority: Opinion Dynamics in SOM<sup>-</sup>. Triangles represent silent agents, circles non-silent ones. Colored areas indicate opinion values within each agent's tolerance radius. Here  $B^0 = (0.1, 0.2, 0.15, 1.0, 0.85, 0.25, 0.3, 0.05)$  and  $\tau = (0.1, 0.05, 0.1, 0.85, 0.6, 0.05, 0.1, 0.05)$  and majority thresholds are set to 0.5.

This configuration results in the vocal minority (Group 2) disproportionately influencing the network. The silent majority (Group 1) remains quiet for most of the update process, only becoming active when opinions have already shifted closer to the perceived majority view. This example demonstrates how the SOM<sup>-</sup> models can simulate scenarios where a minority opinion, through strategic positioning and persistent vocalization, can shape the overall opinion landscape, even when numerically outnumbered.

## 4.2 Hidden Consensus in SOM<sup>+</sup> Models

SOM<sup>+</sup> models reveal a noteworthy phenomenon: the possibility of reaching a consensus that remains undetected by the agents themselves. Figure 5 illustrates this scenario using a clique graph with four agents.

Initially, the agents hold diverse opinions ( $\mathbf{B}^0 = (1.0, 0.9, 0.1, 0.0)$ ). However, due to the graph's influence structure, all agents become silent after  $t = 0$ . Despite this silence, their private opinions converge to a common value over time. Crucially, this convergence occurs without any further public expression of opinions, leaving each agent unaware of the emerging consensus. This hidden consensus phenomenon mirrors real-world scenarios in social media where individuals may unknowingly share common views while perceiving disagreement due to outdated public expressions [18].

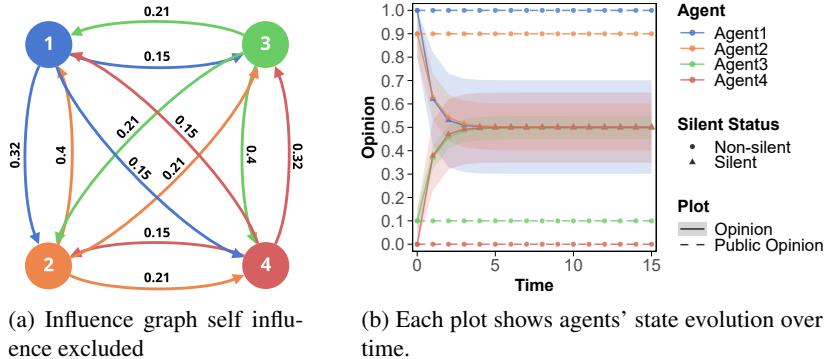


Fig. 5: Silent Consensus Scenario. Triangles represent silent agents, circles non-silent ones. Colored areas indicate opinion values within each agent's tolerance radius. Initial state vector:  $\mathbf{B}^0 = (1.0, 0.9, 0.1, 0.0)$ ; tolerance radii  $\tau = (0.2, 0.1, 0.05, 0.15)$ ; majority thresholds equal to 0.5 for each agent.

## 4.3 Scaling to Large-Scale Network Simulations

While small-scale simulations illustrate core mathematical behaviors, they do not capture the complexity of real-world social networks. To bridge this gap, we developed a scalable simulation platform supporting networks of over 2.1 million agents ( $2^{21}$ ), enabling analysis of emergent phenomena in realistic settings.

Our implementation combines a modified preferential attachment algorithm [1] to generate networks with small-world properties [29] and power-law (rich-get-richer) degree distributions. To address computational intensity, we employ parallel processing via Scala and Akka Actors [20,19], with results persisted in PostgreSQL for efficient querying. For each experiment, we generated networks with *density* (minimum neighbor count) ranging from 1 to 15, executing 1,024 simulations per parameter combination. Consensus rates were then calculated across all runs for statistical significance. All

agents have their tolerance radius and majority threshold set to 0.1 and 0.5, respectively, and their initial beliefs as random uniformly distributed values from 0 to 1. The code is available [here](#).

**Contrasting Silence Opinion Models.** The  $\text{SOM}^+$  and  $\text{SOM}^-$  models exhibit dramatically opposing behaviors in consensus formation as seen in Fig. 6. For  $\text{SOM}^+$  models, consensus becomes increasingly rare beyond 512 agents ( $2^9$ ), occurring in only 0.097% of simulations (1/1024 runs) at this threshold, exclusively for density 4. Beyond this scale, no simulations reached consensus. This aligns with social media dynamics where persistent disagreement emerges, as agents' historical opinions create perception gaps between private consensus and public expression [18]. Even a minimal notion of memory (retaining only most recent public opinions) proves sufficient to sustain dissensus.

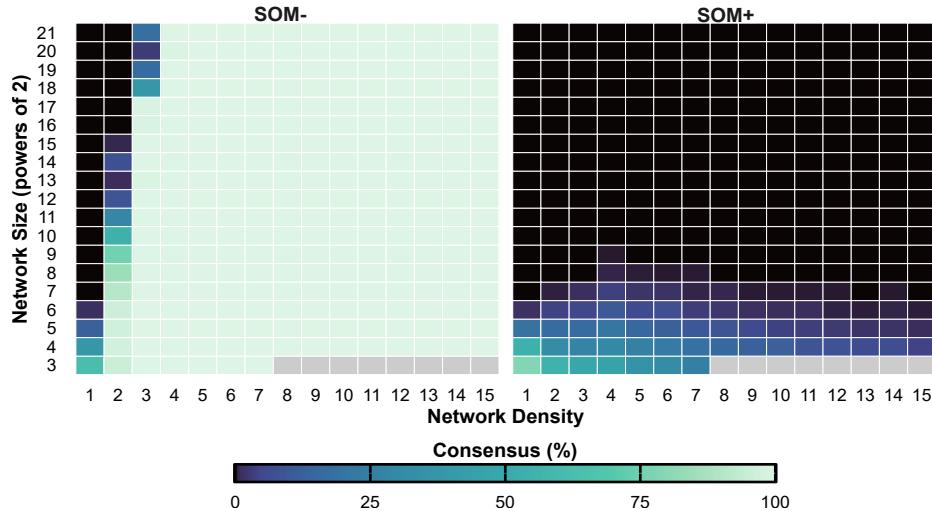


Fig. 6: Heatmap indicating consensus % across different network sizes (y-axis) and densities (x-axis) for both  $\text{SOM}^+$  and  $\text{SOM}^-$  models. Lighter hues indicate a higher rate of consensus, while darker hues indicate a lower rate. All tolerance radii and majority thresholds are equal to 0.1 and 0.5 resp.

Conversely,  $\text{SOM}^-$  models exhibit the opposite scalability behavior. Once a certain density threshold is reached, consensus consistently emerges in all simulations. Larger networks require slightly higher connectivity: for 2,048 agents ( $2^{11}$ ), a density of 3 suffices, while for 2.1 million agents ( $2^{21}$ ), a density of 4 is needed. This highlights how the spiral of silence depends on network structure. In sparse networks, some agents may remain silent indefinitely due to critical bridge nodes (see Section 3.3). In contrast, denser networks make it increasingly unlikely for bridges to be disconnected, since disrupting consensus would require silencing more than  $d$  agents (for density  $d$ )—a rare occurrence in networks of this type.

## 5 Conclusions and Related Work

We extended the classical DeGroot framework to incorporate key social dynamics described by the Spiral of Silence. Our contributions highlight how the addition of memory, even in a limited form, as in  $SOM^+$  models, introduces significant complexity to consensus-building processes. We demonstrated that consensus, while achievable for cliques in  $SOM^-$  models, is no longer guaranteed, in contrast with DeGroot models, in arbitrary strongly-connected aperiodic graphs. This points to the impact that silence and memory can have on opinion formation in social networks. It also offers insights into the challenges of converging to consensus in real-world scenarios. We also discuss simulations reflecting predictions of the Spiral of Silence, such as the reinforcement of dominant views and hidden consensus. Finally, we discussed the simulation in our models of large-scale graphs reflecting the small-world topology of real social networks.

*Future Work.* Experimentation and validation on real social networks typically involve continuously collecting user opinions and inferring influence via methods such as stance detection and influencer identification from natural language processing, machine learning and graph mining. We defer this challenging task to future work.

*Related Work.* To our knowledge, no prior work has extended DeGroot-based models to incorporate the Spiral of Silence. Some recent studies have examined the Spiral of Silence in agent-based networks. E.g., [24] models manipulative actors in social networks using agents with fixed binary opinions (agree/disagree) who choose whether to express them based on the prevailing opinion climate. In [8], the authors explore how the number of communities and their connectivity influence perceived opinion climates. However, these studies do not address opinion updates, convergence to consensus, or memory, which are central aspects of this paper.

In [12,33] the stabilization property was studied in majority-based opinion models, where individuals adopt the most prevalent opinion in their social circles (their neighbors). Although these works focus on majority-based opinion dynamics, similar to models inspired by Spiral of Silence, their models are exclusively memoryless with only two-value discrete opinions.

Recent studies [7,25,32] explored conflict, disagreement, and polarization using the Friedkin and Johnsen (FJ) model. This model extended DeGroot's by adding a fixed internal opinion and modeling stubbornness, where agents' expressed opinions reflected their internal beliefs. While silent agents in  $SOM^+$  and stubborn agents in the FJ model shared some similarities, their behavior differed. In  $SOM^+$ , silent agents could later become active, allowing their public opinions to change. In contrast, stubborn agents in the FJ model did not shift in this way. Thus, although neither model guaranteed consensus, their opinion dynamics diverged.

The exclusion of silent agents in  $SOM^-$  models amounts to having edges (influences) disappearing and reappearing during opinion evolution which reflects the dynamic influence nature of this model. There are several works studying dynamic influence in opinion formation. The work [15] introduces a version of the DeGroot model in which self-influence changes over time while the influence on others remains the same. The works [10,11] explore convergence and stability, respectively, in models where influences change over time. The work [5] demonstrates how asynchronous communi-

cation, when combined with dynamic influence, can prevent consensus. None of this work deals with the dynamics derived from the Spiral of Silence.

## References

1. Albert, R., Barabási, A.L.: Statistical mechanics of complex networks. *Reviews of Modern Physics* **74**(1), 47–97 (2002). <https://doi.org/10.1103/RevModPhys.74.47>
2. Alvim, M.S., Amorim, B., Knight, S., Quintero, S., Valencia, F.: A Multi-agent Model for Polarization Under Confirmation Bias in Social Networks. In: 41th International Conference on Formal Techniques for Distributed Objects, Components, and Systems (FORTE) (2021), <https://inria.hal.science/hal-03740263>
3. Alvim, M.S., Amorim, B., Knight, S., Quintero, S., Valencia, F.: A Formal Model for Polarization under Confirmation Bias in Social Networks. *Logical Methods in Computer Science* (Mar 2023). [https://doi.org/10.46298/lmcs-19\(1:18\)2023](https://doi.org/10.46298/lmcs-19(1:18)2023), <https://hal.science/hal-03872692>
4. Alvim, M.S., Gaspar da Silva, A., Knight, S., Valencia, F.: A multi-agent model for opinion evolution in social networks under cognitive biases. In: Castiglioni, V., Francalanza, A. (eds.) *Formal Techniques for Distributed Objects, Components, and Systems*. pp. 3–19. Springer Nature Switzerland, Cham (2024)
5. Aranda, J., Betancourt, S., Díaz, J.F., Valencia, F.: Fairness and consensus in an asynchronous opinion model for social networks. In: Majumdar, R., Silva, A. (eds.) 35th International Conference on Concurrency Theory, CONCUR 2024, September 9–13, 2024, Calgary, Canada. LIPIcs, vol. 311, pp. 7:1–7:17. Schloss Dagstuhl - Leibniz-Zentrum für Informatik (2024). <https://doi.org/10.4230/LIPICS.CONCUR.2024.7>
6. Atanesyan, A., Hakobyan, A., Reynolds, B.: Communicating covid-19 on social media: The effects of the spiral of silence. University of Helsinki Research Portal (2023), <https://researchportal.helsinki.fi/en/publications/communicating-covid-19-on-social-media-the-effects-of-the-spiral>
7. Biondi, E., Boldrini, C., Passarella, A., Conti, M.: Dynamics of opinion polarization. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* **53**(9), 5381–5392 (2023)
8. Cabrera, B., Ross, B., Röchert, D., Brünker, F., Stieglitz, S.: The influence of community structure on opinion expression: an agent-based model. *Journal of Business Economics* **91**, 1–25 (11 2021). <https://doi.org/10.1007/s11573-021-01064-7>
9. Chandrasekhar, A.G., Larreguy, H., Xandri, J.P.: Testing models of social learning on networks: Evidence from a lab experiment in the field. Working Paper 21468, National Bureau of Economic Research (August 2015). <https://doi.org/10.3386/w21468>
10. Chatterjee, S., Seneta, E.: Towards consensus: Some convergence theorems on repeated averaging. *Journal of Applied Probability* **14**(1), 89–97 (1977), <https://doi.org/10.2307/3213262>
11. Chen, Z., Qin, J., Li, B., Qi, H., Buchhorn, P., Shi, G.: Dynamics of opinions with social biases. *Automatica* **106**, 374–383 (2019). <https://doi.org/10.1016/j.automatica.2019.04.035>
12. Chistikov, D., Lisowski, G., Paterson, M., Turrini, P.: Convergence of opinion diffusion is pspace-complete. In: Proceedings of the AAAI conference on artificial intelligence. vol. 34, pp. 7103–7110 (2020)
13. Dandekar, P., Goel, A., Lee, D.: Biased assimilation, homophily and the dynamics of polarization. *Proceedings of the National Academy of Sciences of the United States of America* **110** (03 2013). <https://doi.org/10.1073/pnas.1217220110>
14. DeGroot, M.H.: Reaching a consensus. *Journal of the American Statistical Association* **69**(345), 118–121 (1974)
15. DeMarzo, P.M., et al.: Persuasion bias, social influence, and unidimensional opinions. *The Quarterly Journal of Economics* **118**(3), 909–968 (2003), <http://www.jstor.org/stable/25053927>

16. Diestel, R.: Graph Theory. Springer, 5th edn. (2017)
17. Golub, B., Sadler, E.: Learning in social networks. Available at SSRN 2919146 (2017)
18. Hampton, K.N., Rainie, L., Lu, W., Dwyer, M., Shin, I., Purcell, K.: Social media and the 'spiral of silence'. Tech. rep., Pew Research Center, Washington, DC (2014), <https://www.pewresearch.org/internet/2014/08/26/social-media-and-the-spiral-of-silence/>
19. Hewitt, C.: Actor model for discretionary, adaptive concurrency. CoRR **abs/1008.1459** (2010), <http://arxiv.org/abs/1008.1459>
20. Hewitt, C., Bishop, P., Steiger, R.: A universal modular ACTOR formalism for artificial intelligence. In: Proceedings of the 3rd International Joint Conference on Artificial Intelligence. pp. 235–245. Morgan Kaufmann Publishers Inc., Stanford, CA (August 1973)
21. Matthes, J., Knoll, J., von Sikorski, C.: The "spiral of silence" revisited: A meta-analysis on the relationship between perceptions of opinion support and political opinion expression. *Communication Research* **45**(1), 3–33 (2018). <https://doi.org/10.1177/0093650217745429>
22. Mossel, E., Tamuz, O.: Opinion exchange dynamics. *Probability Surveys* **14**(none), 155 – 204 (2017). <https://doi.org/10.1214/14-PS230>
23. Noelle-Neumann, E.: The spiral of silence: A theory of public opinion. *Journal of Communication* **24**(2), 43–51 (1974). <https://doi.org/10.1111/j.1460-2466.1974.tb00367.x>
24. Ross, B., Pilz, L., Cabrera, B., Brachten, F., Neubaum, G., Stieglitz, S.: Are social bots a real threat? an agent-based model of the spiral of silence to analyse the impact of manipulative actors in social networks. *European Journal of Information Systems* **28**(4), 394–412 (2019). <https://doi.org/10.1080/0960085X.2018.1560920>
25. Shirzadi, M., Zehmakan, A.N.: Do stubborn users always cause more polarization and disagreement? a mathematical study. arXiv preprint arXiv:2410.22577 (2024)
26. Sohn, D.: Spiral of silence in the social media era: A simulation approach to the interplay between social networks and mass media. *Communication Research* **49**(1), 139–166 (2022). <https://doi.org/10.1177/0093650219856510>
27. Sohrab, H.H.: Basic Real Analysis. Birkhauser Basel, 2nd edn. (2014). <https://doi.org/10.1007/0-8176-4441-5>
28. Travers, J., Milgram, S.: An experimental study of the small world problem. *Sociometry* **32**(4), 425–443 (1969)
29. Watts, D.J., Strogatz, S.H.: Collective dynamics of ‘small-world’ networks. *Nature* **393**(6684), 440–442 (1998). <https://doi.org/10.1038/30918>
30. X, C., P, T., J, L., T, D.B.: Opinion dynamics with backfire effect and biased assimilation. *PLoS ONE* **16**(9) (2021). <https://doi.org/https://doi.org/10.1371/journal.pone.0256922>
31. Xia, W., Ye, M., Liu, J., Cao, M., Sun, X.M.: Analysis of a nonlinear opinion dynamics model with biased assimilation. *Automatica* **120**, 109113 (2020). <https://doi.org/10.1016/j.automatica.2020.109113>
32. Xu, W., Zhu, L., Guan, J., Zhang, Z., Zhang, Z.: Effects of stubbornness on opinion dynamics (2022), <https://arxiv.org/abs/2208.04160>
33. Zehmakan, A.N.: Majority opinion diffusion in social networks: An adversarial approach. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 35, pp. 5611–5619 (2021)

## A Appendix

**Lemma 1.** Let  $(G, \mathbf{B}^0, \mathbf{S}^0, \mu_G)$  be an  $SOM^-$  model with  $G = (A, E, I)$ . For any  $t \in \mathbb{N}$ , if  $\mathbf{S}_i^t = 0$  for all  $i \in A$  then for all  $i \in A$ ,  $\mathbf{S}_i^{t+1} = 1$ .

**Proof.** Let  $\mathbf{S}^t$  such that  $\mathbf{S}_i^t = 0$  for all  $i \in A$ . Then, by applying Eq. 3, we can conclude that  $\mathbf{S}_i^{t+1} = 1$  for all  $i \in A$ , given that  $|N_i^t| = 0$  for all  $i \in A$ ,  $\square$

**Lemma 2 (Opinion Bounds).** Let  $(G, \mathbf{B}^0, \mathbf{S}^0, \mu_G)$  be an  $SOM^-$  model with  $G = (A, E, I)$ . For any  $t \in \mathbb{N}$ ,  $\min(\mathbf{B}^t) \leq \mathbf{B}_i^{t+1} \leq \max(\mathbf{B}^t)$  for all  $i \in A$ .

**Proof.** From Definition 3 and Eq. 4 we know that:

$$\mathbf{B}_i^{t+1} = (1 - \sum_{j \in N_i} I_{ji} \cdot \mathbf{S}_j^t) \cdot \mathbf{B}_i^t + \sum_{j \in N_i} I_{ji} \cdot \mathbf{S}_j^t \cdot \mathbf{B}_j^t$$

Now, as  $\mathbf{B}_k^t \leq \max(\mathbf{B}^t)$  for all  $k \in A$ , we can conclude that:

$$\mathbf{B}_i^{t+1} \leq (1 - \sum_{j \in N_i} I_{ji} \cdot \mathbf{S}_j^t) \cdot \max(\mathbf{B}^t) + \sum_{j \in N_i} I_{ji} \cdot \mathbf{S}_j^t \cdot \max(\mathbf{B}^t) = \max(\mathbf{B}^t)$$

Thus,  $\mathbf{B}_i^{t+1} \leq \max(\mathbf{B}^t)$  as wanted. The proof that  $\min(\mathbf{B}^t) \leq \mathbf{B}_i^{t+1}$  is similar.  $\square$

**Lemma 3 (Convergent Extremes Differences).** Let  $(G, \mathbf{B}^0, \mathbf{S}^0, \mu_G)$  be an  $SOM^-$  model with an  $n$ -agent influence graph  $G = (A, E, I)$  where  $G$  is a clique with  $n \geq 3$ . For all  $m \in \mathbb{N}$ , there exists  $t \in \mathbb{N}$  such that  $R_t \leq R_0 \cdot (1 - I_{\min})^m$ .

**Proof.** First, we prove that at each time unit  $t$  where there is at least one non-silent agent, the distance between the maximum and minimum opinions will be reduced by at least a factor of  $(1 - I_{\min})$  at time  $t + 1$ , i.e.  $R_{t+1} \leq (1 - I_{\min}) \cdot R_t$  where  $1 - I_{\min}$  is a positive constant below one.

We consider the following two cases for each unit time  $t$  where there is at least one non-silent agent:

- At least one agent is silent. In this case, we consider a non-silent agent  $k$ , the updated opinion of every agent  $i$  influenced by agent  $k$  will have an upper bound, thus:

$$\mathbf{B}_i^{t+1} = (1 - \sum_{j \in N_i} I_{ji} \cdot \mathbf{S}_j^t) \cdot \mathbf{B}_i^t + \sum_{j \in N_i} I_{ji} \cdot \mathbf{S}_j^t \cdot \mathbf{B}_j^t \leq (1 - I_{\min}) \cdot \max(\mathbf{B}^t) + I_{\min} \cdot \mathbf{B}_k^t$$

And the lower bound would be the following:

$$\mathbf{B}_i^{t+1} = (1 - \sum_{j \in N_i} I_{ji} \cdot \mathbf{S}_j^t) \cdot \mathbf{B}_i^t + \sum_{j \in N_i} I_{ji} \cdot \mathbf{S}_j^t \cdot \mathbf{B}_j^t \geq (1 - I_{\min}) \cdot \min(\mathbf{B}^t) + I_{\min} \cdot \mathbf{B}_k^t$$

Therefore:

$$(1 - I_{min}) \cdot \max(\mathbf{B}^t) + I_{min} \cdot \mathbf{B}_k^t \geq \mathbf{B}_i^{t+1} \geq (1 - I_{min}) \cdot \min(\mathbf{B}^t) + I_{min} \cdot \mathbf{B}_k^t \quad (7)$$

As the graph is a clique and there is at least one silent agent, every agent in the graph is influenced by agent  $k$ , including itself; notice that as there is at least one silent agent, call it  $l$ , then  $(1 - \sum_{j \in N_i} I_{ji})$  from Eq. 4, corresponding to the self-influence on agent  $i$ , is at least  $I_{li} \geq I_{min}$ . Therefore, all agents in the graph satisfy Eq. 7; then, there are both an upper and lower bound for all opinions in unit time  $t + 1$ , including the maximum and the minimum opinions, thus:

$$\max(\mathbf{B}^{t+1}) \leq (1 - I_{min}) \cdot \max(\mathbf{B}^t) + I_{min} \cdot \mathbf{B}_k^t \quad (8)$$

$$\min(\mathbf{B}^{t+1}) \geq (1 - I_{min}) \cdot \min(\mathbf{B}^t) + I_{min} \cdot \mathbf{B}_k^t \quad (9)$$

Then, as  $R_{t+1} = \max(\mathbf{B}^{t+1}) - \min(\mathbf{B}^{t+1})$ ,  $R_{t+1} \leq (1 - I_{min}) \cdot \max(\mathbf{B}^t) + I_{min} \cdot \mathbf{B}_k^t - (1 - I_{min}) \cdot \min(\mathbf{B}^t) - I_{min} \cdot \mathbf{B}_k^t = (1 - I_{min}) \cdot (\max(\mathbf{B}^t) - \min(\mathbf{B}^t)) = (1 - I_{min}) \cdot R_t$ . Then  $R_{t+1} \leq (1 - I_{min}) \cdot R_t$ .

2. All agents speak at time  $t$ : as the graph is a clique with  $n \geq 3$ , there must be at least one non-silent agent  $k$  that is not the unique maximum or minimum opinion agent in the graph<sup>7</sup>; agent  $k$  influences all the other agents such that their opinions are bounded as follows:

$$\mathbf{B}_i^{t+1} = (1 - \sum_{j \in N_i} I_{ji} \cdot \mathbf{S}_j^t) \cdot \mathbf{B}_i^t + \sum_{j \in N_i} I_{ji} \cdot \mathbf{S}_j^t \cdot \mathbf{B}_j^t \leq (1 - I_{min}) \cdot \max(\mathbf{B}^t) + I_{min} \cdot \mathbf{B}_k^t$$

$$\mathbf{B}_i^{t+1} = (1 - \sum_{j \in N_i} I_{ji} \cdot \mathbf{S}_j^t) \cdot \mathbf{B}_i^t + \sum_{j \in N_i} I_{ji} \cdot \mathbf{S}_j^t \cdot \mathbf{B}_j^t \geq (1 - I_{min}) \cdot \min(\mathbf{B}^t) + I_{min} \cdot \mathbf{B}_k^t$$

As agent  $k$  is influenced by an agent with maximum opinion and an agent with minimum opinion at time  $t$ , its opinion at time  $t + 1$  will be bounded as follows:

$$\mathbf{B}_k^{t+1} \geq (1 - I_{min}) \cdot \min(\mathbf{B}^t) + I_{min} \cdot \max(\mathbf{B}^t)$$

$$\mathbf{B}_k^{t+1} \leq (1 - I_{min}) \cdot \max(\mathbf{B}^t) + I_{min} \cdot \min(\mathbf{B}^t)$$

Therefore:

$$\mathbf{B}_k^{t+1} \geq (1 - I_{min}) \cdot \min(\mathbf{B}^t) + I_{min} \cdot \max(\mathbf{B}^t) \geq (1 - I_{min}) \cdot \min(\mathbf{B}^t) + I_{min} \cdot \mathbf{B}_k^t$$

and

---

<sup>7</sup> Notice that this case is not applicable in the clique in Remark 3, as such a clique does not have a non-silent agent that is not the unique maximum or minimum opinion agent in the graph.

$$\mathbf{B}_k^{t+1} \leq (1 - I_{min}) \cdot \max(\mathbf{B}^t) + I_{min} \cdot \min(\mathbf{B}^t) \leq (1 - I_{min}) \cdot \max(\mathbf{B}^t) + I_{min} \cdot \mathbf{B}_k^t$$

Therefore, every agent in the graph, including  $k$ , holds the following bounds:

$$(1 - I_{min}) \cdot \max(\mathbf{B}^t) + I_{min} \cdot \mathbf{B}_k^t \geq \mathbf{B}_i^{t+1} \geq (1 - I_{min}) \cdot \min(\mathbf{B}^t) + I_{min} \cdot \mathbf{B}_k^t \quad (10)$$

Then, as  $R_{t+1} = \max(\mathbf{B}^{t+1}) - \min(\mathbf{B}^{t+1})$ ,  $R_{t+1} \leq (1 - I_{min}) \cdot \max(\mathbf{B}^t) + I_{min} \cdot \mathbf{B}_k^t - (1 - I_{min}) \cdot \min(\mathbf{B}^t) - I_{min} \cdot \mathbf{B}_k^t = (1 - I_{min}) \cdot (\max(\mathbf{B}^t) - \min(\mathbf{B}^t)) = (1 - I_{min}) \cdot R_t$ . Then  $R_{t+1} \leq (1 - I_{min}) \cdot R_t$ .

Thus, we have proved that  $R_{t+1} \leq (1 - I_{min}) \cdot R_t$  for any time unit  $t$  where there is at least one non-silent agent.

Finally, by using the above and Cor. 1, it can be proved, by induction, that for all  $m \in \mathbb{N}$ , there exists  $t \in \mathbb{N}$  such that  $R_t \leq R_0 \cdot (1 - I_{min})^m$ .  $\square$

**Theorem 2 (Consensus in SOM<sup>-</sup> Cliques).** *Let  $(G, \mathbf{B}^0, \mathbf{S}^0, \mu_G)$  be an SOM<sup>-</sup> model with an  $n$ -agent influence graph  $G = (A, E, I)$  where  $G$  is a clique. If  $n \geq 3$ , then the agents in  $A$  converge to consensus.*

**Proof.** From Th. 1, there exist  $U, L \in [0, 1]$  such that  $U = \lim_{t \rightarrow \infty} \max(\mathbf{B}^t)$  and  $L = \lim_{t \rightarrow \infty} \min(\mathbf{B}^t)$ . We now prove  $U = L$  by contradiction. Suppose that  $U \neq L$ . As  $U$  can not be lower than  $L$ , therefore,  $U - L > 0$ . From  $\lim_{m \rightarrow \infty} R_0 \cdot (1 - I_{min})^m = 0$ <sup>8</sup> and Lem. 3 there must exists a time  $t$  where  $R_t < (U - L)$ , but it is not possible as, from Th. 1 and Cor. 2,  $\min(\mathbf{B}^t) \leq L$  and  $\max(\mathbf{B}^t) \geq U$ , which is a contradiction. Therefore,  $U = L$  and the model converges to consensus.  $\square$

---

<sup>8</sup> Notice that  $R_0$  and  $I_{min}$  are positive constants such that  $1 > (1 - I_{min}) > 0$ .