Learning Communities from Equilibria of Nonlinear Opinion Dynamics

Yu Xing, Anastasia Bizyaeva, and Karl H. Johansson

Abstract— This paper studies community detection for a nonlinear opinion dynamics model from its equilibria. It is assumed that the underlying network is generated from a stochastic block model with two communities, where agents are assigned with community labels and edges are added independently based on these labels. Agents update their opinions following a nonlinear rule that incorporates saturation effects on interactions. It is shown that clustering based on a single equilibrium can detect most community labels (i.e., achieving almost exact recovery), if the two communities differ in size and link probabilities. When the two communities are identical in size and link probabilities, and the inter-community connections are denser than intra-community ones, the algorithm can achieve almost exact recovery under negative influence weights but fails under positive influence weights. Utilizing fixed point equations and spectral methods, we also propose a detection algorithm based on multiple equilibria, which can detect communities with positive influence weights. Numerical experiments demonstrate the performance of the proposed algorithms.

I. Introduction

Learning networks from group dynamics has gained significant interest in various disciplines [1], [2], due to its wide applications in influence maximization [3] and recommender systems [4]. Community detection is one of the central topics among network inference [5], as real networks often comprise communities that are sparsely connected to each other. Recently, an increasing amount of research has focused on community detection based on observations from dynamical systems (e.g., [6], [7], [8]). However, most studies address the detection problem for linear dynamics, and there is still a need to understand how nonlinearity affects detection performance. Since nonlinear models can behave differently from linear averaging dynamics [9], [10], it is important to investigate how to adapt and apply traditional methods, such as spectral clustering, to complex nonlinear dynamics.

A. Related Work

Community detection has been extensively investigated for over two decades [5]. There are three major approaches to the problem. The most common approach involves optimizing quality functions. The Louvain method [11] is a widely-used fast algorithm based on greedy optimization of the

modularity, which measures the extent a given network partition implying dense connections within subgroups. The Infomap method [12] represents the approach based on dynamics. The method aims to compress random walks over networks, by looking for partitions that reduce information required for describing the trajectories. The approach based on statistical inference has become well-established in recent years. The methods infer generative network models that presuppose community structure. The stochastic block model (SBM) is a crucial example, where edges exist with probability depending on pre-assigned community labels. This framework facilitates theoretical analysis of community detectability [13], [14], [15].

When only data from dynamics over a network are available, rather than direct edge information, a standard method is to construct a network based on state similarity and then apply detection algorithms to that network. The paper [7] introduces a Bayesian hierarchical model for time series, and demonstrates that model-based approaches can outperform traditional methods. Maximum likelihood methods for cascade dynamics are explored in [16], and nonparametric Bayesian methods for epidemics and an Ising model are proposed in [17]. A blind community detection approach is developed in [6], [18], [19]. The method applies spectral clustering to sample covariance matrices derived from a single snapshot from multiple trajectories. The papers [8], [20] propose algorithms leveraging transient and asymptotic behaviors of gossip opinion dynamics.

Most of the above studies focus on community detection for linear dynamics, and the problem for nonlinear dynamics remains unexplored. In this paper, we study a nonlinear opinion dynamics model [10]. In the model, individuals update their opinions according to a saturation interaction rule, which is also found in neural and biological systems [21], [22]. This type of models can capture the transition from opinion consensus to polarization without the presence of external influence, unlike the DeGroot, Friedkin-Johnsen, or Hegselmann-Krause models [23]. In particular, [24] uses the nonlinear model to explain political polarization dynamics. Additionally, [9] proposes a model combining the saturation rule with homophily, and reproduces the echo chamber phenomenon on social media. Therefore, investigating community detection from these nonlinear dynamics can help understand the structure of real-world dynamics.

B. Contributions

We study community detection based on equilibria of a nonlinear opinion dynamics. It is assumed that the dynamics evolve over networks generated from an SBM with two com-

munities. For the case with a single equilibrium available, we propose a community detection algorithm based on k-means (Algorithm 1). The algorithm can detect most community labels with high probability (i.e., achieving almost exact recovery), if a non-consensus equilibrium is used, and the two communities differ in size and link probabilities (Theorem 1 (i.a)). When the communities are identical in size and link probabilities, and inter-community connections are denser than intra-community ones, the algorithm can achieve almost exact recovery if interpersonal influence weights are negative, but fails if the weights are positive (Theorem 1 (i.b) and (ii)). For multiple equilibria with external inputs, another detection algorithm (Algorithm 2) is developed by leveraging fixed point equations and spectral clustering methods. Its performance is validated by numerical experiments.

By studying a typical nonlinear model [25], the results demonstrate how community detectability is affected by nonlinearity. Stationary states under external excitation or transient states are necessary for detection in linear dynamics (e.g., [6], [18], [20]). In contrast, it is found that community detection is possible by using only equilibria of nonlinear dynamics without excitation. However, community structure information may not be preserved in the equilibria when agents are strongly influence by nonlinear interactions. These findings provide key insight into the design of community detection methods for real-world complex dynamics.

C. Outline

Section II introduces the nonlinear dynamics and the SBM. Section III formulates the problem. Two detection algorithms are proposed in Section IV, and numerical experiments presented in Section V. Section VI concludes the paper.

Notation. Denote the set of positive integers by $\mathbb{N}_+$ and the set of positive real numbers by $\mathbb{R}_+$. Denote the n-dimensional all-one vector by $\mathbf{1}_n$. I_n is the identity matrix, and $\mathbf{1}_{m,n}$ $(\mathbf{0}_{m,n})$ is the $m \times n$ all-one (all-zero) matrix. Denote the Euclidean norm of a vector and the spectral norm of a matrix by $\|\cdot\|$. A vector is denoted by a boldface letter, e.g., $\mathbf{x}$, and its i-th entry by x_i. For a matrix $A \in \mathbb{R}^{n \times n}$, a_{ij} or $[A]_{ij}$ denotes its (i,j)-th entry. For a symmetric $A \in \mathbb{R}^{n \times n}$, denote its smallest and largest eigenvalues by $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$. For a vector-valued function $\mathbf{f}(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}^m$, its Jacobian matrix is an $m \times n$ matrix, denoted by $D_{\mathbf{x}}\mathbf{f}$, whose (i,j)-th entry is $[D_{\mathbf{x}}\mathbf{f}]_{ij} = \partial f_i/\partial x_j$. The Jacobian of $\mathbf{f}$ at a point $\mathbf{x}_0 \in \mathbb{R}^n$ is written as $D_{\mathbf{x}}\mathbf{f}(\mathbf{x}_0)$. For real numbers $a(n), b(n) > 0$, $n \in \mathbb{N}$, denote $a(n) = \Theta(b(n))$, if there exist $C_1, C_2 > 0$ such that $C_1 b(n) \leq a(n) \leq C_2 b(n)$. The function $\mathbb{I}_{[\text{property}]}$ is the indicator function, which is one if the property in the bracket holds, and zero otherwise. A sequence of events $\{A_n\}$ happens with high probability (w.h.p.) if $\lim_{n \to \infty} \mathbb{P}\{A_n\} = 1$. An undirected graph is denoted by $\mathcal{G} = (\mathcal{V}, \mathcal{E}, A)$, where $\mathcal{V}$ is the agent set, $\mathcal{E}$ is the edge set, and $A = [a_{ij}]$ is the adjacency matrix such that $a_{ij} = 1$ $(a_{ij} = 0)$ if $\{i,j\} \in \mathcal{E}$ $(\{i,j\} \notin \mathcal{E})$.

II. PRELIMINARIES

In this section, we introduce the nonlinear opinion dynamics and the SBM, and briefly discuss their properties.

A. Nonlinear Opinion Dynamics

The nonlinear opinion dynamics model takes place over an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, A)$ with $\mathcal{V} = \{1, \ldots, n\}$ and no self-loops ($a_{ii} = 0$). Each agent $i \in \mathcal{V}$ has a state $x_i(t)$, $t \in \mathbb{R}_+$, and the model evolves in continuous time according to the update rule

$$\dot{x}_i = -dx_i + uS\Big(\alpha x_i + \gamma \sum_{k \in \mathcal{V}} a_{ik} x_k \Big) + b_i, \qquad (1)$$

where $d > 0$ is the damping coefficient, u is the agent attention parameter to the nonlinear network interaction, and S is an odd saturating function satisfying $S(0) = 0$, $S'(0) = 1$, and $\text{sgn}(S''(z)) = -\text{sgn}(z)$. Here we assume $S = \tanh$, i.e., the hyperbolic tangent. In the nonlinear term, $\alpha \geq 0$ is the self weight and $\gamma \in \mathbb{R}$ is the influence weight of other agents. Lastly, $b_i \in \mathbb{R}$ is an additive input, which can be seen as individual prejudice or external influence.

The compact form of (1) can be written as

$$\dot{\mathbf{x}} = -d\mathbf{x} + u\mathbf{S}(\alpha\mathbf{x} + \gamma A\mathbf{x}) + \mathbf{b}, \qquad (2)$$

where $\mathbf{S}(\mathbf{x}) := [S(x_1), \ldots, S(x_n)]^{\text{T}}$ for $\mathbf{x} \in \mathbb{R}^n$.

The model and its extensions have been thoroughly studied, focusing on their bifurcation and steady-state behavior [10], [25], [26]. Here we consider the case where the parameters are homogeneous (i.e., d, u, α, and γ are identical for all agents) and each agent has a single opinion. The following result [26, Theorem 1] demonstrates that two new equilibria of the model (2) without inputs emerge, due to bifurcation from the origin as the attention parameter u increases beyond specific thresholds.

Proposition 1: Suppose that $\mathcal{G}$ is connected, $u \geq 0$, and $\mathbf{b} = \mathbf{0}$.

(i) If $\gamma > 0$, the origin $\mathbf{x} = \mathbf{0}$ is a locally exponentially stable equilibrium for $0 < u < u_1$ and unstable for $u > u_1$, where $u_1 := d/(\alpha + \gamma\lambda_{\max}(A))$. At $u = u_1$, branches of equilibria $\tilde{\mathbf{x}} \neq \mathbf{0}$ emerge in a steady-state bifurcation off of $\mathbf{x} = \mathbf{0}$ along the eigenspace corresponding to $\lambda_{\max}(A)$, where the entries of $\tilde{x}$ have the same sign.

(ii) If $\gamma < 0$, the origin $\mathbf{x} = \mathbf{0}$ is a locally exponentially stable equilibrium for $0 < u < u_2$ and unstable for $u > u_2$, where $u_2 := d/(\alpha + \gamma\lambda_{\min}(A))$. At $u = u_2$, branches of equilibria $\tilde{\mathbf{x}} \neq \mathbf{0}$ emerge in a steady-state bifurcation off of $\mathbf{x} = \mathbf{0}$ along the eigenspace corresponding to $\lambda_{\min}(A)$, where the entries of $\tilde{\mathbf{x}}$ have different signs. ∎

The real-valued opinion x_i represents the agent i's level of support for two options. The sign $\text{sgn}(x_i)$ indicates which of the two options the agent supports, and $x_i = 0$ represents a neutral position. Proposition 1 shows that agreement steady states with all agents having the same sign emerge from the neutral state as u increases, if the influence weight is positive.

In contrast, if the influence is negative, disagreement steady states emerge.

B. Stochastic Block Model

Assume that the agent set $\mathcal{V}$ consists of two disjoint communities $\mathcal{V}_1$ and $\mathcal{V}_2$. Let the community structure vector be $\mathcal{C} \in \{1,2\}^n$, satisfying that $\mathcal{C}_i = 1$ if $i \in \mathcal{V}_1$ and $\mathcal{C}_i = 2$ if $i \in \mathcal{V}_2$ (i.e., agents in $\mathcal{V}_1$ (in $\mathcal{V}_2$) have the label 1 (label 2)). The two-community SBM is defined as follows.

Definition 1 (SBM): Let $n \in \mathbb{N}_+$ be the network size, $\mathbf{n} = [n_1 \ n_2]^T \in \mathbb{N}_+^2$ be the community size vector with $n_1 + n_2 = n$, and

$$\boldsymbol{\ell} = \begin{bmatrix} \ell_{11} & \ell_{12} \\ \ell_{21} & \ell_{22} \end{bmatrix} \in [0,1]^{2 \times 2}$$

be the link probability matrix with $\ell_{12} = \ell_{21}$. In $\mathrm{SBM}(\mathbf{n}, \boldsymbol{\ell})$, agents $1, \ldots, n_1$ are assigned with community label 1 and agents $n_1 + 1, \ldots, n$ with label 2. Then the SBM generates an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, A)$ without self-loops, by independently adding $\{i, j\}$ with $i \neq j$ to $\mathcal{E}$ with probability $\ell_{\mathcal{C}_i, \mathcal{C}_j}$. If $n_1 = n_2 = n/2$ and $\ell_{11} = \ell_{22}$, the SBM is called symmetric SBM, denoted by $\mathrm{SSBM}(n, \boldsymbol{\ell})$. In this case, denote $\ell_s := \ell_{11} = \ell_{22}$ and $\ell_d := \ell_{12} = \ell_{21}$. ∎

The following assumption on link probability ℓ_{ij} of an SBM is given to ensure the random graph is connected w.h.p. A technical assumption for the SSBM is also introduced.

Assumption 1: For the $\mathrm{SBM}(\mathbf{n}, \boldsymbol{\ell})$, assume that $\ell_{ij} = \omega(\log n / n)$, $\forall i, j \in \{1, 2\}$. If the SBM is $\mathrm{SSBM}(n, \boldsymbol{\ell})$ and $\ell_s > \ell_d$, further assume that $\ell_d = \omega(\sqrt{\ell_s \log n})$.

To measure the performance of an algorithm detecting communities of the SBM from observations, we introduce the accuracy of an estimate $\hat{\mathcal{C}}$ given by the algorithm,

$$\mathrm{Acc}(\mathcal{C}, \hat{\mathcal{C}}) := \frac{1}{n} \max \left\{ \sum_{i=1}^n \mathbb{I}_{[\mathcal{C}_i = \hat{\mathcal{C}}_i]}, \sum_{i=1}^n \mathbb{I}_{[\mathcal{C}_i = 3 - \hat{\mathcal{C}}_i]} \right\}. \quad (3)$$

Now we define almost exact recovery of an algorithm detecting communities in SBMs as follows [13], [15].

Definition 2: For an SBM with n agents and a community structure $\mathcal{C}$, suppose that a detection algorithm outputs an estimate $\hat{\mathcal{C}}$. The algorithm achieves almost exact recovery, if $\mathbb{P}\{\mathrm{Acc}(\mathcal{C}, \hat{\mathcal{C}}) = 1 - o(1)\} = 1 - o(1)$. ∎

Almost exact recovery means that the algorithm can correctly detect most community labels w.h.p.

III. PROBLEM FORMULATION

We investigate how to detect communities from equilibria of the model (2). It is assumed that a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, A)$ is generated from an SBM and then fixed. The opinion model (2) evolves over this graph and reaches steady state. Since we will consider cases where the network size is large, we set $\gamma = \pm 1/\Delta$, where $\Delta := \max_i \mathbb{E}\{\sum_j a_{ij}\}$ is the maximum expected degree of $\mathcal{G}$. Hence, $\gamma A \mathbf{x}$ is a weighted average of the opinions $\mathbf{x}$. The problem studied in this paper is described as follows.

Problem. Given a single equilibrium $\mathbf{x}^*$ or multiple input-equilibrium pairs $\{[\mathbf{b}^{(1)} \ \mathbf{x}^{(1)}], \ldots, [\mathbf{b}^{(m)} \ \mathbf{x}^{(m)}]\}$ of the model (2) over an SBM, design algorithms to detect the communities of the SBM and analyze their performance.

Algorithm 1 (Detection Based on Single Equilibrium)

Input: Community number $k = 2$.
Output: Community estimate $\hat{\mathcal{C}}$.
1: Obtain an equilibrium $\mathbf{x}^*$ of the model (2).
2: Apply k-means to $\mathbf{x}^*$ to get an estimate of the community structure $\mathcal{C}$.

IV. DETECTION ALGORITHMS AND MAIN RESULTS

In this section, we first address the community detection problem based on a single equilibrium. The performance of the proposed algorithm is analyzed theoretically for several SBMs (Theorem 1). We then design a detection algorithm based on multiple equilibria, by approximating the adjacency matrix using fixed point equations of the system and applying spectral clustering techniques to the estimated matrix.

A. Detection from Single Equilibrium

Given only one equilibrium, we employ the k-means method to cluster the states, as shown in Algorithm 1. It is assumed that the exact equilibrium is obtained. We will study detection from noisy observations in the future.

The neutral steady state $\mathbf{x} = \mathbf{0}$ provides no information about the community structure, similar to the consensus of linear dynamics such as the DeGroot model. However, the equilibria emerging at the bifurcation from the origin as stated in Proposition 1 can reveal information of the network, as shown in the following theorem.

Theorem 1: Suppose that Assumption 1 holds, $\mathbf{b} = \mathbf{0}$, and the equilibrium $\mathbf{x}^*$ emerges at the bifurcation from the origin as stated in Proposition 1.

(i) Assume that $\gamma = 1/\Delta > 0$, and $u - \bar{u}_1 > 0$ is small enough, where $\bar{u}_1 = d/(\alpha + \gamma \lambda_{\max}(\mathbb{E}\{A\}))$.

(i.a) For $\mathrm{SBM}(\mathbf{n}, \boldsymbol{\ell})$, if $n_2 = o(n_1)$, $\ell_{11} n_1 = \Theta(\ell_{22} n_2)$, and $\ell_{12} = \Theta(\sqrt{\ell_{11} \ell_{22}})$, then Algorithm 1 achieves almost exact recovery.

(i.b) For $\mathrm{SSBM}(n, \boldsymbol{\ell})$, it holds that $\|\mathbf{x}^* - \theta c(u) \mathbf{1}_n / \sqrt{n}\| = o(c(u))$ w.h.p., for some $\theta \in \{1, -1\}$. Here $c(u) \in \mathbb{R}$ depends on u and $c(u) \to 0$ as $u \to \bar{u}_1$. If $\liminf \ell_d / \ell_s > 3$, then Algorithm 1 cannot achieve almost exact recovery.

(ii) Assume that $\gamma = -1/\Delta < 0$, and $u - \bar{u}_2 > 0$ is small enough, where $\bar{u}_2 = d/(\alpha + \gamma \lambda_{\min}(\mathbb{E}\{A\}))$. For $\mathrm{SSBM}(n, \boldsymbol{\ell})$, if $\ell_s < \ell_d$, then Algorithm 1 achieves almost exact recovery. ∎

Remark 1: The result (i.a) states that the communities of the SBM can be detected, if the influence weights are positive and the two communities differ in size and link probabilities. The result (i.b) shows that the entries of the equilibrium $\mathbf{x}^*$ are close to each other in the SSBM. As a result, the communities cannot be detected by clustering. However, almost exact recovery in the disassortative SSBM (inter-community connections are denser than intra-community) can be achieved, if the influence weight is negative.

From Proposition 1, the equilibrium $\mathbf{x}^*$ reflects the eigenvector centrality if the influence weight is positive. Note that agents in the SSBM have similar centrality. Conditions

in (i.a) imply a leader-follower structure, where the community $\mathcal{V}_2$ has much fewer agents but denser connections, compared with the community $\mathcal{V}_1$. This structure ensures different centrality in communities, making detection possible. In the case of negative influence, the equilibrium $\mathbf{x}^*$ is close to the eigenvector corresponding to $\lambda_{\min}(A)$. For the disassortative SSBM, the sign of entries of that eigenvector corresponds to community labels, so almost exact recovery can be achieved.

In the theorem, we assume that u is close to the bifurcation threshold $\bar{u}_1$ or $\bar{u}_2$. To ensure detectability, u has to be neither too small or too large. When u is below the bifurcation threshold, only the neutrality equilibrium exists, providing no information about the graph. As u increases, agents become more attentive to nonlinear interactions, and equilibria move away from the eigenvector centrality. See Section V for how u and different sigmoid functions influence detection. ∎

Proof Sketch: We briefly explain how the results are obtained. The detailed proof is given in the extended version [27]. First, leveraging the Lyapunov-Schmidt reduction, we can calculate an explicit form of the equilibrium. The equilibrium is close to (up to a sign flip) the eigenvector corresponding to the largest or smallest eigenvalue of the adjacency matrix A. Next, from matrix perturbation theory and concentration inequalities, it follows that the aforementioned eigenvector is close to the eigenvector of the expected graph. Analyzing the properties of the expected graph and eigenvectors completes the proof. ∎

We conclude this subsection by discussing several extensions of Theorem 1. The result in (i.a) holds as long as agents in different communities have different eigenvector centrality. The disassortative condition for ℓ_d in (i.b) is technical, and simulation given in Section V shows that Algorithm 1 also fails to achieve almost exact recovery for the assortative SBM. The theorem only considers all-positive ($\gamma > 0$) or all-negative ($\gamma < 0$) relationships. A natural extension is that agents within the same community have positive edges, whereas agents between the two communities have negative edges. In this signed-graph case, equilibria of different communities have opposite signs (see e.g., [28, Theorem 1]), so communities are detectable.

B. Detection from Multiple Equilibria

In this subsection, we assume that the system has external inputs, and the resulting equilibria from multiple trajectories are available. Consider the dataset of input-equilibrium pairs $\{[\mathbf{b}^{(1)} \ \mathbf{x}^{(1)}], \dots, [\mathbf{b}^{(m)} \ \mathbf{x}^{(m)}]\}$ with $m \in \mathbb{N}_+$. We will design a community detection algorithm based on fixed point equations of the model.

The multiple trajectory case is a commonly studied scenario, in which different discussions are observed (e.g., [2], [6], [18]). The inputs can be seen as individual prejudice towards different topics or external information that is given for each discussion.

Note that an input-equilibrium pair $[\mathbf{b}^{(k)} \ \mathbf{x}^{(k)}]$, $1 \le k \le m$, satisfies the following fixed point equation

$$\mathbf{0} = -d\mathbf{x}^{(k)} + u\mathbf{S}(\alpha\mathbf{x}^{(k)} + \gamma A\mathbf{x}^{(k)}) + \mathbf{b}^{(k)},$$

Algorithm 2 (Detection Based on Multiple Equilibria)

Input: Multiple input-equilibrium pairs $\{[\mathbf{b}^{(1)} \ \mathbf{x}^{(1)}], \dots, [\mathbf{b}^{(m)} \ \mathbf{x}^{(m)}]\}$ with $m \in \mathbb{N}_+$, model parameters α, γ, d, u, and community number $k = 2$.

Output: Community estimate $\hat{\mathcal{C}}$.

1: Calculate for $1 \le k \le m$

$$\mathbf{y}^{(k)} = \frac{1}{\gamma}\Big(\mathbf{S}^{-1}\Big(\frac{1}{u}(d\mathbf{x}^{(k)} - \mathbf{b}^{(k)})\Big) - \alpha\mathbf{x}^{(k)}\Big).$$

2: Calculate the estimate $\hat{A}$ of the adjacency matrix A:

$$\tilde{A} = \mathbf{Y}\mathbf{X}^\dagger,$$
$$\hat{A} = (\tilde{A} + \tilde{A}^\mathrm{T})/2,$$

where

$$\mathbf{X} = [\mathbf{x}^{(1)} \ \cdots \ \mathbf{x}^{(m)}],$$
$$\mathbf{Y} = [\mathbf{y}^{(1)} \ \cdots \ \mathbf{y}^{(m)}].$$

3: Find the eigenvector corresponding to the second largest eigenvalue of $\hat{A}$, denoted by $\hat{\mathbf{v}}$. Apply k-means to $\hat{\mathbf{v}}$ and obtain an estimate of community labels $\hat{\mathcal{C}}$.

---

which implies that

$$\mathbf{y}^{(k)} := \frac{1}{\gamma}\Big(\mathbf{S}^{-1}\Big(\frac{1}{u}(d\mathbf{x}^{(k)} - \mathbf{b}^{(k)})\Big) - \alpha\mathbf{x}^{(k)}\Big) = A\mathbf{x}^{(k)},$$

where $\mathbf{S}^{-1}(\mathbf{x}) := [S^{-1}(x_1), \dots, S^{-1}(x_n)]^\mathrm{T}$ for $\mathbf{x} \in \mathbb{R}^n$. Hence, $\mathbf{Y} := [\mathbf{y}^{(1)} \ \cdots \ \mathbf{y}^{(m)}] = A[\mathbf{x}^{(1)} \ \cdots \ \mathbf{x}^{(m)}]$. It is expected that $\mathbf{X} := [\mathbf{x}^{(1)} \ \cdots \ \mathbf{x}^{(m)}]$ is invertible when m is large and $\{\mathbf{b}^{(k)}\}$ sufficiently excites the system. Then the estimate of the adjacency matrix A can be given by $\mathbf{Y}\mathbf{X}^{-1}$. When the number of samples is much less than the network size, $\hat{A} = \mathbf{Y}\mathbf{X}^\dagger$ gives an approximation of the matrix A, where $\mathbf{X}^\dagger$ is the pseudo-inverse of $\mathbf{X}$.

Inspired by this observation, we propose Algorithm 2 which utilizes spectral clustering techniques to recover the communities for the SSBM. Line 1 of the algorithm calculates the data matrix $\mathbf{Y}$. In Line 2, $\mathbf{Y}\mathbf{X}^\dagger$ is calculated and then projected to the set of symmetric matrices, since A is symmetric. For the adjacency matrix of an SSBM, its expectation has a block structure, and the eigenvector corresponding to its second largest eigenvalue satisfies that the entries in different communities have different signs. Leveraging this property, Line 3 applies the spectral clustering method. Knowing the model and parameters is a strong assumption, and future work will explore how to address this limitation.

As shown in Theorem 1 (i.b), the SSBM may not be recovered by using a single equilibrium without external inputs. However, when multiple input-equilibrium pairs are available, the recovery is possible, as shown in Section V.

V. Numerical Experiments

In this section, we conduct numerical experiments for community detection in the model (2). For all experiments, we set the damping coefficient $d = 1$, the agent self weight
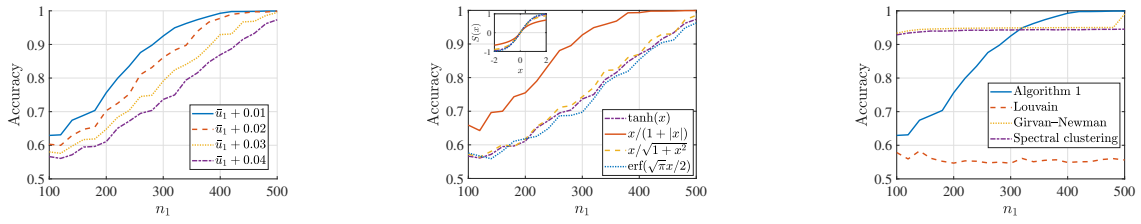
(a) Accuracy of Algorithm 1 increases with n but decreases with u.

(b) Accuracy of Algorithm 1 for the model (2) with $S(x)$ being different nonlinear functions, where $u = \bar{u}_1 + 0.04$.

(c) Accuracy of Algorithm 1 with $u = \bar{u}_1 + 0.01$, the Louvain, the Girvan-Newman, and spectral clustering methods.

Fig. 1. Detection accuracy for an SBM with unequal-sized communities and $\gamma > 0$.

$\alpha = 1$, the positive influence weight $\gamma = 1/\Delta$, and the negative weight $\gamma = -1/\Delta$, where Δ is the maximum expected degree of the SBM. The equilibria are obtained using the `ode45` solver in MATLAB. The results are consistent with other solvers such as `ode78`, `ode89`, and `ode15s`.

We first study the SBM that has two communities different in size and link probabilities, as in Theorem 1 (i.a). In this experiment, we consider positive interpersonal influence ($\gamma > 0$), and calculate the averaged accuracy of the algorithm based on 50 random graph samples for the network size n_1 from 100 to 500 and $n_2 = 0.05n_1$. The link probabilities are set to be $\ell_{11} = 0.05$, $\ell_{12} = 0.1$, and $\ell_{22} = 0.5$. Additionally, we set the attention parameter $u = \bar{u}_1 + 0.01, \ldots, \bar{u}_1 + 0.04$, where $\bar{u}_1$ is given in Theorem 1. Fig. 1(a) shows that the averaged accuracy of Algorithm 1 increases with the network size. Additionally, as u grows, the performance becomes worse, indicating the impact of nonlinearity on the structure of the equilibria. To study the influence of the nonlinear interaction on detectability, we examine detection performance for the model (2) with the saturating function $S(x) = x/(1+|x|)$, $x/\sqrt{1+x^2}$, $\tanh(x)$, and $\text{erf}(\sqrt{\pi}x/2)$, where erf is the Gauss error function. Fig. 1(b) illustrates that the detection performance worsens as agents become more easily saturated. Next, we compare Algorithm 1 with classic methods which are applied directly to the adjacency matrices. As shown in Fig. 1(c), the Girvan–Newman method (where the partition with two communities is selected) and the spectral clustering have high accuracy because the SBM has two blocks. For the Louvain method, we set its resolution parameter to produce a partition with two communities. It has difficulty identifying the smaller community, and hence does not perform well. Utilizing only state observations instead of graphs, Algorithm 1 can achieve high accuracy when n is large enough while $u - \bar{u}_1$ is not too large.

For SSBM with positive weights, Algorithm 1 cannot detect the communities w.h.p. This is shown in Fig. 2(a), where the averaged accuracy of the algorithm is close to 0.5, similar to random guess. This validates Theorem 1 (i.b). Here the attention parameter is chosen as above. The link probabilities are $\ell_s = 0.3$ and $\ell_d = 0.05$.

We then consider the disassortative SSBM with negative influence weights, to validate Theorem 1 (ii). In the experiment, we set $u = \bar{u}_2 + 0.01, \ldots, \bar{u}_2 + 0.04$, and link probabilities to be $\ell_s = 0.005$ and $\ell_d = 0.03$. The

link probabilities here are much smaller than previous ones, but the algorithm still has high detection accuracy and the attention parameter has less impact (see Fig. 2(b)).

Finally, we study the performance of Algorithm 2 based on multiple input-equilibrium pairs $\{[\mathbf{b}^{(1)} \; \mathbf{x}^{(1)}], \ldots, [\mathbf{b}^{(m)} \; \mathbf{x}^{(m)}]\}$. We consider positive influence weights and the same SSBM as the second experiment. The network size n is set to be $20, 60$, and 100, and u to be $\bar{u}_1 + 0.01$. For each network, we generate 10 graph samples, and for each graph sample we collect 10 sets of input-equilibrium pairs where $\mathbf{b}^{(k)}$ are independently generated from the standard Gaussian distribution. To test the effect of sample size, we set $m = 0.1n, 0.2n, \ldots, n$. Fig. 3(a) shows that the performance of the algorithm increases with both sample and network sizes. We then compare the proposed algorithm with classic methods for community detection from dynamical observations: the spectral clustering applied to the sample covariance matrix $\frac{1}{m}\sum_{k=1}^{m}(\mathbf{x}^{(k)} - \bar{\mathbf{x}})(\mathbf{x}^{(k)} - \bar{\mathbf{x}})^\mathrm{T}$ (e.g., [6], [18]), the Girvan-Newman and the Louvain algorithms applied to correlation matrices [29]. The accuracy of these methods is below 0.6, as shown in Fig. 3(b). Algorithm 2 performs better because it exploits the nonlinear structure and utilizes model parameter information.

VI. CONCLUSION

We studied community detection for a nonlinear opinion dynamics model over a stochastic block model. Two algorithms based on a single or multiple equilibria were investigated. Future work includes studying multi-dimensional generalizations [30], labeled SBMs [13], and joint learning of communities and model parameters.

REFERENCES

[1] R. Ramakrishna, H.-T. Wai, and A. Scaglione, "A user guide to low-pass graph signal processing and its applications: Tools and applications," *IEEE Signal Processing Magazine*, vol. 37, no. 6, pp. 74–85, 2020.

[2] C. Ravazzi, F. Dabbene, C. Lagoa, and A. V. Proskurnikov, "Learning hidden influences in large-scale dynamical social networks: A data-driven sparsity-based approach, in memory of Roberto Tempo," *IEEE Control Systems Magazine*, vol. 41, no. 5, pp. 61–103, 2021.

[3] D. Kempe, J. Kleinberg, and É. Tardos, "Maximizing the spread of influence through a social network," in *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 137–146, 2003.

[4] R. Burke, "Hybrid recommender systems: Survey and experiments," *User Modeling and User-Adapted Interaction*, vol. 12, pp. 331–370, 2002.
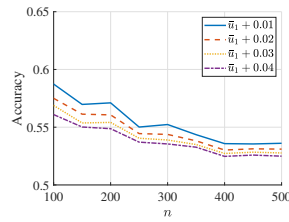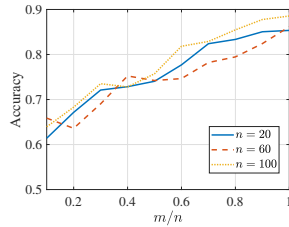
(a) Positive interpersonal influence. The accuracy is similar to random guess (close to 0.5).
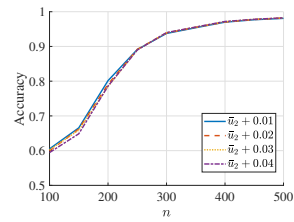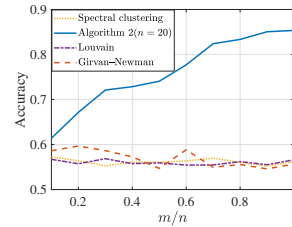

(b) The dissassortative SSBM with negative interpersonal influence. The accuracy increases with n.

Fig. 2.   Detection accuracy for the SSBM.

(a) The accuracy of Algorithm 2 increases with the sample size m and the network size n.

(b) The accuracy of Algorithm 2 with $n = 20$, the Louvain, the Girvan–Newman, and spectral clustering methods.

Fig. 3.   Detection accuracy from multiple input-equilibrium pairs.

[5] S. Fortunato and D. Hric, "Community detection in networks: A user guide," *Physics Reports*, vol. 659, pp. 1–44, 2016.

[6] H.-T. Wai, S. Segarra, A. E. Ozdaglar, A. Scaglione, and A. Jadbabaie, "Blind community detection from low-rank excitations of a graph filter," *IEEE Transactions on Signal Processing*, vol. 68, pp. 436–451, 2019.

[7] T. Hoffmann, L. Peel, R. Lambiotte, and N. S. Jones, "Community detection in networks without observing edges," *Science Advances*, vol. 6, no. 4, p. eaav1478, 2020.

[8] Y. Xing, X. He, H. Fang, and K. H. Johansson, "Community structure recovery and interaction probability estimation for gossip opinion dynamics," *Automatica*, vol. 154, p. 111105, 2023.

[9] F. Baumann, P. Lorenz-Spreen, I. M. Sokolov, and M. Starnini, "Modeling echo chambers and polarization dynamics in social networks," *Physical Review Letters*, vol. 124, no. 4, p. 048301, 2020.

[10] A. Bizyaeva, A. Franci, and N. E. Leonard, "Nonlinear opinion dynamics with tunable sensitivity," *IEEE Transactions on Automatic Control*, vol. 68, no. 3, pp. 1415–1430, 2022.

[11] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2008, no. 10, p. P10008, 2008.

[12] M. Rosvall and C. T. Bergstrom, "Maps of random walks on complex networks reveal community structure," *Proceedings of the National Academy of Sciences*, vol. 105, no. 4, pp. 1118–1123, 2008.

[13] S.-Y. Yun and A. Proutiere, "Community detection via random and adaptive sampling," in *Conference on Learning Theory*, pp. 138–175, 2014.

[14] L. Massoulié, "Community detection thresholds and the weak ramanujan property," in *Proceedings of the Forty-Sixth Annual ACM Symposium on Theory of Computing*, pp. 694–703, 2014.

[15] E. Abbe, "Community detection and stochastic block models: Recent developments," *The Journal of Machine Learning Research*, vol. 18, no. 1, pp. 6446–6531, 2017.

[16] L. Prokhorenkova, A. Tikhonov, and N. Litvak, "When less is more: Systematic analysis of cascade-based community detection," *ACM Transactions on Knowledge Discovery from Data*, vol. 16, no. 4, pp. 1–22, 2022.

[17] T. P. Peixoto, "Network reconstruction and community detection from dynamics," *Physical Review Letters*, vol. 123, no. 12, p. 128301, 2019.

[18] M. T. Schaub, S. Segarra, and J. N. Tsitsiklis, "Blind identification of stochastic block models from dynamical observations," *SIAM Journal on Mathematics of Data Science*, vol. 2, no. 2, pp. 335–367, 2020.

[19] H.-T. Wai, Y. C. Eldar, A. E. Ozdaglar, and A. Scaglione, "Community inference from partially observed graph signals: Algorithms and analysis," *IEEE Transactions on Signal Processing*, vol. 70, pp. 2136–2151, 2022.

[20] Y. Xing and K. H. Johansson, "Almost exact recovery in gossip opinion dynamics over stochastic block models," in *IEEE Conference on Decision and Control*, pp. 2421–2426, 2023.

[21] E. Eldar, J. D. Cohen, and Y. Niv, "The effects of neural gain on attention and learning," *Nature Neuroscience*, vol. 16, no. 8, pp. 1146–1153, 2013.

[22] R. Bogacz, M. Usher, J. Zhang, and J. L. McClelland, "Extending a biologically inspired model of choice: Multi-alternatives, nonlinearity and value-based multidimensional choice," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 362, no. 1485, pp. 1655–1670, 2007.

[23] A. V. Proskurnikov and R. Tempo, "A tutorial on modeling and analysis of dynamic social networks. Part I," *Annual Reviews in Control*, vol. 43, pp. 65–79, 2017.

[24] N. E. Leonard, K. Lipsitz, A. Bizyaeva, A. Franci, and Y. Lelkes, "The nonlinear feedback dynamics of asymmetric political polarization," *Proceedings of the National Academy of Sciences*, vol. 118, no. 50, p. e2102149118, 2021.

[25] A. Bizyaeva, T. Sorochkin, A. Franci, and N. E. Leonard, "Control of agreement and disagreement cascades with distributed inputs," in *IEEE Conference on Decision and Control*, pp. 4994–4999, 2021.

[26] A. Bizyaeva, A. Matthews, A. Franci, and N. E. Leonard, "Patterns of nonlinear opinion formation on networks," in *American Control Conference*, pp. 2739–2744, 2021.

[27] Y. Xing, A. Bizyaeva, and K. H. Johansson, "Learning communities from equilibria of nonlinear opinion dynamics," *arXiv preprint arXiv:2409.08004*, 2024.

[28] A. Bizyaeva, G. Amorim, M. Santos, A. Franci, and N. E. Leonard, "Switching transformations for decentralized control of opinion patterns in signed networks: Application to dynamic task allocation," *IEEE Control Systems Letters*, vol. 6, pp. 3463–3468, 2022.

[29] D. J. Fenn, M. A. Porter, P. J. Mucha, M. McDonald, S. Williams, N. F. Johnson, and N. S. Jones, "Dynamical clustering of exchange rates," *Quantitative Finance*, vol. 12, no. 10, pp. 1493–1520, 2012.

[30] A. Bizyaeva, A. Franci, and N. E. Leonard, "Multi-topic belief formation through bifurcations over signed social networks," *arXiv preprint arXiv:2308.02755*, 2023.