Influence measurement in a complex dynamical model: an information theoretic approach

Pierre-Alain Toupance, Laurent Lefèvre, Bastien Chopard

Innfluence measurement in a complex dynamical model: an information theoretic approach

Pierre-Alain Toupance[a,b], Laurent Lefèvre[b], and Bastien Chopard[a]

[a] *University of Geneva, Geneva, Switzerland*

[b] *Univ. Grenoble Alpes, LCIS, F-26902, Valence, France*

Abstract

We address the link between the controllability or observability of a stochastic complex system and concepts of information theory. We show that the most influential degrees of freedom can be detected without acting on the system, by measuring the *time-delayed multi-information*. Numerical and analytical results support this claim, which is developed in the case of a simple stochastic model on a graph, the so-called voter model. The importance of the noise when controlling the system is demonstrated, leading to the concept of *control length*. The link with classical control theory is given, as well as the interpretation of controllability in terms of the capacity of a communication canal.

Keywords: reachability and observability analysis, Mutual and multi-information, voter model

1. Introduction

Causality is an important concept in many areas of science [20]. It helps to better understand the behavior of complex dynamical systems. In particular, it reveals how the different degrees of freedom of a system influence each other. In this paper we investigate how causality (considered in

a pragmatic and intuitive way) can be used to discover efficient control strategies in a complex system.

The key ingredient of our approach is the concept of the most influential components in a complex system. This notion is defined here as the impact of controlling a given variable on the behaviour of the other variables. For instance, one can measure the change in the joint probability distribution (Kulback-Leibler divergence) when the value of a selected variable is imposed. Alternatively, we can measure the variation of an average quantity when a perturbation is applied. The variable for which this change is the most important is labelled as the most influential. Following this procedure we can rank the degrees of freedom of a system from the most to the less influential. Arguably the notion of influence depends on the quantity used to measure the effect of forcing the variable. Then, to control this quantity in the system, it will be more effective to act on the corresponding most influential nodes.

The aforementioned procedure is intrusive in the sense that it requires to act on the system to be able to determine the effects of a perturbation. Here we would like to consider a non-intrusive approach, essentially based on the observation of the system. The non-intrusive approach we proposed is based on a time delayed multi-information on the free system, measure that Schreiber [23] has named transfer information. This procedure can be performed by simple sampling on the system variables, even if the underlying dynamics is unknown, like for instance in financial systems.

In order to show that the metrics of information theory are good tools to find the influencers of a dynamic system, we have chosen here as a dynamic system a voting model where the vote of an agent depends on the vote of its neighbours. Voting models have been studied in many articles. For example Castellano and all [7] have defined a q-voter model in which an agent votes like its neighbours if the opinion is unanimous; otherwise the vote is random. This model has been used by Nycska [18]. Our model is closer to those used by Mobilia [15] or Masuda [13]. In our case, we want to detect the influencers of the system. Morone [16] has proposed a numerical method to find them by using the adjacency matrix of the graph. Here we want to study the system without necessarily knowing its topology, using information theory. To verify that the theory of information is a good way to evaluate the influence of an agent, we first make the agent play the role of zealots (an agent that does not change its vote, as defined by Mobilia [14], or inflexible as defined by Galam [10]). In the literature, there are many articles that deal with the effects of zealots on the voter model dynamic: for instance, in [15], Mobilia studied the role of zealots on the result of a vote and Masuda [13] showed the link between the role of zealots and their degree.

The most influential nodes can be determined by controlling successively each variables and measuring the impact on the average opinion of the entire group. We will show that the same ranking of influence can also be obtained by monitoring the time-delayed multi-infomation.

The determination of the most influential variables has a clear connection with the well developed theory of control, in which observability and controllability of a system are defined and explored. In section 4.7 we make the link between the standard concepts of control theory and our present approach. A important element of our discussion is related to the effect of noise on the possibility to control a system. The voter model shows that in presence of noise the influential nodes cannot force the opinion of the far enough agents, despite the existence of a connecting path. This result shows the limit of some previous approaches about the controllability of systems on a complex network [11].

The paper is organized as follows: section 2 introduces our voter model, then section 3 demonstrates the link between influence and time-delayed multi-information. Section 4 solves the 1D voter model analytically, in the mean-field regime and gives a formal link between influence and delayed multi-information. The link between the control length and the capacity of a communication channel is also given. Section 4.7 proposes a formulation of the 1D voter model in the usual framework of control theory. Loss of controllability is related to the noise intensity and the cost of controllability is expressed with a Gramian.

2. Voter Model

Simple models that abstracts the process of opinion formation have been proposed by many researchers [6, 9]. The version we consider here is an agent-based model defined on a graph of arbitrary topology, whether directed or not.

A binary agent occupies each node of the network. The dynamics is specified by assuming that each agent i looks at every other agent in its neighborhood, and counts the percentage ρ_i of those which are in the state $+1$ (in case an agent is linked to itself, it obviously belongs to its own neighborhood). A function f is specified such that $0 \leq f(\rho_i) \leq 1$ gives the probability for agent i to be in state $+1$ at the next iteration. For instance, if f would be chosen as $f(\rho) = \rho$, an agent for which all neighbors are in state $+1$ will turn into state $+1$ with certainty. The update is performed synchronously over all n agents.

Formally, the dynamics of the voter model can be express as

$$s_i(t+1) = \begin{cases} 1 & \text{with probability } f(\rho_i(t)) \\ 0 & \text{with probability } 1 - f(\rho_i(t)) \end{cases} \tag{2.1}$$

where $s_i(t) \in \{0, 1\}$ is the state of agent i at iteration t, and

$$\rho_i(t) = \frac{1}{|N_i|} \sum_{j \in N_i} s_j(t). \tag{2.2}$$

The set N_i is the set of agents j that are neighbors of agent i, as specified by the network topology.

The global density of all n agents with opinion 1 is obviously obtained as

$$\rho(t) = \frac{1}{n} \sum_{i=1}^{n} s_i(t) \tag{2.3}$$

According to the law of total probability, the probability p_i that agent i votes $+1$ is

$$\begin{aligned} p_i(t+1) &= (1-\epsilon) p_{V_i}(t)) + \epsilon(1 - p_{V_i}(t)) \\ &= (1-2\epsilon) p_{V_i}(t) + \epsilon \end{aligned}$$

where ϵ is the probability to take a decision different from that of the neighborhood and $p_{V_i}(t)$ is the probability that the majority of neighbours of agent i vote 1 at time t. In what follows, we will use a particular function f, (see Fig. 1)

$$f(\rho) = (1-\epsilon)\rho + \epsilon(1-\rho) = (1-2\epsilon)\rho + \epsilon \tag{2.4}$$

From now on, the quantity ϵ is called the noise. We limit the noise in the range $0 \leq \epsilon \leq 1/2$. The upper value $\epsilon = 1/2$ corresponds to a blind vote, with probability $1/2$ for each outcome.

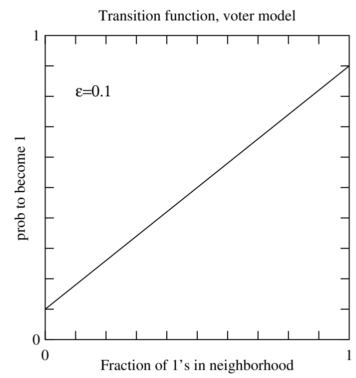

Figure 1: The probablility $f(\rho)$ used in this study. The noise ϵ is visible as the values $f(0)$ and $1 - f(1)$.

To illustrate the behavior of this model, we consider a random scale-free [2] graph G, as simple instance of a social network, as proposed by Newman in the paper Random graph models of social networks [17]. In a scale-free network, a small number of nodes have many connections. These nodes are the leaders of the social network. And most nodes have very few connections. The majority of voters are represented by these nodes. The scale free graph structure is based on communities built around a leader (see the paper by Andrew Wu [26]). We use the algorithm of Bï¿œla Bollobï¿œs [3] to generate a random graph scale free.

Figure 2 shows the corresponding density of agents with opinion 1, as a function of time. We can see that there is a lot of fluctuations due to the fact that states "all 0's" or "all 1's" are no longer absorbing states when $\epsilon \neq 0$.

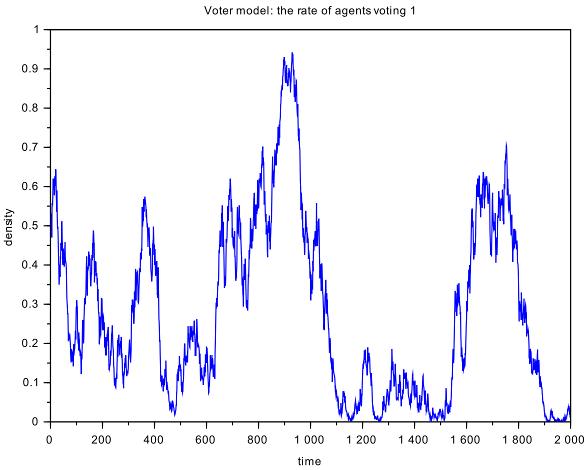

Figure 2: Graph of the time evolution of the density of opinion 1 with noise $\epsilon = 0.001$ and $n = 200$ agents connected through a scale-free network.

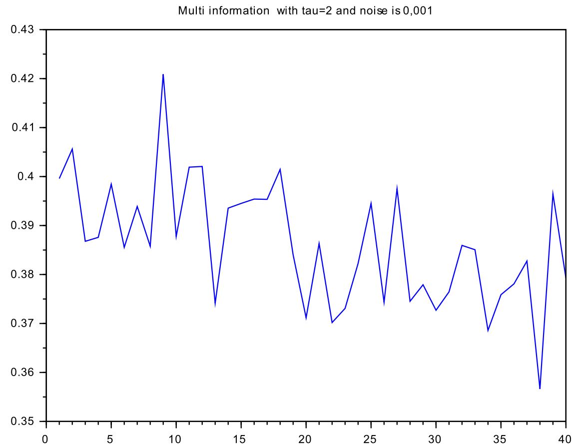

Figure 3: τ-delayed multi-information $w_i(\tau)$ $(\tau = 3)$ as a function of i, for graph G with $n = 50$ agents and noise level $\epsilon = 0.001$.

3. Characterisation of the influence of an agent

In this study we would like to characterize how the opinion of one agent influences that of its neighbors and that of the entire system. We will first propose an approach based on information theory, and then measure the influence directly by forcing (or controlling) the opinion of one agent. We will show that both characterizations are strongly correlated. The information theoretic quantities that will be considered are the time-delayed mutual information and the time-delayed multi-information. The purpose of considering a time delay is to capture the causal effect of one element on another.

3.1. Delayed mutual- and multi-information

Let us consider a set of random variables $X_i(t)$ associated with each agent i, taking their values in a set A. For instance, $X_i(t) = s_i(t)$ would be the opinion of agent i at iteration t.

To measure the influence between agents i and j, we define the τ-delayed mutual information $w_{i,j}$ as

$$
\begin{aligned}
w_{i,j}(t,\tau) &= I(X_i(t), X_j(t+\tau)) && (3.1)\\
&= \sum_{(x,y)\in A^2} p_{xy} \log\left(\frac{p_{xy}}{p_x p_y}\right) && (3.2)
\end{aligned}
$$

with

$$p_{xy} = \mathbb{P}(X_i(t) = x, X_j(t+\tau) = y)$$

$$p_x = \mathbb{P}(X_i(t) = x) \text{ and } p_y = \mathbb{P}(X_j(t+\tau) = y)$$

We also define the τ-delayed multi-information w_i to measure the influence of one agent i on all the others

$$w_i(t,\tau) = I(X_i(t), Y_i(t+\tau)) \quad (3.3)$$

$$Y_i(t+\tau) = \sum_{k\neq i} X_k(t+\tau) \quad (3.4)$$

These information metrics can be computed by the method of sampling. We consider $N = 10^5$ instances of the system in order to perform an ensemble average. According to the central limit theorem, we know that, with this number of instances, we obtain a precision of 3×10^{-2} with a risk of 5% for the approximate values of the probabilities that we compute (see Appendix B for details).

3.2. Non-intrusive characterisation of the nodes influences: delayed multi-information

The τ-delayed multi-information can be used as a measure of the influence of opinion of each node i on the vote of the other agents. For instance, Fig. 3 shows $w_i(\tau = 2)$ in a steady state, where the origin of time is arbitrary. We observe that some agents i exhibit a more pronounced peak of multi-information towards the rest of the system, suggesting that the opinion of these agents may affect the global opinion of all agents. Note that this results is obtained only by probing the systems, without modifying any of its components. For this reason, we describe this approach as "non-intrusive". The algorithms used throughout this paper to numerically evaluate the delayed mutual- and multi-informations in the voter model example are described in Appendix C.

As we can see in Fig. 4, the influence of each agent decreases strongly with noise. We deduce that with a high noise it will be difficult to control the system. Brede [5] obtained similar results about the effect of the noise on the control.

3.3. Intrusive characterization: forcing

In this section, we consider another way to measure the influence of an agent on the system. We call this approach

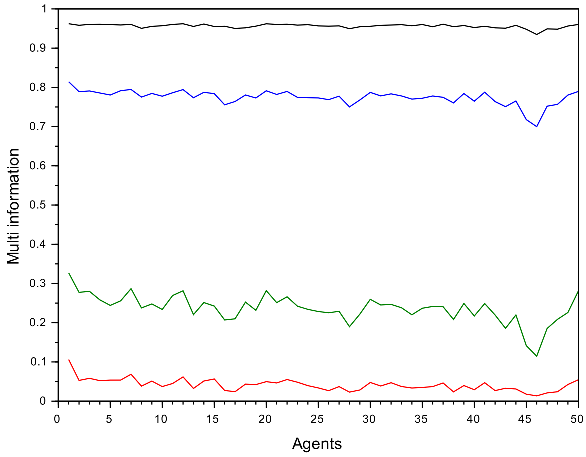

Figure 4: τ-delayed multi-information $w_i(\tau)$ ($\tau = 3$) as a function of i, for graph G with $n = 50$ agents and with different noise level $\epsilon = 0.0001$ (black curve) , $\epsilon = 0.001$ (blue curve), $\epsilon = 0.01$ (green curve) and $\epsilon = 0.05$ (red curve).

"intrusive" as it implies a perturbation, and no longer just an observation.

To measure the influence of agent i, its opinion is forced to a chosen value, for instance the value 1. As a result the density (2.3) of opinions 1 in the system

$$\rho(t) = \frac{1}{n}\sum_{j=1}^{n} s_j(t) \tag{3.5}$$

can be averaged over a large number N of independent realizations, to give a quantity $\langle\rho(t)\rangle_i$, where the subscript i indicate which agent has been forced to 1. If t is large enough, $\langle\rho\rangle_i$ no longer depends on t.

The influence can be measured in a steady state, or from the initial state where all agents are initialized uniformly to 0 or 1 with probability 1/2, respectively.

The color representations of the graphs (Figures 5 and 6) show that the multi-information give some information about the controllability and the observability of the system. In the case the multi-information is calculated from the initial state, these figures emphasize the link between the multi-information and the influence of an agent. We can then identify the agents that allow the best control of the system when their vote is forced.

With the multi-information, we get the same results as Masuda [13]. Choosing hubs as zealots is a good strategy to control the system and the effect of the control is strongly correlated with the hub's degree.

The measurement obtained in the steady state for the delayed multi-information is different from that observed in the transient regime. Low-impact agents can get a high multi-information by being a proxi of an influential neighbor. In this case, the multi-information rather evaluates the observability than the controlability.

4. The 1D Voter model

The previous section gave an illustration of the link between influence defined by intrusive forcing and the influence measured by observing the time-delayed multi-information. In this section, we propose an analytical meanfield solution of the voter model, in a one-dimensional topology. This solution will formally specify the proposed links. In particular we will introduce a characteristic control length.

4.1. Presentation

We consider the case of n voters organized along a line so that voter i looks at voter $i-1$ and itself to take its decision.

Chain network have been already studied. Brede [4] has demontred that the influence of indirect control on an agent decreases with the distance from the controlled agent when there is noise.

Agent $i = 0$ has no left neighbor and will have a controlled dynamics. For instance its opinion will be always 1. The other agents are initialized randomly in $\{0, 1\}$.

Since agent 1 is looking at agent 0, its next state will likely to be 1. And so on for agent $2, 3, \ldots, n$. Intuitively, we could expect that the entire system will become 1, due to the control imposed by agent 1. But noise is changing this conclusion.

If $p_i(t)$ is the probability that agent i is 1 at time t, we can write the equation

$$p_i(t+1) = p_i(t)W_{1\to1}(t) + (1 - p_i(t))W_{0\to1}(t) \tag{4.1}$$

where $W_{a\to b}$ is the probability that the state evolves from a to b. In a meanfield approximation, we can write,

$$\begin{aligned} W_{0\to1} &= p_{i-1}(t)f(1/2) + (1 - p_{i-1}(t))f(0) \\ W_{1\to1} &= p_{i-1}(t)f(1) + (1 - p_{i-1}(t))f(1/2) \end{aligned} \tag{4.2}$$

Before attempting to solve the above system analytically, we can observe its behavior numerically. We can see in Fig. 7 that if the noise is absent ($\epsilon = 0$), the entire system is indeed controlled by the left-most agent whose state is always 1. But, as sson as the noise is increased ($\epsilon = 0.01$) the control is not effective anymore. There is a critical noise $\epsilon = \epsilon_c(n)$ below which a system of size n can be controlled by the first node, and above which the influence of the driving node is diluted by the noise. Figure 8 shows the density of agents with opinion 1, as a function of time, for different intensities of noise, ϵ. We observe in this figure the effect of the system size. For smaller systems, the effect of controlling agent $i = 1$ is more effective than for larger n.

4.2. Probability distribution of opinions

We can determine the probability distribution in the case of the linear voter model. We have

$$\begin{aligned} p_i(t+1) &= p_i(t)W_{1\to1}(t) + (1 - p_i(t))W_{0\to1}(t) \\ &= p_i(t)(W_{1\to1}(t) - W_{0\to1}(t)) + W_{0\to1}(t) \end{aligned}$$

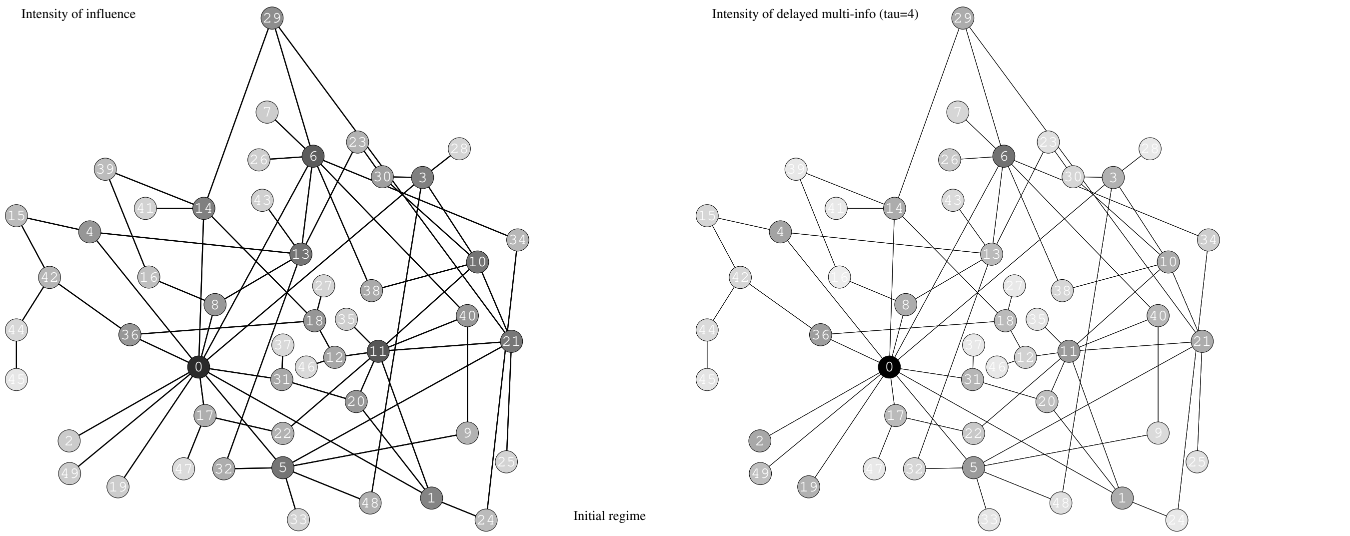

Figure 5: Scale free graph colored as a function of the values of the influence (left) and the τ-delay multi-information (right), for $\tau = 4$. In this case, the multi-information is computed from the initial state.

With

$$\begin{aligned} W_{1\to 1}(t) &= p_{i-1}(t)f(1) + (1 - p_{i-1}(t))f(1/2) \\ &= p_{i-1}(t)(1-\epsilon) + (1 - p_{i-1}(t))\frac{1}{2} \end{aligned}$$

and

$$\begin{aligned} W_{0\to 1}(t) &= p_{i-1}(t)f(1/2) + (1 - p_{i-1}(t))f(0) \\ &= p_{i-1}(t)\frac{1}{2} + (1 - p_{i-1}(t))\epsilon \end{aligned}$$

we obtain

$$p_i(t+1) = \left(\frac{1}{2} - \epsilon\right) p_i(t) + \left(\frac{1}{2} - \epsilon\right) p_{i-1}(t) + \epsilon \quad (4.3)$$

As $p_0(t) = 1$, we obtain

$$p_1(t+1) = \left(\frac{1}{2} - \epsilon\right) p_1(t) + \frac{1}{2} \quad (4.4)$$

Let $P(t)$ be the vector of probability defined by

$$P(t) = \begin{pmatrix} p_1(t) \\ p_2(t) \\ \vdots \\ p_n(t) \end{pmatrix}$$

With this notation, the system can be expressed in a matrix form

$$P(t+1) = AP(t) + B \quad (4.5)$$

with

$$A = (\frac{1}{2} - \epsilon) \begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ 1 & 1 & 0 & \ddots & \vdots \\ 0 & 1 & 1 & \ddots & 0 \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 & 1 \end{pmatrix}$$

and

$$B = \begin{pmatrix} 1/2 \\ \epsilon \\ \vdots \\ \epsilon \end{pmatrix}$$

This equation can be solved recursively and gives

$$P(t) = A^t P(0) + \big(\sum_{j=0}^{t-1} A^j\big) B \quad (4.6)$$

The explicit forms for power matrices A^j are given in Appendix A.

4.3. Stationnary system

Let us write

$$\Pi = \begin{pmatrix} \pi_1 \\ \pi_2 \\ \vdots \\ \pi_n \end{pmatrix}$$

the stationary distribution. From relation (4.5), it obeys

$$A\Pi + B = \Pi \Leftrightarrow \begin{cases} (\frac{1}{2} + \epsilon)\pi_1 = \frac{1}{2} \\ \forall i \in [\![2; n]\!],\ \pi_i = \dfrac{1-2\epsilon}{1+2\epsilon}\pi_{i-1} + \dfrac{2\epsilon}{1+2\epsilon} \end{cases} \quad (4.7)$$

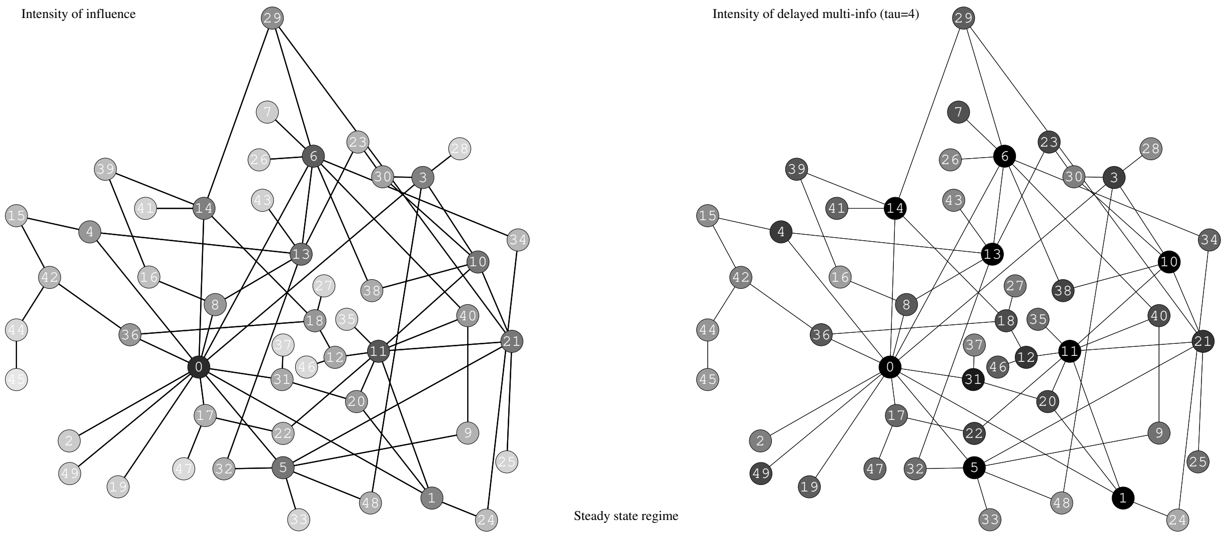

Figure 6: Scale free graph colored as a function of the values of the influence (left) and the τ-delay multi-information (right), for $\tau = 4$. In this case, the multi-information is computed when the system is in a steady state regime.

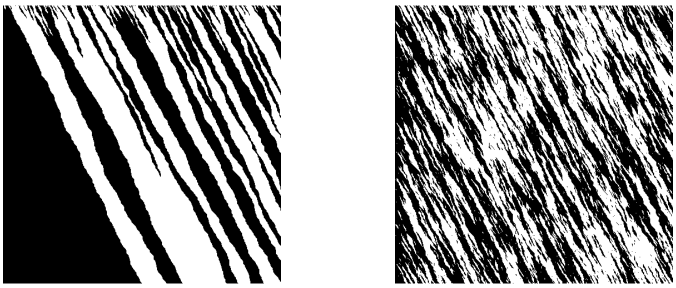

Figure 7: Space-time diagram of the evolution of the states of a $n = 500$ agents of a voter model organized in a line. Line t of the figure depicts the configuration of the n voter at iteration t. We can see the first 500 iterations. Left: $\epsilon = 0$. Right: $\epsilon = 0.01$.

It is an arithmetico-geometric sequence which can be solved for all agents i as

$$\begin{cases} \pi_1 = \dfrac{1}{1+2\epsilon} \\ \pi_i = \dfrac{1}{2} + \Big(\dfrac{1-2\epsilon}{1+2\epsilon}\Big)^{i-1}(\pi_1 - \dfrac{1}{2}) \end{cases} \tag{4.8}$$

with

$$\pi_1 - \frac{1}{2} = \frac{1}{2}\left(\frac{1-2\epsilon}{1+2\epsilon}\right)$$

Further, we can write eq.(4.8) as

$$\begin{aligned} \pi_i &= \frac{1}{2}\left[1 + \left(\frac{1-2\epsilon}{1+2\epsilon}\right)^i\right] \\ &= \frac{1}{2} + \frac{1}{2}\exp\left[-i\ln\left(\frac{1+2\epsilon}{1-2\epsilon}\right)\right] \\ &= \frac{1}{2} + \frac{1}{2}\exp\left[-\frac{i}{\ell_c}\right] \end{aligned} \tag{4.9}$$

where ℓ_c is defined as

$$\ell_c = \frac{1}{\ln\left(\frac{1+2\epsilon}{1-2\epsilon}\right)} \tag{4.10}$$

and referred to as the *control length* as it gives a value for i above which the exponential falls quickly to zero. It is a characteristic distance from the controlled agent where its inluence starts to fade.

We see that, when ϵ approaches $1/2$, the length of control ℓ_c converges to 0, which corresponds to a total loss of the controlability of the system. Figure 9 shows that ℓ_c decreases very quickly to 0 when ϵ increases to $1/2$.

4.4. Average vote of the system

In the case of a stationnary system, we can calculate the average density of agents with vote 1.

$$S = \frac{1}{n}\sum_{i=1}^{n}\pi_i$$

with n is the number of free agents.

According to (4.9), we have

$$S = \frac{1}{n}\left[\frac{n}{2} + \frac{1}{2}\sum_{i=1}^{n}\left(\frac{1-2\epsilon}{1+2\epsilon}\right)^i\right] \tag{4.11}$$

When $\epsilon \neq 0$, we have

$$\sum_{i=1}^{n}\left(\frac{1-2\epsilon}{1+2\epsilon}\right)^i = \left(\frac{1-2\epsilon}{1+2\epsilon}\right)\sum_{i=0}^{n-1}\left(\frac{1-2\epsilon}{1+2\epsilon}\right)^i \tag{4.12}$$

$$= \left(\frac{1-2\epsilon}{1+2\epsilon}\right)\frac{1-\left(\frac{1-2\epsilon}{1+2\epsilon}\right)^n}{1-\left(\frac{1-2\epsilon}{1+2\epsilon}\right)} \tag{4.13}$$

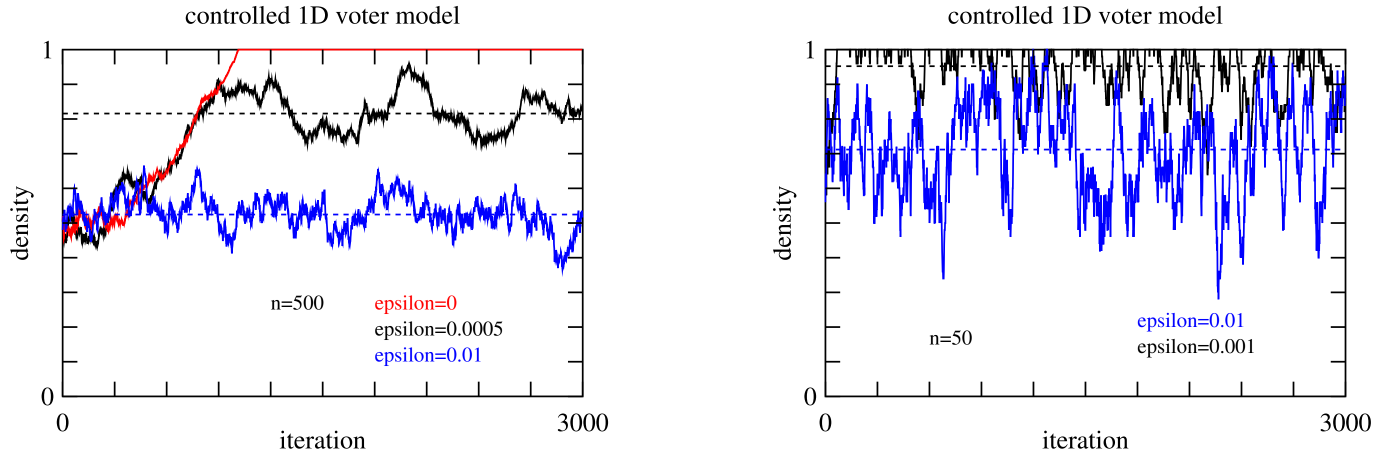

Figure 8: Density of agents with value 1 as a function of time, for different noise levels, and two different system sizes, $n = 500$ and $n = 50$. The dashed lines are the predictions of the meanfield analytical approach, see eq. (4.14).

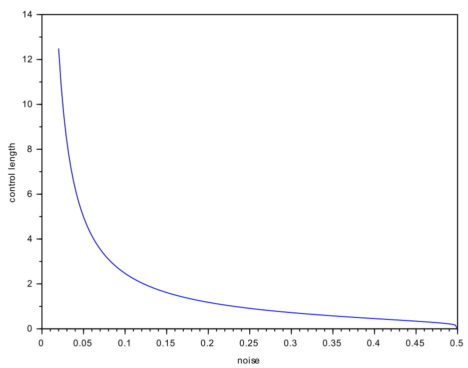

Figure 9: Control length ℓ_c as a function of the noise ϵ, according to eq. (4.10).

and we obtain

$$\begin{aligned} S &= \frac{1}{2} + \frac{1}{2n}\left[\left(\frac{1-2\epsilon}{1+2\epsilon}\right)\frac{1-\left(\frac{1-2\epsilon}{1+2\epsilon}\right)^n}{1-\left(\frac{1-2\epsilon}{1+2\epsilon}\right)}\right] \\ &= \frac{1}{2} + \frac{1}{2n}\left(\frac{1-2\epsilon}{4\epsilon}\right)\left[1-\left(\frac{1-2\epsilon}{1+2\epsilon}\right)^n\right] \end{aligned} \tag{4.14}$$

In Figure 8, we see that the simulations are in agreement with this theoretical result. Indeed, the density of agents who votes 1 oscillates around the mean value S represented by the dashed lines in the figure.

4.5. Delayed mutual information

In this section we will compute the influence of an agent based on the τ-delayed mutual information, $w_{i,j}(\tau)$, between agents i and j, as defined in eq. (3.2). These values are obtained by a sampling of the simulation of the 1D voter model, with $n = 50$ agents. Measurements are performed when the system has reached a stationnary state, that is after t iterations such that all the probabilities $A^t P(0)$ are smaller than a certain threshold. In our case, we take the threshold at 10^{-4}.

In Fig. 10, we notice that the mutual information $w_{i,j}(\tau)$ is zero if $j < i$, has a plateau for $j < i + \tau$, shows a peak for $j = i+\tau$, and decreases for $j > i+\tau$. This observation reflects the fact that agent i can only influence agents on its right as the voting decision of an agent is based on the state of its left neighbor. The plateau shows the influence of the past $j - i$ iterations. The influence of i over j is maximum for $j = i+\tau$ as it takes τ iterations for the vote of i to travel from i to j. For $j > i+\tau$ the influence is due to the steady state regime.

The results that we have are consistent with those that Brede [4] obtained recently when he studied the dependence between the vote of the first agent and that of the others. He computed analytically the dependance of the average stationary vote of a node on a distance i to the node controlled, and he obtained that this dependence decreases exponentially with i, as obtained from the mutual information.

In Fig. 11 we consider the behavior of $w_{i,j}(j - i)$. It suggests the following relationship

$$\forall j > i, w_{i,j}(j-i) = \alpha_i \exp\left[-\lambda_i(j-i)\right] \tag{4.15}$$

where α_i and λ_i depend on the noise level, ϵ.

The coefficients of correlation between $\ln(w_{i,j}(j - i))$ and j, for different values of the noise are found to be between -1 and -0.99, thus confirming the relation proposed in eq. (4.15). The value of α_i and λ_i can be determined with a least squares method.

Consequently, the value of the delayed mutual information $w_{i,j}(j - i)$ decreases quickly as j departs from i. This reflects the difficulty to control agent j from agent i.

This interpretation is confirmed by Fig. 12 which shows the relation between the values of $\lambda \equiv \lambda_1$ and the control

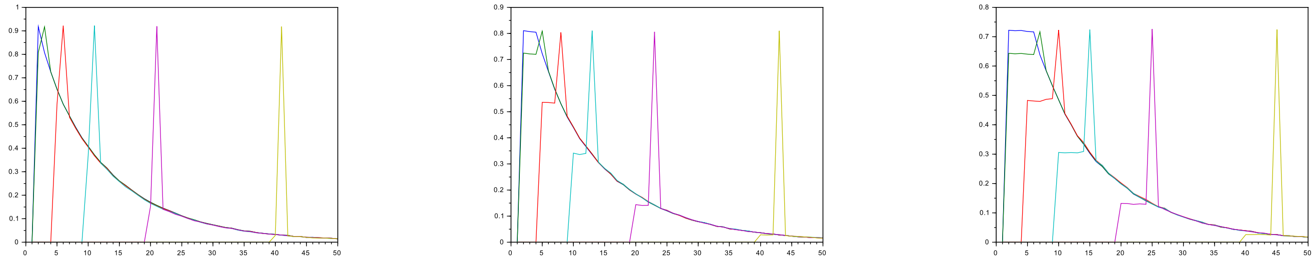

Figure 10: Delayed mutual information, $w_{i,j}(\tau)$, as a function of agent j, for different values of i. The different curves correspond to $i = 1,\ 2,\ 5,\ 10,\ 20$ et 40, from left to right, respectively. The vote of agent $i = 0$ is forced to 1 and the noise is $\epsilon = 0.01$. The delay is $\tau = 1$ (left panel), $\tau = 3$ (middle panel) and $\tau = 5$ (right panel).

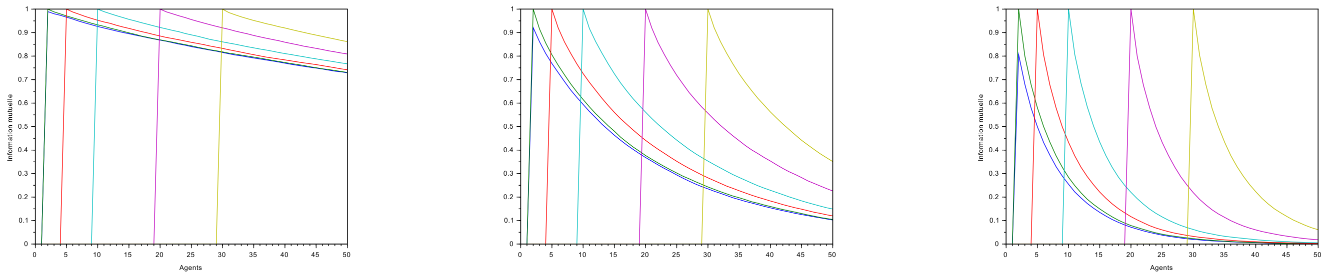

Figure 11: Delayed mutual information $w_{i,j}(j-i)$ as a function of j, for agents $i = 1,\ 2\ 5,\ 10,\ 20$ and 30 (curves from left to right, respectively). The vote of agent $i = 0$ is forced to 1 and the noise is $\epsilon = 0.001$ (left panel), $\epsilon = 0.01$ (middle panel) and $\epsilon = 0.05$ (right panel).

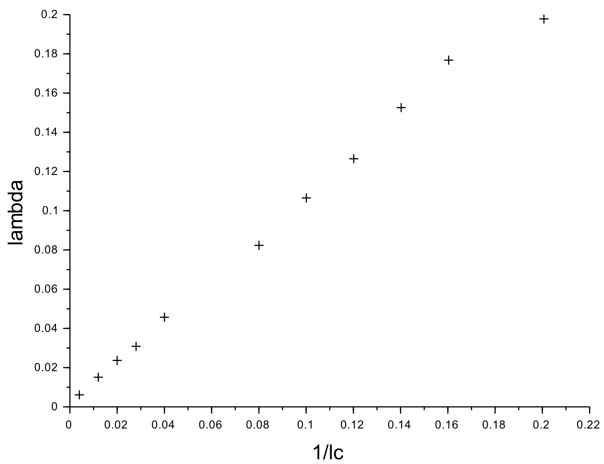

Figure 12: λ as a function of $1/\ \ell_c$

length ℓ_c defined in eq. (4.10). Each point in this figure corresponds to a different value of the noise. The relation can be fitted by

$$\lambda(\epsilon) = a\frac{1}{\ell_c(\epsilon)} + b \tag{4.16}$$

with $a = 0.973$ and $b = -0.003$, independent of the value of ϵ. The coefficient of correlation is 0.997, in agreement with the proposed linear link between λ and $1/\ell_c$.

4.6. Control of the density of vote 1

The previous section suggests that the influence of an agent decreases exponentially with the distance to others, with a characteristic length which decreases as the noise increases. This result follows both from studying an intrusive action on the system, or by simply observing it. In this section we exploit this result to find a strategy to control the full system by acting on more than one agent. In practice we consider the situation where agents $i = kd$ are forced to vote 1, where $k \in \{0, 1, 2, \ldots\}$ and d is given by the control length $\lfloor \ell_c \rfloor$ or $\lfloor \frac{1}{\lambda_1} \rfloor$.

Fig. 13 shows the simulation results for d chosen as $\lfloor \frac{1}{\lambda_1} \rfloor$. The density of agents voting 1 increases significantly. The quantity $n\lambda_1$ is the number of controlled agents and is a good indicator to evaluate the cost to control the system.

4.7. Noise and information capacity

In the previous section, by evaluating the mutual information, we found that the cost of control increased greatly when the noise increases. This result can be related to the notion of *capacity*, as defined in the standard theory of information. In the linear voter model, agent $i+1$ can be considered as a channel of communication where the input message is the vote of agent i at time t and the output message is the vote of agent $i+2$ at time $t+2$.

The information channel capacity C_2 is defined as (see

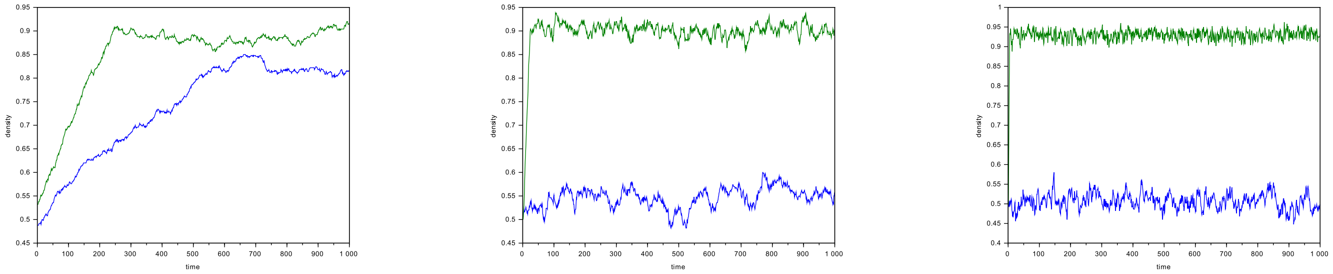

Figure 13: Evolution of the density of vote 1 during the time when only the vote of the agent 1 is forced to 1 (blues curves) and when every vote of the agents $j \propto \lfloor \frac{1}{\lambda} \rfloor$ is forced to 1 (greens curves). left $\epsilon = 0.001$. middle :$\epsilon = 0.01$, right: $\epsilon = 0.05$.

[8])

$$C_2 = \sup_{\mathbb{P}_{X_i}} I(X_i(t), X_{i+2}(t+2)) = \sup_{\beta \in [0;1]} I(X_i(t), X_{i+2}(t+2)) \tag{4.17}$$

With $\begin{pmatrix} \mathbb{P}(X_i(t) = 1) \\ \mathbb{P}(X_i(t) = 0) \end{pmatrix} = \begin{pmatrix} \beta \\ 1-\beta \end{pmatrix}$, we obtain

$$\begin{cases} \mathbb{P}(X_{i+1}(t+1) = 1) = (1-\epsilon)\mathbb{P}(X_i(t) = 1) + \epsilon \mathbb{P}(X_i(t) = 0) \\ \mathbb{P}(X_{i+1}(t+1) = 0) = \epsilon \mathbb{P}(X_i(t) = 1) + (1-\epsilon)\mathbb{P}(X_i(t) = 0) \end{cases}$$

which we write as

$$\begin{pmatrix} \mathbb{P}(X_{i+1}(t+1) = 1) \\ \mathbb{P}(X_{i+1}(t+1) = 0) \end{pmatrix} = A \begin{pmatrix} \beta \\ 1-\beta \end{pmatrix} \text{ oï¿œ } A = \begin{pmatrix} 1-\epsilon & \epsilon \\ \epsilon & 1-\epsilon \end{pmatrix}$$

Therefore,

$$\begin{pmatrix} \mathbb{P}(X_{i+2}(t+1) = 1) \\ \mathbb{P}(X_{i+2}(t+1) = 0) \end{pmatrix} = A^2 \begin{pmatrix} \beta \\ 1-\beta \end{pmatrix}$$

where

$$A^2 = \begin{pmatrix} (1-\epsilon)^2 + \epsilon^2 & 2\epsilon(1-\epsilon) \\ 2\epsilon(1-\epsilon) & (1-\epsilon)^2 + \epsilon^2 \end{pmatrix} \tag{4.18}$$

$$= \begin{pmatrix} 1 - 2\epsilon(1-\epsilon) & 2\epsilon(1-\epsilon) \\ 2\epsilon(1-\epsilon) & 1 - 2\epsilon(1-\epsilon) \end{pmatrix} \tag{4.19}$$

$$\tag{4.20}$$

This corresponds to a binary symmetric channel, with a probability of error

$$p_e = 2\epsilon(1-\epsilon)$$

. We know that the capacity of a binary symmetric channel with a probabiliy of error p_e is

$$C = 1 - H_2(p_e) \tag{4.21}$$

with $H_2(p_e) = -p_e \log_2(p_e) - (1-p_e)\log_2(1-p_e)$
Therefore

$$C_2 = 1 - H_2(2\epsilon(1-\epsilon)) \tag{4.22}$$

Now, we consider all agents from $i+1$ to $i+m-1$ as a channel of communication between agents i et $i+m$. We denote C_m the capacity of this channel (it depends only on m, the length of the channel). Following the same derivation as before, we obtain

$$\begin{pmatrix} \mathbb{P}(X_{i+m}(t+m) = 1) \\ \mathbb{P}(X_{i+m}(t+m) = 0) \end{pmatrix} = A^m \begin{pmatrix} \beta \\ 1-\beta \end{pmatrix}$$

Since A is a symmetric matrix it can be cast in a diagonal form with an orthonormal basis. The eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = 1 - 2\epsilon$. Thus, A^m can be expressed as

$$A^m = P^T \begin{pmatrix} 1 & 0 \\ 0 & (1-2\epsilon)^m \end{pmatrix} P$$

with $P = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ Thus

$$A^m = \frac{1}{2} \begin{pmatrix} 1 + (1-2\epsilon)^m & 1 - (1-2\epsilon)^m \\ 1 - (1-2\epsilon)^m & 1 + (1-2\epsilon)^m \end{pmatrix}$$

and we obtain a symmetric binary channel of length m with a probabily of error

$$\epsilon_m = \frac{1}{2}(1 - (1-2\epsilon)^m)$$

Therefore, the capacity of this channel is

$$C_m = 1 + \epsilon_m \log_2(\epsilon_m) + (1-\epsilon_m)\log_2(1-\epsilon_m)$$

We know that the capacity is an upper bound of the mutual information for each value of ϵ. In Fig. 14, the capacity C_m is shown as a function of its length m, for different values of the noise. The fact that the capacity C_m decreases with m and with the noise, gives another confirmation of the increasing difficulty to control agent $i+m$ by forcing the vote of agent i.

5. Observability analysis in the case of a voter model with linear topology

The linear voter model analysis given above may be interpreted in terms of reachability or observability using

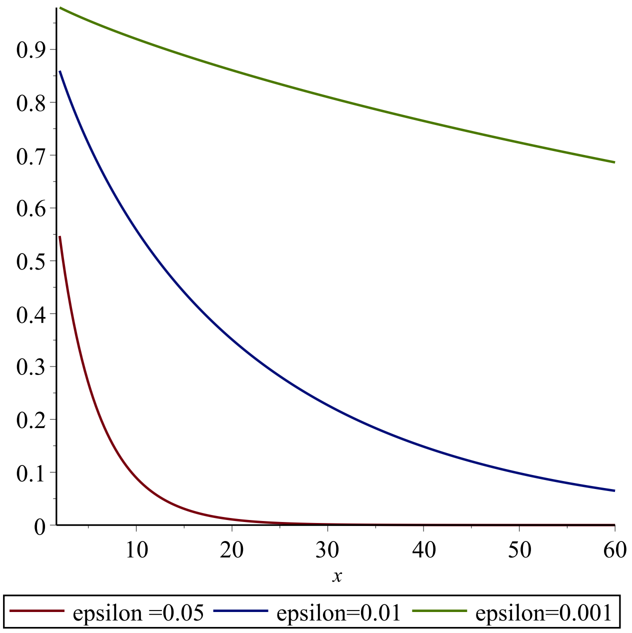

Figure 14: Capacity C_m of the channel between agent i and $i+m$, as a function of m, for different noise levels ϵ.

classical tools from system (control) theory (see [1], chapter 4 for an introduction to the control notions used hereafter). Roughly speaking, reachability denotes the posssibility to drive the dynamics of the whole network to a desired value for the state of all nodes by choosing the values of some of them (the inputs) at every time instant. Observability is the dual problem (see e.g. [1]) and denotes the ability to recover the initial state value of all agents from the observation of some of nodes (the outputs). Many results exist related to the reachability analysis of dynamical network systems. Structural analysis aims at analysing the reachability based on the local or global interaction topology within the network [21, 11, 12]. The drawbacks of such analyses is that they do not provide any quantitative information about the amount of energy which may be observed in the ouputs or which is required from the inputs. Such a quantitative information is requested when trying to understand the effects of the distance or noise on the reachability and observability of network systems. Therefore, numerous metrics have been defined related to these energies [19], including the reachability and observability Gramians which are precisely measures for the control or observation energies and wich are analyzed hereafter in this paper. The computation of the reachability Gramians in the case of *target control* problems - where inputs are designed to manipulate a group of target nodes rather than the whole network - is the subject of many recent papers [25, 24]. Qualitative properties and bounds for their eigenvalues have been established [22]. These results apply for general interconnexion topologies and many network dynamical processes with a diffusive structure, hence also for most voter models. However, in the case of very simple topologies and dynamics - such as in the case of the linear voter model investigated in this section - it is possible to compute analytically the observability matrix and Gramian and deduce anlytical results on the decay of the output signal energy with the control length and with the noise. It is proved hereafter that one recovers then exactly the behaviour exhibited by the delayed mutual information.

Let us consider as previously a linear topology with $n+1$ voting agents. In section 4, we mostly considered the case where agent was forced to vote 1. Here we consider a more general case. For $l, m \in \{0, \ldots, n\}$, forcing the vote of agent l may be considered as a control action, while observing the vote of agent m may be considered as an output measurement. Since we are interested in the deviation from 1/2 of the probability to vote 1 (thus measuring the influence of a forcing action, for instance), we define these deviations as state space variables

$$\tilde{p}_i(t) := p_i(t) - \frac{1}{2} \in \left[-\frac{1}{2}; \frac{1}{2}\right] \tag{5.1}$$

for all $t \geq 0$ and $i \in \{0, \ldots, n\}$. We will consider in the sequel, with no loss of generality, a forcing of agent 0 vote and an observation of agent n vote, since the influence in the considered linear voter is unidirectional (from left to right). Therefore, the input variable, $\tilde{u}$, and output variable, $\tilde{y}$, will be defined as

$$\tilde{u}(t) := p_0(t) - \frac{1}{2} \ ; \ \tilde{y}(t) := p_n(t) - \frac{1}{2} \tag{5.2}$$

Using these state space, input and output variables, the dynamical voter model (4.5) transforms into the state space system

$$\begin{aligned} \tilde{\mathbf{p}}(t+1) &= A\tilde{\mathbf{p}}(t) + \mathbf{b}\tilde{u}(t) \\ \tilde{y}(t) &= \mathbf{c}^T \tilde{\mathbf{p}}(t) \end{aligned} \tag{5.3}$$

with the state vector $\tilde{\mathbf{p}}(t) := [\tilde{p}_1(t), \ldots, \tilde{p}_n(t)]^T \in \mathbb{R}^n$ and the internal dynamics matrix (generator) A defined as

$$A := \left(\frac{1}{2} - \epsilon\right) \begin{pmatrix} 1 & 0 & \ldots & \ldots & 0 \\ 1 & 1 & 0 & \ldots & 0 \\ 0 & 1 & 1 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & 1 & 1 \end{pmatrix} \tag{5.4}$$

The control column matrix $\mathbf{b}$ and observation row matrix $\mathbf{c}^T$ are defined respectively as

$$\mathbf{b} := \left(\frac{1}{2} - \epsilon\right) \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{and} \quad \mathbf{c}^T := \begin{pmatrix} 0 & \dots & 0 & 1 \end{pmatrix} \tag{5.5}$$

For any time $t \geq 0$, any initial probability distribution $\tilde{\mathbf{p}}(0) := \tilde{\mathbf{p}}_0 \in \left[-\frac{1}{2}; \frac{1}{2}\right]^n$ and any control (forcing) signal values $\tilde{u}(t) \in \left[-\frac{1}{2}; \frac{1}{2}\right]$, the solution $\phi(\tilde{u}; \tilde{\mathbf{p}}_0; t)$ of the state space equations (5.3) may be written

$$\phi(\tilde{u}; \tilde{\mathbf{p}}_0; t) = A^t \tilde{\mathbf{p}}_0 + \sum_{j=0}^{t-1} A^{(t-1)-j} \mathbf{b} \tilde{u}(j) \tag{5.6}$$

Note that the matrix A has a unique eigenvalue $\lambda(A) = (\frac{1}{2} - \epsilon)$, with multiplicity n and such that $|\lambda(A)| < 1$ (since the noise ϵ satisfies $0 \leq \epsilon < \frac{1}{2}$). Therefore the trajectory (5.6) is bounded when $t \to \infty$ and the dynamical system (5.3) is said *stable*.

A state $\tilde{\mathbf{p}} \in \left[-\frac{1}{2}; \frac{1}{2}\right]^n$ is said *unobservable* if the corresponding output can not be distinguished from the output associated with the zero state, that is if

$$y(t) = \mathbf{c}^T \phi(0; \tilde{\mathbf{p}}; t) = 0 \tag{5.7}$$

for all $t \geq 0$ (in the observability analysis, only the free response dynamics is analyzed and $\tilde{u}$ is set to zero). The whole state space system (5.3) is said *observable* if the set of unobservable states reduces to $\{0\}$. With the solution (5.6) and Cayley theorem, it is easy to prove that this is the case if and only if the *observability matrix*

$$\mathcal{O}_n = \left[\mathbf{c} \; ; \; A^T \mathbf{c} \; ; \; \dots \; ; \; (A^{n-1})^T \mathbf{c}\right]^T \tag{5.8}$$

is full rank or when the *infinite observability Gramian*

$$W^o := \lim_{n \to \infty} \mathcal{O}_n^T \mathcal{O}_n = \sum_{k=0}^{\infty} (A^k)^T \mathbf{c}\mathbf{c}^T A^k \tag{5.9}$$

is strictly positive definite.

The infinite observability Gramian gives additionnal quantitative information about how much the system or a particular state is observable. Indeed, the largest observation energy (i.e. the maximum energy for the output signal) is reached when $t \to \infty$ and equals

$$\|\tilde{y}\|_2^2 := \lim_{t \to \infty} \sum_{k=0}^{t} |\tilde{y}(k)|^2 = \tilde{\mathbf{p}}^T W^o \tilde{\mathbf{p}} \tag{5.10}$$

for any given state space trajectory $\phi(0; \tilde{\mathbf{p}}; t)$. Therefore, with the appropriate change of state space coordinates, the components of the initial condition (or subspaces) may be re-ordered, from the less to the most observable ones. If some of the infinite horizon observability Gramian eigenvalue are zero, then the corresponding vector spaces are unobservable. If some of these eigenvalues are small, then initial conditions variations in the corresponding subspaces will cause low energy variations in the output signal.

In the linear voter model example, rather than measuring the influence of forcing permanently the agent 0 to vote 1 (with a constant input signal $\tilde{u}(t) = \frac{1}{2}$, $\forall t \geq 0$) on the vote of agent n, we could instead analyze to effect of considering the initial probability distribution

$$\tilde{\mathbf{p}} := [1, 0, \dots, 0]^T \in \mathbb{R}^{n+1} \tag{5.11}$$

on agent $n+1$, by measuring the corresponding observation energy. We will consider a long range time horizon $k > n$ for which the influence of the initial state of agent 1 has reached agent $n+1$ in the line. The last row of matrix A^k may be written (see Appendix A):

$$A^k_{(n+1,\cdot)} := \begin{cases} (\frac{1}{2} - \epsilon)^k \left[\binom{k}{n} \; \binom{k}{n-1} \; \dots \; \binom{k}{2} \; k \; 1\right] & \text{when } k \geq n \\ (\frac{1}{2} - \epsilon)^k \left[0 \; \dots \; 0 \; \binom{k}{k} \; \dots \; \binom{k}{2} \; k \; 1\right] & \text{when } k < n \end{cases} \tag{5.12}$$

According to definition (5.9), since we are measuring the vote of agent $n+1$, we get for the components of the infinite observability Gramian

$$W^o_{i,j} := \sum_{k=0}^{\infty} A^k_{(n+1,i)} A^k_{(n+1,j)} \tag{5.13}$$

for all $i, j \in \{1, \dots, n\}$. Measuring the influence of the initial vote of agent 1, we start with the initial probability distribution (5.11) and get, for the agent $n+1$, the observation energy

$$\|\tilde{y}\|_2^2 := \sum_{k=0}^{\infty} \left(A^k_{(n+1,1)}\right)^2 = \sum_{k=0}^{\infty} (\frac{1}{2} - \epsilon)^{2k} \left(A^k_{(n+1,1)}\right)^2 \tag{5.14}$$

With equation (5.12), one gets

$$\begin{aligned} \|\tilde{y}\|_2^2 &= \sum_{k=n}^{\infty} (\frac{1}{2} - \epsilon)^{2k} \binom{k}{n}^2 \qquad (5.15) \\ &= \frac{1}{(n!)^2} (\frac{1}{2} - \epsilon)^{2n} \sum_{p=0}^{\infty} (\frac{1}{2} - \epsilon)^{2p} \left(\frac{(p+n)!}{p!}\right)^2 \end{aligned}$$

Using the lower bound

$$(p+1)^n < \frac{(p+n)!}{p!} \tag{5.16}$$

one gets

$$\frac{1}{((n-1)!)^2} (\frac{1}{2} - \epsilon)^{2n} \frac{4(3 - 2\epsilon)}{(1 + 2\epsilon)^3} < \|\tilde{y}\|_2^2 \tag{5.17}$$

On the other hand, since

$$\begin{aligned}
&\sum_{p=0}^{\infty}(\frac{1}{2}-\epsilon)^{2p}\left(\frac{(p+n)!}{p!}\right)^2 \\
&\leq \left(\sum_{p=0}^{\infty}(\frac{1}{2}-\epsilon)^{p}\frac{(p+n)!}{p!}\right)^2 \\
= &\left(\sum_{k\geq n}(\frac{1}{2}-\epsilon)^{k-n}k(k-1)\dots(k-(n-1))\right)^2 \\
= &\left(\frac{n!2^{n+1}}{(1+2\epsilon)^{n+1}}\right)^2 \qquad (5.18)
\end{aligned}$$

we get the following upper bound for the observation energy

$$\|\tilde{y}\|_2^2 < \frac{4(1-2\epsilon)^{2n}}{(1+2\epsilon)^{2n+2}} \qquad (5.19)$$

It is worthwile to notice how this upper bound behaves with the number of agents along the line and with the noise ϵ. For instance, the upper bound (5.19) decreases with the number of agents and the corresponding observation energy is divided by two when k supplementary agents are added in the line, with

$$k \geq \frac{1}{2}\frac{1}{\log_2\left(\frac{1+2\epsilon}{1-2\epsilon}\right)} = \frac{\ell_c}{2} \qquad (5.20)$$

When the noise increases, the observation energy upper bound decreases

$$\|\tilde{y}\|_2^2 = \mathcal{O}\left((1-2\epsilon)^{2n}\right) \to 0 \text{ as } \epsilon \to \left(\frac{1}{2}\right)^- \qquad (5.21)$$

The lower bound (5.17) decreases similarly, with the same order, when the noise decreases. However it decreases much faster with the number of agents in the voter line since this lower bound for the observation energy is divided by $\left(\frac{1-2\epsilon}{2n}\right)^2$ when only one agent is added to the n previous ones.

Note that we performed the observability analysis on the linear voter model. We could as well develop the dual reachability analysis for the same example. In this analysis, the initial condition is assumed to be zero and one analyzes the forced solution of the state space model (5.3). More specifically, one could be interested in its reachability property. A state $\tilde{\mathbf{p}}$ is said *reachable* when there is a an input signal $\tilde{u}(t)$ such that

$$\lim_{t\to\infty}\phi\left(\tilde{u},\mathbf{0},t\right) = \tilde{\mathbf{p}}$$

It may be proved (see, e.g. [1]) that, among those input signals which can reach the state $\tilde{\mathbf{p}}$ from a zero intial condition, the one with minimum energy may be written as

$$\|\tilde{u}^*\|_2^2 = \tilde{\mathbf{p}}^T(W^c)^{-1}\tilde{\mathbf{p}}$$

where the *infinite reachability Gramian* W^c is defined as

$$W^c := \sum_{k=0}^{\infty}A^k\mathbf{b}\mathbf{b}^T(A^T)^k \qquad (5.22)$$

Therefore, a reachability Gramian analysis may be used to compute the forcing of agent 1 with minimal energy requested to reach a state $\tilde{\mathbf{p}}$ where all agents in the line vote 1, that is such that $\tilde{p}_i = 1$, for all $i \in 1,\dots,n+1$. However, in this case, it would be necessary to compute the sum of all the elements in $(W^c)^{-1}$, which is a much more involved computation than the one performed for the observability analysis. Besides, the duality between reachability and observability for linear systems [1] and the particular topology of the linear voter model lead us to the conjecture that the reachability analysis would not bring any new result fundamentally different from the ones obtained through the observability analysis.

6. Conclusions

In this paper we show that time delayed mutual- and multi-informations are promising tools to better grasp the behavior of a dynamical system on complex networks. In particular it can be used to determine the most influential degrees of freedom and the most observable variables. This knowledge can be obtained without perturbing the system, by just probing its behavior.

We claim that influential nodes are those that are the most interesting to control or monitor to (i) force a system to reach a given target, or (ii) to have a proxy giving an information on the state of the entire system.

We illustrated our approach in a simple stochastic dynamical model on a graph, a so-called voter model, where agents iteratively adapt their opinion to that of the majority of their neighbors, with however a given noise level. We first discussed the case of a general scale-free topology, where only numerical results can be obtained. Then we consider a 1D topology for which analytical results can be obtained. There, we rigorously showed that the influence of an agent on the entire system can be equivalently measured by actually forcing its behavior, or, in a non-intrusive way, by measuring the time delayed multi-information of this agent with respect to the rest of the system. In particular, we proposed the concept of a control length, which indicates a characteristic distance above which the influence of a controlled agent fades exponentially.

The link with classical control theory has been proposed and the control length has been related to the reachability Gramian, thus indicating that the cost of control becomes intractable at large distance. The importance of the noise is clearly shown as being a central element in the possibility of observing or controlling a system, as opposed to previous literature that claimed that a causality path was sufficient to achieve control [11].

As an additional link of our approach to existing concepts, we showed that controllability can also be considered in the framework of the capacity of communication channel, as defined in information theory by Shannon. We showed that this capacity drops as agent are separated by a distance above the control length.

In a forthcoming paper we will apply our approach to other complex systems, in particular those for which the underlying dynamics and topology of interaction are not known. We already obtained (not shown here) that the time delayed multi-information can be used to infer the topology of the graph of Fig. 5. Further, we want to use the concept of observability as a way to detect early warning signal of possible tipping points in a complex dynamical system. In simple words, we want to analyze the idea that the most influential degree of freedom is the best variable to observe to know in advance if a given system is likely to move to another regime. These nodes being the most influential ones, we can argue that their evolution will dictate the evolution of the other variables.

Acknowledgment

We thank Gregor Chliamovitch and Alex Dupuis for initiating several of the ideas developed in this paper, during the FP7 project SOPHOCLES (2012-2015), and for the reading of the manuscript.

References

[1] Athanasios C Antoulas. *Approximation of large-scale dynamical systems*, volume 6. Siam, 2005.

[2] Albert-László Barabási, Réka Albert, and Hawoong Jeong. Scale-free characteristics of random networks: the topology of the world-wide web. *Physica A: statistical mechanics and its applications*, 281(1-4):69–77, 2000.

[3] Béla Bollobás and Oliver M Riordan. Mathematical results on scale-free random graphs. *Handbook of graphs and networks: from the genome to the internet*, pages 1–34, 2003.

[4] Markus Brede, Valerio Restocchi, and Sebastian Stein. Resisting influence: how the strength of predispositions to resist control can change strategies for optimal opinion control in the voter model. *Frontiers in Robotics and AI*, 5:34, 2018.

[5] Markus Brede, Valerio Restocchi, and Sebastian Stein. Transmission errors and influence maximization in the voter model. *Journal of Statistical Mechanics: Theory and Experiment*, 2019(3):033401, 2019.

[6] Claudio Castellano, Santo Fortunato, and Vittorio Loreto. Statistical physics of social dynamics. *Rev. Mod. Phys.*, 81:591–646, May 2009.

[7] Claudio Castellano, Miguel A Muñoz, and Romualdo Pastor-Satorras. Nonlinear q-voter model. *Physical Review E*, 80(4):041129, 2009.

[8] Thomas M Cover and Joy A Thomas. *Elements of information theory*. John Wiley & Sons, 2012.

[9] S. Galam, B. Chopard, A. Masselot, and M. Droz. Competing species dynamics: Qualitative advantage versus geography. *Eur. Phys. J. B*, 4:529–531, 1998.

[10] Serge Galam and Frans Jacobs. The role of inflexible minorities in the breaking of democratic opinion dynamics. *Physica A: Statistical Mechanics and its Applications*, 381:366–376, 2007.

[11] Y.-Y. Liu, J.-J. Slotine, and A.-L. Barabasi. Controllability of complex networks. *Nature*, 473:167, 2011.

[12] Yang-Yu Liu, Jean-Jacques Slotine, and Albert-László Barabási. Observability of complex systems. *Proceedings of the National Academy of Sciences*, 110(7):2460–2465, 2013.

[13] Naoki Masuda. Opinion control in complex networks. *New Journal of Physics*, 17(3):033031, 2015.

[14] Mauro Mobilia. Does a single zealot affect an infinite group of voters? *Physical review letters*, 91(2):028701, 2003.

[15] Mauro Mobilia, Anna Petersen, and Sidney Redner. On the role of zealotry in the voter model. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(08):P08029, 2007.

[16] Flaviano Morone and Hernán A Makse. Influence maximization in complex networks through optimal percolation. *Nature*, 524(7563):65, 2015.

[17] Mark EJ Newman, Duncan J Watts, and Steven H Strogatz. Random graph models of social networks. *Proceedings of the National Academy of Sciences*, 99(suppl 1):2566–2572, 2002.

[18] Piotr Nyczka and Katarzyna Sznajd-Weron. Anticonformity or independence? insights from statistical physics. *Journal of Statistical Physics*, 151(1-2):174–202, 2013.

[19] Fabio Pasqualetti, Sandro Zampieri, and Francesco Bullo. Controllability metrics, limitations and algorithms for complex networks. *IEEE Transactions on Control of Network Systems*, 1(1):40–52, 2014.

[20] J. Pearl. *Causality: Models, Reasoning and Inference*. Cambridge University Press, 2009.

[21] Amirreza Rahmani, Meng Ji, Mehran Mesbahi, and Magnus Egerstedt. Controllability of multi-agent systems from a graph-theoretic perspective. *SIAM Journal on Control and Optimization*, 48(1):162–186, 2009.

[22] Sandip Roy and Mengran Xue. Controllability-gramian submatrices for a network consensus model. *arXiv preprint arXiv:1903.09125*, 2019.

[23] Thomas Schreiber. Measuring information transfer. *Physical review letters*, 85(2):461, 2000.

[24] Henk J Van Waarde, M Kanat Camlibel, and Harry L Trentelman. A distance-based approach to strong target control of dynamical networks. *IEEE Transactions on Automatic Control*, 62(12):6266–6277, 2017.

[25] Amirkhosro Vosughi, Charles Johnson, Mengran Xue, Sandip Roy, and Sean Warnick. Target control and source estimation metrics for dynamical networks. *Automatica*, 100:412–416, 2019.

[26] Andrew Y Wu, Michael Garland, and Jiawei Han. Mining scale-free networks using geodesic clustering. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 719–724. ACM, 2004.

Appendix A. Explicit evaluation of A^p

In 4.6, we need to calculate A^p with

$$A = (\frac{1}{2} - \epsilon) \begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ 1 & 1 & 0 & \ddots & \vdots \\ 0 & 1 & 1 & \ddots & 0 \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 & 1 \end{pmatrix}$$

We have $A = I_n + C$ with

$$C = \begin{pmatrix} 0 & 0 & \dots & \dots & 0 \\ 1 & 0 & 0 & \ddots & \vdots \\ 0 & 1 & 0 & \ddots & 0 \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 & 0 \end{pmatrix}$$

C is a nilpotent matrix and $\forall k \in \mathbb{N},\ A^k = \sum_{p=0}^{k} \binom{k}{p} C^p$. Therefore, for $p \geqslant n-1$:

$$A^p = (\frac{1}{2} - \epsilon)^p \begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ \binom{p}{1} & \ddots & \ddots & & \vdots \\ \binom{p}{2} & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \binom{p}{n-1} & \dots & \binom{p}{2} & \binom{p}{1} & 1 \end{pmatrix}$$

and for $p < n-1$:

$$A^p = (\frac{1}{2} - \epsilon)^p \begin{pmatrix} 1 & 0 & \dots & \dots & \dots & \dots & \dots & 0 \\ \binom{p}{1} & 1 & \ddots & & & & & \vdots \\ \binom{p}{2} & \binom{p}{1} & \ddots & \ddots & & & & \vdots \\ \vdots & \binom{p}{2} & \ddots & \ddots & \ddots & & & \vdots \\ \binom{p}{p} & & \ddots & & \ddots & \ddots & & \vdots \\ 0 & \ddots & & \ddots & \ddots & \ddots & 0 & \vdots \\ \vdots & \ddots & \ddots & & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \binom{p}{p} & \dots & \binom{p}{2} & \binom{p}{1} & 1 \end{pmatrix}$$

Appendix B. Accuracy and confidence for the numerical evaluation of probability distributions

Let us consider an attribute of the members of a population which appears with probability p. For a sample of size n drawn in this population, let F_n be the random variable equal to the proportion of those elements having this attribute. According to the Moivre-Laplace theorem, the quantity $\frac{F_n - p}{\sqrt{p(1-p)/n}}$ converges in distribution to a Gaussian distribution

$$\begin{aligned} \mathbb{P}(\mid F_n - p \mid \leqslant \epsilon) = 1 - \alpha &\Leftrightarrow \mathbb{P}\Big(\frac{\mid F_n - p \mid}{\sqrt{\frac{p(1-p)}{n}}} \leqslant \frac{\epsilon}{\sqrt{\frac{p(1-p)}{n}}}\Big) = 1 - \alpha \\ &\Leftrightarrow \frac{\epsilon\sqrt{n}}{\sqrt{p(1-p)}} = t_{1-\alpha/2} \\ &\Leftrightarrow \epsilon = t_{1-\alpha/2}\sqrt{\frac{p(1-p)}{n}} \end{aligned}$$

where $t_{1-\alpha/2}$ is the real number defined by $\mathbb{P}(X \leqslant t_{1-\alpha/2}) = \alpha$ with $X \sim \mathcal{N}(0,1)$. As $p \in [0,1]$ and $p(1-p) \leqslant 0.25$, therefore $\epsilon \leqslant t_{1-\alpha/2}\frac{1}{2\sqrt{n}}$. For $N = 10^5$ and $\alpha = 0,05$, we have $t_{1-\alpha/2} = 1.96$, we obtain an approximation value of p with a precision of 0.03, with a risk of 5%.

Appendix C. Algorithms for the computations of mutual and multi-information

Appendix C.1. Mutual information

We consider a scale free graph G, with n agents. To compute the τ-delayed mutual information between 2 agents i and j, we generate N runs. For every run, we have a matrix S defined by: for $i \in [\![1,n]\!]$, and for $j \in [\![1,t+\tau]\!]$, such that $S[i][j]$ is the state of the agent i at the moment j.

We use 4 $n \times n$ matrix (N00, N01, N10 and N11), initialized to zeros. For every run, we compare the vote of the agent i at the time t and the vote of the agent j at the time $t+\tau$

for i from 1 to n
for j from 1 to n
if $S[i][t] = 0$ and $S[j][t+\tau] = 0$ then $N00[i][j] ++$ endif
if $S[i][t] = 0$ and $S[j][t+\tau] = 1$ then $N01[i][j] ++$ endif
if $S[i][t] = 1$ and $S[j][t+\tau] = 0$ then $N10[i][j] ++$ endif
if $S[i][t] = 1$ and $S[j][t+\tau] = 1$ then $N00[i][j] ++$ endif
end for
end for

We then compute $w_{i,j}(t,\tau)$, the τ-delayed mutual information at time t between agents i and j, according to definition (3.2). We obtain

$$\begin{aligned}\forall(i,j),\ w_{i,j}(t,\tau) \quad = \quad & \frac{N00[i][j]}{N}\log_2\Big(\frac{N00[i][j]\times N}{(N00[i][j]+N10[i][j])\times(N00[i][j]+N10[i][j])}\Big)\\ &+\frac{N01[i][j]}{N}\log_2\Big(\frac{N01[i][j]\times N}{(N00[i][j]+N01[i][j])\times(N01[i][j]+N11[i][j])}\Big)\\ &+\frac{N00[i][j]}{N}\log_2\Big(\frac{N10[i][j]\times N}{(N10[i][j]+N[i][j])\times(N00[i][j]+N10[i][j])}\Big)\\ &+\frac{N11[i][j]}{N}\log_2\Big(\frac{N11[i][j]\times N}{(N10[i][j]+N11[i][j])\times(N01[i][j]+N11[i][j])}\Big)\end{aligned} \tag{C.1}$$

Appendix C.2. Multi-information

To compute the delayed multi-information, as for the delayed mutual information, we execute N runs, and for every run, we compute the state matrix S.
We use two $n \times n$ matrix, $N0$ and $N1$ (n is the number of agents) defined by : $\forall(i,j),\ N0[i][j]$ is equal to the number of runs where the vote of the agent i is 0 and the number of agents who voted 1 is $j-1$ at time $t+\tau$.
$\forall(i,j),\ N1[i][j]$ is equal to the number of runs where the vote of the agent i is 1 and the number of agents (without the agent i) who voted 1 is $j-1$.
These matrices give us the probability distribution of the couple of random variable $(X_i(t), Y_i(t+\tau))$ that we have in the definition of the delayed multi-information, eq. (3.3).