9th North American Conference on Industrial Engineering and Operations Management
Washington D.C., USA, June 4-6, 2024
Publisher: IEOM Society International, USA DOI: 10.46254/NA09.20240263
Published: June 4, 2024

Small Data and Sparse Bayesian Learning for Sequential Inference of Network Connectivity

Jinming Wan, Jun Kataoka, Hiroki Sayama and Changqing Cheng
Department of Systems Science and Industrial Engineering
Binghamton University
Binghamton, NY 13902, USA
jwan8@binghamton.edu, jkataok1@binghamton.edu, sayama@binghamton.edu, ccheng@binghamton.edu

Jayanth Sivakumar
Citibank, N.A. company
New York, NY 10013, USA
jsivaku1@binghamton.edu

Eric Peña
Motional company
Boston, MA 02210, USA
epena4@binghamton.edu

Yiming Che
School of Computing and Augmented Intelligence
Arizona State University
Tempe, AZ 85281, USA
yche14@ase.edu

Abstract

While significant efforts have been attempted in the design, control, and optimization of complex networks, most existing works assume the network structure is known or readily available. However, the network topology can be radically recast after an adversarial attack and may remain unknown for subsequent analysis. In this work, we propose a novel Bayesian sequential learning approach to reconstruct network connectivity adaptively: A sparse Spike and Slab prior is placed on connectivity for all edges, and the connectivity learned from reconstructed nodes will be used to select the next node and update the prior knowledge. Central to our approach is that most realistic networks are sparse, in that the connectivity degree of each node is much smaller compared to the number of nodes in the network. Sequential selection of the most informative nodes is realized via the between-node expected improvement. We corroborate this sequential Bayesian approach in connectivity recovery for the IEEE-118 power grid system. Results indicate that only a fraction (~50%) of the nodes need to be interrogated to reveal the network topology.

Keywords
Network Reconstruction, Inverse Problem, Network Inference, Spike and Slab, Sequential Node Selection.

1. Introduction

The past decades have witnessed the expanding complexity of interconnected engineering systems to accomplish sophisticated design functions. For instance, manufacturing systems are becoming more complicated in the context of globalized operations and the infiltration of renewable energy sources has compounded the control of power grid systems. As those systems are increasingly interconnected, the interdependency in conjunction with the bewildering complexity has brought network science into the spotlight. Yet, the vast majority of network research is focused on the forward problem: given the network topology and interaction between constituent components, what emergent behaviors will the system exhibit? What is the shortest path to traverse the network? However, direct access to network structures remains elusive, leaving us with only a restricted set of observable data (Hempel et al., 2011). For example, in contingencies of malicious attack, natural disaster, or human misoperation, connectivity of significant portions of the power grid or communication networks remains unknown during unexpected events, hindering rescue efforts. Thus, network connectivity could be seriously altered in such extreme events, underscoring the imperative for reconstruction approaches to unveil the intricate network structures with the wide availability of big data. The task of network reconstruction is inherently formidable because the structural information is obscured within the measurable data in an enigmatic fashion. Moreover, the solution space encompassing all conceivable structural configurations is characterized by an exceedingly high dimensionality, as in most inverse problem settings.

This study aims to accurately infer network topology and connectivity from a paucity of sensing data in a timely and efficient fashion. This inverse problem plays a quintessential role in anomaly detection, root cause diagnosis, and timely deployment of corrective actions. It is noted that most realistic networks possess sparse connectivity, in that each node is only connected to a small subset of the nodes. This natural sparsity transforms the connectivity reconstruction into a sparse representation problem. Accordingly, sparse learning or compressive sensing techniques developed in the machine learning community have been attempted in network reconstruction using small samples of measurement data (Basiri et al., 2018; Han et al., 2015; Napoletani and Sauer, 2008; Timme and Casadiego, 2014; Zhu and Abur, 2006). However, these methods reconstruct nodal connectivity iteratively without considering redundant or conflicting information, hindering their applicability to large-scale networks.

In this study, a novel sequential Bayesian analysis framework is developed to reconstruct network connectivity from short time series data recorded at the node level. The connectivity recovered from one node serves as prior information for unexplored nodes, updated through a Bayesian framework to address conflicting information, especially for 0-1 connectivity in unweighted networks (Han et al., 2015). Further, leveraging the redundant information, a sequential sampling scheme based on expected improvement is adopted to adaptively select the most informative node for recovery and maximally reduce the uncertainty in network inference. This allows efficient reconstruction with only a subset of pivotal nodes and has the potential to scale to massive networks.

2. Literature Review

Mathematically, network topology inference is a high-dimensional inverse problem. Most attempts assume prior knowledge of network structure and emphasize concerns regarding directed or undirected interactions, the existence of a link, and the strength of interactions along with their temporal and spatial scales. For instance, Timme and Casadiego (2014) revealed the interaction topology of a network from the collective dynamics of its constituent nodes driven by a given external force. Yu et al. (2006) developed a network copy whose topology can be continuously updated to mimic the dynamical behaviors of the target network. Similarly, Ren et al. (2010) found that noise in dynamical correlation between nodal dynamics leads to more accurate topology estimation, as the interaction between nodes vanishes during synchronization, and the noise brings an additional dimension to study the network topology.

A thorough review of data-driven reconstruction of complex networks and dynamical processes is provided in (Wang et al., 2016). Nevertheless, a “small data” issue is also omnipresent in the scenario of rare or extreme events (Shamsan et al., 2020; Wu et al., 2020). For instance, the power grid network connectivity is vulnerable to perturbations owing to cascading failures. To rapidly roll out the rescue measures and dispatch frontline staff, only a short duration of phasor data can be collected (Che et al., 2023). This has posed a tremendous challenge to the accuracy of network topology inference, particularly considering the uncertainty and ambient noise associated with the data recording. Indeed, standard measures (e.g., information theory and correlation) may not decipher the network interdependency with extremely short time series.

In contrast, a recent line of research exploits compressive sensing on the premise that realistic complex networks are sparse. In other words, the connectivity degree of each node is considerably low compared to the total number of nodes in the network. Thus, the connectivity vector to be reconstructed is sparse with only a few non-zero entities. Correspondingly, the observational or data collection requirements can be relaxed, and compressive sensing is powerful to reconstruct a sparse signal from small data. For instance, capitalizing on compressive sensing theory, Shen et al. (2014) reconstructed the epidemic spreading networks with highly stochastic dynamics from binary time series (e.g., infected or not). Likewise, the complex network topology is reconstructed under game interactions from short sensing data in discrete time (Wang et al., 2011). In the same vein, the sparse learning approach was adopted in the reconstruction of power grid networks with node-level sensing data (Basiri et al., 2018). The reduction in measurement of nodal variables is achieved through compressed sensing that makes use of structural properties of the power grid network.

Improvements are possible when entering the Bayesian domain. In (Huang et al., 2019), compressed sensing and the Bayesian approach registered robust and accurate results in network reconstruction from potentially incomplete and noisy data. Huang et al. (2020) incorporated a hierarchical prior model in Bayesian learning for network reconstruction on evolutionary game data, and the learning parameters were updated iteratively as the reconstruction of nodes progresses. It is noteworthy that the Bayesian approach is appealing in connectivity inference, which offers a distribution as opposed to the point estimate in the Lasso framework (Xu et al., 2020). However, accurate and robust inference of large-scale complex networks is still a confounding pursuit, especially considering the short and limited measurements corrupted by the ambient or measurement noise and other artifacts.

3. Methodology

The network reconstruction is in essence a sparse signal recovery problem: given nodal observation $\boldsymbol{Y}$ and a measurement matrix $\boldsymbol{\phi}$, the network connectivity matrix $\boldsymbol{X}$ is sought after. More specifically, for node i, $\boldsymbol{Y_i} = \boldsymbol{\phi_i X_i}$, where $\boldsymbol{Y_i} = [y_i(t_1), \dots, y_i(t_M)]^T$ is the observation for t_M time steps, and the measurement matrix $\boldsymbol{\phi_i} = \begin{bmatrix} \phi_{i1}(t_1) & \cdots & \phi_{iN}(t_1) \\ \vdots & \ddots & \vdots \\ \phi_{i1}(t_M) & \cdots & \phi_{iN}(t_M) \end{bmatrix}$ signifies the potential dynamic interactions between node i with all N nodes (including self-interaction) in the network for t_M time steps. $\boldsymbol{X_i} = [x_{i1}, \dots, x_{iN}]^T$ is the connectivity vector for node i, which is sparse with only a few nonzero entities. For unweighted networks, $x_{ij} \in \{0,1\}$, $x_{ij} = 0$ implies that node i and j are not connected and $x_{ij} = 1$ represents that node i and j are connected (with dynamic interactions).

Our network reconstruction framework consists of two parts: sparse Bayesian learning for recovery of node connectivity and sequential retrieval to select the next most informative node to investigate.

3.1. Sparse Bayesian Learning for Node Connectivity Recovery

Sparse Bayesian learning has garnered tremendous traction recently to account for the uncertainty associated with sparse solutions (Huang et al., 2019; Huang et al., 2020). The key is the sparsity-promoting prior formulation, and different prior distributions have been investigated in literature, including normal product (Zhou et al., 2015), Laplace (Babacan et al., 2010), and Gaussian-inverse Gamma (Tipping, 2001). Nonetheless, it is not straightforward to encode *a priori* knowledge of the sparsity pattern of $\boldsymbol{X}_i$ into the Bayesian specification with these distributions. In contrast, as a flexible shrinkage method, Spike and Slab prior allows user-specified sparsity (Ishwaran and Rao, 2005). Here, the spike refers to the distribution with spike at $\boldsymbol{X}_i = \boldsymbol{0}$, and the slab determines the distribution of non-zero entities of $\boldsymbol{X}_i$. On a side remark, the sparsity induced by the Spike and Slab model has also been widely used in feature selection (Hernández-Lobato et al., 2013; Ishwaran and Rao, 2005). In this study, it is adopted to specify the prior distribution of connectivity $\boldsymbol{X}_i$:

$$p(\boldsymbol{X_i}|\boldsymbol{z_i}) = \prod_{j=1}^{N} p\left(x_{ij}|z_{ij}\right) = \prod_{j=1}^{N}\left[\left(1 - z_{ij}\right)\delta\left(x_{ij}\right) + z_{ij}\mathcal{N}\left(x_{ij}|m_{ij}, v_{ij}\right)\right] \qquad (1)$$

where $\delta(\cdot)$ is the point probability mass centered at the spike 0, and m_{ij} and v_{ij} are the mean and variance of the slab distribution (normal here). $\boldsymbol{z_i} = [z_{i1}, \dots, z_{iN}]$, and $z_{ij} \in \{0,1\}$ is the latent binary variable indicating whether x_{ij} attains the deterministic value 0 $\left(z_{ij} = 0\right)$ or is drawn from the slab distribution $\left(z_{ij} = 1\right)$. In a hierarchical prior setting, z_{ij}

follows a Bernoulli distribution, $z_{ij} \sim Bernoulli(z_{ij}|\gamma_{ij})$. The parameter $\gamma_{ij} \in [0,1]$ is the prior probability that x_{ij} deviates from zero. Here, the parameters m_{ij}, v_{ij}, and γ_{ij} are initially set to values 0, 1, and 0.5, respectively.

Given observation $\boldsymbol{Y_i}$ and measurement matrix $\boldsymbol{\phi_i}$, the posterior distribution $p(\boldsymbol{X_i}, \boldsymbol{z_i}|\boldsymbol{Y_i}, \boldsymbol{\phi_i})$ is

$$p(\boldsymbol{X_i}, \boldsymbol{z_i}|\boldsymbol{Y_i}, \boldsymbol{\phi_i}) = \frac{p(\boldsymbol{Y_i}|\boldsymbol{X_i}, \boldsymbol{\phi_i})p(\boldsymbol{X_i}|\boldsymbol{z_i})p(\boldsymbol{z_i})}{p(\boldsymbol{Y_i}|\boldsymbol{\phi_i})} \tag{2}$$

Numerical approximations such as expectation propagation (EP) (Hernández-Lobato et al., 2013; Hernández-Lobato et al., 2015; Ishwaran and Rao, 2005) can be used to estimate the posterior. Further, we have the following representation for the joint distribution of $\boldsymbol{X_i}$, $\boldsymbol{z_i}$, and $\boldsymbol{Y_i}$ given $\boldsymbol{\phi_i}$: $p(\boldsymbol{X_i}, \boldsymbol{z_i}, \boldsymbol{Y_i}|\boldsymbol{\phi_i}) = p(\boldsymbol{Y_i}|\boldsymbol{X_i}, \boldsymbol{\phi_i})p(\boldsymbol{X_i}|\boldsymbol{z_i})p(\boldsymbol{z_i}) = \prod_{j=1}^{3} g_j(\boldsymbol{X_i}, \boldsymbol{z_i})$.

For ease of representation, we denote $g_1(\boldsymbol{X_i}, \boldsymbol{z_i}) = p(\boldsymbol{Y_i}|\boldsymbol{X_i}, \boldsymbol{\phi_i})$, $g_2(\boldsymbol{X_i}, \boldsymbol{z_i}) = p(\boldsymbol{X_i}|\boldsymbol{z_i})$, and $g_3(\boldsymbol{X_i}, \boldsymbol{z_i}) = p(\boldsymbol{z_i})$. g_j can be approximated as a simpler form $\hat{g}_j$ by EP method. Therefore, the joint distribution can be represented as $p(\boldsymbol{X_i}, \boldsymbol{z_i}, \boldsymbol{Y_i}|\boldsymbol{\phi_i}) = \prod_{j=1}^{3} g_j(\boldsymbol{X_i}, \boldsymbol{z_i}) \approx \prod_{j=1}^{3} \hat{g}_j(\boldsymbol{X_i}, \boldsymbol{z_i}) = Q(\boldsymbol{X_i}, \boldsymbol{z_i})$. In this context, $\hat{g}_j(\boldsymbol{X_i}, \boldsymbol{z_i})$ adheres to the exponential distribution family (Jordan, 2003), ensuring that their product $Q(\boldsymbol{X_i}, \boldsymbol{z_i})$ is also exponential, as dictated by the closure property of the exponential distribution (Hernández-Lobato et al., 2015). Consequently, the posterior distribution $p(\boldsymbol{X_i}, \boldsymbol{z_i}|\boldsymbol{Y_i}, \boldsymbol{\phi_i})$ can be obtained through the normalized $Q(\boldsymbol{X_i}, \boldsymbol{z_i})$ with normalizing constant $p(\boldsymbol{Y_i}|\boldsymbol{\phi_i})$ based on Eq. (2).

Next, $\hat{g}_1, \hat{g}_2$, and $\hat{g}_3$ are iteratively updated to minimize the Kullback-Leibler (KL) divergence between $\hat{g}_j(\boldsymbol{X_i}, \boldsymbol{z_i})\, Q^{\setminus j}(\boldsymbol{X_i}, \boldsymbol{z_i})$ and $g_j(\boldsymbol{X_i}, \boldsymbol{z_i})\, Q^{\setminus j}(\boldsymbol{X_i}, \boldsymbol{z_i})$. Here, $Q^{\setminus j}(\boldsymbol{X_i}, \boldsymbol{z_i})$ denotes the current approximation of the joint distribution with $\hat{g}_j$ removed, i.e., $Q^{\setminus j}(\boldsymbol{X_i}, \boldsymbol{z_i}) = Q(\boldsymbol{X_i}, \boldsymbol{z_i})/\hat{g}_j(\boldsymbol{X_i}, \boldsymbol{z_i})$. The resulting approximate posterior distribution also takes the form of the exponential family and can be represented as (Hernández-Lobato et al., 2015):

$$p(\boldsymbol{X_i}, \boldsymbol{z_i}|\boldsymbol{Y_i}, \boldsymbol{\phi_i}) = \prod_{j=1}^{N} \mathcal{N}(x_{ij}|m_{ij}, v_{ij}) Bernoulli(z_{ij}|\gamma_{ij}) \tag{3}$$

where $\boldsymbol{m_i} = [m_{i1}, \dots, m_{iN}]$, $\boldsymbol{v_i} = [v_{i1}, \dots, v_{iN}]$, and $\boldsymbol{\gamma_i} = [\gamma_{i1}, \dots, \gamma_{iN}]$ are the free parameters estimated through the EP. With the estimated posterior distribution obtained from Eq. (3), an edge between node i and j exists if $p(z_{ij}|\boldsymbol{Y_i}, \boldsymbol{\phi_i}) > \varrho$, where $0 \le \varrho \le 1$ is a prescribed threshold. $\varrho = 0.5$ is used in this study. Next, the updated parameters m_{ij}, v_{ij}, and γ_{ij} derived from node i are subsequently treated as the priors m_{ji}, v_{ji}, and γ_{ji} for x_{ji} of node j in Eq. (1), which will be further updated to obtain $\boldsymbol{X_j}$ given observation $\boldsymbol{Y_j}$ and measurement matrix $\boldsymbol{\phi_j}$. This effectively eschews the pitfall of conflicting results in the conventional sparse learning approaches as reported in (Basiri et al., 2018; Han et al., 2015).

3.2. Sequential Retrieval of Node Connectivity

Nonetheless, it is computationally daunting to retrieve connectivity $\boldsymbol{X_i}$ within a short time interval for networks with large N. Ideally, we can evaluate only a subset of the nodes to uncover the network connectivity with acceptable reconstruction error, leveraging proceeding reconstruction efforts. To select the optimal subset, the submodularity property will be investigated.

For a network with a finite set of nodes $\mathbb{Q} = 1,2, \dots, N$, let θ_k denote the selected set of k nodes for connectivity reconstruction, then the reconstruction error is manifested in terms of the sum of square error (SSE), $SSE(\theta_k) = \left\|\boldsymbol{Y_i} - \boldsymbol{\phi_i}\widehat{\boldsymbol{X}}_i\right\|_2^2$. $SSE(\theta_k)$ embodies the sum of the discrepancy between the nodal observation $\boldsymbol{Y_i}$ and its estimate, and $\widehat{\boldsymbol{X}}_i$ is the approximation of $\boldsymbol{X_i}$ for node i via the proposed sparse Bayesian learning approach. It is evident that the null set $\emptyset$ results in the maximum SSE without any knowledge of the connectivity. Herein, utility over a set θ_k is defined as $f(\theta_k) = SSE(\emptyset) - SSE(\theta_k)$. It is noted that $f(\theta_k)$ is a monotonic submodular function and possesses the following two properties: (1) *Monotonicity*: $f(\theta_1) < f(\theta_2)$ for all $\theta_1 \subseteq \theta_2 \subseteq \mathbb{Q}$. (2) *Submodularity*: For $\theta_1 \subseteq \theta_2 \subseteq \mathbb{Q}$ and any element $s \in \mathbb{Q}\setminus\theta_2$, $f(\theta_1) - f(\theta_1 \cup s) \ge f(\theta_2) - f(\theta_2 \cup s)$, where $\mathbb{Q}\setminus\theta_2$ are the nodes in $\mathbb{Q}$ but not contained in the set θ_2.

Simple greedy algorithms prove effective for near-optimal subset selection for maximization of monotonic submodular functions (Zhu et al., 2019). Starting with a set θ_k (here, $\theta_0 = \emptyset$), the next node s_{k+1} is selected via $s_{k+1} = \underset{s \in \mathbb{Q}\setminus\theta_k}{\operatorname{argmax}} f(\theta_k \cup s)$ and $\theta_{k+1} = \theta_k \cup \{s_{k+1}\}$. Here, the node with the largest observation $\boldsymbol{Y_i}$ is selected initially without any connection information. To find the best subsequent node s_{k+1} to maximize $f(\theta_k \cup s)$, it is necessary to estimate $f(\theta_k \cup s)$ while incorporating the uncertainty associated with the estimation for all $s \in \mathbb{Q}\backslash\theta_k$, without computing the objective function. We define improvement by a utility function over selecting new node s given the retrieval node set θ_k: $I_m(s|\theta_k) = \max(0, \epsilon_e - \epsilon_{min})$, where $\epsilon_{min} = \min_i \left(\left\| \boldsymbol{Y}_i - \boldsymbol{\phi}_i \widehat{\mathbf{X}}_i^{\theta_k} \right\|_2^2 \right)$ represents the minimum estimated error for $i \in \theta_k$ and $\epsilon_e = \left\| \boldsymbol{Y}_s - \boldsymbol{\phi}_s \widehat{\mathbf{X}}_s^{\theta_k} \right\|_2^2$ represents the predictive error of node $s \in \mathbb{Q}\backslash\theta_k$ given the retrieved information of the set θ_k. In such a manner, the most informative node s that causes the largest predictive reconstruction error for the whole network can be identified and then the sparse Bayesian learning algorithm can be applied to significantly reduce the network reconstruction error.

In concreteness, we initially set all columns and rows as zero and one for weight matrix $\boldsymbol{m}$ and variance matrix $\boldsymbol{v}$ ($m_{ij} = 0, v_{ij} = 1$ for $i, j \in [1 \quad ,\cdots, \quad N]$), in order to determine the priors for $\boldsymbol{X}$ as in Eq. (1). Only the node i in the retrieved set θ_k are fully estimated by the sparse Bayesian learning algorithm and the unretrieved node $s \in \mathbb{Q}\backslash\theta_k$ are partially recovered due to the symmetric character of the network. Hence, the fully estimated weight and variance result for node i are $\widehat{\boldsymbol{m}}_i^{\theta_k} = [\widehat{m}_{i1} \quad \cdots \quad \widehat{m}_{iN}]^{\boldsymbol{T}}$ and $\widehat{\mathbf{v}}_i^{\theta_k} = [\widehat{v}_{i1} \quad \cdots \quad \widehat{v}_{iN}]^{\boldsymbol{T}}$; the partially recovered (predictive) weight and variance result for node s are $\widehat{\boldsymbol{m}}_s^{\theta_k} = [0 \quad \cdots \quad \widehat{m}_{si} \quad \cdots \quad 0]^{\boldsymbol{T}}$ and $\widehat{\mathbf{v}}_s^{\theta_k} = [1 \quad \cdots \quad \widehat{v}_{si} \quad \cdots \quad 1]^{\boldsymbol{T}}$. We could balance $I_m(s|\theta_k)$ between its estimated weight and variance. This yields the so-called expected improvement (EI) acquisition function

$$EI(s) = \epsilon \boldsymbol{\psi} \left(\frac{\epsilon}{\left\| \widehat{v}_s^{\theta_k} \right\|_2^2} \right) + \left\| \widehat{v}_s^{\theta_k} \right\|_2^2 \boldsymbol{\Gamma} \left(\frac{\epsilon}{\left\| \widehat{v}_s^{\theta_k} \right\|_2^2} \right) \tag{4}$$

where $\epsilon = \left\| \boldsymbol{Y}_s - \boldsymbol{\phi}_s \widehat{\mathbf{m}}_s^{\theta_k} \right\|_2^2 - \min_i \left(\left\| \boldsymbol{Y}_i - \boldsymbol{\phi}_i \widehat{\boldsymbol{m}}_i^{\theta_k} \right\|_2^2 \right)$, $\boldsymbol{\psi}(\cdot)$ and $\boldsymbol{\Gamma}(\cdot)$ are the CDF and PDF of standard normal distribution, respectively. The first term in Eq. (4) is the exploitation and the second term corresponds to the exploration term. The most informative node $s_{k+1} = \underset{s \in \mathbb{Q}\backslash\theta_k}{\operatorname{argmax}} EI(s)$ to be selected is determined based on $EI(s)$ for $s \in \mathbb{Q}\backslash\theta_k$.

4. Numerical Results

In this study, we demonstrate the performance of our reconstruction framework within the context of a power grid system employing the IEEE-118 bus network, with data recorded at a time step of $M = 60$. The application of the IEEE-118 bus network in the power grid system involves considering it as a weighted network, reflecting varying susceptances of the transmission lines.

For comparison, we include a baseline model Random Spike and Slab (RandomSS), which uses the Spike and Slab prior in the sparse Bayesian learning approach to reveal the node connectivity but only randomly selects the next node for reconstruction. By contrast, the proposed Sequential Spike and Slab model (SeqSS) selects the next most informative node via expected improvement.

We compare the performance of the RandomSS algorithm with our proposed SeqSS algorithm by averaging 30 experiments (each with randomly added noise) using the following metrics: the Frobenius norm for the connectivity discrepancy $Error_{\widehat{A}} = \frac{\|A - \widehat{A}\|_F}{\|A\|_F}$, $\boldsymbol{A}$ and $\widehat{\boldsymbol{A}}$ denotes the true and estimated connectivity matrices, respectively; and the Frobenius norm for the observation discrepancy $Error_{\widehat{Y}}$ and weight discrepancy $Error_{\widehat{m}}$ given by $Error_{\widehat{Y}} = \frac{\|Y - \widehat{Y}\|_F}{\|Y\|_F}$ and $Error_{\widehat{m}} = \frac{\|\boldsymbol{m} - \widehat{\boldsymbol{m}}\|_F}{\|\boldsymbol{m}\|_F}$. Here, $\widehat{\boldsymbol{m}}$ and $\widehat{\boldsymbol{A}}$ with entry z_{ij} are estimated from Eq. (2). We assume that $z_{ij} = 1$ (has a connection between node i and j) if $\gamma_{ij} > 0.5$ otherwise $z_{ij} = 0$ (has no connection between node i and j). However, the $\widehat{\mathbf{Y}}$ is indirectly estimated from Eq. (2), we calculate $\widehat{\mathbf{Y}}_{\boldsymbol{i}} = \boldsymbol{\phi}_{\boldsymbol{i}} \widehat{\mathbf{m}}_{\boldsymbol{i}}$ for weighted network.

The topology of the power grid system could remain unknown at the onset of blunt perturbations, such as inclement weather or operation glitches. With the simplified direct current (DC) approximation, the power flow P_{ij} from node i

to j (e.g., generator or load) over the transmission line with reactance r_{ij} is given as: $P_{ij} = \frac{|V_i|\cdot|V_j|}{r_{ij}}\sin(\varphi_i - \varphi_j)$, where $|V_i|$ and $|V_j|$ are the voltage magnitudes, and φ_i and φ_j are the phase angles of node i and j, respectively. Here, we simulate the phase angle variation on the power grid system: the phase angle $\varphi_i(t) = (\omega + \Delta\omega_i)t$, where $\omega = 2\pi \times 50$ is the angular frequency of grid operation and $\Delta\omega_i \sim \mathcal{N}(0, 20)$ is the random frequency perturbation for node i. Moreover, the voltage magnitude $|V_i|$ is set to the unit value for all nodes for simplicity. Thereby, the effective power balance at node i is delineated $y_i = \sum_{j\in U_i} P_{ij} = \sum_{j\in U_i} \frac{\sin(\varphi_i - \varphi_j)}{r_{ij}}$, where U_i is the set of connected nodes with i. The sensing data from the phasor measurement units for node i include the phase angle φ_i and power flow y_i during $t \in [t_1, t_M]$.

Considering noise in data collection, we include noise $\boldsymbol{\varepsilon} = [\varepsilon_{t_1} \quad \cdots \quad \varepsilon_{t_m} \quad \cdots \quad \varepsilon_{t_M}]^T$ to power flow $\boldsymbol{Y}$. Here, $\varepsilon_{t_m} \sim \mathcal{N}(\varepsilon_{t_m} \mid 0, (\beta \times \max(\boldsymbol{Y}))^2)$, $t_m = t_1, \ldots, t_M$, where $\beta = \{0.01, 0.02, 0.03\}$ is the scale to the maximum absolute value in $\boldsymbol{Y}$. Thus, $\boldsymbol{Y_i} = \boldsymbol{\phi_i X_i} + \boldsymbol{\varepsilon}$:

$$\begin{bmatrix} y_i(t_1) \\ \vdots \\ y_i(t_M) \end{bmatrix} = \begin{bmatrix} \phi_{i1}(t_1) & \cdots & \phi_{iN}(t_1) \\ \vdots & \ddots & \vdots \\ \phi_{i1}(t_M) & \cdots & \phi_{iN}(t_M) \end{bmatrix} \begin{bmatrix} x_{i1} \\ \vdots \\ x_{iN} \end{bmatrix} + \begin{bmatrix} \varepsilon_{t_1} \\ \vdots \\ \varepsilon_{t_M} \end{bmatrix} \tag{5}$$

Here, $\phi_{ij}(\mathrm{t}) = \sin(\varphi_i(t) - \varphi_j(t))$ and susceptance $x_{ij} = \frac{1}{r_{ij}} > 0$ if $j \in U_i$, otherwise $r_{ij} = \infty$ and $x_{ij} = 0$. Thus, x_{ij} is the weight of edge in the grid network. $\boldsymbol{Y_i}$ is generated by randomized repetitions of 30 experiments at each scale level β for the IEEE-118 network structure.

We compare performance of SeqSS and RandomSS in terms of the connectivity discrepancy $Error_{\hat{A}}$, the observation discrepancy $Error_{\hat{Y}}$, and weight discrepancy $Error_{\hat{m}}$ with different cardinality $|\theta_k|$ of the selected node set θ_k. The connectivity discrepancy of the power grid system under IEEE-118 network is depicted in Figure 1. An error bar is also provided at each point denoting the standard deviation from the 30 experiments. SeqSS registers the same accuracy as RandomSS with edge construction for only ~50% of the nodes.

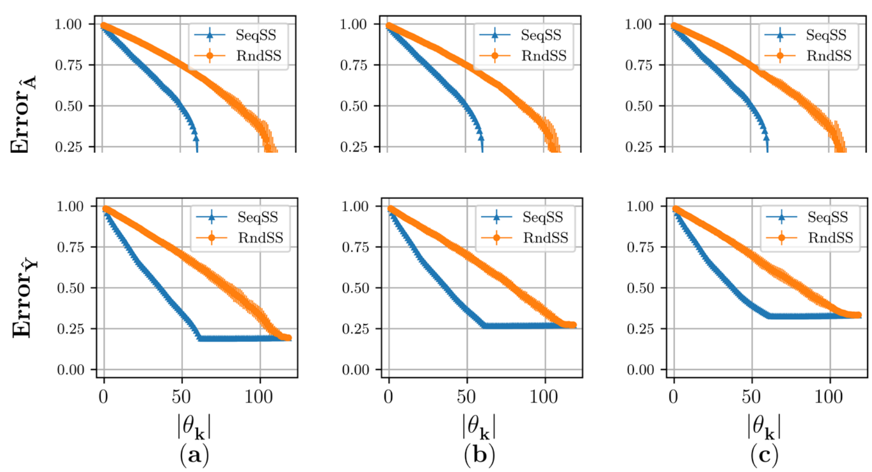

Figure 1. Observation discrepancy of the power grid system for SeqSS and RandomSS algorithms with different noise levels (a) $\beta = 0.01$, (b) $\beta = 0.02$, and (c) $\beta = 0.03$ under IEEE-118 network.

The observation discrepancy $Error_{\hat{Y}}$ is shown in Figure 2. As the cardinality $|\theta_k|$ increases, the $Error_{\hat{Y}}$ edges down continuously. While a higher noise level renders a larger $Error_{\hat{Y}}$, SeqSS is overall robust to such perturbations. It bears mentioning that the discrepancy variance (represented by the error bar) in SeqSS is notably smaller than that observed in RandomSS. This minor variability is attributed to the EI algorithm in SeqSS that consistently selects the most informative nodes.

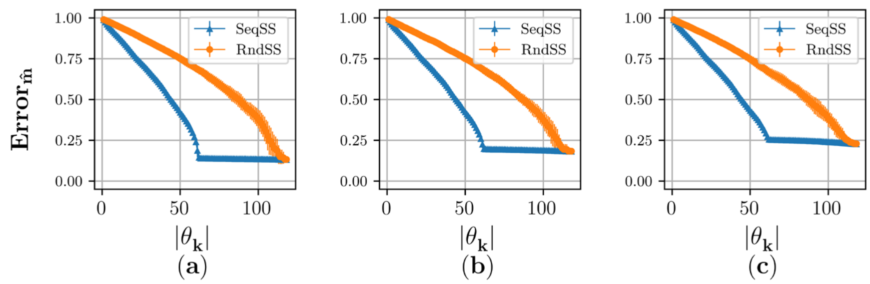

Figure 2. Weight discrepancy of the power grid system for SeqSS and RandomSS algorithms with different noise levels (a) $\beta = 0.01$, (b) $\beta = 0.02$, and (c) $\beta = 0.03$ under IEEE-118 network.

In Figure 2, the $Error_{\hat{m}}$ of the IEEE-118 network structure is exhibited. Here, with only a fraction of the recovered nodes, the weight discrepancy of SeqSS approaches the same level of RandomSS. Unlike the unweighted network or connectivity matrix, estimate of the weight matrix in the grid network entails more nodes to converge, and the accuracy is sensitive to the noise level. The lower the noise level, the lower the level of $Error_{\hat{m}}$.

5. Conclusion and Discussion

This study has the potential to significantly scale up the reconstruction algorithms to accommodate massive networks, radically transform the monitoring and operation for various realistic networked systems, including the power grid, transportation and communication, and provide a new robust and efficient foundation for military operational decision making. Particularly, this proposed research can find critical applications in a host of military operation scenarios, such as the monitoring and maintenance of the transportation/communication network or power grid system.

In this present study, we develop a sparse Bayesian approach based on the Spike and Slab modeling for sequential reconstruction of network connectivity. We corroborate this approach on a power grid system under the weighted IEEE-118 network. The simulation results imply that our proposed SeqSS algorithm identifies the network connectivity in an efficient fashion (only ~50% nodes are interrogated) compared to RandomSS. Owing to the EI algorithm to select the most informative node, the SeqSS algorithm notches smaller discrepancy variance than that of RandomSS

References

Babacan, S. D., Molina, R., and Katsaggelos, A. K., Bayesian compressive sensing using laplace priors, *IEEE Transactions on Image Processing*, vol. 19, pp. 53–63, 2010.

Basiri, F., Casadiego, J., Timme, M., and Witthaut, D., Inferring power-grid topology in the face of uncertainties, *Physical Review E*, vol. 98, pp. 012305, 2018.

Che, Y., Zhang, Z. (John), and Cheng, C., Physical–statistical learning in resilience assessment for power generation systems, *Physica A: Statistical Mechanics and its Applications*, vol. 615, pp. 128584, 2023.

Han, X., Shen, Z., Wang, W.-X., and Di, Z., Robust reconstruction of complex networks from sparse data, *Physical Review Letters*, vol. 114, pp. 028701, 2015.

Hempel, S., Koseska, A., Kurths, J., and Nikoloski, Z., Inner composition alignment for inferring directed networks from short time series, *Physical Review Letters*, vol. 107, pp. 054101, 2011.

Hernández-Lobato, D., Hernández-Lobato, J. M., and Dupont, P., Generalized Spike-and-Slab priors for Bayesian group feature selection using expectation propagation, *Journal of Machine Learning Research*, vol. 14, pp. 1891–1945, 2013.

Hernández-Lobato, J. M., Hernández-Lobato, D., and Suárez, A., Expectation propagation in linear regression models with Spike-and-Slab priors, *Machine Learning*, vol. 99, pp. 437–487, 2015.
Huang, K., Deng, W., Zhang, Y., and Zhu, H., Sparse Bayesian learning for network structure reconstruction based on evolutionary game data, *Physica A: Statistical Mechanics and its Applications*, vol. 541, pp. 123605, 2020.
Huang, K., Jiao, Y., Liu, C., Deng, W., and Wang, Z., Robust network structure reconstruction based on Bayesian compressive sensing, *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 29, pp. 093119, 2019.
Ishwaran, H., and Rao, J. S., Spike and Slab variable selection: Frequentist and Bayesian strategies, *The Annals of Statistics*, vol. 33, pp. 730–773, 2005.
Jordan, M. I., An Introduction to probabilistic graphical models, Available: https://people.eecs.berkeley.edu/~jordan/prelims/, December 27, 2023.
Napoletani, D., and Sauer, T. D., Reconstructing the topology of sparsely connected dynamical networks, *Physical Review E*, vol. 77, pp. 026103, 2008.
Ren, J., Wang, W.-X., Li, B., and Lai, Y.-C., Noise bridges dynamical correlation and topology in coupled oscillator networks, *Physical Review Letters*, vol. 104, pp. 058701, 2010.
Shamsan, A., Wu, X., Liu, P., and Cheng, C., Intrinsic recurrence quantification analysis of nonlinear and nonstationary short-term time series, *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 30, pp. 093104, 2020.
Shen, Z., Wang, W.-X., Fan, Y., Di, Z., and Lai, Y.-C., Reconstructing propagation networks with natural diversity and identifying hidden sources, *Nature Communications*, vol. 5, pp. 4323, 2014.
Timme, M., and Casadiego, J., Revealing networks from dynamics: an introduction: Journal of Physics A, *Mathematical and Theoretical*, vol. 47, pp. 343001, 2014.
Tipping, M. E., Sparse Bayesian learning and the relevance vector machine, *The Journal of Machine Learning Research*, vol. 1, pp. 211–244, 2001.
Wang, W.-X., Lai, Y.-C., and Grebogi, C., Data based identification and prediction of nonlinear and complex dynamical systems, *Physics Reports*, vol. 644, pp. 1–76, 2016.
Wang, W.-X., Lai, Y.-C., Grebogi, C., and Ye, J., Network reconstruction based on evolutionary-game data via compressive sensing, *Physical Review X*, vol. 1, pp. 021021, 2011.
Wu, X., Zheng, Y., Che, Y., and Cheng, C., Pattern recognition and automatic identification of early-stage atrial fibrillation, *Expert Systems with Applications*, vol. 158, pp. 113560, 2020.
Xu, S., Zhang, C., Wang, P., and Zhang, J., Variational Bayesian weighted complex network reconstruction, *Information Sciences*, vol. 521, pp. 291–306, 2020.
Yu, D., Righero, M., and Kocarev, L., Estimating topology of networks, *Physical Review Letters*, vol. 97, pp. 188701, 2006.
Zhou, Z., Liu, K., and Fang, J., Bayesian compressive sensing using normal product priors, *IEEE Signal Processing Letters*, vol. 22, pp. 583–587, 2015.
Zhu, J., and Abur, A., Identification of network parameter errors, *IEEE Transactions on Power Systems*, vol. 21, pp. 586–592, 2006.
Zhu, R., Yao, B., Yang, H., and Leonelli, F., Optimal sensor placement for space–time potential mapping and data fusion, *IEEE Sensors Letters*, vol. 3, pp. 1–4, 2019.

Biographies

Jinming Wan received the B.E. degree in Forest Chemical Engineering from the Northwest Agriculture & Forestry University, Yangling, China, in 2015, and the M.S. degrees in Engineering and Technology Management from Morehead State University, Morehead, KY, U.S, in 2018. His research focuses on modeling and analysis, uncertainty quantification, machine learning, dynamics of complex systems, and optimal control techniques.

Jun Kataoka received his B.S. (2016) at Chuo University, Tokyo, Japan, MSc (2017) at University of Warwick, Coventry, UK and is currently a Ph.D. candidate in the Systems Science and Industrial Engineering department at SUNY Binghamton. He is working on physics-informed machine learning and domain adaptation applications in smart manufacturing as part of his Ph.D. dissertation.

Jayanth Sivakumar received an M.S. in Computer Science and a Ph.D. in Systems Science from Binghamton University. His research focuses on synthetic data generation methodologies for structured datasets to tackle small dataset problem. He is currently a senior data scientist at Citibank N.A. He leads the optimization of the target communication channel for Citi's retail offers.

Eric Peña received a B.S. in Engineering Physics from The Ohio State University and an M.S. degree in Systems Science from Binghamton University. He was a Clifford D. Clark Fellow, received an Advanced Graduate Certificate in Complex Systems modeling, and was a member of the Binghamton Center of Complex Systems (CoCo). He is currently a senior data scientist in the autonomous systems engineering team at Motional validating the safety of autonomous vehicle robotaxis.

Yiming Che received the B.S. degree in Industrial Engineering from Capital University of Economics and Business, Beijing, China, in 2017, the M.S. degree in Industrial Engineering from the Binghamton University, Binghamton, NY, USA, in 2018, and the Ph.D. degree in Systems Science from the Binghamton University, Binghamton, NY, USA, in 2023. He joined the School of Computing and Augmented Intelligence, Arizona State University, Tempe, AZ, as a post-doctoral research fellow supervised by Dr. Teresa Wu, in 2023. His current research interests include computer vision in medical imaging, Bayesian inference and active learning.

Hiroki Sayama is a Professor in the Department of Systems Science and Industrial Engineering, and the Director of the Binghamton Center of Complex Systems, at Binghamton University, State University of New York, USA. He also serves as a non-tenure-track Professor in the School of Commerce at Waseda University, Japan, as well as an External Faculty member of the Vermont Complex Systems Center at the University of Vermont, USA. He received his B.Sc., M.Sc. and D.Sc. in Information Science, all from the University of Tokyo, Japan. His research interests include complex dynamical networks, human and social dynamics, collective behaviors, artificial life/chemistry, interactive systems, and complex systems education, among others. He is an expert on mathematical/computational modeling and analysis of various complex systems.

Changqing Cheng is an associate professor in the Department of Systems Science and Industrial Engineering at Binghamton University. His research interests include nonlinear dynamics, sensing and data-driven modeling, simulation and analytics for process monitoring, quality control, and performance optimization of complex systems. He serves as associate editor for the journal of IISE Transactions on Healthcare Systems Engineering.