Learning influence structure in sparse social networks

Chiara Ravazzi, *Member, IEEE*, Roberto Tempo, *Fellow, IEEE*, and Fabrizio Dabbene, *Senior Member, IEEE*

***Abstract*—Aim of this paper is to propose a novel methodology for estimating the social influence among agents interacting in a sparse social network described by the Friedkin and Johnsen's model. In this classical model, n agents discuss $m \ll n$ topics, are influenced by the others' opinions, but are not completely open-minded, being persistently driven by their prejudices. We reconstruct the social network topology and the strength of the interconnections starting from observations of the initial and final opinions' profile only. The intrinsic sparsity of the graph is exploited via an ℓ_0/ℓ_1 minimization. Different from the techniques previously proposed in the literature, no partial knowledge of the social graph is assumed, and there is no need of optimally placing external stubborn agents injecting prescribed inputs, thus changing the terminal behavior of the opinion dynamics. Under suitable assumptions, we derive theoretical conditions that guarantee that the problem is well posed and sufficient requirements on the number of topics under discussion that ensure perfect recovery. Extensive simulations on synthetic and real networks corroborate theoretical results.**

I. Introduction

The interest of the control community towards the study of opinion formation in social networks and belief systems has been constantly increasing in the recent years, stimulated by the numerous and challenging system-theoretic problems arising in this field. The notion of social network refers to a structure defined by agents (individual entities) and connections among them, describing relations, such as friendships, collaboration or trust. These relationships can exhibit different level of trust. As described in the recent survey [2], the key challenges in this framework are related to the issues of *modeling* (finding a meaningful description of the social interactions), *analysis* (studying the evolution of the social dynamics), and *control* (identifying the most influential players and introducing external agents to influence the final opinions).

Various dynamical models, based on several communication mechanisms, have been proposed in the literature to describe certain features in the opinion dynamics, such as the emergence of consensus as in French-DeGroot models [3], [4], [5], [6], or the persistent disagreement in social systems when stubborn agents are present [7], [8], [9]. Theoretical studies have focused on characterizing the convergence, ergodicity of the opinion dynamics with stubborn agents, actors who never change opinion, and with random interactions [9], [10], developing techniques to reach fastest convergence and controlling the steady-state of the opinions [11].

The advent of social media, such as Facebook and Twitter, has made a huge volume of data easily available, and political organizations and business firms have been interested in analyzing these massive data for decision making and action planning [12], [13]. Advanced tools have been developed to extract low-dimensional structures from large social networks that characterize the social behavior of individuals and groups [14], [15]. Using data describing the relationships, we are able to detect communities [16], identify social leaders, who influence the behavior of others in the network, using various notions of centrality measures [17], [18], [19], and on the other hand, to determine which people are most affected by other network participants. The final key problem addressed in this literature is related to the so-called minimum controlling set, which is the smallest set of opinion makers who can suitably adjust their network intercommunication to control other individuals in social networks [20]. Various strategies of optimal placement in social networks have been proposed in [11], [21].

However, all these tasks require a complete knowledge of the relationships in the network. Motivated by this consideration, in this paper we are interested in the estimation of the social influence network among n agents that interact in a social network. This important problem has received relatively less attention in the control community than the problems mentioned above. In [22] some methods to infer the network structure in French-DeGroot models with stubborn agents have been proposed.

In this paper, we adopt the classical Friedkin and Johnsen's model [23], in which the agents discuss $m \ll n$ independent topics, are influenced by the others' opinions, but are not completely open-minded, and are persistently driven by their initial opinions. More precisely, at each round of communication the agents update their beliefs by taking a convex combination of the opinions coming from the neighbors, weighted with respect to an influence matrix, and their prejudices. It should be noted that this model extends the French-DeGroot model with stubborn agents, which is included as a particular case. A key feature of the Friedkin and Johnsen's model is that it has been experimentally validated for small and medium size groups [7], [24].

Chiara Ravazzi and Fabrizio Dabbene are with the Institute of Electronics, Computer and Telecommunication Engineering, National Research Council of Italy (CNR-IEIIT), c/o Politecnico di Torino, Corso Duca Degli Abruzzi, 10129, Italy. E-mail: chiara.ravazzi@ieiit.cnr.it, fabrizio.dabbene@ieiit.cnr.it.

Roberto Tempo is a former member of the CNR-IEIIT. He suddenly passed away on January 14, 2017. Chiara Ravazzi and Fabrizio Dabbene confirm that he has made significant intellectual contributions to the research, has been involved in the first stages of the manuscript writing process, and agreed to be accountable for the work.

The authors' research has been partially supported by the International Bilateral Joint CNR Lab COOPS.

A preliminary version of the paper was accepted for publication in the Proceeding of IEEE Conference on Decision and Control 2017 (see [1]).

In general, given a group of individuals, whose opinions on m independent issues are supposed to evolve according to a specific model, we can distinguish two strategies to estimate the interactions in the network, that can be referred to as *finite-horizon* and *infinite-horizon* identification procedures. In the experiments of the first kind the opinions are observed after T rounds of conversation and the influence matrix is estimated as the matrix best fitting the dynamics for $0 \leq k < T$. The drawback of these methods is that they require the knowledge of the time-sequence for observations made, and to store a sufficiently long subsequence of opinions $x(k),\ x(k+1), .., \ x(k+M-1)$. This knowledge may be difficult to obtain in general and the collection may involve a large amount of data. The loss of data from one of the agents in general requires to restart the experiment. Moreover, the system is usually updated with an unknown interaction rate and the interactions time between agents are not observable in most scenarios, as in [25], [26]. In this paper we adopt the second method, where the history of agents' opinions is not required and the interactions are not limited to any prescribed number of rounds. Similar to the experiments from [22], the agents interact until their opinions stabilize and the identification problem considers only the initial and the final opinions.

Compared to the methods previously introduced in literature, our inference method is innovative in the following sense. First, it is applicable to large networks since it uses a number of topics that is strictly smaller than the size of the graph, exploiting the sparsity of the graph structure. The strategy does not assume partial knowledge of the support of the social graph and it does not consider an optimized placement of stubborn agents injecting inputs that affect the natural behavior of the opinion dynamics as in [22].

More precisely, we formulate the network sensing problem as a sparse recovery problem using the popular ℓ_0/ℓ_1 minimization technique. Moreover, we prove theoretical conditions guaranteeing the well-poseness of the problem (see Proposition 2) and, under the assumption that the initial opinions are distributed according to a Gaussian distribution, we derive sufficient conditions for perfect recovery using some tools coming from compressed sensing (CS) theory [27]. It is worth noticing that, in the classical framework of CS, the sensing matrix is generally chosen by the user and is independent of the signal to be recovered. As it will be shown, in the considered problem the sensing matrix depends on the unknowns and is dictated by the evolution of the opinions. Although classical properties usually exploited to guarantee perfect recovery, as Restricted Isometry Properties (RIP, [28]), can be violated with high probability in some regimes (see Proposition 4), we show that perfect recovery is still possible with high probability, using the less known Restricted Eigenvalue Condition (REC, [29]). More precisely, in Theorem 1 we prove that the number of observed topics sufficient for reconstruction scales only logarithmically in the size of the network, linearly on the maximum degree of the network, and depends on the maximum level of agents' trust. In Theorem 2, we show that this result can be extended to models where the interactions are random and pairwise, exploiting the ergodicity of the random process.

A preliminary version of some of the results appears in [1]. In this paper, we extend the results in [1] from both a theoretical and a practical point of view: the results are rigorously proved for general graphs and for any level of agents' stubbornness. Moreover, the analysis is extended to models where the dynamics are random and the interactions are pairwise [10]. The case where the level of agents' stubborness is not known is also discussed. Besides the extensive simulations on synthetic networks, theoretical results and our findings are corroborated through experiments on a real Facebook network.

This paper is organized as follows. Preliminary facts on Friedkin and Johnsen's model are introduced in Section II. In Section III the social influence identification problem is formulated and its relation to Compressed sensing (CS) theory is discussed. Then in Section IV the main contribution of the paper is presented and in Section V the inference problem is then extended to models with gossip-interactions. Numerical results and some examples are shown in Section VI. Finally, some concluding remarks complete the paper (see Section VII).

II. Preliminaries

Throughout this paper, we use the following notation. Natural, nonnegative integer, and real numbers are denoted by $\mathbb{N}$, $\mathbb{Z}_{\geq 0}$, and $\mathbb{R}$ respectively. The sequence of integers from 1 to $n \in \mathbb{N}$ is summarized by notation $[n]$. The symbol $|\cdot|$ denotes the cardinality of a set. The vector with all entries equal to 1 is denoted with $\mathbb{1}$. Given the vector $v \in \mathbb{R}^n$, we define $\|v\|_p = \left(\sum_{i=1}^n |v_i|^p\right)^{1/p}$ with $p \in \{1,2\}$ and we denote the number of nonzero elements in v with $\|v\|_0$. We say that $v \in \mathbb{R}^n$ is k-sparse if $\|v\|_0 \leq k$. Given $M \in \mathbb{R}^{n \times n}$ and a set $S \subseteq [n]$, we use the notation $M_{SS} \in \mathbb{R}^{|S| \times |S|}$ to identify the matrix obtained by taking the rows and columns of M indexed by S. Analogously, the vector $v_S \in \mathbb{R}^{|S|}$ is the vector obtained by $v \in \mathbb{R}^n$ by taking the elements in the set $S \in [n]$. Given $M \in \mathbb{R}^{n \times n}$, we denote the minimum and maximum eigenvalue of M with the symbols $\mu_{\min}(M)$ and $\mu_{\max}(M)$, respectively. A matrix W is row-stochastic when its entries are nonnegative and $W\mathbb{1} = \mathbb{1}$. A matrix W is said to be Schur stable if the absolute value of all its eigenvalues is strictly smaller than 1. We denote with $\mathrm{Diag}: \mathbb{R}^n \to \mathbb{R}^{n \times n} / \mathrm{diag}: \mathbb{R}^{n \times n} \to \mathbb{R}^n$ the diagonal operator that map a vector/square matrix to a square matrix/vector. With this notation $\mathrm{Diag}(\mathrm{diag}(W))$ is the square matrix with only diagonal elements in W. We write $\mathrm{off}(W)$ to denote the square matrix with only off-diagonal elements in W, i.e. $\mathrm{off}(W) = W - \mathrm{Diag}(\mathrm{diag}(W))$. Given the matrix W, the inequality $W \geq 0$ is intended element-wise.

A directed graph is a pair $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the set of nodes and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of edges. We say that $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is an undirected graph if $(u, v) \in \mathcal{E}$ implies that (v, u) is also an edge in $\mathcal{E}$. For each agent i, we denote its neighborhood with $\mathcal{N}_i := \{j \in \mathcal{V} : (i, j) \in \mathcal{E}\}$. A directed graph $\mathcal{G}$ is called strongly connected if there is a path from each vertex in the graph to every other vertex. A graph is said d-regular if $|\mathcal{N}_i| = d$ for all $i \in \mathcal{V}$. To any matrix $W \in \mathbb{R}^{\mathcal{V} \times \mathcal{V}}$

with non-negative entries, we can associate a directed graph $\mathcal{G}_W = (\mathcal{V}, \mathcal{E}_W)$ by putting $(i,j) \in \mathcal{E}_W$ if and only if $W_{ij} > 0$. The matrix W is said to be adapted to graph $\mathcal{G}$ if $W_{ij} = 0$ if and only if $(i,j) \notin \mathcal{E}$. A path from u to v in $\mathcal{G}$ is an ordered sequence of edges $((u, w_1), (w_1, w_2), \ldots, (w_{\ell-1}, v))$. If such a path exists, we say that v can be reached from u. We say that $v \in \mathcal{V}$ is a globally reachable node, it if can be reached from every other node.

A. Friedkin and Johnsen's opinion dynamics model

We introduce the Friedkin and Johnsen's (FJ) model formulated in the pioneering work [7]. We consider a set of n agents $\mathcal{V}$, whose interactions are described by a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, which we refer to as the *social network*. To avoid trivialities, we assume that $n = |\mathcal{V}| > 1$. We suppose that there are m different discussions in the social network and we assume that these issues are completely unrelated. We assume that each agent $i \in \mathcal{V}$ is endowed with a state $x_i^\ell(k)$ which represents the opinion of agent $i \in \mathcal{V}$ on ℓ-th issue at time $k \in \mathbb{Z}_{\geq 0}$. For instance if $x_i^\ell(k) \in [-1, 1]$, 1 may represent extreme positive, 0 neutral and -1 negative *belief or opinion*, respectively. We draw an edge $(i,j) \in \mathcal{E}$ if agent j directly influences the belief of agent i. Let $W \in \mathbb{R}^{\mathcal{V} \times \mathcal{V}}$ be a nonnegative matrix adapted to the graph $\mathcal{G}$ which defines the strength of the interactions and $\Lambda = \text{Diag}(\lambda)$ be a diagonal matrix with $\lambda \in \mathbb{R}^n$ describing how sensitive each agent is to the opinions of the others, based on interpersonal influences. We assume $\lambda_i \in [0,1]$ for all $i \in [n]$ and that W is nonnegative, row-stochastic, and sparse, in the sense that $\|W\|_0 \ll n^2$.

The dynamics of opinions $x^\ell(k)$ proposed in [7] evolves as follows

$$x^\ell(k+1) = \Lambda W x^\ell(k) + (I - \Lambda) u^\ell, \tag{1}$$

with $x^\ell(0) = u^\ell$, $u^\ell \in \mathbb{R}^{\mathcal{V}}$ and for any $\ell \in [m]$. The values $u_i^\ell = x_i^\ell(0)$ are interpreted as the agents prejudices.

The limit behavior of the opinions is described in the following result.

Proposition 1 (Opinion convergence). *The following facts hold:*

1) *if $\Lambda = I$ and the graph is aperiodic and has a globally reachable node, then the dynamics in (1) leads asymptotically to a consensus*

$$x^\ell(\infty) := \lim_{k \to +\infty} x^\ell(k) = \mathbb{1}\pi^\top u^\ell,$$

where π is the invariant probability of matrix W, (i.e. *$\pi^\top W = \pi^\top$*).

2) *if $\Lambda \neq I$ and for any node $v \in \mathcal{V}$ there exists a path from v to a node m such that $\lambda_m < 1$. Then, the opinions converge to*

$$x^\ell(\infty) := \lim_{k \to +\infty} x^\ell(k) = (I - \Lambda W)^{-1}(I - \Lambda)u^\ell.$$

It should be noticed that if $\mathcal{G}$ is strongly connected, then the assumptions in Proposition 1 are satisfied.

Remark 1. *Since W is stochastic and nonnegative, it can be easily seen that if the initial opinions are at consensus, $u^\ell = \beta\mathbb{1}$ then also $x^\ell(k) = \beta\mathbb{1}$ for all $k \in \mathbb{Z}_{\geq 0}$. Moreover, under the assumption of Proposition 1, also the total effects matrix $V := (I - \Lambda W)^{-1}(I - \Lambda)$ is stochastic: this means that the limit opinion of each agent is a convex combination of the preconceived opinions of the group $x_i^\ell(\infty) = \sum_j V_{ij} u_j^\ell$. As a special case, the asymptotic opinion profile is a consensus.*

III. Sparse influence identification problem

Given the prejudices $U = [u^1, \ldots, u^m]$ and final opinions $X(\infty) = [x^1(\infty), \ldots, x^m(\infty)]$, our goal is to find W from the data.

As already noted in Remark 1, if the initial opinions are at consensus, then also the final opinions are at consensus. In this case the problem is not well posed, since any stochastic matrix W is consistent with the data. Motivated by this consideration, from now on we assume that for all $\ell \in [m]$ there exists $i, j \in \mathcal{V}$ such that $u_i^\ell \neq u_j^\ell$.

From Proposition 1 we know that when there are no stubborn agents ($\Lambda = I$) and under suitable requirements on the graph associated to the matrix W, all agents' opinions converge to a common value, which is a convex combination of the initial opinions with weights given by the invariant probability of W. Then, it should be easily noticed that the problem is not well posed in this case, *i.e* there are infinitely many matrices that lead the dynamics in (1) to the same value of consensus. This fact is illustrated in the following example.

Example 1. *Let $\Lambda = I$, and let $W \in \mathbb{R}^{n \times n}$ be any doubly stochastic matrix such that $\mathcal{G}_W$ is strongly connected and aperiodic. Then, Perron Frobenius theorem [30] guarantees that*

$$X(\infty) = \mathbb{1}\mathbb{1}^\top X(0)/n = \mathbb{1}\mathbb{1}^\top U/n,$$

where the last inequality follows from the assumption $X(0) = U$.

On the other hand, if $\Lambda = 0$, then $X(\infty) = U$ and all doubly stochastic matrices are consistent to the data. If $\lambda_i = 0$, then agent i is totally stubborn and is not influenced by any other agent. Without loss of generality, in the rest of the paper we suppose that $\lambda_i \neq 0$ for all $i \in \mathcal{V}$, $\Lambda \neq I$, and for any node $v \in \mathcal{V}$ there exists a path from v to a node m such that $\lambda_m < 1$ (each agent is influenced by at least one stubborn agent). Then for any initial profile the opinion dynamics leads asymptotically to an equilibrium point that can be computed from the weights, the obstinacy levels, and the initial opinions.

Also in this case the following system of equations

$$\begin{cases} (I - \Lambda W)x^\ell(\infty) = (I - \Lambda)u^\ell, \\ W\mathbb{1} = \mathbb{1}, \\ W \geq 0, \Lambda \geq 0 \end{cases} \tag{2}$$

admits an implicit ambiguity, as it is shown in the following results.

Proposition 2. *Let (Λ, W) be a solution of (2) then, for any non-negative diagonal matrix D with $D_{ii} \in [0,1]$, the couple*

(Λ', W') *such that*

$$\Lambda' = I - D(I - \Lambda)$$
$$\text{off}(\Lambda' W') = D\text{off}(\Lambda W)$$
$$\text{diag}(\Lambda' W') = \mathbb{1} - D((I - \Lambda)\mathbb{1} + \text{off}(\Lambda W)\mathbb{1}).$$

is a solution of (2).

The proof is given in Appendix A. It is worth mentioning that if $\lambda_i \in \{0, 1\}$, Proposition 2 is equivalent to Lemma 1 in [22]. In absence of noise, the system in (2) is consistent then admits at least one solution. As already pointed out in [22], Proposition 2 shows that the identification problem is not well defined. In fact, the ambiguity described is due to the fact that the information about the rate of social interactions is missing. This ambiguity can not be removed without making any additional assumptions. In particular, we have the following result.

Corollary 1. *If Λ is known, $m \geq n$, and the system in* (2) *is full rank then the problem in* (2) *admits a unique solution.*

Moreover the problem in (2) can be solved, e.g using linear programming or any solver for convex optimization.

From now on we suppose that Λ is known and we consider the more interesting case when $m \ll n$. It should be remarked that, if also the matrix Λ is part of the learning, we must define an invariant quantity among the ambiguous solutions, by defining equivalence classes and resolve the ambiguity for instance by imposing constraints on $\text{diag}(W)$. This is further discussed at the end of Section IV.

A. Sparse identification problem

We start from the observation that a social network is typically sparse, in the sense that the interactions among the agents are few if compared to the network dimension. Given Λ, U, and $X(\infty)$, this leads us to estimate the social influence networks solving the following ℓ_0-minimization [31]:

$$\min_{W \in \mathbb{R}^{n \times n}} \|W\|_0, \qquad \text{s.t.} \begin{cases} AW^\top = B^\top, \\ W\mathbb{1} = \mathbb{1}, \\ W \geq 0. \end{cases} \tag{3}$$

with $A \doteq X(\infty)^\top$, $B \doteq \Lambda^{-1}[X(\infty) - (I - \Lambda)U]$.

It should be noticed that this problem is separable into n subproblems, since each row of $W = [w_1^\top, \ldots, w_n^\top]^\top$ can be learned independently from the others. More precisely,

$$\min_{w_j \in \mathbb{R}^n} \|w_j\|_0, \qquad \text{s.t.} \begin{cases} Aw_j = b_j, \\ \mathbb{1}^\top w_j = 1, \\ w_j \geq 0. \end{cases} \tag{4}$$

where b_j equal to j-th row of B for every $j \in [n]$.

The reachability of each node from a stubborn node is an assumption that the true network must satisfy to guarantee the stability of the affine dynamics in (1) and the existence of the final opinion profile. However, as it should be noticed in the optimization problem (4), this constraint is not imposed in the recovery problem.

B. Relation to compressed sensing

The optimization problems in (4) can be seen as a sparse recovery problem starting from compressed measurements [31], a problem also known as Compressed Sensing (CS) under positivity constraints. It should be noticed that in the considered regime ($m < n$), the linear system is underdetermined and admits infinitely many solutions. However, a sufficient condition for recovery can be derived exploiting the sparsity of the desired solution and using the notion of *spark* of a matrix [32].

Definition 1 (Spark of a matrix)**.** *The spark of a given matrix A, denoted with $\text{spark}(A)$, is the smallest number of columns from A that are linearly dependent.*

When dealing with sparse vectors, the spark provides a complete characterization of when sparse recovery is possible.

Proposition 3. *For any vector b, there exists at most one vector w such that $b = Aw$ if and only if $\text{spark}(A) > 2\|w\|_0$.*

The interested reader can refer to [32, Theorem 1.1] for the proof.

Computing the spark of a matrix involves checking the dependence of combinations of columns. Testing the condition in Proposition 3 is computationally expensive for practical purposes, as it requires a combinatorial search. Moreover each of the n problems of the form (4) is NP-hard. For this reason we consider the ℓ_1-based relaxation.

IV. Recovery via convex optimization

The convex relaxation of (4), i.e.

$$\min_{w_j \in \mathbb{R}^n} \|w_j\|_1, \qquad \text{s.t.} \begin{cases} Aw_j = b_j, \\ \mathbb{1}^\top w_j = 1, \\ w_j \geq 0, \end{cases} \tag{5}$$

has been extensively studied and the literature proposes a large amount of algorithms to solve it efficiently [32]. It is well known that under certain conditions on the matrix A, the number of measurements m, and the sparsity of w_j, both (4) and (5) have the same unique solution [31]. Much of the theory concerning explicit performance bounds for CS considers the Restricted Isometry Property (RIP) that characterizes matrices which are nearly orthonormal, at least when acting on sparse vectors [28].

Definition 2 (Restricted Isometry Property)**.** *Let $A \in \mathbb{R}^{m \times n}$. Suppose that there exists a constant $\delta_s \in (0, 1)$ such that*

$$(1 - \delta_s)\|z\|_2^2 \leq \|Az\|_2^2 \leq (1 + \delta_s)\|z\|_2^2.$$

for all $z \in \Sigma_s = \{z \in \mathbb{R}^n : \|z\|_0 \leq s\}$. Then, the matrix A is said to satisfy the s-restricted isometry property with restricted isometry constant δ_s.

Denote with A_S the matrix with columns indexed by $S \subseteq [n]$. It can be shown (see [28]) that, if a given matrix satisfies the RIP of order $2s$ with a constant $\delta_{2s} \in (0, 1/(\sqrt{2} + 1))$, then one can uniquely recover a s-sparse vector from (5). It is straightforward to see that

$$(1 - \delta_s) \leq \mu_{\min}(A_S^\top A_S) \leq \mu_{\max}(A_S^\top A_S) \leq (1 + \delta_s).$$

and, consequently, it must hold

$$1 \approx \frac{\mu_{\max}(A_S^\top A_S)}{\mu_{\min}(A_S^\top A_S)} \leq 2.$$

for all $S \subseteq [n]$ with $|S| \leq s$.

It is well known [28] that sub-Gaussian random matrices with i.i.d. entries satisfy the RIP of order $2s$ with constant δ_{2s} with probability close to 1 if

$$m \geq \frac{cs}{\delta_{2s}^2} \log\left(\frac{n}{\delta_{2s} s}\right)$$

where $c > 0$ is a positive constant. It is worth noticing that, in the classical framework of CS the sensing matrix is generally chosen by the user and is independent of the signal to be recovered. Here, on the contrary, the sensing matrix depends on the unknowns and is dictated by the evolution of the opinions. In fact, we notice that

$$A = X(\infty)^\top = U^\top (I - \Lambda)(I - \Lambda W)^{-\top}.$$

If the initial opinions are i.i.d Gaussian random variables, i.e. $u^\ell \sim \mathcal{N}(0, m^{-1}I)$, then also A is a random variable. Intuitively, if Λ has small diagonal entries, $\lambda_{\max} := \max_{i \in \mathcal{V}} \lambda_i \approx 0$, and W is very sparse, one could in principle consider $A \approx U^\top$ that satisfies the RIP with high probability. However for not vanishing values of $\lambda_{\max} < 1$, A is a random variable whose entries are coupled across rows in a peculiar way. This makes highly nontrivial to apply standard proof techniques from the compressive sensing literature.

From now on we suppose the following assumption holds.

Assumption 1. *The initial opinions u^ℓ on topic ℓ are i.i.d according to $\mathcal{N}(0, I)$, for all $\ell \in [m]$.*

It is worth mentioning that the hypothesis on the Gaussian-like distribution of the initial condition is a common assumption in opinion dynamics literature [33]. This assumption can be explained also by the following consideration: the initial opinions are pre-averaged opinions, or several criteria which are seen as independent random variables. Thus, the distribution of initial opinions has a Gaussian-like shape due to the central limit theorem [34].

Moreover, the assumption on the zero expected value and identity covariance matrix is not restrictive. In fact we can always rescale and shift the initial opinion in order to fit the data to Assumption 1. If u^ℓ have nonzero expected value, then we can consider $u^\ell - \overline{u}^\ell \mathbb{1}$ and $z^\ell(\infty) = V(u^\ell - \overline{u}^\ell \mathbb{1})$ where V is the total effects matrix. Since the total effects matrix is stochastic (see Remark 1) then $z^\ell(\infty) = x^\ell(\infty) - \overline{u}^\ell \mathbb{1}$.

In next proposition we show that the RIP condition can be violated with probability close to one also for very simple graphs. Formally, we have the following result, which is proved in Appendix B.

Proposition 4. *Let $u^\ell \sim \mathcal{N}(0, I)$ for all $\ell \in [m]$. Then, for an arbitrary set $S \subseteq [n]$ of size $|S| = s$, defined $\widehat{\Sigma}_{SS} = A_S^\top A_S / m$, with probability greater than $1 - 2\mathrm{e}^{-m/32}$ we have*

$$\frac{\mu_{\max}(\widehat{\Sigma}_{SS})}{\mu_{\min}(\widehat{\Sigma}_{SS})} \geq \frac{1}{3} \frac{\mu_{\max}(\Sigma_{SS})}{\mu_{\min}(\Sigma_{SS})}$$

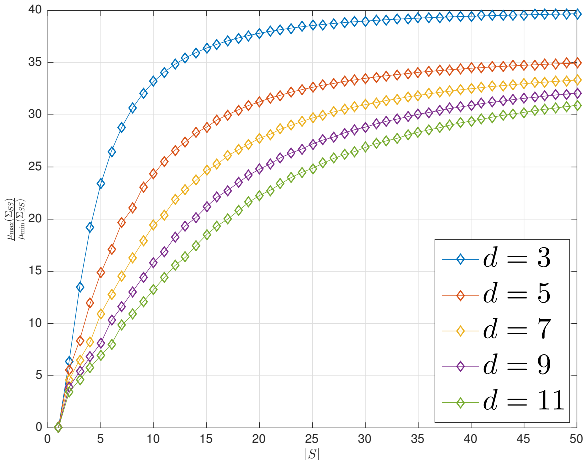

Fig. 1. Lower bounds on condition number of Cayley graphs: different curves correspond to graphs with different degree d. The sensitivity parameter is set to $\lambda = 0.8$ and the size of the nertwork is $n = 150$.

where $\Sigma = (I - \Lambda W)^{-1}(I - \Lambda)^2 (I - \Lambda W)^{-\top}$.

In Figure 1 we show the condition number of Σ_{SS} with $\Lambda = \lambda I$, $\lambda = 0.8$, and W corresponding to the weighted adjacency matrix of Cayley graphs for several degrees. One can notice that the condition number of Σ_{SS} is above 3 for all values of $|S|$. As a consequence, this result points out that a necessary condition to hold the RIP is not satisfied for high values of λ. For this reason, we need to resort to more powerful tools. It is shown [31] that a necessary and sufficient condition in order to guarantee the recovery via (5) is the nullspace property. More formally,

Definition 3 (Null Space Property (NSP))**.** *The matrix $A \in \mathbb{R}^{m \times n}$ satisfies the NSP of order s if, given*

$$\mathcal{C}(S) = \{\eta \in \mathbb{R}^n : \|\eta_{S^c}\|_1 \leq \|\eta_S\|_1\},$$
$$\mathcal{C}(S) \cap \mathsf{Ker}(A) = \{0\}.$$

for all index set S with $|S| \leq s$.

Proposition 5 (Theorem 1 in [31])**.** *Let $A \in \mathbb{R}^{m \times n}$. The optimization problem $\{\min \|z\|_1 : Az = b\}$ uniquely recovers all s-sparse vectors $z^\star$ from measurements $b = Az^\star$ if and only if A satisfies the nullspace property with order $2s$.*

Definition 4 (Restricted Eigenvalue Condition (REC))**.** *We say that a matrix A satisfies the REC of order s if there exists a $\delta_s > 0$ such that*

$$\frac{1}{m}\|Az\|_2^2 \geq \delta_s^2 \|z\|_2^2$$

for all $z \in \mathcal{C}(S)$, uniformly for all index sets $S \subseteq [n]$ with $|S| \leq s$.

It is straightforward to see that REC is equivalent to NSP.

For random matrices A with i.i.d. entries drawn from particular distributions or for unitary matrices, it can be shown that, if enough measurements are available, then REC condition is satisfied with the prescribed δ_s with probability close to 1 [35].

With these definitions at hand, we are ready to state the first of the main theoretical results of this paper, whose proof is reported in Appendix C.

Theorem 1. *Let $u^\ell \sim \mathcal{N}(0, I)$ for all $\ell \in [m]$, W be row-stochastic. If the number of considered topics satisfies the following condition*

$$m \geq 4c\frac{(1+\lambda_{\max})^2(1-\lambda_{\min})^2}{(1-\lambda_{\max})^4} d_{\max}\log n \tag{6}$$

then the solution to (5) is unique and coincides with that of (4) with probability at least $1 - c'\mathrm{e}^{-c''m}$, where c, c' and c'' are positive constants, $d_{\max} = \max_{v\in\mathcal{V}} |\mathcal{N}(v)|$, $\lambda_{\max} = \max_j \lambda_j$, and $\lambda_{\min} = \min_j \lambda_j$.

Remark 2. *The condition in (6) emphasizes that the sensitivity trust to other opinions cannot be too high: if $\lambda_{\max} \to 1$ then the number of measurements needed for recovery diverges to infinity. This is reasonable, since the final opinions are a function of preconceived opinions and the network sensing performance should depend on the strength of the influencing power of the prejudices.*

Remark 3. *Our result shows that, as conjectured in [22], another important issue that affects the reconstruction performance in Theorem 1 is the degree distribution in the social network. In this sense, if we have a fixed total number of edges, it is easier to recover a network with a concentrated degree distribution (e.g., the Watts-Strogatz network [36]) while a network with power law degree distribution (e.g., the Barabasi-Albert network [37]) is more difficult to recover.*

Proposition 2 shows that if Λ is not known than the identification problem is not well-posed. The ambiguity is due to the missing information about the rate of social interactions. This ambiguity can not be removed without making additional assumptions. However, we can determine an invariant quantity among the ambiguous solutions by defining equivalence classes and resolve the ambiguity by imposing constraints on $\mathrm{diag}(W)$.

Corollary 2 (Learning sensitivity matrix Λ)**.** *If Λ is unknown, then the influence estimation can be cast as in (5) with $A \doteq [X(\infty)^\top, u_j - x_j(\infty)]$ and $B = U$, where $x_j(\infty), u_j$ are the column vector corresponding to j-th row of $X(\infty)$ and U, respectively, with the additional constraint that $w_{jj} = 0$.*

V. Sparse Identification in FJ Model under Gossip-Based Communication

The pioneering version of FJ model is essentially based on the assumption of synchronous rounds of interaction, i.e. agents can communicate and revise their opinions in a synchronous fashion. This is not a realistic assumption in many contexts. As pointed out in [7], *it is obvious that interpersonal influences do not occur in the simultaneous way and there are complex sequences of interpersonal influences in a group*. A more realistic opinion dynamics has been introduced in [10], [38], assuming that only two agents interact during each step, and extended to cases with correlated topics in [39], [40]. On each step a communication edge is randomly activated with uniform distribution from the interaction graph $\mathcal{G} = (V, E)$ with influence matrix W and sensitivity matrix $\Lambda = \mathrm{Diag}(\lambda)$. If (i, j) is selected at time k, then the i-th and j-th agent update their opinions according to

$$\begin{aligned} x_i^\ell(k+1) &= \lambda_i((1-W_{ij})x_i^\ell(k) + W_{ij}x_j^\ell(k)) \\ &\quad + (1-\lambda_i)u_i^\ell(k), \\ x_j^\ell(k+1) &= \lambda_j((1-W_{ji})x_i^\ell(k) + W_{ji}x_j^\ell(k)) \\ &\quad + (1-\lambda_j)u_j^\ell(k). \\ x_s^\ell(k+1) &= x_s^\ell(k) \quad \forall s \neq \{i, j\}. \end{aligned} \tag{7}$$

Proposition 6 (Theorem 1 in [38])**.** *If $\Lambda \neq I$ and for any node $j \in \mathcal{V}$ there exists a path from j to a node m such that $\lambda_m < 1$, it holds that*

1) *the expected dynamics converges and*

$$\begin{aligned} x^{\star,\ell} &:= \lim_{k\to\infty} \mathbb{E}[x^\ell(k)] \\ &= ((I-\Lambda) + D^{-1}\Lambda(I-W))^{-1}(I-\Lambda)u^\ell; \end{aligned}$$

2) *the dynamics (7) is* mean-square-ergodic, *that is,*

$$\lim_{k\to+\infty} \mathbb{E}[\|\bar{x}^\ell(k) - x^{\star,\ell}\|_2^2] = 0, \tag{8}$$

where

$$\bar{x}(k) = \frac{1}{k+1}\sum_{k'=0}^{k} x^\ell(k').$$

where D is the degree matrix of $\mathcal{G}$, i.e., a diagonal matrix whose diagonal entry is equal to the degree $d_i = |\mathcal{N}_i|$.

In this case, the influence identification problem becomes of the form in (3) with

$$\begin{aligned} &A \doteq \overline{X}(\infty)^\top, \\ &B \doteq (I + \Lambda^{-1}D - \Lambda^{-1}D\Lambda)\overline{X}(\infty) - \Lambda^{-1}D(I-\Lambda)U. \end{aligned} \tag{9}$$

We have the following results, whose proof is reported in Appendix D.

Theorem 2. *Let $u^\ell \sim \mathcal{N}(0, I)$ for all $\ell \in [m]$ and W be row-stochastic. If the number of considered topics satisfies the following condition*

$$m \geq 4\chi\frac{(1-\lambda_{\min})^2(1-\lambda_{\min}+\lambda_{\max}/d_{\min})^2}{(1-\lambda_{\max})^4} d_{\max}\log n \tag{10}$$

then the solution to (5) with A and B defined in (9) is unique and coincides with that of (4) with probability at least $1 - \chi'\mathrm{e}^{-\chi''m}$, where χ, χ' and χ'' are positive constants.

VI. Numerical Results

In this section, we evaluate the performance of the proposed method. In order to validate our predictions, we conduct several experiments, considering synthetic networks. The Monte-Carlo simulations were obtained by averaging over 100 instances.

A. Learning sparse synthetic networks

We focus on regular graphs with degree d but similar results can be obtained with the Erdös-Renyi graph with connectivity $p = d/n$. At each instance of the Monte-Carlo method a

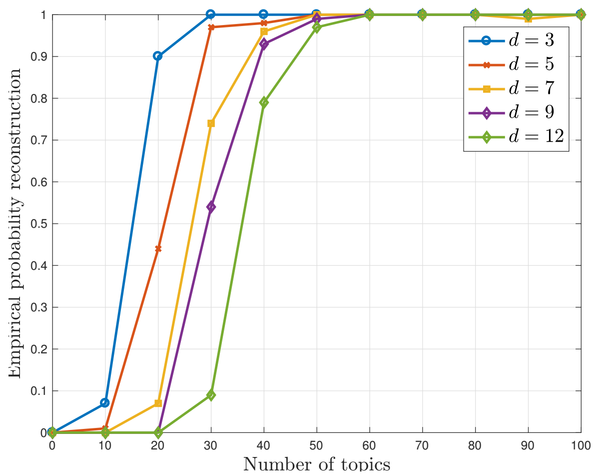

Fig. 2. Regular graphs: Empirical reconstruction probability of regular graphs with uniform weights as a function of number of topics under discussions. Different curves correspond to different values of degree $d \in \{3, 5, 7, 9, 12\}$. The value of sensitivity is fixed to $\lambda = 0.6$.

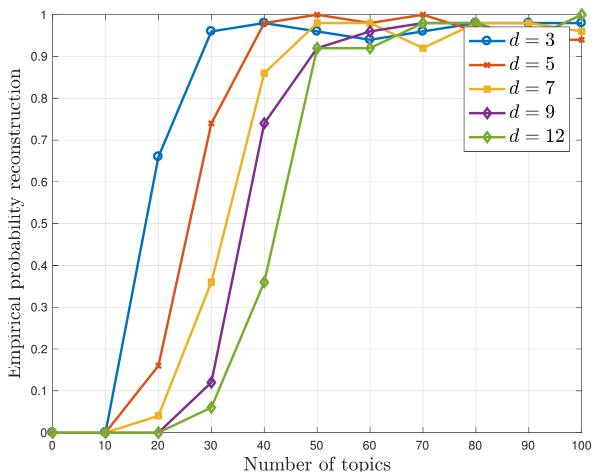

Fig. 3. Regular graphs: Empirical reconstruction probability of regular graphs with uniform weights as a function of number of topics under discussions. Here, the sensitivity parameters λ_i are distributed according to a uniform distribution on the interval [0.5,0.9] and are supposed unknown. Different curves correspond to different values of degree d..

random graph $\mathcal{G}_{n,d} = (\mathcal{V}, \mathcal{E})$ is selected from the ensembles of d-regular graph with size $n = 100$ uniformly at random [41]. The *true* influence matrix $W^\star = W^\star_{\mathcal{G}_{n,d}}$ to be learned is obtained by fixing $W^\star_{ij} = 1/d$ for all $(i, j) \in \mathcal{E}$ and $W^\star_{ij} = 0$ otherwise.

To assess the quality of the reconstructed influence matrix W we evaluate the normalized mean square error defined as follows

$$\mathsf{NMSE} = \|W - W^\star\|_F^2 / \|W^\star\|_F^2$$

The recovery is considered successfully when the NMSE is below 10^{-6}. As a first experiment we consider $\Lambda = 0.6I$. In Figure 2 the empirical probability of reconstruction is depicted as a function of the number of topics considered in the inference problem. Several curves correspond to different values of degree of the regular graph. Moreover, as to be expected, and correctly predicted by (6), the recovery performance degrades as the degree of the graph grows. In all curves we can clearly identify a transition phase: if the number of topics is larger than a threshold, depending on the degree, then the recovery will be successful. In [1], several experiments have been carried out to study the effect of the sensitivity λ in the reconstruction. As to be expected from (1), for fixed value of degree the empirical probability of reconstruction shows a small degradation of the performance as λ increases.

We now consider the case when the sensitivity parameters are not known. The regular network and the influence matrix is generated as before with $n = 100$ nodes and different values of degree $d \in \{3, 5, 7, 9, 12\}$. The sensitivity parameters λ_i are distributed according to a uniform distribution on the interval $[0.5, 0.9]$. Figure 3 shows the empirical probability of reconstruction as a function of number of measurement $m \in [1, 100]$. As pointed out in Discussion 1, there are many solutions consistent to the data. In this case we define equivalence classes and resolve the ambiguity by imposing constraints on $\mathrm{diag}(W) = 0$. As it can be seen, the equivalence class can be determined with high probability using a small amount of observations, despite the lack of information on λ_i.

B. Reconstruction of real networks

We test now the prediction capability of our theoretical results on a real large-scale online social network. We consider the topology of the online social network, extracted from Stanford Large Network Dataset Collection (see the database in https://snap.stanford.edu/data/egonets-Facebook.html). This dataset contains anonymized personal networks of connections between friends of survey participants and was collected by Facebook app. Such personal networks represent friendships of a focal node, known as "ego" node, and such networks are therefore called "ego" networks. We consider Ego Network of node 107, displayed in Figure 4. The size of the network is $|\mathcal{V}| = 1017$, the number of connections is equal to $|\mathcal{E}| = 26734$. Then the total level of sparsity is given by $k = |\mathcal{E}|/|\mathcal{V}|^2 = 0.0257$ and the maximum level of sparsity of the rows is $d_{\max}/|\mathcal{V}| = 0.2495$. The color of the nodes reflects the degree centrality. More precisely, the maximum degree is $d_{\max} = 253$ and the average degree is 53. The distribution of the degree follows a power law decay and few people have a large number of degrees and the majority have small number of degrees.

Also in this case the influence matrix $W^\star$ is obtained by fixing $W^\star_{ij} = 1/d_i$ for all $(i, j) \in \mathcal{E}$ and $W^\star_{ij} = 0$ otherwise, where d_i is the degree of node $i \in \mathcal{V}$. The sensitivity matrix is fixed $\Lambda = \lambda I$. In order to recover a strongly sparse graph structure, we threshold the small entries of the learned matrix, by discarding the entries of the recovered adjacency matrix that are smaller than a predefined threshold of 10^{-3}. Figure 5 depicts the recovered network obtained with $m = 600$ observations. In this case $\lambda = 0.7$ and the $\mathsf{NMSE} = 6.06 \cdot 10^{-9}$.

In Figure 6 the empirical probability of reconstruction is depicted as a function of the number of topics considered in the inference problem. As before, the recovery is considered successfully when the NMSE is below 10^{-6}. Several

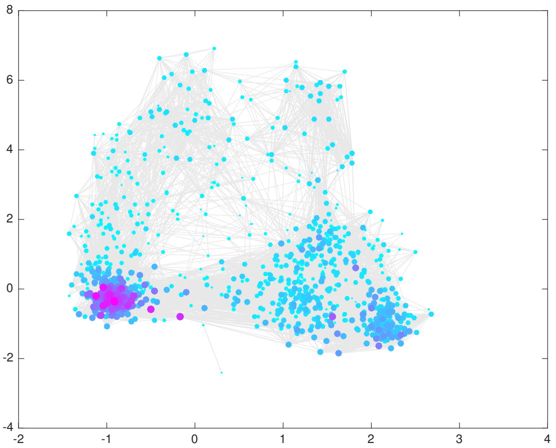

Fig. 4. Ego Network of node 107: Original network with size $n = 1019$. The color of the nodes reflects the degree centrality.

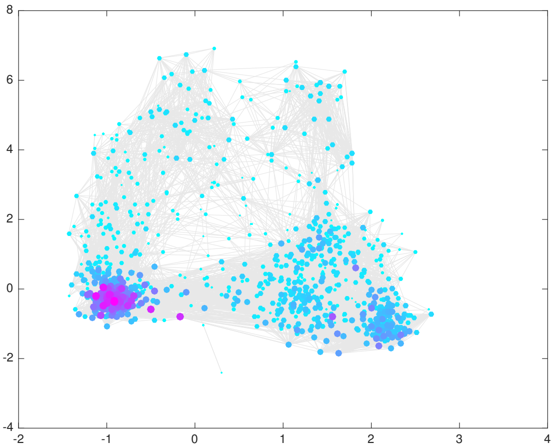

Fig. 5. Ego Network of node 107: Reconstructed network from from $m = 600$ observations. The sensitivity matrix $\Lambda = 0.7I$ and $\mathsf{NMSE} = 6.06{\cdot}10^{-9}$.

curves correspond to different values of sensitivity parameter $\lambda \in \{0.5, 0.6, 0.7, 0.8\}$. In all curves we can clearly identify a transition phase: if the number of topics is larger than a threshold, depending on the degree, then the recovery will be successful. For example, if $\lambda \leq 0.7$, the observation of opinions on 600 (corresponding to around $2.6d_{\max}$) topics are sufficient for perfect recovery with probability one. If $\lambda = 0.6$, then we can recover perfectly with probability equal to 0.7.

VII. Concluding remarks

In this paper, we have presented a framework for learning the sparse influence matrix starting from the observation of the opinions in a social networks, where the agents are not completely open-minded and are persistently influenced by some prejudices. Under the assumption that the initial opinions are independent and identically distributed as a Gaussian distribution, we derive conditions on parameters of the FJ model that guarantee perfect recovery. Experimental results confirm the usefulness of the proposed algorithm in recovering

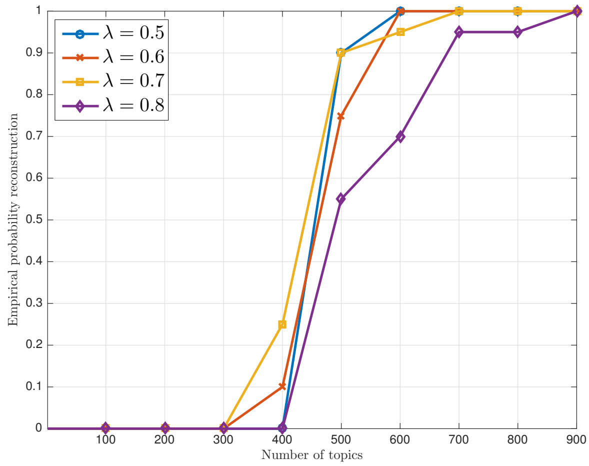

Fig. 6. Ego Network of node 107: Size of network $n = 1019$. Empirical reconstruction probability with uniform weights as a function of number of topics under discussions. Here, the topics are independent. Different curves correspond to different values of sensitivity $\lambda \in \{0.5, 0.6, 0.7, 0.8\}$

a meaningful graph topology and in leading to better data understanding and inference.

In [39], [40] the FJ model has been extended in order to take into account interdependent issues under discussions. We point out that also in that setting the identification problem can be cast as a compressed sensing problem. However, proving the REC condition for the sensing matrix A in that case is not an easy task. In fact, even if the entries of U are i.i.d. with Gaussian distribution, the ensuing sensing matrix A follows a so-called matrix normal distribution. This complicated distribution makes it difficult to apply standard techniques, usually exploited in CS literature, to provide theoretical reconstruction guarantees. The study of specific instances in which reconstruction properties can be theoretically guaranteed will be the subject of future work.

Acknowledgements

The authors are grateful to Noah E. Friedkin and A. Proskurnikov for their valuable comments on a preliminary version of this paper. The authors would also like to thank Giacomo Como, Fabio Fagnani, Costantino Lagoa and Mario Sznaier for interesting conversations on the topics of this paper.

Appendix

A. *Proof of Proposition 2*

The existence of (Λ, W) is guaranteed from Proposition 1. The first constraint in (2) is satisfied by (Λ', W') for an arbitrary diagonal matrix D:

$$\begin{aligned}
&\Lambda' W' x(\infty) + (I - \Lambda')u \\
&\quad = \Lambda' W' x(\infty) + D(I - \Lambda)u \\
&\quad = [\mathrm{Diag}(\mathrm{diag}(\Lambda' W')) + \mathrm{off}(\Lambda' W')]\, x(\infty) + D(I - \Lambda)u \\
&\quad = [\mathrm{Diag}(\mathbb{1} - D((I - \Lambda)\mathbb{1} + \mathrm{off}(\Lambda W)\mathbb{1})) \\
&\quad + D\mathrm{off}(\Lambda W)]\, x(\infty) + D(I - \Lambda)u
\end{aligned}$$

$$= [I - D((I-\Lambda) + \Lambda(I-W)]\, x(\infty) + D(I-\Lambda)u$$
$$= [I - D + D\Lambda W]x(\infty) + D(I-\Lambda)u$$
$$= (I-D)x(\infty) + D(\Lambda W x(\infty) + (I-\Lambda)u) = x(\infty).$$

where the last equality follows from $\Lambda W x(\infty) + (I-\Lambda)u = x(\infty)$. The second constraint also is satisfied: if $\Lambda'_{ii} \neq 0$ for all $i \in [n]$, then

$$\begin{aligned} W'\mathbb{1} &= [\mathrm{Diag}(\mathrm{diag}(W')) + \mathrm{off}(W')]\, \mathbb{1} \\ &= \big[\Lambda'^{-1}\mathrm{Diag}(\mathbb{1} - D((I-\Lambda)\mathbb{1} \\ &+ \mathrm{off}(\Lambda W)\mathbb{1})) + \Lambda'^{-1} D\mathrm{off}(\Lambda W)\big]\, \mathbb{1} \\ &= \big[\Lambda'^{-1}\mathrm{Diag}(\mathbb{1} - D((I-\Lambda)\mathbb{1} + \mathrm{off}(\Lambda W)\mathbb{1})) \\ &+ \Lambda'^{-1} D\Lambda \mathrm{off}(W)\big]\, \mathbb{1} \\ &= \big[\Lambda'^{-1}\mathrm{Diag}(\mathbb{1} - D(I-\Lambda)\mathbb{1})\big]\, \mathbb{1} \\ &= \mathbb{1}. \end{aligned}$$

It is also obvious that the third condition $W' \geq 0$ follows from $\Lambda \geq 0$ and $W \geq 0$. □

B. Proof of Proposition 4

If $u^\ell \sim \mathcal{N}(0, I)$ then x^ℓ is an affine transformation of x^ℓ. Therefore, $x^\ell \sim \mathcal{N}(0, \Sigma)$ with covariance matrix $\Sigma = (I - \Lambda W)^{-1}(I-\Lambda)^2(I-\Lambda W)^{-\top}$. Let now u and v be the eigenvectors corresponding to $\mu_{\min}(\Sigma_{SS})$ and $\mu_{\max}(\Sigma_{SS})$ with unitary norm, respectively, and build the corresponding vectors $\widetilde{u}, \widetilde{v} \in \mathbb{R}^n$ such that $\widetilde{u}_S = u$, $\widetilde{u}_{S^c} = 0$ and $\widetilde{v}_S = v$, $\widetilde{v}_{S^c} = 0$. Since the rows of $A = [a_1^\top, \ldots, a_m^\top]^\top$ are such that $a_i = \Sigma^{1/2}\zeta_i$ with $\zeta_i \sim \mathcal{N}(0, I)$, then

$$\begin{aligned} \frac{1}{m}\|A\widetilde{u}\|_2^2 &= \frac{1}{m}\sum_{i=1}^m (a_i^\top \widetilde{u})^2 = \frac{1}{m}\sum_{i=1}^m (\zeta_i^\top \Sigma^{1/2}\widetilde{u})^2 \\ &= \mu_{\min}(\Sigma_{SS}) \left(\frac{1}{m}\sum_{i=1}^m y_i^2\right) \end{aligned} \tag{11}$$

where $y_i = \zeta_i^\top \widetilde{u} \sim \mathcal{N}(0,1)$. Equivalently, it can be shown that

$$\frac{1}{m}\|A\widetilde{v}\|_2^2 = \mu_{\max}(\Sigma_{SS}) \left(\frac{1}{m}\sum_{i=1}^m y_i^2\right), \tag{12}$$

with $y_i \sim \mathcal{N}(0,1)$. From Bernstein's inequality [42] we have

$$\mathbb{P}\left(\left|\frac{1}{m}\sum_{i=1}^m y_i^2 - 1\right| \geq t\right) \leq 2\mathrm{e}^{-mt^2/8}$$

from which, choosing $t = 1/2$,

$$\mathbb{P}\left(\frac{1}{2} \leq \frac{1}{m}\sum_{i=1}^m y_i^2 \leq 3/2\right) \geq 1 - 2\mathrm{e}^{-m/32}.$$

Combining this bound with (11) and (12) we conclude that at least with probability $1 - 2\mathrm{e}^{-m/32}$ the following inequality holds

$$\frac{\mu_{\max}(A_S^\top A_S)}{\mu_{\min}(A_S^\top A_S)} = \frac{\frac{1}{m}\|A\widetilde{v}\|_2^2}{\frac{1}{m}\|A\widetilde{u}\|_2^2} \geq \frac{1}{3}\frac{\mu_{\max}(\Sigma_{SS})}{\mu_{\min}(\Sigma_{SS})}.$$

□

C. Proof of Theorem 1

The following lemma will be useful to prove Theorem 1. In particular, it guarantees that, given a Gaussian matrix with i.i.d. Gaussian rows, whenever the covariance matrix satisfies REC of order s, then also the sample covariance matrix satisfies the same property if the number of rows is sufficiently large.

Lemma 1 (Restricted Eigenvalue Property). *Let $A \in \mathbb{R}^{m\times n}$ be a matrix with i.i.d. $\mathcal{N}(0,\Sigma)$ rows. Suppose that Σ satisfies the REC of order s with parameter δ and let $\rho^2(\Sigma) = \max_j \Sigma_{jj}$. Then, for universal constants c, c', c'', if the sample size satisfies*

$$m > c\frac{4\rho^2(\Sigma)}{\delta^2} s \log n,$$

then $\widehat{\Sigma} = A^\top A/m$ satisfies REC with parameter $\delta/8$ with probability at least $1 - c'\mathrm{e}^{-c''m}$.

We refer the reader to [29] for the proof of this claim. We are now ready to prove the recovery theorem.

Proof of Theorem 1: If $u^\ell \sim \mathcal{N}(0, I)$ then x^ℓ is an affine transformation of u^j. Therefore, $x^\ell \sim \mathcal{N}(0,\Sigma)$ with covariance matrix $\Sigma = (I - \Lambda W)^{-1}(I-\Lambda)^2(I-\Lambda W)^{-\top}$. It should be noticed that Σ is real, symmetric and definite positive. Moreover, for any $S \subseteq [n]$ we have $\mu_{\min}(\Sigma_{SS}) \geq \mu_{\min}(\Sigma)$ and

$$\begin{aligned} \mu_{\min}(\Sigma) &\geq \min_{\|z\|=1} z^\top (I-\Lambda W)^{-1}(I-\Lambda)^2(I-\Lambda W)^{-\top} z \\ &\geq (1-\lambda_{\max})^2 \min_{\|z\|=1} z^\top (I-\Lambda W)^{-1}(I-\Lambda W)^{-\top} z \\ &= (1-\lambda_{\max})^2 \gamma_n \end{aligned}$$

with

$$\begin{aligned} \gamma_n^{-1} &= \max_{\|v\|_2=1} v^\top (I-\Lambda W)^\top (I-\Lambda W) v \\ &= \max_{\|v\|_2=1} \|(I-\Lambda W)v\|^2 \leq \|I-\Lambda W\|^2 \\ &\leq (1+\lambda_{\max}\|W\|)^2 \leq (1+\lambda_{\max})^2. \end{aligned}$$

Therefore, $\mu_{\min}(\Sigma) \geq (1-\lambda_{\max})^2/(1+\lambda_{\max})^2 =: \delta^2$ Moreover, the maximum variance is bounded as follows

$$\begin{aligned} \rho^2(\Sigma) &= \max_j \Sigma_{jj} \\ &= \max_j \sum_{\ell=1}^n (1-\lambda_\ell)^2 [(I-\Lambda W)^{-1}]_{j\ell} [(I-\Lambda W)^{-\top}]_{\ell j} \\ &\leq (1-\lambda_{\min})^2 \max_j \sum_{\ell=1}^n [(I-\Lambda W)^{-1}]_{j\ell} [(I-\Lambda W)^{-\top}]_{\ell j} \\ &= (1-\lambda_{\min})^2 \max_j \sum_{\ell=1}^n [(I-\Lambda W)^{-1}]_{j\ell} [(I-\Lambda W)^{-1}]_{j\ell} \\ &= (1-\lambda_{\min})^2 \max_j \sum_{\ell=1}^n \left(\sum_{k=0}^\infty [(\Lambda W)^k]_{j\ell}\right)^2 \\ &\leq (1-\lambda_{\min})^2 \max_j \sum_{\ell=1}^n \sum_{k=0}^\infty \sum_{k'=0}^\infty \lambda_{\max}^{k+k'} [W^k]_{j\ell} [W^{k'}]_{j\ell} \\ &= (1-\lambda_{\min})^2 \max_j \sum_{k=0}^\infty \sum_{k'=0}^\infty \lambda^{k+k'} \sum_{\ell=1}^n [W^k]_{j\ell} [W^{k'}]_{j\ell}. \end{aligned}$$

As W is a nonnegative and row-stochastic matrix, then W^p is row-stochastic for all $p \in \mathbb{Z}_{\geq 0}$ and we obtain

$$\rho^2(\Sigma) \leq (1-\lambda_{\min})^2 \max_j \left\{ \sum_{k=0}^{\infty} \sum_{k'=0}^{\infty} \lambda_{\max}^{k+k'} \right\} = \frac{(1-\lambda_{\min})^2}{(1-\lambda_{\max})^2}.$$

Applying Lemma 1 with $s = d_{\max}$ and using bound on $\rho^2(\Sigma)$ we conclude the thesis.

D. Proof of Theorem 2

If $u^\ell \sim \mathcal{N}(0, I)$ then x^ℓ is an affine transformation of u^ℓ and $x^\ell \sim \mathcal{N}(0, \Sigma)$ with covariance matrix given by

$$\Sigma = [I-\Lambda+D^{-1}\Lambda(I-W)]^{-1}(I-\Lambda)^2[I-\Lambda+D^{-1}\Lambda(I-W)]^{-\top}.$$

It should be noticed that Σ is a real, symmetric and positive definite matrix. Let us denote its eigenvalues $\gamma_1 > \ldots > \gamma_n > 0$. We thus have for any $S \subseteq [n]$

$$\mu_{\min}(\Sigma_{SS}) \geq \mu_{\min}(\Sigma) \geq (1-\lambda_{\max})^2 \gamma_n^{-1}$$

$$\begin{aligned} \gamma_n^{-1} &= \max_{\|v\|_2=1} \|(I-\Lambda+D^{-1}\Lambda(I-W))v\|^2 \\ &\leq \|(I-\Lambda+D^{-1}\Lambda(I-W))\|^2 \\ &\leq (1-\lambda_{\min}+2\lambda_{\max}/d_{\min})^2 \end{aligned}$$

We thus have

$$\mu_{\min}(\Sigma) \geq (1-\lambda_{\max})^2/(1-\lambda_{\min}+2\lambda_{\max}/d_{\min})^2 =: \delta^2.$$

Moreover the covariance matrix has maximum variance

$$\begin{aligned} \rho^2(\Sigma) &= \max_j \Sigma_{jj} \\ &= \max_j \sum_{\ell=1}^{n} (1-\lambda_\ell)^2 \left([(I-\Lambda+D^{-1}\Lambda(I-W))^{-1}]_{j\ell} \right)^2 \\ &\leq (1-\lambda_{\min})^2 \max_j \sum_{\ell=1}^{n} \left(\sum_{k=0}^{\infty} (\Lambda(I-D^{-1}(I-W)))^k]_{j\ell} \right)^2 \\ &\leq (1-\lambda_{\min})^2 \max_j \sum_{k=0}^{\infty} \sum_{k'=0}^{\infty} \lambda_{\max}^{k+k'} \sum_{\ell=1}^{n} [M^k]_{j\ell}[M^{k'}]_{j\ell} \end{aligned}$$

with $M = I - D^{-1}(I-W)$ Since W is a nonnegative matrix and row-stochastic, then also M^k is row-stochastic and $\rho^2(\Sigma) \leq (1-\lambda_{\min})^2/(1-\lambda_{\max})^2$. Applying Lemma 1 we conclude the thesis.

References

[1] C. Ravazzi, R. Tempo, and F. Dabbene, "Influence estimation in sparse social networks," Proc. of IEEE Conference on Decision and Control 2017, to appear.

[2] A. V. Proskurnikov and R. Tempo, "A tutorial on modeling and analysis of dynamic social networks. part i," *Annual Reviews in Control*, vol. 43, no. Supplement C, pp. 65 – 79, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1367578817300172

[3] J. French, "A formal theory of social power," *Psychological Review*, vol. 63, pp. 181–194, 1956.

[4] M. H. Degroot, "Reaching a Consensus," *Journal of the American Statistical Association*, vol. 69, no. 345, pp. 118–121, 1974.

[5] S. Chatterjee and E. Seneta, "Towards Consensus: Some Convergence Theorems on Repeated Averaging," *J. Appl. Prob.*, vol. 14, pp. 159–164, 1977.

[6] C. Altafini, "Consensus problems on networks with antagonistic interactions," *IEEE Transactions on Automatic Control*, vol. 58, no. 4, pp. 935–946, April 2013.

[7] N. E. Friedkin and E. C. Johnsen, "Social influence networks and opinion change," *Advances in Group Processes*, vol. 16, pp. 1–29, 1999.

[8] R. Hegselmann and U. Krause, "Opinion dynamics and bounded confidence, models, analysis and simulation," *Journal of Artificial Societies and Social Simulation*, vol. 5, no. 3, p. 2, 2002.

[9] D. Acemoğlu, G. Como, F. Fagnani, and A. Ozdaglar, "Opinion fluctuations and disagreement in social networks," *Math. Oper. Res.*, vol. 38, no. 1, pp. 1–27, Feb. 2013. [Online]. Available: http://dx.doi.org/10.1287/moor.1120.0570

[10] C. Ravazzi, P. Frasca, R. Tempo, and H. Ishii, "Ergodic randomized algorithms and dynamics over networks," *IEEE Transactions on Control of Network Systems*, vol. 2, no. 1, pp. 78–87, March 2015.

[11] E. Yildiz, A. Ozdaglar, D. Acemoglu, A. Saberi, and A. Scaglione, "Binary opinion dynamics with stubborn agents," *ACM Trans. Econ. Comput.*, vol. 1, no. 4, pp. 19:1–19:30, Dec. 2013.

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

[13] D. Knoke, *Political Networks: The Structural Perspective*, ser. Structural Analysis in the Soc. Cambridge University Press, 1994. [Online]. Available: https://books.google.it/books?id=9djTlBLaOccC

[14] S. Wasserman and K. Faust, *Social network analysis : methods and applications*, ser. Structural analysis in the social sciences. Cambridge, New York, Melbourne: Cambridge University Press, 1997, 1994.

[15] L. C. Freeman, *The Development of Social Network Analysis: A Study in the Sociology of Science*. BookSurge Publishing, 2004.

[16] S. Fortunato and D. Hric, "Community detection in networks: A user guide," *Physics Reports*, vol. 659, pp. 1 – 44, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0370157316302964

[17] N. E. Friedkin, "Theoretical Foundations for Centrality Measures," *The American Journal of Sociology*, vol. 96, no. 6, pp. 1478–1504, 1991.

[18] S. Brin and L. Page, "The anatomy of a large-scale hypertextual web search engine," *Comput. Netw. ISDN Syst.*, vol. 30, no. 1-7, pp. 107–117, Apr. 1998. [Online]. Available: http://dx.doi.org/10.1016/S0169-7552(98)00110-X

[19] M. E. J. Newman, "Analysis of weighted networks," *Phys. Rev. E*, vol. 70, no. 5, p. 056131, Nov. 2004. [Online]. Available: http://pre.aps.org/abstract/PRE/v70/i5/e056131

[20] V. Nicosia, R. Criado, M. Romance, G. Russo, and V. Latora, "Controlling centrality in complex networks," *Scientific Reports*, vol. 2, pp. 218 EP –, 01 2012. [Online]. Available: http://dx.doi.org/10.1038/srep00218

[21] Y.-Y. Liu, J.-J. Slotine, and A.-L. Barabasi, "Controllability of complex networks," *Nature*, vol. 473, no. 7346, pp. 167–173, May 2011. [Online]. Available: http://dx.doi.org/10.1038/nature10011

[22] H.-T. Wai, A. Scaglione, and A. Leshem, "Active sensing of social networks." *IEEE Trans. Signal and Information Processing over Networks*, vol. 2, no. 3, pp. 406–419, 2016.

[23] N. Friedkin and E. Johnsen, "Social-influence and opinions," *Journal of Mathematical Sociology*, vol. 15, no. 3-4, p. 193–205, 1990.

[24] N. Friedkin, "A formal theory of reflected appraisals in the evolution of power," *Administrative Science Quarterly*, vol. 56, no. 4, pp. 501– 529, Dec 2011.

[25] M. Timme, "Revealing Network Connectivity from Response Dynamics," *Physical Review Letters*, vol. 98, no. 22, pp. 224 101+, May 2007.

[26] X. Wang, X. Chen, and L. Wang, "Evolutionary dynamics of fairness on graphs with migration," *Journal of Theoretical Biology*, vol. 380, pp. 103 – 114, 2015.

[27] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theor.*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[28] E. Candes, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathematique*, vol. 346, no. 9-10, pp. 589–592, May 2008.

[29] G. Raskutti, M. J. Wainwright, and B. Yu, "Restricted eigenvalue properties for correlated gaussian designs," *J. Mach. Learn. Res.*, vol. 11, pp. 2241–2259, Aug. 2010.

[30] G. Strang, *Introduction to Linear Algebra*. Wellesley-Cambridge Press, 2009.

[31] E. J. Candès, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Comm. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, Aug. 2006.

[32] Y. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*, ser. Compressed Sensing: Theory and Applications. Cambridge University Press, 2012.

[33] A. Mirtabatabaei, P. Jia, and F. Bullo, "Eulerian opinion dynamics with bounded confidence and exogenous inputs," *Siam Journal on Applied Dynamical Systems*, vol. 13, no. 1, p. 425–446, 2014.
[34] J. Lorenz, "Universality in movie rating distributions," *The European Physical Journal B: Condensed Matter and Complex Systems*, vol. 71, no. 2, pp. 251–258, 2009. [Online]. Available: https://EconPapers.repec.org/RePEc:spr:eurphb:v:71:y:2009:i:2:p:251-258
[35] E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theor.*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.
[36] D. J. Watts and S. H. Strogatz, "Collective dynamics of'small-world'networks." *Nature*, vol. 393, no. 6684, pp. 409–10, 1998.
[37] R. Z. Albert, "Statistical mechanics of complex networks," Ph.D. dissertation, Notre Dame, IN, USA, 2001.
[38] P. Frasca, H. Ishii, C. Ravazzi, and R. Tempo, "Distributed randomized algorithms for opinion formation, centrality computation and power systems estimation: A tutorial overview," *European journal of control*, vol. 24, pp. 2–13, July 2015.
[39] S. E. Parsegov, A. V. Proskurnikov, R. Tempo, and N. E. Friedkin, "Novel multidimensional models of opinion dynamics in social networks," *IEEE Transactions on Automatic Control*, vol. PP, no. 99, pp. 1–1, 2016.
[40] N. E. Friedkin, A. V. Proskurnikov, R. Tempo, and S. E. Parsegov, "Network science on belief system dynamics under logic constraints," *Science*, vol. 354, no. 6310, pp. 321–326, October 2016.
[41] B. Bollobas, *Random Graphs*, W. Fulton, A. Katok, F. Kirwan, P. Sarnak, B. Simon, and B. Totaro, Eds. Cambridge University Press, 2001.
[42] J. A. Tropp, "An introduction to matrix concentration inequalities," *Foundations and Trends in Machine Learning*, vol. 8, no. 1-2, pp. 1–230, 2015.

Chiara Ravazzi (M'13) received the B.Sc., M.Sc., and Ph.D. degrees from the Politecnico di Torino,Italy, in 2005, 2007, and 2011, respectively, all in applied mathematics. In 2010, she was a Visiting Member with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge. From 2011 to 2012, she was a Research Assistant with the Department of Mathematics, Politecnico di Torino. From 2012 to 2016, she held a post-doctoral position with the Department of Electronics and Telecommunications, Politecnico di Torino, where she was involved in developing algorithms for sparse recovery problems, within the project Towards Compressive Information Processing Systems funded by the European Union as ERC starting grant (2012-2016). From 2013 to 2016, she was a Research Associate with the National Research Council of Italy (CNR), Institute of Electronics, Computer and Telecommunication Engineering (IEIIT). Since 2017, she has been a Tenured Researcher with the CNR-IEIIT, working in the Systems Modeling & Control Group. Her current research interests include signal processing, optimization and control of network systems.

Roberto Tempo (F'00) was born in Cuorgne', Italy, in 1956. In 1980 he graduated in Electrical Engineering from Politecnico di Torino, Italy. After a period spent at Politecnico di Torino, he joined the National Research Council of Italy (CNR) at the research institute IEIIT, Torino, where he was a Director of Research of Systems and Computer Engineering since 1991. He held visiting and research positions at Tsinghua University and Chinese Academy of Sciences in Beijing, Kyoto University, The University of Tokyo, University of Illinois at Urbana-Champaign, German Aerospace Research Organization in Oberpfaffenhofen and Columbia University in New York.

His research activities were focused on the analysis and design of complex systems with uncertainty, and various applications within information technology, distributed localization of wireless sensor networks and opinion dynamics in social networks.

Dr. Tempo was a Fellow of the IEEE for Contributions to Robust Identification and Control of Uncertain Systems and a Fellow of the International Federation of Automatic Control (IFAC) for Contributions to the Analysis and Control of Uncertain Systems, for Pioneering the Probabilistic Approach to Robustness. He was a recipient of the "IEEE Control Systems Magazine Outstanding Paper Award", of the "Automatica Outstanding Paper Prize Award", and of the Distinguished Member Award from the IEEE Control Systems Society. He was a Corresponding Member of the Academy of Sciences, Institute of Bologna, Italy, Class Engineering Sciences.

In 2010 Dr. Tempo was President of the IEEE Control Systems Society. He served as Editor-in-Chief of Automatica. He was a Senior Editor of the IEEE Transactions on Automatic Control in 2011-2014 and Editor for Technical Notes and Correspondence of the same journal in 2005-2009. He was a member of the Advisory Board of Systems & Control: Foundations & Applications, Birkhauser. He was General Co-Chair for the IEEE Conference on Decision and Control, Florence, Italy, 2013 and Program Chair of the first joint IEEE Conference on Decision and Control and European Control Conference, Seville, Spain, 2005.

On January 14, 2017, Roberto Tempo passed away during a skiing excursion at the Alps near his home town in Northwestern Italy.

Fabrizio Dabbene (SM'89) received the Laurea degree in Electrical Engineering in 1995 and the Ph.D. degree in Systems and Computer Engineering in 1999, both from Politecnico di Torino, Italy. He is currently holding a Senior Researcher position at the institute CNR IEIIT of the National Research Council of Italy, where he is the coordinator of the Systems & Control Technologies (SCT) Group. He has held visiting and research positions at The University of Iowa, at Penn State University and at the RAS (Russian Academy of Sciences) Institute of Control Science, Moscow.

His research interests include probabilistic and randomized methods for systems and control, robust control and identification of complex systems, convex optimization and modeling of environmental systems. On these topics, he has published more than 100 research papers, which include 30 articles published in international journals. He is co-author of the book Randomized Algorithms for Analysis and Control of Uncertain Systems, Springer-Verlag, published in two editions, and editor of the book Probabilistic and Randomized Methods for Design under Uncertainty, also published by Springer. He has been co-organizer of several invited sessions and special courses, and of three Conference Workshops at CDC-00, MSC-08 and MSC-10. He is a recipient of the Outstanding Paper Award from EurAgeng in 2010. He serves as Associate Editor for Automatica and as Chair of the IFAC Technical Committee on Robust Control.

Dr. Dabbene is a Senior Member of the IEEE, and he has taken various responsibilities within the IEEE Control Systems Society. He has been an Associate Editor for the IEEE Transactions on Automatic Control (2008-2012), Program Chair for the CACSD Symposium of the 2010 IEEE Multi-conference on Systems and Control, Chair of the IEEE Technical Committee on CACSD (2010-present), member of CEB (2002-2008), IPC member of various IEEE conferences, and member of several CDC Best Student Paper Award Committees. He served as elected member of the IEEE-CSS Board of Governors for the years 2014-2016, and as IEEE-CSS Vice-President for Publication Activities for the years 2015-2016.