

Preserving Privacy of the Influence Structure in Friedkin-Johnsen Systems

Jack Liell-Cock, Ian R. Manchester* and Guodong Shi*

Abstract

The nature of information sharing in common distributed consensus algorithms permits network eavesdroppers to expose sensitive system information. An important parameter within distributed systems, often neglected under the scope of privacy preservation, is the influence structure – the weighting each agent places on the sources of their opinion pool. This paper proposes a local (i.e. computed individually by each agent), time varying mask to prevent the discovery of the influence structure by an external observer with access to the entire information flow, network knowledge and mask formulation. This result is produced through the auxiliary demonstration of the preserved stability of a Friedkin-Johnsen system under a set of generalised conditions. The mask is developed under these constraints and involves perturbing the influence structure by decaying pseudonoise. This paper provides the information matrix of the best influence structure estimate by an eavesdropper lacking a priori knowledge and uses stochastic simulations to analyse the performance of the mask against ranging system hyperparameters.

1 Introduction

In distributed systems, many independent, sparsely connected agents work together to achieve a common goal. Due to advancements in communication and automation technology, many systems have begun to adopt a decentralised approach such as intelligent transportation networks [1], smart grids [2] and the Internet of Things [3]. In large distributed networks, local controllers do not guarantee the optimisation or stability of the system, while centralised controllers cannot be suitably scaled or are physically impractical to implement. Thus emerged the branch of distributed control, the origin of which can be traced from the work of distributed decision-making [4] and parallel computation [5].

A simple tool developed in distributed control theory are consensus algorithms [6–9] – local algorithms which require minimal computation, communication and synchronisation to aggregate the parameters within a distributed system. In fact, common optimisation and estimation problems such as least squares, sensor calibration, vehicle coordination and Kalman filtering can be formulated as averages of some system parameters [10], and thus can be solved using consensus algorithms.

Many consensus algorithms stem from the DeGroot model [6], which conveys how a set of agents may reach a consensus by opinion broadcasting. In the context of distributed systems, an opinion is an agent's

*I. R. Manchester and G. Shi are with Australian Centre for Field Robotics, The University of Sydney, NSW 2006, Australia. (email: ian.manchester@sydney.edu.au, guodong.shi@sydney.edu.au)

evaluation of a parameter within the system. At each iteration, the opinions are broadcast, and each agent updates their opinion to a weighted aggregate of their local neighbours' ones. The weightings that each agent places on the opinions of other agents is called the influence structure.

An extension of the DeGroot model which employs heterogeneity is the Friedkin-Johnsen (FJ) model [11]. This scheme was originally developed to model the propagation of opinions through social networks, because positively interacting networks can still persistently disagree and cluster if their agents are diverse [12, 13]. This model generalises the DeGroot model with each agent retaining a constant prejudice towards an external bias – commonly chosen as the agent's initial opinion. The weighting the agents place on their biased opinions versus the opinions of other agents is called their susceptibility. A low susceptibility indicates the agent is stubborn in their original opinion whereas a high susceptibility means the agent is more gullible to the opinions of others. We define a distributed system which uses the FJ model to reach a consensus as an FJ system.

While consensus algorithms enjoy system scalability, these approaches are vulnerable to network eavesdroppers due to their information sharing nature. Considerable results have been established to preserve the privacy of the initial opinions of the agents [14–17] by injecting noise or randomly delaying the broadcasts of the true opinions. However, literature still fails to recognise the complications arising from the exposure of the influence structure by an eavesdropping agent. For example, in smart grid technology, a range of electricity producers and consumers are interconnected to form an efficient energy exchange system [2]. It is likely that an electricity producer would use an influence structure to aggregate their consumer's needs and generate power accordingly. The exposure of this influence structure would unveil an unfair bargaining advantage for the consumers by indicating how much the producer relies on their energy consumption. Similar privacy concerns are prevalent in the exchange of information in the smart car industry, and the home automation industry with the Internet of Things.

The aim of this paper is to develop a mask for the influence structure of an FJ system. The term *mask* was coined in [17] and describes a local (in the sense it can be computed on an agent by agent basis) function which induces noise into some system parameter. The purpose of a mask is to preserve the privacy of the parameter, without affecting the limiting conditions of the system. Our mask development process originates by establishing a range of perturbations applicable to the parameters of an FJ system which don't affect the resulting aggregated opinions. The mask is then formed within these restrictions and consists of randomly offsetting the true influence structure by noise from a decaying normal distribution. The extent to which a network eavesdropper with access to the entire information flow, network knowledge and mask formulation can discover the influence structure is then carefully explored. The key results from this investigation are the information matrix produced by the eavesdropper's maximum likelihood estimate, and numerical simulations outlining the dependencies of the mask's performance with respect to FJ system hyperparameters.

The remainder of this paper is organised as follows. In Section 2, we outline notation and revisit the problem formulation. In Section 3, we extend the stability conditions of the FJ system to allow for non-constant influence structures, susceptibilities and external biases. Section 4 uses this result to develop a mask for the influence structure in an FJ system. Section 5 precisely defines the system eavesdropper and performs maximum likelihood estimation on the influence structure from its perspective. Numerical

results are also produced to aid the discussions on the mask's performance. Finally, Section 6 ends the paper with some concluding remarks and potential future directions.

2 Preliminaries and Problem Definition

2.1 Notation

Discrete, time-varying sequences of vectors and matrices are a common theme throughout this paper. To maintain consistency across index notation, a superscript index is an index with respect to the elements (or agents) in a vector or matrix, while a subscript index is an index with respect to time. For example, given a discrete-time vector sequence of evolving opinions $(x_t)_{t \in \mathbb{N}}$, the value x_t^i is the opinion of the i -th agent at timestep t .

The notation $\text{diag}(A_1, \dots, A_m)$ denotes the block diagonal matrix constructed using matrices $\{A_k \mid k = 1, \dots, m\}$ with zeros elsewhere. A non-negative matrix that has all rows sum to 1 is called a row stochastic matrix. For any square matrix A , we say $A \succ 0$ if and only if A is positive definite. All other vector or matrix inequalities, for example $A \geq B$, are taken elementwise for B the same size as A , or B a scalar. For matrix Q , $Q \in N(A)$ if its columns form an orthonormal basis for the nullspace of A .

Let $\|v\|_p = \sqrt[p]{\sum_k |v^k|^p}$ be the l_p -norm of the vector v with $\|v\| = \|v\|_2$. We define $\mathbf{1}_m$ as the vector of length m with 1 as all its entries, and I_m as the $m \times m$ identity matrix. If m is not specified, the size of the vector or matrix is assumed from the context of the equation. Let $\mathcal{N}(\mu, \sigma^2)$ be the normal distribution with mean μ and variance σ^2 .

Consider a set of agents \mathcal{V} with communication links $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ whose interactions are described by a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. This is referred to as a network. Let the total number of agents in the network be $|\mathcal{V}| = n < \infty$. An edge is drawn $(i, j) \in \mathcal{E}$ if agent j directly influences the opinion of agent i . A path from i to j in \mathcal{G} is an ordered set of edges $((i, k_1), (k_1, k_2), \dots, (k_l, j))$. If such a path exists, we say j indirectly influences i . The neighbourhood $\mathcal{N}^i = \{j \mid (i, j) \in \mathcal{E}\}$ is the set of agents which directly influence agent i , and the degree $d^i = |\mathcal{N}^i|$ is the size of this neighbourhood. If there exists a \tilde{d} such that $d^i = \tilde{d}$ for all $i \in \mathcal{V}$, then the network or associated system is said to have degree \tilde{d} . Finally, W is said to be a matrix adapted to the graph \mathcal{G} if and only if $(i, j) \notin \mathcal{E}$ implies $W^{ij} = 0$. The remaining elements W^{ij} where $(i, j) \in \mathcal{E}$ are denoted influence elements.

2.2 The FJ Model

Consider a network of agents $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. Let each agent hold a state $x_t \in \mathbb{R}^n$ which represents the opinions of the agents at timestep $t \in \mathbb{N}$. The influence structure W is the row stochastic matrix adapted to the network \mathcal{G} which holds the relative weightings each agent places on the other agents. We call the i -th row of W the influence structure of agent i . Define the diagonal matrix $0 \leq \Lambda \leq I_n$ holding the ordered susceptibilities $\lambda \in \mathbb{R}^n$ of each agent as the susceptibility matrix, and its non-negative complement $\bar{\Lambda} = I_n - \Lambda$ as the stubbornness matrix. An agent $i \in \mathcal{V}$ is said to be oblivious if $\lambda^i = 1$. Let $u \in \mathbb{R}^n$ be the external biases for each agent. Then the opinions in an FJ system evolve as

$$x_{t+1} = \Lambda W x_t + \bar{\Lambda} u. \quad (1)$$

The stability conditions of FJ systems are given in the following result from [18].

Proposition 1. *For every agent $i \in \mathcal{V}$, if some agent $j \in \mathcal{V}$ is not oblivious and indirectly influences i , the FJ system is stable with the opinions converging to*

$$x_\infty = \lim_{t \rightarrow \infty} x_t = (I - \Lambda W)^{-1} \bar{\Lambda} u. \quad (2)$$

Moreover, this is equivalent to ΛW being Schur stable.

2.3 Problem Formulation

The ability for a resourced eavesdropper to expose the influence structure of an FJ system is outlined in the following results.

Lemma 1. *If ΛW is identifiable, then λ and W are identifiable.*

Proof. Let $i \in \mathcal{V}$ be an arbitrary agent. From the assumptions, the values of $\lambda^i W^i$ are known. By the row stochastic nature of W , $\lambda^i = \|\lambda^i W^i\|_1$. The influence structure of agent i can then be identified by $W^i = \frac{1}{\lambda^i} \cdot \lambda^i W^i$. Repeating this for all agents completes the proof. \square

Proposition 2. *For a sufficiently excited FJ system, an eavesdropper with access to the opinion trajectories of all agents after time $T > 0$ can identify the influence structure.*

In this case, sufficiently excited means that no opinion trajectory can be reconstructed as a linear combination of the others. This is satisfied when the biases and initial opinions of the agents have an independent, random aspect to them. Since the agents are considered independent entities, we assume this is true.

Proof. Using the evolution of the FJ system in (1), the dependence on the external bias can be removed by taking the difference between two successive updates,

$$\tilde{x}_t = \Lambda W \tilde{x}_{t-1},$$

where $\tilde{x}_t = x_{t+1} - x_t$. Since the agents are sufficiently excited, ΛW can be identified using linear regression after at least $n + 2$ successive measurements of x_t ,

$$\Lambda W \tilde{X}_{T:T+n-1} = \tilde{X}_{T+1:T+n}, \quad (3)$$

where $\tilde{X}_{s:t} = [\tilde{x}_s, \dots, \tilde{x}_t]$. Then from Lemma 1, W and λ can be individually identified. \square

The following example instantiates the above results by identifying the influence structure and susceptibilities of a simple 3 agent FJ system using observations of the opinion trajectories for $1 \leq t \leq 5$.

Example 1. Define an FJ system by

$$W = \begin{bmatrix} 0 & 0.5 & 0.5 \\ 0.2 & 0.2 & 0.6 \\ 0.5 & 0 & 0.5 \end{bmatrix}, \quad \lambda = \begin{bmatrix} 0.4 \\ 0.5 \\ 0.6 \end{bmatrix}, \quad x_0 = u = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

The opinion trajectories for $1 \leq t \leq 5$ are

$$\begin{bmatrix} 1.6 & 1.52 & 1.5 & 1.4916 & 1.48676 \\ 2.2 & 2.1 & 2.082 & 2.071 & 2.0651 \\ 2.4 & 2.4 & 2.376 & 2.3628 & 2.35632 \end{bmatrix},$$

which produces

$$\begin{aligned} \tilde{X}_{1:3} &= \begin{bmatrix} -0.08 & -0.02 & -0.0084 \\ -0.1 & -0.018 & -0.011 \\ 0 & -0.024 & -0.0132 \end{bmatrix}, \\ \tilde{X}_{2:4} &= \begin{bmatrix} -0.02 & -0.0084 & -0.00484 \\ -0.018 & -0.011 & -0.0059 \\ -0.024 & -0.0132 & -0.00648 \end{bmatrix}. \end{aligned}$$

We can identify ΛW by

$$\Lambda W = \tilde{X}_{2:4} \left(\tilde{X}_{1:3} \right)^{-1} = \begin{bmatrix} 0 & 0.2 & 0.2 \\ 0.1 & 0.1 & 0.3 \\ 0.3 & 0 & 0.3 \end{bmatrix}.$$

Finally using Lemma 1, we can individually calculate λ and W on a row by row basis,

$$\lambda = \|W\|_1 = \begin{bmatrix} 0.4 \\ 0.5 \\ 0.6 \end{bmatrix}, \quad W = \Lambda W / \lambda = \begin{bmatrix} 0 & 0.5 & 0.5 \\ 0.2 & 0.2 & 0.6 \\ 0.5 & 0 & 0.5 \end{bmatrix}.$$

We define a new metric to quantify the performance of an influence structure mask.

Definition 1. The estimate error is the square root of the largest eigenvalue of the covariance for a maximum likelihood estimate of the influence structure. This is equivalent to the inverted square root of the E-optimal experiment design criterion.

Similar privacy metrics using maximum likelihood estimation are common for analysing differential privacy schemes such as in [15]. This metric is a valid measure of privacy because it is the expected error between the elements from a best estimate of W given no a priori knowledge. Since the range of permissible values of W is between 0 and 1, if the expectation of the estimate error is greater than 1, the influence structure is said to be undiscoverable. Otherwise, we call the influence structure discoverable. The goal of this paper is to develop a mask to make the influence structure of an FJ system undiscoverable.

3 A Family of Masked FJ Models

Popular schemes for privacy preservation perturb the system behaviour in various ways, so there is an inherent tension between secrecy and maintaining normal function. Therefore, it is useful to characterise the stability of FJ systems under a wide range of distortions. This establishes a “development environment” outlining a set of rules a mask must follow to preserve the limiting conditions of the system. Theorem 1 provides this classification for a range of disturbances to the parameters of an FJ system.

Theorem 1. *Given an FJ system (1) which is stable with limit x_∞ , and an arbitrary set of element-wise matrix and vector sequences*

$$\left. \begin{array}{l} W_t \rightarrow W \\ \Lambda_t \rightarrow \Lambda \\ u_t \rightarrow u \end{array} \right\} \text{as } t \rightarrow \infty,$$

with u_t converging exponentially, the system

$$z_{t+1} = \Lambda_t W_t z_t + \bar{\Lambda}_t u_t \quad (4)$$

converges to x_∞ independent of the initial opinions.

Proof. From Proposition 1, an FJ system is stable if and only if the matrix ΛW is Schur stable. Given that ΛW is Schur stable, for some symmetric positive definite matrix $R \in \mathbb{R}^{n \times n}$, there exists a symmetric positive definite matrix $P \in \mathbb{R}^{n \times n}$ such that [19]

$$P - (\Lambda W)' P (\Lambda W) = R.$$

As P and R are both positive definite, there exists some $\tilde{\alpha} \in (0, 1)$ such that

$$R - \tilde{\alpha} P \succ 0,$$

which implies

$$(1 - \tilde{\alpha})P - (\Lambda W)' P (\Lambda W) \succ 0. \quad (5)$$

Define the set of matrices

$$\mathcal{M} := \{M \in \mathbb{R}^{n \times n} \mid (1 - \tilde{\alpha})P - M' P M \succ 0\}.$$

Since the set of positive definite matrices is open, and the function $(1 - \tilde{\alpha})P - M' P M$ is continuous with respect to M , \mathcal{M} is an open set. Also, from (5), $\Lambda W \in \mathcal{M}$. By the multiplication rule for limits $\Lambda_t W_t \rightarrow \Lambda W \in \mathcal{M}$, hence there exists a $t_0 \in \mathbb{N}$ such that for all $t \geq t_0$, $\Lambda_t W_t \in \mathcal{M}$.

Define a discrete time system as

$$x_{t+1} = \Lambda_t W_t x_t := f_t(x_t). \quad (6)$$

As P is positive definite, the function $V : \mathbb{R}^n \rightarrow \mathbb{R}$ defined by $V(x) = \sqrt{x' P x}$ is a vector norm. V is a continuous, positive definite, radially unbounded function, and for all $t \geq t_0$ and all $x \in \mathbb{R}^n \setminus \{0\}$,

$$\begin{aligned} V(f_t(x)) &= \sqrt{x' (\Lambda_t W_t)' P (\Lambda_t W_t) x} \\ &< \sqrt{x' (1 - \tilde{\alpha}) P x} \\ &= \alpha V(x), \end{aligned}$$

where $\alpha := \sqrt{1 - \tilde{\alpha}} < 1$. Hence, V is a well-defined Lyapunov function for (6) when $t \geq t_0$.

For z_t defined by (4), z_{t_0} will remain finite provided t_0 is finite. Moreover, the convergence properties of a sequence are independent of the first finite terms. Therefore, for the remainder of the proof, the analysis of system (4) is reduced to the analysis on $(z_t)_{t \geq t_0}$. Using the triangle inequality,

$$V(z_{t+1}) = V(f_t(z_t) + \bar{\Lambda}_t u_t) \leq V(f_t(z_t)) + V(\bar{\Lambda}_t u_t). \quad (7)$$

By the multiplication rule for limits $\bar{\Lambda}_t u_t \rightarrow \bar{\Lambda}u$, hence there exists an $m_1 > 0$ such that for all $t \in \mathbb{N}$, $V(\bar{\Lambda}_t u_t) < m_1$. Continuing from (7) gives

$$V(z_{t+1}) < \alpha V(z_t) + m_1. \quad (8)$$

Therefore, for all $V(z_t) > \frac{m_1}{1-\alpha}$,

$$V(z_{t+1}) < \alpha V(z_t) + (1-\alpha)V(z_t) = V(z_t). \quad (9)$$

So $V(z_t)$ is strictly decreasing when it is larger than $\frac{m_1}{1-\alpha}$, implying that z_t is bounded.

Define the sequence $y_t = z_{t+1} - z_t$. By the triangle inequality,

$$\begin{aligned} V(y_t) &= V(\Lambda_t W_t z_t + \bar{\Lambda}_t u_t - \Lambda_{t-1} W_{t-1} z_{t-1} - \bar{\Lambda}_{t-1} u_{t-1}) \\ &= V(\Lambda_t W_t z_t - \Lambda_t W_t z_{t-1} \\ &\quad + \Lambda_t W_t z_{t-1} - \Lambda_{t-1} W_{t-1} z_{t-1} \\ &\quad + \bar{\Lambda}_t u_t - \bar{\Lambda}_{t-1} u_{t-1}) \\ &\leq V(f_t(y_{t-1})) + V([\Lambda_t W_t - \Lambda_{t-1} W_{t-1}]z_{t-1}) \\ &\quad + V(\bar{\Lambda}_t u_t - \bar{\Lambda}_{t-1} u_{t-1}). \end{aligned}$$

As z_t is bounded and $\Lambda_t W_t$ contracts exponentially under V , $V([\Lambda_t W_t - \Lambda_{t-1} W_{t-1}]z_{t-1}) \rightarrow 0$ exponentially. Additionally, $V(\bar{\Lambda}_t u_t - \bar{\Lambda}_{t-1} u_{t-1}) \rightarrow 0$ exponentially from the exponential convergence of u_t . Therefore, there exists some $\beta < 1$ and $m_2 > 0$ such that,

$$\begin{aligned} m_2 \beta^t &> V([\Lambda_t W_t - \Lambda_{t-1} W_{t-1}]z_{t-1}) \\ &\quad + V(\bar{\Lambda}_t u_t - \bar{\Lambda}_{t-1} u_{t-1}). \end{aligned}$$

Combining these results gives

$$V(y_t) < \alpha V(y_{t-1}) + m_2 \beta^t.$$

When $V(y_{t-1}) > \frac{2m_2 \beta^t}{1-\alpha}$,

$$V(y_t) < \alpha V(y_{t-1}) + \frac{1-\alpha}{2} V(y_{t-1}) = \frac{1+\alpha}{2} V(y_{t-1}).$$

So $V(y_t)$ falls below $m_2 \beta^t$ at an exponential rate. Since $m_2 \beta^t \rightarrow 0$ exponentially, $V(y_t) = V(z_{t+1} - z_t) \rightarrow 0$ exponentially. Thus z_t converges to some $z_\infty \in \mathbb{R}^n$. Taking the limit of both sides of (4) gives

$$z_\infty = \Lambda W z_\infty + \bar{\Lambda}u \implies z_\infty = (I - W\Lambda)^{-1} \bar{\Lambda}u = x_\infty,$$

completing the proof. □

4 Influence Structure Mask

In this section, we present a specific realisation of the class defined in Theorem 1 to mask the values of the influence structure in an FJ system. As indicated by the theorem, this mask offsets the true influence structure to produce confidentiality. The susceptibilities and the external biases remain unchanged because

we assume the worst-case scenario that the eavesdropper is aware of these values. Additionally, the non-influence elements of W are unaltered because the existence of the corresponding connections in the underlying graph is not certain.

By motivation from differential privacy ideas explored in [15,17], the mask we developed is

$$W_t = W + e^{-\varphi t} V_t, \quad (10)$$

where $\varphi > 0$ is called the decay rate and is publicly known. V_t is a matrix adapted to the network \mathcal{G} with the remaining influence elements independently chosen from $\mathcal{N}(0, 1)$. Each row of V_t is chosen locally and privately by the corresponding agent, allowing the algorithm to be compatible with distributed systems. Compared to [15,17], we propose to add a mask mechanism to the weights instead of the dynamic states of the system. While conceptually similar, the significant difference is that our mask leads to a time-varying system, and the others a time-invariant system with time-varying noises.

At each time step, the following actions are taken:

- (i) Each agent populates their elements of V_t with independently selected random variables chosen from $\mathcal{N}(0, 1)$.
- (ii) The decoy influence structure is set as $W_t = W + e^{-\varphi t} V_t$.
- (iii) The opinions are pooled: $x_{t+1} = \Lambda W_t x_t + \bar{\Lambda} u$.

W_t converges to W by the decay term, hence Theorem 1 enforces that the system

$$x_{t+1} = \Lambda W_t x_t + \bar{\Lambda} u$$

converges to the same consensus as (1), provided (1) is stable.

The idea behind the mask is that the decay rate φ , is slower than the convergence of the agent's opinions to their consensus value, so the opinions convergence is dominated by the mask fading rather than the dynamics of the FJ system. Figure 1 illustrates the application of this mask to a standard 5 agent FJ system.

5 System Discoverability

We aim to prevent the influence structure of any agent from being directly measured or inferred by observers who gain access to the data flow and structure of the distributed system during the opinion pooling. In this section, we introduce the eavesdropper of concern which enjoys a worst-case knowledge set of the distributed system parameters. We determine the information matrix for the maximum likelihood estimate by this eavesdropper and use numerical results to discuss threats to the security of the distributed system. The maximum likelihood estimation problem is an extension of [20] where the influence structure is estimated from a range of initial biases and limiting opinions resulting from multiple opinion poolings over the same FJ network.

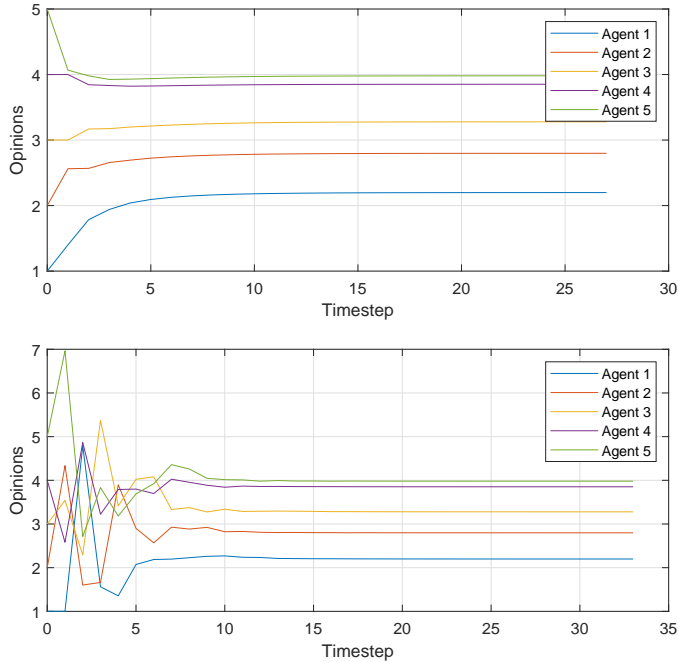


Figure 1: Comparison of FJ system without (top) and with (bottom) the implementation of the mask

5.1 The Eavesdropper

The eavesdropper attempts to estimate the parameter W through observations of x_t and the remaining parameters of the FJ system. We assume the formula of the mask and the distribution of the randomly generated elements of V_t are public knowledge. The knowledge space of the eavesdropper is therefore

$$\mathcal{K} = \{\mathcal{G}, u, \Lambda, \{x_0, \dots, x_T\}\},$$

where T is the timestep when a consensus is reached, and the agents cease their opinion broadcasts.

With the mask, the trajectory of opinions in the FJ system is given by (10). Rearranging this produces

$$W_t x_t = \Lambda^{-1} (x_{t+1} - \bar{\Lambda} u). \quad (11)$$

Without loss of generality, we assume the eavesdropper is seeking the influence structure of the first agent. Specifically, the eavesdropper must compute the values in the first row of W . W is adapted from the underlying network \mathcal{G} , so any elements of W representing disconnections are known by the eavesdropper to be zero. Therefore, exposing the influence structure of the first agent reduces to the problem of calculating the influence elements in the first row of W .

To accommodate notation, let w , w_t and v_t be vectors holding the influence elements of the first rows of W , W_t and V_t respectively. Consequently,

$$w_t = w + e^{-\varphi t} v_t. \quad (12)$$

Let A_t be the row vector containing the opinions x_t which directly influence the opinion of the first agent, and $B_t = \Lambda^{-1} (x_{t+1} - \bar{\Lambda} u)$ be the right-hand side of (11). Taking b_t as the first element of B_t ,

$$A_t w_t = b_t. \quad (13)$$

The number of influence elements in the first row of W is equivalent to the degree of the first agent d^1 , which we denote d for ease. It follows that the vectors w , w_t , v_t and A_t all have d elements.

5.2 Maximum Likelihood Estimation

The eavesdropper understands that w_t is generated by (12), where the elements of v_t are selected from $\mathcal{N}(0, 1)$. Therefore, the elements of $e^{-\varphi t}v_t$ can be considered to be selected from the distribution $\mathcal{N}(0, e^{-2\varphi t})$. The probability density function of this vector is

$$p(e^{-\varphi t}v_t) = \left(\frac{1}{e^{-\varphi t}\sqrt{2\pi}} \right)^d \exp\left(-\frac{\|e^{-\varphi t}v_t\|^2}{2e^{-2\varphi t}} \right).$$

This can be treated as a likelihood function of w given an observation w_t using (12).

$$L(w | w_t) = \left(\frac{1}{e^{-\varphi t}\sqrt{2\pi}} \right)^d \exp\left(-\frac{\|w_t - w\|^2}{2e^{-2\varphi t}} \right)$$

Ignoring the scalar term, the log-likelihood of w given w_t is

$$l(w | w_t) = -\frac{\|w_t - w\|^2}{2e^{-2\varphi t}}.$$

Therefore, the log-likelihood function of w for a range of observations is

$$l(w | w_0, \dots, w_{T-1}) = -\sum_{t=0}^{T-1} \frac{\|w_t - w\|^2}{2e^{-2\varphi t}}.$$

The values of w_t cannot be directly measured by the eavesdropper. Instead, each observation provides a restriction for w_t expressed by the constraint (13). Additional constraints on w are $\mathbf{1}'w = 1$ and $w \geq 0$ because W is row stochastic.

Using all information available to the eavesdropper, the maximum likelihood estimate for w can be formulated as the solution to a quadratic optimisation problem.

$$\begin{aligned} & \underset{w, w_t}{\text{maximize}} && -\sum_{t=0}^{T-1} \frac{\|w_t - w\|^2}{2e^{-2\varphi t}} \\ & \text{subject to} && A_t w_t = b_t, \\ & && \mathbf{1}'w = 1, \\ & && w \geq 0 \end{aligned} \tag{14}$$

Define the vector $\tilde{w} = [w'_0, \dots, w'_{T-1}, w']'$ to hold the variables of (14), and define $A = \text{diag}(A_0, \dots, A_{T-1}, \mathbf{1}')$. Similarly, define the vector $b = [b_0, \dots, b_{T-1}, 1]'$. Therefore, the set of linear equality constraints can be expressed as $A\tilde{w} = b$.

Define the diagonal matrix $H_k = \text{diag}(H_{0,k}, \dots, H_{T-1,k})$, where $H_{t,k} = e^{2\varphi t}I_k$ are the identity matrices iteratively multiplied by the scaling factors of the cost function in (14). Also define

$$Y = \begin{bmatrix} I_d & & -I_d \\ & \ddots & \vdots \\ & & I_d & -I_d \end{bmatrix} \in \mathbb{R}^{Td \times (T+1)d}.$$

It follows that $l(\tilde{w}) = -1/2 \cdot \tilde{w}'Y'H_dY\tilde{w}$, so the maximum likelihood problem can be reduced to:

$$\begin{aligned} & \underset{\tilde{w}}{\text{maximize}} && -\frac{1}{2} \cdot \tilde{w}'Y'H_dY\tilde{w} \\ & \text{subject to} && A\tilde{w} = b, \\ & && w \geq 0. \end{aligned}$$

Define $Q = \text{diag}(Q_0, \dots, Q_{T-1}, Q_{\mathbb{1}})$ where $Q_t \in N(A_t)$ and $Q_{\mathbb{1}} \in N(\mathbf{1}')$. Therefore, $Q \in N(A)$. The information matrix of the maximum likelihood estimate is the negative curvature of the cost function restricted to the constraints, which is given by

$$\mathcal{I} = \frac{\partial^2}{\partial s^2} \left(\frac{1}{2} \cdot sQ'Y'H_dYQs \right) = Q'Y'H_dYQ.$$

Although \mathcal{I} gives the Fisher information for the estimate quality of \tilde{w} , only the information for w is of importance because the variables w_t were merely constructed to aid in its recovery. It is of no concern to the eavesdropper if the information about the variables w_t is minimal, provided the estimate for w is accurate.

By block matrix inversion [21], the information matrix of w is the Schur compliment of the bottom right $(d-1) \times (d-1)$ block of \mathcal{I} . Expanding and simplifying the expression for \mathcal{I} yields

$$\mathcal{I} = \begin{bmatrix} H_{d-1} & K \\ K' & R \end{bmatrix}$$

where $R = \sum_{t=0}^{T-1} e^{2t\varphi} I_{d-1}$, and

$$K = \begin{bmatrix} -Q'_0 Q_{\mathbb{1}} \\ -e^{2\varphi} Q'_1 Q_{\mathbb{1}} \\ \vdots \\ -e^{2\varphi(T-1)} Q'_{T-1} Q_{\mathbb{1}} \end{bmatrix}.$$

Thus, the information matrix for w is

$$\begin{aligned} \mathcal{I}_w &= R - K'(H_{d-1})^{-1}K \\ &= \sum_{t=0}^{T-1} e^{2t\varphi} (I_{d-1} - Q'_1 Q_t Q'_t Q_{\mathbb{1}}). \end{aligned}$$

Claim 1. Let $A \in \mathbb{R}^{1 \times n}$ be a row matrix and define $\hat{A} = \frac{A}{\|A\|}$ as the unit vector in the direction of A . Set $\hat{Q} \in N(A)$. Then

$$\hat{Q}\hat{Q}' = I_n - \hat{A}'\hat{A}. \quad (15)$$

Proof. Let $x \in \mathbb{R}^n$ be an arbitrary vector. By definition of the nullspace, the columns of \hat{Q} and \hat{A}' form an orthonormal basis for \mathbb{R}^n . Hence, there exists vectors x_Q and x_A such that

$$x = \hat{Q}x_Q + \hat{A}'x_A \quad (16)$$

This gives

$$\begin{aligned} \hat{Q}\hat{Q}'x &= \hat{Q}\hat{Q}'\hat{Q}x_Q + \hat{Q}\hat{Q}'\hat{A}'x_A \\ &= \hat{Q}x_Q. \end{aligned}$$

On the other hand,

$$\begin{aligned} (I_n - \hat{A}'\hat{A})x &= \hat{Q}x_Q + \hat{A}'x_A - \hat{A}'\hat{A}\hat{Q}x_Q - \hat{A}'\hat{A}\hat{A}'x_A \\ &= \hat{Q}x_Q. \end{aligned}$$

Since $\hat{Q}\hat{Q}'x = (I_n - \hat{A}'\hat{A})x$ for all $x \in \mathbb{R}^n$, it follows that $\hat{Q}\hat{Q}' = I_n - \hat{A}'\hat{A}$. \square

Let $\hat{A}_t = \frac{A_t}{\|A_t\|}$ be the unit vector in the direction of A_t . From the claim it follows

$$\begin{aligned} \mathcal{I}_w &= \sum_{t=0}^{T-1} e^{2t\varphi} (I_{d-1} - Q'_1 (I_d - \hat{A}'_t \hat{A}_t)' Q_1) \\ &= \sum_{t=0}^{T-1} e^{2t\varphi} Q'_1 \hat{A}'_t \hat{A}_t Q_1. \end{aligned} \quad (17)$$

The inverse of this matrix is the covariance for the maximum likelihood estimate of w . Thus, the square root of the largest principal component of $\Sigma_w = (\mathcal{I}_w)^{-1}$ is the estimate error for the eavesdropper.

The value of the estimate error cannot be cleanly extracted from this information matrix formulation. While the exponentially growing term appears concerning, the auxiliary principal component contributions of the matrices $Q'_1 \hat{A}'_t \hat{A}_t Q_1$ exponentially shrink. The rate of this reduction is approximately equal to the unmasked system's convergence. The performance of the mask thus results as expected from the tension between the decay rate of the mask and convergence rate of the system. Figure 2 illustrates the best possible estimate errors for 1,000,000 randomly generated, masked FJ systems of 100 agents with degree 10, and decay rate 1. These errors are finite, so the influence structure is identifiable, however the expected estimate error is much greater than 1, so the influence structure is undiscoverable. The extent to which it remains undiscoverable over varying system parameters is discussed using further numerical results in the following subsection.

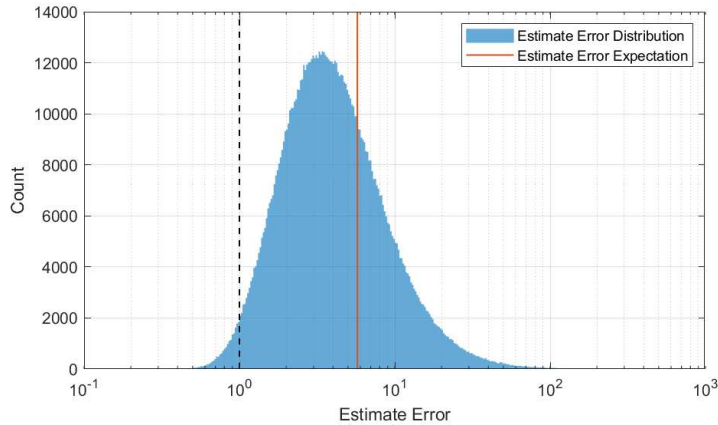


Figure 2: Distribution of the estimate error for random masked FJ systems

5.3 Discoverability Dependence on System Parameters

Due to the complex and stochastic nature of \mathcal{I}_w , a deterministic lower bound on the estimate error could not be calculated. Therefore, we moved our analysis away from an analytical approach, and towards a numerical one.

To determine the dependencies of the estimate error, we simulated many masked FJ systems holding the hyper-parameters constant, except the one under analysis. We tuned this parameter over a range of practical values and recorded the results. When each of the global parameters were not being altered, we set the number of agents to 100, the degree of the system to 10, the decay rate to 0.3, the convergence tolerance to 10^{-4} , and the initial opinions and susceptibilities were selected from the uniform distribution $[0,1]$. The convergence tolerance is the maximum allowable change of the agents' opinions over a single iteration for the system to be considered converged.

A comparison of the spreads of the estimate error against the decay rate of the mask is shown in Figure 3. Interestingly, there is an optimal decay rate for privacy preservation. For an FJ system defined by the previous global parameters, the box plots show that the decay rate which maximises the privacy of this system is $\varphi = 1$. The expected estimate error at this decay rate was 5.19, thus we concluded in this case the influence structure was undiscoverable.

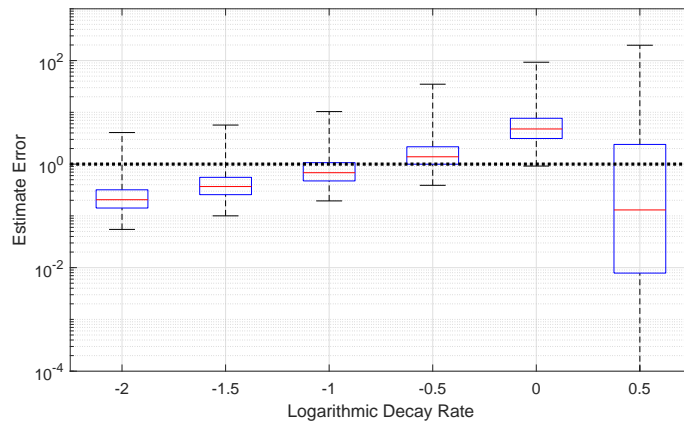


Figure 3: Spread of estimate errors against logarithmic decay rate of the mask ($\log_{10} \varphi$)

The decline in privacy as the decay rate decreased was due to the slower convergence of the system as the mask perturbed the opinions longer. This delayed the opinions from converging to within the thresholds, so the eavesdropper had access to more timesteps to extract system information. On the other hand, as the decay rate increased beyond unity, the noise injected by the mask decayed too fast revealing the true trajectories before they had sufficiently converged.

A comparison of the spreads of the estimate error against the centre point of the distribution from which the agent susceptibilities were selected is shown in Figure 4. In this case, the susceptibilities were randomly selected from a uniform distribution of width 0.1 with midpoints ranging from 0.05 to 0.95.

The estimate error of the system decreased as the susceptibilities of the agents increased. Additionally, the decrease in privacy occurred at a decreasing rate until the susceptibilities neared a value of 1, where the worst-case estimate error cascaded towards zero. As the susceptibilities decreased, the eigenvalues of ΛW grew smaller, hence the unmasked FJ system converged faster. Therefore, the convergence of the FJ system was primarily dominated by the mask for lower susceptibilities, so it was difficult for the eavesdropper to isolate the opinion trends resulting from the true influence structure. As the susceptibilities grew, the convergence of the masked system became increasingly dependent on the underlying influence structure, making it more susceptible to discovery.

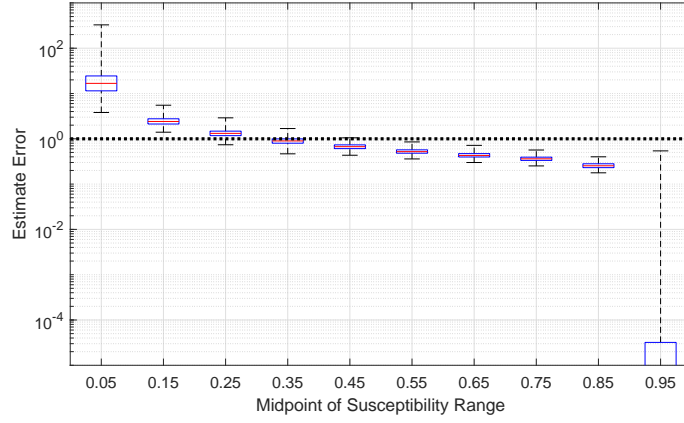


Figure 4: Spread of estimate errors against the susceptibility midpoint

Figure 5 illustrates that the estimate error increased exponentially with the degree of the system. The largest exponential rate of increase occurred when the degree was less than 15, where the median of the estimate errors rapidly approached 1. Therefore, a larger degree was more beneficial to prevent the discovery of the influence structure.

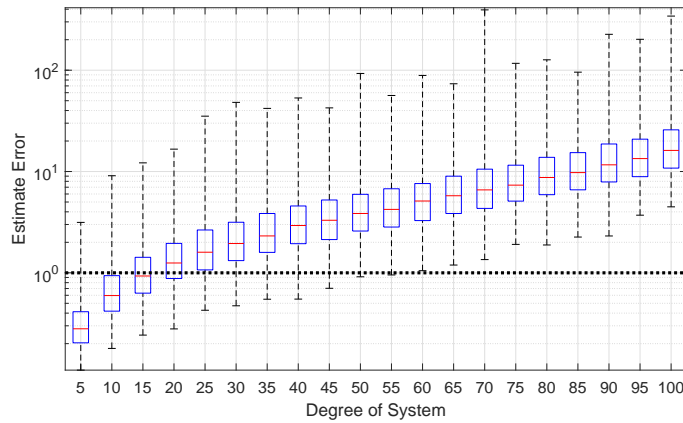


Figure 5: Spread of estimate errors against the degree of the system

Figure 5 also displays the importance of the availability of the underlying network to the eavesdropper when it performs the maximum likelihood estimate. Without knowledge of the graph, the problem simplification in Section 5.1 where the non-influence elements of W were discarded could not be performed. This is because the influence elements of W are specified by the underlying network. Therefore, the eavesdropper would have no option but to extend the “degree” of the estimate of the influence structure from d to n . Provided the estimate of the influence structure \hat{w} was accurate, the entries of \hat{w} close to zero could be assumed disconnections in the network. However, the number of agents in a distributed system is commonly orders of magnitude larger than the degree of the system, so the estimate error would rapidly increase, rendering the influence structure undiscoverable.

6 Conclusions

In this paper, we showed that the convergence behaviour of an FJ system was preserved under a much broader set of conditions than what is currently appreciated in literature. Rather than requiring the agents' susceptibilities, biases and influence structure to be constant with respect to time, we demonstrated that provided these parameters converged to their true values, the limiting opinions of the agents were preserved. We used this result to develop a mask for the influence structure of an FJ system. The privacy preserving performance of the mask was dependent on global system parameters, so numerical simulations were required to determine its validity on a case by case basis. Nevertheless, a high degree, concealed underlying network, low convergence tolerance, or low susceptibility range gave strong indicators that the privacy of the influence structure was preserved by the mask.

In future work, the formulation of the mask could be refined to improve the numerical results or enforce strict bounds on the estimate error. Additionally, other masks for the susceptibilities, external biases, or initial opinions of the agents could be developed using the liberties granted by Theorem 1. The compatibility of these masks could be analysed to determine if they work in unison to preserve the privacy of multiple parameters of an FJ system.

References

- [1] C. C. de Wit, F. Morbidi, L. L. Ojeda, A. Y. Kibangou, I. Bellicot, and P. Bellemain, "Grenoble traffic lab: An experimental platform for advanced traffic monitoring and forecasting," *IEEE Control Systems Magazine*, vol. 35, no. 3, pp. 23–39, Jun. 2015.
- [2] D. J. Hill, T. Liu, and G. Verbic, "Smart grids as distributed learning control," in *IEEE Power and Energy Society General Meeting*, Jul. 2012, pp. 1–8.
- [3] J.-Y. Chang, "A distributed cluster computing energy-efficient routing scheme for internet of things systems," *Wireless Personal Communications*, vol. 82, no. 2, pp. 757–776, May 2015.
- [4] J. N. Tsitsiklis, D. P. Bertsekas, and M. Athans, "Distributed asynchronous deterministic and stochastic gradient optimization algorithms," *IEEE Transactions on Automatic Control*, vol. 31, no. 9, pp. 803–812, Sep. 1986.
- [5] N. A. Lynch, *Distributed Algorithms*. San Francisco, CA: Morgan Kaufmann, Mar. 1996.
- [6] M. H. DeGroot, "Reaching a Consensus," *Journal of the American Statistical Association*, vol. 69, no. 345, pp. 118–121, 1974.
- [7] 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, Jul. 2002.
- [8] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2508–2530, Jun. 2006.

- [9] V. Amelkin, F. Bullo, and A. K. Singh, “Polar opinion dynamics in social networks,” *IEEE Transactions on Automatic Control*, vol. 62, no. 11, pp. 5650–5665, Nov. 2017.
- [10] F. Garin and L. Schenato, *A Survey on Distributed Estimation and Control Applications Using Linear Consensus Algorithms*. London: Springer, 2010, pp. 75–107.
- [11] N. E. Friedkin and E. C. Johnsen, *Social Influence Network Theory: A Sociological Examination of Small Group Dynamics*. New York: Cambridge University Press, 2011.
- [12] W. Xia and M. Cao, “Clustering in diffusively coupled networks,” *Automatica*, vol. 47, no. 11, pp. 2395–2405, Nov. 2011.
- [13] D. Aeyels and F. D. Smet, “Cluster formation in a time-varying multi-agent system,” *Automatica*, vol. 47, no. 11, pp. 2481–2487, Nov. 2011.
- [14] X. Duan, J. He, P. Cheng, Y. Mo, and J. Chen, “Privacy preserving maximum consensus,” in *54th IEEE Conference on Decision and Control (CDC)*, Dec. 2015, pp. 4517–4522.
- [15] Y. Mo and R. M. Murray, “Privacy preserving average consensus,” *IEEE Transactions on Automatic Control*, vol. 62, no. 2, pp. 753–765, Feb. 2017.
- [16] Y. Liu, J. Wu, I. R. Manchester, and G. Shi, “Gossip algorithms that preserve privacy for distributed computation part 1: The algorithms and convergence conditions,” in *57th IEEE Conference on Decision and Control (CDC)*, Dec. 2018, pp. 4499–4504.
- [17] C. Altafini, “A system-theoretic framework for privacy preservation in continuous-time multiagent dynamics,” Apr. 2019, arXiv:1904.11246.
- [18] 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. 62, no. 5, pp. 2270–2285, May 2017.
- [19] J. P. Hespanha, *Linear Systems Theory*. Princeton: Princeton University Press, Feb. 2018.
- [20] C. Ravazzi, R. Tempo, F. Dabbene, “Influence estimation in sparse social networks,” in *56th IEEE Conference on Decision and Control (CDC)*, Dec. 2017, pp. 775–780.
- [21] T.-T. Lu and S.-H. Shiou, “Inverses of 2 x 2 block matrices,” *Computers and Mathematics with Applications*, vol. 43, no. 1, pp. 119–129, Jan. 2002.