Title

Graph-Signal-to-Graph Matching for Network De-Anonymization Attacks

Authors

Liu, Hang
Scaglione, Anna
Peisert, Sean

Publication Date

2024

DOI

10.1109/tifs.2024.3483669

Peer reviewed

Graph-Signal-to-Graph Matching for Network De-anonymization Attacks

Hang Liu, *Member, IEEE*, Anna Scaglione, *Fellow, IEEE*, and Sean Peisert, *Senior Member, IEEE*

***Abstract*—Graph matching over two given graphs is a well-established method for re-identifying obscured node labels within an anonymous graph by matching the corresponding nodes in a reference graph. This paper studies a new application, termed the graph-signal-to-graph matching (GS2GM) problem, where the attacker observes a set of filtered graph signals originating from a hidden graph. These signals are generated through an unknown graph filter activated by certain input excitation signals. Our goal is to match their components to a labeled reference graph to reveal the labels of asymmetric nodes in this unknown graph, where the excitations can be either known or unknown to the attacker. To this end, we integrate the existing blind graph matching algorithm with techniques of graph filter inference and covariance-based eigenvector estimation. Furthermore, we establish sufficient conditions for perfect node de-anonymization through graph signals, showing that graph signals can leak substantial private information on the concealed labels of the underlying graph. Experimental results validate our theoretical insights and demonstrate that the proposed attack effectively reveals many of the hidden labels, particularly when the graph signals are adequately uncorrelated and sampled.**

***Index terms*—** Graph matching, network de-anonymization, network privacy attack, node identification, graph signal processing.

I. Introduction

The emergence of expansive networks, such as social media, infrastructure systems, and the Internet of Things, has led to an era of immense data proliferation. This surge emphasizes the critical need to protect the private information of network users. Although data publishers often de-identify or randomize names and other identifying details to safeguard personal information, recent findings indicate that these conventional methods are insufficient to avoid re-identification of individual records. Adversaries can potentially infer a target user's identity or other sensitive labels from its local network connections by leveraging auxiliary data or contextual information.

This work was supported in part by the DoD-ARO under Grant No. W911NF2210228. This research was supported in part by the Director, Cybersecurity, Energy Security, and Emergency Response (CESER) office of the U.S. Department of Energy, via the Privacy-Preserving, Collective Cyberattack Defense of DERs project, under contract DE-AC02-05CH11231. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect those of the sponsors of this work. An earlier version of this paper was presented in part at the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), April 2024 [1].

Hang Liu and Anna Scaglione are with the Department of Electrical and Computer Engineering, Cornell Tech, Cornell University, New York, NY, 10044 USA (e-mails: {hl2382, as337}@cornell.edu). Sean Peisert is with the Computing Sciences Research, Lawrence Berkeley National Laboratory, Berkeley, CA 94720 USA (e-mail: sppeisert@lbl.gov).

This paper has supplementary downloadable material available at http://ieeexplore.ieee.org., provided by the authors. The material includes the value of the system admittance matrix for the IEEE 22-bus system. Contact hl2382@cornell.edu for further questions about this work.

A. Related Work

One notable attack method within this framework is known as *graph de-anonymization* or *node re-identification*. The aim is to infer labels within an anonymized network by aligning target nodes with a labeled reference graph sourced from public datasets, topology snapshots, etc. This de-anonymization strategy was first introduced by Narayanan and Shmatikov [2], where IMDB data served as a reference to identify a supposedly Netflix dataset. The same authors expanded this approach to user de-anonymization in large social networks [3]. Another pioneering work [4] also studied address anonymization for IP networks.

Graph de-anonymization was initially explored in scenarios involving 'seeds' [3], [5], where the labels of certain key nodes in the target graph are assumed to be known and used to infer other labels. In scenarios lacking prior seed information, *seedless* graph de-anonymization comes into play, aiming to identify arbitrary labels in the undisclosed graph without any seed. Pedarsani and Grossglauser [6] framed this seedless de-anonymization challenge as a *graph matching* problem, aiming to align nodes of the target and reference graphs by minimizing edge mismatches subject to a node permutation. Techniques such as maximum-a-posterior (MAP) estimation have been investigated for seedless social network de-anonymization, especially in networks with community structures [7], [8]. Studies have also derived sufficient conditions for perfect node de-anonymization in specific graph models, including Erdös-Rényi (ER) random graphs [6] and stochastic block models [7]. Additionally, Miao *et al.* [9] introduced a metric to quantify the number of potentially de-anonymizable nodes through graph matching, termed graph de-anonymizability.

Current research relies on the full topology information of the anonymous graph for effective graph de-anonymization or matching. However, acquiring precise topological information proves resource-intensive or unattainable in many real-world applications [10]. More frequently, attackers might directly observe interactions between nodes in an undisclosed graph, known as graph signals [11]. These are evident in various contexts, such as opinion exchanges in social networks or nodal measurements in infrastructure systems and power grids. Recent studies have demonstrated that graph signals carry a plethora of information that can be leveraged for network analysis [12]. Graph signal processing has gained widespread

adoption in the fields of graph learning and graph neural networks [13]–[16]. Particularly, studies [10], [17] have shown the feasibility of inferring unknown graph topology and Laplacian matrices from graph signals. A related challenge arises in power grid networks, where the use of grid measurements to estimate graph topology has been explored in several studies [18]–[21]. Similarly, Wai *et al.* [22] explored the use of graph signals for community detection, bypassing the need for direct access to the graph topology. Studies [23]–[27] investigated blind source separation using graph signals. Moreover, Liu *et al.* [28] proposed a blind graph matching algorithm to match nodes from two unknown graphs using their graph signals.

B. Our Contributions

Motivated by the aforementioned discussions, we focus on privacy attacks in seedless graph de-anonymization using graph signals instead of relying on graph topology information. Our goal is to establish node matching from graph signals originating from the undisclosed graph to a known reference graph, a process we define as the graph-signal-to-graph matching (GS2GM) problem. Successfully accomplishing this task can reveal user identities in the hidden nodes by associating them with the labels of the reference [6].

In this work, we consider a system where the observed graph signals are produced by processing input excitation signals through an unknown graph filter, which encapsulates the node interactions within the network. Our study assumes that this graph filter is a polynomial function of the graph Laplacian matrix and maintains a known reshuffling order of the Laplacian spectrum. We explore the GS2GM problem in both contexts—where the excitation signals are either known or unknown to the attacker. To tackle this challenge, we integrate the blind graph matching algorithm from [28] with techniques for graph filter inference and eigenvector estimation based on sample covariance. Furthermore, we analyze sufficient conditions for perfect de-anonymization and substantiate our theoretical results with experimental results. The contributions of this work are summarized as follows.

- We formulated the GS2GM problem as a task of computing node matches between a set of observed graph signals and a reference graph. We show that symmetric structures within the graph present challenges in evaluating node identification performance. Consequently, we focus on identifying all the asymmetric nodes of the undisclosed graphs and present a metric to measure the accuracy of graph de-anonymization by the success probability of correctly matching asymmetric nodes.
- We developed a unified framework to solve the GS2GM problem in various settings, including cases where the excitation signals are known or unknown. The method adapts the blind graph matching algorithm from [28] to align the eigenmodes of the unknown graph Laplacian estimated from the graph signals with those of the reference graph. Furthermore, we conducted a theoretical analysis to quantify the successful matching probability in relation to the eigenmode estimation accuracy. The analysis enables direct comparison of various methodologies under a consistent metric of matching error probability and offers concise expressions to quantify the impact of key system parameters on de-anonymization performance, including the sample size, noise power, and spectral gaps of the graph filter. Additionally, to facilitate numerical assessment of de-anonymization performance, we present a low-complexity mechanism for detecting symmetric nodes using the reference graph.
- For settings with inaccessible excitation signals, we use sample covariance matrices to estimate the empirical eigenvectors in the GS2GM algorithm. Its performance is then theoretically validated by establishing a bound on the de-anonymization error probability. In contrast, in situations where excitation signals are known, we present an eigenmode estimation algorithm rooted in sparse graph filter inference. Our findings highlight the benefits of utilizing sparsity in filter inference.

We conducted experiments on synthetic datasets and real-world applications to verify our theoretical findings. The results demonstrate that our method achieves accurate node de-anonymization with adequately sampled and uncorrelated signals. Conversely, in many real-world scenarios where only highly correlated signals are accessible, node identification manages to accurately match only a limited subset of nodes. This is particularly evident when the excitation signals remain undisclosed, suggesting that these networks inherently offer a higher degree of privacy protection for user identities.

C. Organization and Notations

The remainder of this paper is organized as follows. We introduce the system model in Section II. In Section III, we formulate the GS2GM problem and the evaluation metric. In Section IV, we present the overall solution to the GS2GM problem, while detailed approaches for different scenarios are expanded in Sections V and VI. In Section VII, we present experimental results to evaluate the proposed method. Finally, this paper concludes in Section VIII.

Throughout, we use regular, bold small, and bold capital letters to denote scalars, vectors, and matrices, respectively. We use $(\mathbf{X})^t$ to denote the t-th power of matrix $\mathbf{X}$, $\mathbf{X}^T$ to denote the transpose of $\mathbf{X}$, $\mathbf{X}^H$ to denote the conjugate transpose, $\overline{\mathbf{X}}$ to denote the matrix containing the absolute value of the entries of $\mathbf{X}$, $\mathrm{tr}(\mathbf{X})$ to denote the trace, and $\mathrm{vec}(\mathbf{X})$ to denote the column-wise vectorization of $\mathbf{X}$. We use x_i to denote the i-th entry of vector $\mathbf{x}$, x_{ij} or $[\mathbf{X}]_{ij}$ interchangeably to denote the (i,j)-th entry of matrix $\mathbf{X}$, and $\mathbf{x}_j$ to denote the j-th column of $\mathbf{X}$. The real normal distribution with mean $\boldsymbol{\mu}$ and covariance $\mathbf{C}$ is denoted by $\mathcal{N}(\boldsymbol{\mu}, \mathbf{C})$, and the cardinality of set $\mathcal{S}$ is denoted by $|\mathcal{S}|$. We use $\|\cdot\|_p$ to denote the ℓ_p norm, $\|\cdot\|_F$ (resp. $\|\cdot\|_2$) to denote the Frobenius (resp. spectral) matrix norm, $\mathbf{I}_N$ to denote the $N \times N$ identity matrix, $\mathbf{1}$ (or $\mathbf{0}$) to denote the all-one (or all-zero) vector with an appropriate size, and $\mathrm{diag}(\mathbf{x})$ to denote a diagonal matrix with the diagonal entries specified by $\mathbf{x}$. For any positive integer N, we denote the factorial of N by $N!$ and define $[N] \triangleq \{1, 2, \cdots, N\}$.

II. System Model

Consider the de-anonymization attack, also known as node re-identification, on an N-node undirected graph $\mathcal{G}_1 = (\mathcal{V}_1, \mathcal{E}_1)$, where $\mathcal{V}_1 = [N]$ and $\mathcal{E}_1$ represent the sets of nodes and edges, respectively. The node labels of $\mathcal{G}_1$ are anonymous to the attacker, whose objective is to identify these labels. The adjacency matrix for $\mathcal{G}_1$ is represented as $\mathbf{A}^{(1)} \in \mathbb{R}^{N\times N}$, where $a^{(1)}_{kl} = a^{(1)}_{lk} > 0$ if and only if an edge (k, l) exists in $\mathcal{E}_1$. Consequently, the graph Laplacian matrix is denoted by $\mathbf{L}^{(1)} = \mathrm{diag}(\mathbf{A}^{(1)}\mathbf{1}) - \mathbf{A}^{(1)}$.

The attacker aims to identify the node labels of $\mathcal{G}_1$ with the aid of a known reference graph with the same number of nodes,[1] denoted by $\mathcal{G}_2 = (\mathcal{V}_2 = [N], \mathcal{E}_2)$. The associated Laplacian matrix of $\mathcal{G}_2$ is denoted by $\mathbf{L}^{(2)}$. The goal is to determine a permutation function $\sigma(\cdot) : [N] \to [N]$, i.e., node matching, which maps the node set $\mathcal{V}_2$ (or its subset) to $\mathcal{V}_1$ such that the permutation of $\mathcal{G}_2$ under $\sigma(\cdot)$ results in a graph closely resembling $\mathcal{G}_1$ [6], [7]. For simplicity, the node permutation is represented interchangeably by $\sigma(\cdot)$ and its equivalent permutation matrix $\mathbf{P} \in \{0, 1\}^{N\times N}$, where $p_{kl} = 1$ if $\sigma(k) = l$ and $p_{kl} = 0$ otherwise. As the nodes of $\mathcal{G}_2$ have known labels, the attacker can leverage $\sigma(\cdot)$ to deduce the labels of the target nodes in $\mathcal{G}_1$. Subsequently, this identification may enable the attacker to infer private information, such as user identities or associated attributes [6].

Most current research on graph de-anonymization assumes a known Laplacian or adjacency matrix of $\mathcal{G}_1$ to the attacker. Within this framework, the attacker determines the node permutation $\sigma(\cdot)$ by aligning the two Laplacian matrices $\mathbf{L}^{(1)}$ and $\mathbf{L}^{(2)}$. This assumption requires comprehensive knowledge of the topology of the anonymous graph. In contrast, our study explores a more challenging scenario where both the adjacency and Laplacian matrices of $\mathcal{G}_1$ remain concealed. In this situation, the attacker utilizes a set of *graph signals* generated on $\mathcal{G}_1$ to de-anonymize its nodes.

A. Graph Signal Model

The matrices $\mathbf{L}^{(1)}$ and $\mathbf{L}^{(2)}$ admit the following eigendecompositions:

$$\mathbf{L}^{(i)} = \mathbf{V}^{(i)}\mathbf{\Gamma}^{(i)}(\mathbf{V}^{(i)})^T, i = 1, 2, \tag{1}$$

where $\mathbf{\Gamma}^{(i)}$ is a diagonal matrix containing eigenvalues arranged in the descending order: $\gamma^{(i)}_1 \geq \gamma^{(i)}_2 \geq \cdots \geq \gamma^{(i)}_n = 0$, and $\mathbf{V}^{(i)} \in \mathbb{R}^{N\times N}$ is an orthogonal matrix containing the corresponding eigenvectors. Since $\mathbf{L}^{(1)}$ is unknown, its eigendecomposition in (1) is subject to unknown permutations.

We observe a set of *filtered graph signals*, denoted by $\{\mathbf{z}_m\}^M_{m=1}$, generated over the nodes of $\mathcal{G}_1$ through an unknown graph filter. This filter can be expressed as a specific matrix polynomial of the Laplacian matrix $\mathbf{L}^{(1)}$ as

$$\mathcal{H}(\mathbf{L}^{(1)}) = \sum_{t=0}^{T_f-1} h_t(\mathbf{L}^{(1)})^t = \mathbf{V}^{(1)}\left(\sum_{t=0}^{T_f-1} h_t(\mathbf{\Gamma}^{(1)})^t\right)(\mathbf{V}^{(1)})^T, \tag{2}$$

where T_f is the degree of the graph filter, and $\{h_t\}$ are the filter coefficients. We assume that the filter has a finite spectral norm, i.e., $\|\mathcal{H}(\mathbf{L}^{(1)})\|_2 \leq \infty$. With (2), each observed signal vector $\mathbf{z}_m \in \mathbb{R}^{N\times 1}, \forall 1 \leq m \leq M$, is the output of $\mathcal{H}(\mathbf{L}^{(1)})$ being excited by an input signal $\mathbf{x}_m \in \mathbb{R}^{N\times 1}$, as

$$\mathbf{z}_m = \mathcal{H}(\mathbf{L}^{(1)})\mathbf{x}_m + \mathbf{w}_m, \tag{3}$$

where $\mathbf{w}_m$ represents the unknown measurement noise for the m-th sample following the Gaussian distribution $\mathcal{N}(\mathbf{0}, \nu^2\mathbf{I}_n)$. We assume that $\mathbf{x}_m$ is independent and identically distributed (i.i.d.) satisfying $\mathbb{E}[\mathbf{x}_m] = \mathbf{0}$ and $\mathbb{E}[\mathbf{x}_m\mathbf{x}^T_m] = \mathbf{C}_x$ with a known covariance $\mathbf{C}_x$ [17], [28], [29].

The generative model in (3) is applicable to a variety of real-world applications. Two illustrative examples follow.

Example 1 (Opinions over social networks)**.** In a social network, users' opinions or beliefs of certain topics can be modeled as filtered graph signals as described in (3). Specifically, the dynamics of users' opinions among their neighbors are characterized through a graph filter that models the propagation of opinions across the network. The opinion of the stubborn agents on the m-th topic, where $1 \leq m \leq M$, can be represented by the excitation signals $\mathbf{x}_m$, which are typically undisclosed. In contrast, the steady-state (equilibrium) opinions, after long-term propagation, can be represented as the filtered graph signals given by

$$\mathbf{z}_m = (\mathbf{I}_N + \zeta\mathbf{L}^{(1)})^{-1}\mathbf{x}_m, \tag{4}$$

where $\zeta > 0$ is a parameter determined by individuals' trust toward others and susceptibility to external influences. For more details on this graph signal model, we refer interested readers to [12], [30], [31]. The expression in (4) corresponds to that in (3) with the low-pass filter $\mathcal{H}(\mathbf{L}^{(1)}) = (\mathbf{I}_N+\zeta\mathbf{L}^{(1)})^{-1}$.

Example 2 (Phasor measurements in power grids)**.** A power grid system can be conceptualized as a weighted graph that connects buses (nodes) with transmission lines (edges). By defining the graph Laplacian matrix $\mathbf{L}^{(1)}$ (also known as the graph shift operator) as the effective *system admittance matrix*, the relationship between current and voltage phasors can be framed in terms of graph signals as in (3). Specifically, denote the current and voltage phasor measurements at the m-th time slot in an N-bus power system as $\boldsymbol{i}_m \in \mathbb{C}^{N\times 1}$ and $\boldsymbol{v}_m \in \mathbb{C}^{N\times 1}$, respectively. According to Ohm's law, the relationship between these phasors is governed by:

$$\boldsymbol{i}_m = \mathbf{Y}\boldsymbol{v}_m + \mathbf{w}_m, m = 1, \cdots, M, \tag{5}$$

where $\mathbf{Y} \in \mathbb{C}^{N\times N}$ is the *complex symmetric* (not Hermitian) system admittance matrix satisfying $\mathbf{Y} = \mathbf{Y}^T$, and $\mathbf{w}_m$ is the measurement noise. Within this framework, the current and voltage phasors are interpreted as the filtered and excitation signals, respectively, with the graph filter represented by $\mathcal{H}(\mathbf{L}^{(1)}) = \mathbf{L}^{(1)}$. For further details on phasor measurement modeling in this context, we refer to [32], [33].

While filtered graph signals are frequently observable across graph nodes, the accessibility of excitation signals differs across various applications. For instance, in Example 1, the excitation signals are often unobtainable, whereas in Example

[1]Here, it is assumed without loss of generality that the reference graph has the same number of nodes as $\mathcal{G}_1$. This condition can be satisfied by adding isolated dummy nodes to the graph with fewer nodes if necessary.

2 the excitations can be measured using phasor measurement units. The de-anonymization framework proposed in Section III accommodates both scenarios, addressing cases where the excitations are either known or unknown to the attacker.

Remark 1. Some graph signal processing models, like the one in Example 2, operate within the complex domain, where both the graph signals and the graph Laplacian are complex-valued. Our analysis and algorithm are easily extendable to complex-valued systems by substituting the eigendecomposition of real matrices with that of complex symmetric matrices, as shown in Section VII-C. However, for simplicity and clarity in our exposition, we focus on real-valued signals unless specified otherwise.

B. Order of Graph Frequency Responses

Combining (1) and (2), the eigenvalues of $\mathcal{H}(\mathbf{L}^{(i)})$, often referred to as the frequency responses, are given by

$$\tilde{h}_k = \sum_{t=0}^{T_f-1} h_t(\gamma_k^{(1)})^t, k \in [N]. \tag{6}$$

While $\{\gamma_k^{(1)}\}_{k=1}^K$ is aligned in descending order, the order of the frequency responses $\{\tilde{h}_k\}_{k=1}^N$ depends on the filter coefficients, which characterize the filter's trends. For instance, low-pass graph filters, like the opinion dynamic filter in Example 1, typically emphasize frequency responses at lower graph frequencies. Conversely, high-pass graph filters, such as the one in Example 2, amplify higher graph frequencies. We refer to [12] and [28, Sect. III-A] for more discussions on these filters.

A notable observation is that the eigenvectors of $\mathcal{H}(\mathbf{L}^{(1)})$ in (2) can be represented by the rearrangement of $\mathbf{V}^{(1)}$ in (1) given by

$$\mathcal{H}(\mathbf{L}^{(1)}) = \widetilde{\mathbf{V}}^{(1)}\mathbf{\Lambda}(\widetilde{\mathbf{V}}^{(1)})^T, \tag{7}$$

where $\mathbf{\Lambda} = \mathrm{diag}([\lambda_1, \ldots, \lambda_N])$ aligns the eigenvalues in descending magnitude, with λ_k representing the reshuffled arrangement of $\tilde{h}_k$, and $\widetilde{\mathbf{V}}^{(1)}$ includes the corresponding reordered eigenvectors from $\mathbf{V}^{(1)}$. Meanwhile, the eigendecomposition of the outer product of $\mathcal{H}(\mathbf{L}^{(1)})$, denoted by $\mathbf{C}_H$, is given by

$$\mathbf{C}_H = \mathcal{H}(\mathbf{L}^{(1)}) \left(\mathcal{H}(\mathbf{L}^{(1)})\right)^T = \widetilde{\mathbf{V}}^{(1)}(\mathbf{\Lambda})^2(\widetilde{\mathbf{V}}^{(1)})^T, \tag{8}$$

where $(\mathbf{\Lambda})^2$ is the square of $\mathbf{\Lambda}$.

Although the knowledge of $\mathbf{L}^{(1)}$ is unavailable, it is possible to utilize the observed graph signals to estimate the eigenmodes of $\widetilde{\mathbf{V}}^{(1)}$ through (7) or (8). However, the reshuffling between $\{\lambda_k\}_{k=1}^N$ and $\{\tilde{h}_k\}_{k=1}^N$ prevents the direct use of the estimated eigenmodes for matching $\mathbf{V}^{(1)}$ and $\mathbf{V}^{(2)}$, unless the shuffling order of the frequency response is known. In this work, we assume that the attacker knows the ordering of the frequency responses, though not their exact values. Specifically, the interchange between $\{\lambda_k\}_{k=1}^N$ and $\{\tilde{h}_k\}_{k=1}^N$ is determined by an index mapping function $\mathrm{ord}(\cdot) : [N] \to [N]$, such that

$$\lambda_k = \tilde{h}_{\mathrm{ord}(k)}, \forall k. \tag{9}$$

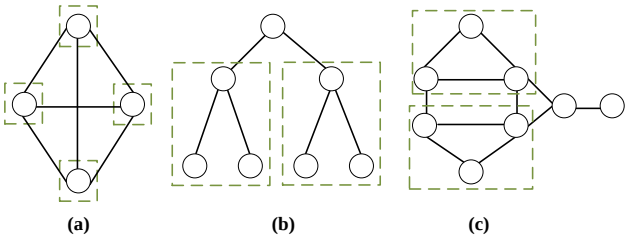

Fig. 1: Examples of symmetric graphs, where circles denote nodes and lines denote edges. The nodes in green boxes represent the symmetric nodes exhibiting identical inner and outer structures.

This assumption is considerably less stringent than knowing the exact values of the responses. In practice, the order of the filter frequency responses can be inferred from the characteristics of the graph filter, like being low-pass or high-pass, which is often known based on models for specific applications. For instance, in the context of de-anonymizing social networks using opinion measurements, as discussed in Example 1, these opinions are generated through a *low-pass* graph filter. Consequently, the index mapping is specified as $\lambda_k = \tilde{h}_{\mathrm{ord}(k)} = \tilde{h}_{N+1-k}$ for all $1 \le k \le N$. This determination of the correct order stems directly from the low-pass nature of the filter, eliminating the need to know its precise value.

C. Impact of Graph Symmetry

Even with complete knowledge of the graph topology of $\mathcal{G}_1$, attackers, in the absence of further information, are unable to conclusively re-identify the true labels of its *symmetric* nodes, i.e., nodes exhibiting identical inner structures and outer connections [9], [34]. The definition of symmetric nodes follows.

Definition 1 (cf. [35])**.** Two nodes i, j are considered symmetric (a.k.a. automorphically equivalent) if a node-swapping function σ with $\sigma(i) = j$, $\sigma(j) = i$, and $\sigma(k) = k, \forall k \neq i, j$, is an automorphism of the graph that yields an isomorphic (equivalent) graph.

Figure 1 shows examples of symmetric nodes. We use the notation $i \sim j$ to indicate that nodes i and j are symmetric. For a graph $\mathcal{G}$ with N nodes, we define the set of all symmetric nodes as

$$\mathcal{S}(\mathcal{G}) \triangleq \{i \in [N] : \exists j \in [N], j \neq i, i \sim j\}. \tag{10}$$

In contrast, the set of asymmetric nodes is denoted by $\mathcal{AS}(\mathcal{G}) \triangleq [N] \setminus \mathcal{S}(\mathcal{G})$, where $\setminus$ represents the set difference operation. A graph with symmetric nodes (often termed a symmetric graph) exhibits at least one non-trivial graph automorphism, inherently introducing permutation ambiguities in computing the node matching $\sigma(\cdot)$ for these nodes.

Existing research on graph de-anonymization and graph matching have explored node matching by either tolerating permutation ambiguities on symmetric nodes in matching outputs [9] or by focusing solely on asymmetric graphs [28]. In contrast, our work aims to achieve precise de-anonymization

of all *asymmetric* nodes within a general graph, which might include symmetric structures. However, we note that by aggregating each set of equivalently symmetric nodes as a singular, uniform entity, a symmetric graph can be converted into an asymmetric subgraph. This transformation allows our proposed method and analysis to extend seamlessly to scenarios where we also want to identify the equivalence set of a symmetric node under the automorphism group in the unknown graph.

III. Problem Statement

We study a general de-anonymization attack problem where the attacker aims to infer the labels of *all the asymmetric nodes* in $\mathcal{AS}(\mathcal{G}_1)$ using the M observed graph signals in (3). This challenge essentially translates to matching the nodes of the known graph $\mathcal{G}_2$ with the signals generated over the unknown graph $\mathcal{G}_1$, namely the GS2GM problem.

Unless stated otherwise, we assume in our analysis that $\mathcal{G}_2$ is a graph isomorphic to $\mathcal{G}_1$, i.e., $\mathcal{G}_1$ and $\mathcal{G}_2$ have an identical Laplacian matrix subject to an unknown optimal node permutation denoted by $\sigma^\star(\cdot)$ and its corresponding permutation matrix by $\mathbf{P}^\star \in \{0,1\}^{N\times N}$. The impact of graph non-isomorphism will be numerically examined in Section VII-B. Due to measurement noise and the limitations of finite signal sampling, the graph matching derived from the graph signals, denoted by $\widehat{\sigma}(\cdot)$ or $\widehat{\mathbf{P}} \in \{0,1\}^{N\times N}$, typically deviates from the true permutation, even with two isomorphic graphs. Our aim is to find an accurate $\widehat{\sigma}(\cdot)$ under the following two scenarios:

Problem 1. Compute $\widehat{\sigma}(\cdot) : [N] \to [N]$ with the given Laplacian matrix $\mathbf{L}^{(2)}$ and only the filtered graph signal $\{\mathbf{z}_m\}_{m=1}^M$ in (3).

Problem 2. Find $\widehat{\sigma}(\cdot)$ given $\mathbf{L}^{(2)}$ and both the filtered and excitation graph signals $\{\mathbf{z}_m, \mathbf{x}_m\}_{m=1}^M$.

While the GS2GM problem primarily focuses on de-anonymizing the labels of asymmetric nodes, the solution $\widehat{\sigma}(\cdot)$ invariably produces an N-node permutation that maps *all* the nodes of $\mathcal{G}_2$ to $\mathcal{G}_1$. Nevertheless, the presence of non-trivial automorphisms related to symmetric nodes renders the accuracy assessment of their de-anonymization less meaningful. Consequently, for the purpose of evaluating matching accuracy, we will discuss a method to exclude symmetric nodes in $\mathcal{G}_2$ when introducing our de-anonymization approach. We emphasize that the process of detecting asymmetric nodes does not impact the fundamental functioning of our de-anonymization algorithm; it merely acts as an optional tool to facilitate the numerical evaluation of matching performance.

As we shall demonstrate in Section IV, our approaches for addressing the problems mentioned above differ in the methodology for estimating the eigenmodes of $\mathbf{L}^{(1)}$, utilizing either the model in (7) or (8). We shall introduce a unified GS2GM framework in Section IV and subsequently detail the specific eigenmode estimation technique for each problem.

Remark 2. We highlight that most of the existing work bases the de-anonymization analysis on specific probabilistic graph models, such as ER random graphs [2], [6] and stochastic block models [7], [8], [36]. As such models might not adequately capture the complexities inherent in real-world networks, our GS2GM framework is applicable to de-anonymizing an arbitrarily fixed graph $\mathcal{G}_1$ without the need of assuming its statistical model.

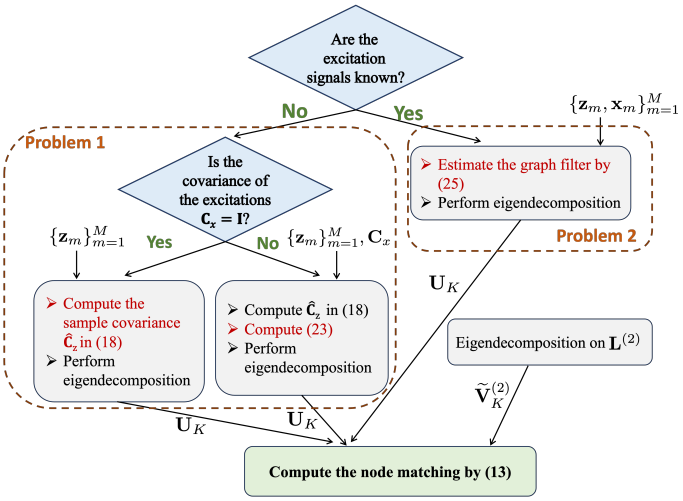

Fig. 2: Schematic view of the GS2GM method.

A. Matching Accuracy Measurement

Before delving into our methodology, we define the metric for measuring the accuracy of the node matching as follows.

Definition 2 (Success probability and asymptotic perfect de-anonymization)**.** Given a set of graph signals and a fixed $\mathbf{L}^{(2)}$, the resultant permutation $\widehat{\sigma}(\cdot)$ successfully re-identifies the asymmetric nodes of $\mathcal{G}_1$ if $\widehat{\sigma}(n) = \sigma^\star(n)$ for $\forall n \in \mathcal{AS}(\mathcal{G}_2)$.

Moreover, for any given possible instance of the graph signals in (3), the *uniform* recovery probability of GS2GM is defined as

$$\Pr(\widehat{\sigma}(n) = \sigma^\star(n), \forall n \in \mathcal{AS}(\mathcal{G}_2)). \tag{11}$$

Furthermore, $\widehat{\sigma}(\cdot)$ is said to achieve asymptotic perfect de-anonymization if

$$\lim_{M\to\infty} \Pr(\widehat{\sigma}(n) = \sigma^\star(n), \forall n \in \mathcal{AS}(\mathcal{G}_2)) = 1, \tag{12}$$

where M denotes the sampling size of the graph signals.

IV. GS2GM Algorithm

A. Overview

Our GS2GM approach is drawn inspiration from [28], [37]. In cases where both $\mathbf{L}^{(1)}$ and $\mathbf{L}^{(2)}$ are known, [37] computes $\widehat{\sigma}(\cdot)$ by maximizing the inner product of the eigenvector matrices $\mathbf{V}^{(1)}$ and $\mathbf{V}^{(2)}$ subject to a node permutation. This technique is referred to as spectral graph matching. Building upon this, [28] extends spectral graph matching to blind matching of two sets of graph signals generated over two unknown graphs, where the matching relies on empirical eigenvectors estimated from sample covariance matrices.

Although the GS2GM problem shares significant similarities with blind graph matching as we described [28], there are two notable distinctions: First, the blind method [28] was initially tailored for matching two unknown asymmetric graphs, presuming an absence of symmetric nodes. As we expound in Section III, it is crucial to exclude symmetric nodes for accurate de-anonymization assessment of asymmetric nodes for general graphs. In alignment with our objective, we incorporate a symmetry detection process using the reference graph $\mathcal{G}_2$ to facilitate numerical evaluation of de-anonymization accuracy. Second, while [28] deals with unknown and uncorrelated excitation signals, our problem formulation broadens this scope to include scenarios with more general excitations that may exhibit spatial correlation and can be either known or unknown. Therefore, our strategies for estimating the eigenmodes of $\mathbf{L}^{(1)}$ are adapted based on the specific information available about the graph signals.

In this section, we present a unified GS2GM approach that identifies node matching using the estimated eigenvectors $\widetilde{\mathbf{V}}^{(1)}$ by either (7) or (8). We outline the algorithmic framework in Figure 2 and the key steps as follows, leaving the details on the eigenmode estimation method for each problem to subsequent sections.

- **Eigenvector Estimation**: The initial step involves computing $\widetilde{\mathbf{V}}^{(1)}$ as specified in (7) and (8) from the observed graph signals. Specifically, for Problem 1, where only $\{\mathbf{z}_m\}_{\forall m}$ are observed, we estimate $\mathbf{C}_H$ in (8) by using $\{\mathbf{z}_m\}_{\forall m}$ and $\mathbf{C}_x$, and then determine its eigenmatrix by (8). In contrast, for Problem 2, where both filtered and excitation graph signals $\{\mathbf{z}_m, \mathbf{x}_m\}_{m=1}^M$ are known, we first estimate $\mathcal{H}(\mathbf{L}^{(1)})$ through *graph filter inference*, followed by performing eigendecomposition on it. We refer to Sections V and VI for more details. In both scenarios, the estimate of $\widetilde{\mathbf{V}}^{(1)}$ is denoted by $\mathbf{U}$. The corresponding estimates of eigenvalues in $\mathbf{\Lambda}$ (or $(\mathbf{\Lambda})^2$) are represented as $\widehat{\mathbf{\Lambda}} = \mathrm{diag}(\hat{\lambda}_1, \ldots, \hat{\lambda}_N)$ (or $(\widehat{\mathbf{\Lambda}})^2 = \mathrm{diag}(\hat{\lambda}_1^2, \ldots, \hat{\lambda}_N^2)$). Additionally, the eigenmatrix $\mathbf{V}^{(2)}$ of $\mathcal{G}_2$ is obtained from the eigendecomposition in (1). To align the two eigenmatrices, we reorder the eigenvectors in $\mathbf{V}^{(2)}$ to match the order in $\mathbf{U}$ according to the reshuffling order of the graph frequency responses as in (9). This reordered eigenmatrix is denoted by $\widetilde{\mathbf{V}}^{(2)} = [\mathbf{v}^{(2)}_{\mathrm{ord}(1)}, \cdots, \mathbf{v}^{(2)}_{\mathrm{ord}(N)}]$.
- **Node matching**: Following [28], we compute the node permutation matrix $\widehat{\mathbf{P}}$ by aligning the first K eigenvectors in $\mathbf{U}$ and $\widetilde{\mathbf{V}}^{(2)}$. The hyper-parameter K can be chosen to maximize the minimum spectral gap in the estimated eigenvalues as suggested in [28, Sect. IV-D], or tuned through trial and error. Furthermore, to address the unknown sign ambiguities inherent in eigendecompositions, we compute the permutation by taking the absolute values of these eigenvectors: (cf. [28, Eq. (14)])

$$\widehat{\mathbf{P}} = \underset{\mathbf{P}\in\mathcal{P}_N}{\arg\max}\, \mathrm{tr}\left(\overline{\mathbf{U}}_K (\overline{\widetilde{\mathbf{V}}}^{(2)}_K)^T \mathbf{P}\right), \tag{13}$$

  where $\mathcal{P}_N$ is the set containing all the $N \times N$ permutation matrices, $\overline{\mathbf{U}}_K$ and $\overline{\widetilde{\mathbf{V}}}^{(2)}_K$ are the matrices containing the absolute values of the first K columns of $\mathbf{U}$ and $\widetilde{\mathbf{V}}^{(2)}$, respectively. The linear assignment problem in (13) can be solved by off-the-shelf algorithms, such as the Hungarian method [38] or the greedy approach in [28, Algorithm 2].
- **Symmetry detection**: Evaluating the accuracy of $\widehat{\mathbf{P}}$ in (13) according to (11) requires the knowledge of the asymmetric node set $\mathcal{AS}(\mathcal{G}_2)$. However, for an N-node graph with as many as $N!$ node permutations in $\mathcal{P}_N$, the exhaustive search for symmetric nodes – a task known as the graph automorphism problem – becomes computationally infeasible for large N. As an alternative, we propose a polynomial-time method to identify a *subset* of $\mathcal{S}(\mathcal{G}_2)$ (or equivalently, a superset of $\mathcal{AS}(\mathcal{G}_2)$). For any node pair i, j, let $\mathbf{P}^{(i,j)}$ be the swapping matrix that swaps the i-th and j-th columns of $\mathbf{I}_N$. Nodes i and j are symmetric if the following holds:

$$\mathbf{L}^{(2)} = (\mathbf{P}^{(i,j)})^T \mathbf{L}^{(2)} \mathbf{P}^{(i,j)}. \tag{14}$$

  We include nodes i and j in the estimated symmetric node set $\widetilde{\mathcal{S}}(\mathcal{G}_2)$ if (14) is satisfied. This involves evaluating a total of $N(N+1)/2$ node pairs. As (14) identifies symmetric nodes subject to only a single swap, $\widetilde{\mathcal{S}}(\mathcal{G}_2)$ is a subset of $\mathcal{S}(\mathcal{G}_2)$. To refine this further, symmetric nodes involved in multiple swaps can be iteratively detected by checking the automorphism condition in (14) using swapping matrices that swap three or more columns. Since symmetry detection only serves the purpose of assessing de-anonymization accuracy, one should strike a balance between computational complexity and accuracy in computing $\widetilde{\mathcal{S}}(\mathcal{G}_2)$. As we demonstrate in Section VII-A, in practice most symmetric nodes in real-world social networks can be effectively detected through the single-swap automorphism check described in (14).

B. Success Probability Analysis

We analyze the success probability in (11) of the proposed method by examining the impact of eigenmode estimation error on the graph matching accuracy. To begin, we use $\overline{\widetilde{\mathbf{V}}}^{(1)}_K$ to represent the matrix containing the absolute values of the left K columns of $\widetilde{\mathbf{V}}^{(1)}$ in (7). Essentially, $\overline{\widetilde{\mathbf{V}}}^{(1)}_K$ is the error-free equivalent of $\overline{\mathbf{U}}_K$ in (13) with perfect eigenvector estimation. Let $X_{i,j}$ be the (i,j)-th entry of $\overline{\widetilde{\mathbf{V}}}^{(1)}_K (\overline{\widetilde{\mathbf{V}}}^{(2)}_K)^T$. Finally, we define an auxiliary variable ρ as

$$\rho \triangleq \min_{n\in\mathcal{AS}(\mathcal{G}_2)} \left(X_{n,\sigma^\star(n)} - \max_{\ell\neq\sigma^\star(n)} X_{n,\ell} \right). \tag{16}$$

As we showed [28], ρ quantifies the cost in the objective value of (13) with respect to the true node matching $\sigma^\star(\cdot)$ in the absence of errors. The next theorem shows that the condition on $\rho > 0$ offers a way to characterize the success matching probability in relation to the eigenvector estimation error.

Theorem 1. Suppose $\rho > 0$ and all the eigenvalues $\{\lambda_k\}_{k=1}^K$ in (7) are distinct. Consider the solution $\widehat{\sigma}(\cdot)$, or equivalently $\widehat{\mathbf{P}}$ in (13), which estimates the eigenvector by (8) for Problem 1, satisfies (15) shown on top of this page. Here, $\mathbf{u}_k$ and $\widetilde{\mathbf{v}}^{(1)}_k$ are the k-th columns of $\mathbf{U}$ and $\widetilde{\mathbf{V}}^{(1)}$, respectively, and δ_k is the k-th spectral gap of the eigenvalues $\{\lambda_k^2\}$ defined as

$$\delta_k = \min\{\lambda_k^2 - \lambda_{k+1}^2, \lambda_{k-1}^2 - \lambda_k^2\}, \forall k \in [K]. \tag{17}$$

$$\Pr\left(\widehat{\sigma}(n)=\sigma^{\star}(n),\forall n\in\mathcal{AS}(\mathcal{G}_2)\right)\geq\prod_{k=1}^{K}\left(1-\Pr\left(|\hat{\lambda}_k^2-\lambda_k^2|\geq\frac{\delta_k}{2}\right)\right)\cdot\left(1-\Pr\left(\rho\leq 2\sqrt{2\sum_{k=1}^{K}(1-(\mathbf{u}_k^T\widetilde{\mathbf{v}}_k^{(1)})^2)}\right)\right). \tag{15}$$

For the proposed solution to Problem 2, the same bound in (15) holds, except that λ_k^2 and $\hat{\lambda}_k^2$ should be substituted with λ_k and $\hat{\lambda}_k$, respectively.

Proof. See Appendix A. □

As discussed in [28], [37], the distinctness of the eigenvalues $\{\lambda_k\}$ is a prerequisite for the spectral graph matching methods. Moreover, the condition $\rho>0$ implies that the true permutation $\sigma^{\star}(\cdot)$ uniquely maximizes the objective in (13) when $\mathbf{V}^{(1)}$ is known precisely. Prompted by this, we use ρ as a metric for assessing the de-anonymizability of graphs in the noiseless setting. The value of ρ is inherently linked to the edge connectivity of the underlying graphs. For instance, [39] examined a variant objective to (16) and found that a similar metric to ρ is positive for large Gaussian models and large ER graphs. Building on this insight, we conjecture that ρ in (16) is also likely to be positive for large ER graphs and large Gaussian models, as corroborated numerically in Section VII. In a general case involving an arbitrary and unknown graph $\mathcal{G}_1$, ρ can be approximately evaluated by substituting $\mathbf{V}^{(1)}$ with the noisy estimate $\mathbf{U}$. This approximation offers a pragmatic way to gauge the de-anonymization success probability, particularly for large-scale systems.

Theorem 1 demonstrates that perfect matching is achievable with our algorithm, provided the spectrum order is correctly estimated and eigenvector perturbations are limited. This robustness against noise stems from two key factors: 1) the approach requires only an accurate estimation of the order of the eigenvalues to distinguish the spectral components, not their exact values, and 2) it can tolerate certain levels of perturbations in eigenvector estimation, where the threshold for acceptable perturbations depends on the spectral gaps of the graph filter. This advantage highlights the efficacy of utilizing the spectral components of the graph filter to deduce topological information for graph matching.

V. Eigenvector Estimation for Problem 1

A. Covariance-Based Eigenvector Estimation

We study the eigenvector estimation problem under the scenarios lacking direct access to the excitation signals $\{\mathbf{x}_m\}_{m=1}^M$ as delineated in Problem 1. We start by considering a situation where the unknown excitation signals are uncorrelated and unit-variance satisfying $\mathbb{E}[\mathbf{x}_m\mathbf{x}_m^T]=\mathbf{C}_x=\mathbf{I}_N$. In this case, the sample covariance of the observed signals $\{\mathbf{z}_m\}_m$ and its eigendecomposition can be computed by

$$\widehat{\mathbf{C}}_z=\frac{1}{M}\sum_{m=1}^{M}\mathbf{z}_m(\mathbf{z}_m)^T-\nu^2\mathbf{I}_N=\mathbf{U}^{(1)}(\widehat{\mathbf{\Lambda}})^2(\mathbf{U}^{(1)})^T, \tag{18}$$

where $\mathbf{U}$ is orthogonal containing the sample eigenvectors, and $(\widehat{\mathbf{\Lambda}})^2=\mathrm{diag}([\hat{\lambda}_1^2,\cdots,\hat{\lambda}_N^2])$ positions the sample eigenvalues in descending order. In (18), we adjust the sample covariance estimation by using the variance of the signal noise to ensure $\widehat{\mathbf{C}}_z$ is an unbiased estimator of $\mathbf{C}_H$ in (8), i.e., $\mathbb{E}[\mathbf{C}_z]=\mathbf{C}_H$. Consequently, with a sufficiently large sampling size M, we anticipate that $\mathbf{C}_z$ is close to $\mathbf{C}_H$ and thereby $\mathbf{U}$ and $\{\hat{\lambda}_n^2\}$ provide an approximation close to $\widetilde{\mathbf{V}}^{(1)}$ and $\{\lambda_n^2\}$ in (8), respectively. Finally, the sample eigenvector $\mathbf{U}$ is used for computing the graph matching in (13).

The complexity of the associated matrix multiplication in (18) scales with $\mathcal{O}(N^2M)$. Additionally, the complexity associated with the eigendecomposition and solving the linear assignment problem in (13) is $\mathcal{O}(N^3)$. Consequently, the total computational complexity of the proposed method is $\mathcal{O}(N^3+N^2M)$.

B. Performance Analysis

Applying Theorem 1, we present the following proposition to quantify the success probability of the de-anonymization method using (18).

Proposition 1. Let the following conditions hold:

i. $\rho>0$.
ii. The excitation signals $\{\mathbf{x}_m\}$ follow an i.i.d. zero-mean sub-Gaussian distribution and satisfy $\mathbb{E}[\mathbf{x}_m\mathbf{x}_m^T]=\mathbf{I}_N$.
iii. The filtered graph signals $\{\mathbf{z}_m\}$ are uniformly bounded above almost surely, i.e., $\|\mathbf{z}_m\|_2\leq Z$ for some $Z<\infty$.

Then, there exists a constant M_0 such that for any $M\geq M_0$ we have

$$\begin{aligned}&\Pr\left(\widehat{\sigma}(n)=\sigma^{\star}(n),\forall n\in\mathcal{AS}(\mathcal{G}_2)\right)\\&\quad\geq 1-\frac{4KZ^4}{M\delta^2}-2Ne^{-\frac{M\rho^2\delta^2}{32C^2(\nu^2+Z)KN}},\end{aligned} \tag{19}$$

where ν^2 is the signal noise variance, C is a constant given in (40), and $\delta=\min_{k\in[K]}\delta_k$ is the minimum spectral gap defined in (17).

Proof. See Appendix B. □

As suggested by [28, Section IV-D], the value of K can be chosen such that δ^2/K remains constant. Consequently, it follows from Proposition 1 that the error probability grows at the rate of $\mathcal{O}(Ne^{-M/N})$ for a large N. Accordingly, a sample size M proportional to $N\log N$ suffices for achieving a small de-anonymization error rate.

C. Extension to Cases with Correlated Excitations

The method in (18) is tailored for cases with uncorrelated excitations. When graph signals exhibit significant spatial correlations, the sample covariance $\widehat{\mathbf{C}}_z$ intertwines the graph filter with the excitation covariance and thus fails to precisely

estimate $\mathbf{C}_H$ in (8). To see this, we represent the excitation signal $\mathbf{x}_m$ in (3) as

$$\mathbf{x}_m = \mathbf{C}_x^{1/2}\widetilde{\mathbf{x}}_m, \tag{20}$$

where $\mathbf{C}_x^{\frac{1}{2}}$ is the principal square root of $\mathbf{C}_x$ such that $\mathbf{C}_x = \mathbf{C}_x^{\frac{1}{2}}\mathbf{C}_x^{\frac{1}{2}}$, and $\widetilde{\mathbf{x}}_m$ denotes the whiten signal satisfying $\mathbb{E}[\widetilde{\mathbf{x}}_m] = \mathbf{0}$ and $\mathbb{E}[\widetilde{\mathbf{x}}_m\widetilde{\mathbf{x}}_m^T] = \mathbf{I}_N$ for $\forall m$. Then, we have

$$\mathbf{C}_z = \mathbb{E}[\mathbf{z}_m\mathbf{z}_m^T] = \mathcal{H}(\mathbf{L}^{(1)})\mathbf{C}_x\mathcal{H}(\mathbf{L}^{(1)}). \tag{21}$$

As a result, the eigendecomposition of $\widehat{\mathbf{C}}_z$ in (18) does not lead to a precise estimate of the eigenmodes of $\mathbf{L}^{(1)}$.

To tackle this challenge, we assume that $\mathbf{C}_x$ is known *a prior*, e.g., from the statistical information on the excitations. Following [29], we have

$$\mathbf{C}_x^{\frac{1}{2}}\mathbf{C}_z\mathbf{C}_x^{\frac{1}{2}} = \mathbf{C}_x^{\frac{1}{2}}\mathcal{H}(\mathbf{L}^{(1)})\mathbf{C}_x\mathcal{H}(\mathbf{L}^{(1)})\mathbf{C}_x^{\frac{1}{2}} = (\mathbf{C}_x^{\frac{1}{2}}\mathcal{H}(\mathbf{L}^{(1)})\mathbf{C}_x^{\frac{1}{2}})^2. \tag{22}$$

Note that our aim is to find the eigenvectors of $\mathcal{H}(\mathbf{L}^{(1)})$ subject to sign ambiguities in the eigendecomposition. When $\mathbf{C}_x$ is non-singular, we propose to approximate $\mathbf{C}_z$ in (22) with $\widehat{\mathbf{C}}_z$ in (18) and then estimate the eigenvectors of $\mathcal{H}(\mathbf{L}^{(1)})$ by the eigenvectors of

$$\mathbf{C}_x^{-\frac{1}{2}}\left(\mathbf{C}_x^{\frac{1}{2}}\widehat{\mathbf{C}}_z\mathbf{C}_x^{\frac{1}{2}}\right)^{\frac{1}{2}}\mathbf{C}_x^{-\frac{1}{2}}. \tag{23}$$

When $\mathbf{C}_x$ is singular, it is impossible to find all the eigencomponents of $\mathcal{H}(\mathbf{L}^{(1)})$ since the measurement signals are observed through an ill-conditioned mapping. In this case, we can replace $\mathbf{C}_x^{-\frac{1}{2}}$ with the pseudo inverse of $\mathbf{C}_x^{\frac{1}{2}}$ in (23), and thereby estimate a subset of the eigenvectors.

We emphasize that the approach in (23) is effective because we only need the absolute values of the eigenvectors of the graph filter in (13), rather than the precise values of the filter matrix as explored previously in [29]. Consequently, there is no necessity to determine the exact values of the eigenvalues of the filter matrix, allowing us to disregard sign ambiguities within eigendecompositions without impacting the accuracy of our method.

VI. Eigenvector Estimation for Problem 2

A. Sparse Graph Filter Inference

In this section, we investigate the eigenmode estimation method for Problem 2, which deals with the scenario where both the excitation and filtered graph signals $\{\mathbf{z}_m, \mathbf{x}_m\}_{m=1}^M$ are known. Problem 2 could technically be considered a subset of Problem 1, with the approaches in Section V being applicable. However, those methods primarily leverage statistical information about the excitations rather than their explicit values, necessitating a large number of signal samples to accurately estimate covariance matrices. Alternatively, a heuristic method is to first estimate the graph Laplacian from the graph signals, known as graph topology inference [17], and then to derive the graph filter using the expression of $\mathcal{H}(\cdot)$. However, this two-step method often requires strong conditions on the underlying graph topology and can be susceptible to estimation error propagation.

To address the limitations of the aforementioned methods, we first estimate the graph filter matrix directly from the affine model presented in (3), termed as graph filter inference, and then compute its eigenvectors by (7). To proceed, we define $\mathbf{Z} \triangleq [\mathbf{z}_1, \cdots, \mathbf{z}_M] \in \mathbb{R}^{N\times M}$, $\mathbf{X} \triangleq [\mathbf{X}_1, \cdots, \mathbf{X}_M] \in \mathbb{R}^{N\times M}$, and $\tilde{\mathbf{z}} \triangleq \mathrm{vec}(\mathbf{Z}) \in \mathbb{R}^{NM}$. The model in (3) can be recast as

$$\begin{aligned}\tilde{\mathbf{z}} &= \mathrm{vec}(\mathcal{H}(\mathbf{L}^{(1)})\mathbf{X}) + \tilde{\mathbf{w}} \\ &= \underbrace{(\mathbf{X}^T \otimes \mathbf{I}_N)}_{\triangleq\mathbf{B}}\mathrm{vec}(\mathcal{H}(\mathbf{L}^{(1)})) + \tilde{\mathbf{w}},\end{aligned} \tag{24}$$

where $\tilde{\mathbf{w}} = \mathrm{vec}([\mathbf{w}_1, \cdots, \mathbf{w}_M])$ consists of entries following the distribution of $\mathcal{N}(0, \nu^2)$, and $\otimes$ denotes the Kronecker product. Consequently, $\mathcal{H}(\mathbf{L}^{(1)})$ can be estimated from the affine model in (24) given the known sensing matrix $\mathbf{B}$.

Graph filter inference often exploits structural information to enhance the accuracy of estimation [10], [17]. A frequently leveraged property is the *sparsity* of $\mathcal{H}(\mathbf{L}^{(1)})$, particularly relevant when the graph Laplacian matrix $\mathbf{L}^{(1)}$ is sparse and the polynomial degree order T_f of the filter is small. For example, graph filters like diffusion dynamic systems are typically sparse due to the localized nature of many diffusion processes in large-scale graphs. Inspired by this, we introduce the following assumption about the sparsity of $\mathcal{H}(\mathbf{L}^{(1)})$:

Assumption 1 (s-sparse graph filter)**.** The graph filter $\mathcal{H}(\mathbf{L}^{(1)})$ is s-sparse, i.e., $\|\mathrm{vec}(\mathcal{H}(\mathbf{L}^{(1)}))\|_0 \le s$ for some $s \le N^2$.

Under Assumption 1, inferring $\mathcal{H}(\mathbf{L}^{(1)})$ in (24) essentially becomes a sparse recovery problem. Its solution can be obtained through quadratically constrained ℓ_1-minimization:

$$\widehat{\mathcal{H}}(\mathbf{L}^{(1)}) = \underset{\mathbf{H}\in\mathbb{R}^{N\times N}}{\arg\min} \ \|\mathrm{vec}(\mathbf{H})\|_1 \tag{25a}$$

$$\text{subject to } \|\tilde{\mathbf{z}} - \mathbf{B}\mathrm{vec}(\mathbf{H})\|_2 \le \eta, \tag{25b}$$

$$\mathbf{H} \in \mathcal{F}, \tag{25c}$$

where $\eta > 0$ is a tuning parameter ensuring $\|\tilde{\mathbf{w}}\|_2 \le \eta$ with high probability, and $\mathcal{F}$ is the feasible set of the graph filter. Here, $\mathcal{F}$ incorporates prior knowledge about the graph filter, i.e., $\mathcal{H}(\cdot)$ is some polynomial function of a Laplacian matrix. Leveraging the properties of the Laplacian matrix, a typical choice of $\mathcal{F}$ is (cf. [17, Eq. (25)])

$$\mathcal{F} = \{\mathbf{H} = \mathcal{H}(\mathbf{L}) : \mathbf{L}^T = \mathbf{L}, \mathbf{L} \succeq \mathbf{0}, \mathbf{L}\mathbf{1} = \mathbf{0}\}. \tag{26}$$

Given that Laplacian matrices are symmetric, the optimization variable of (25) effectively has a dimension of $N(N+1)/2$. Since in our problem the graph Laplacian is unknown, we can replace $\mathcal{F}$ with a tractable convex approximation, such as

$$\mathcal{F}' = \{\mathbf{H} \in \mathbb{R}^{N\times N} : \mathbf{H}^T = \mathbf{H}, \mathbf{H} \succeq \mathbf{0}, \mathbf{H}\mathbf{1} = \mathbf{0}, \mathrm{tr}(\mathbf{H}) \le \alpha\}, \tag{27}$$

where $\alpha > 0$ is a pre-defined constant. The convex problem with (27) can be solved by off-the-shelf solvers, such as the interior point method [40]. After computing $\widehat{\mathcal{H}}(\mathbf{L}^{(1)})$, its eigenvectors are obtained by the eigendecomposition in (7).

Employing the interior point method to solve (25), which involves $\mathcal{O}(N^2)$ variables, results in a complexity of $\mathcal{O}(N^7)$ [40]. Moreover, the matrix multiplication, eigendecomposition, and the linear assignment problem collectively exhibit

a complexity of $\mathcal{O}(N^3 + N^2M)$. Consequently, the total complexity of the filter-inference-based method amounts to $\mathcal{O}(N^7 + N^2M)$. This method demonstrates a higher complexity compared to the covariance-based algorithm described in Section IV, reflecting the tradeoff of enhanced matching accuracy through the use of excitation information.

B. Performance Analysis

We analyze the performance of the solution in (25) within the framework of compressed sensing [41]. To simplify the analysis, we consider the computation of $\widehat{\mathcal{H}}(\mathbf{L}^{(1)})$ by minimizing (25) but omitting the constraint (25c), leading to a standard ℓ_2-norm-constrained ℓ_1-minimization problem. We note that removing the constraint (25c) enlarges the optimization dimension from $N(N+1)/2$ to N^2, and thus leads to a generally sub-optimal solution, even when the solution is later projected back onto $\mathcal{F}$. Consequently, the following analysis provides a sufficient condition for the proposed method to achieve asymptotic perfect de-anonymization.

To proceed, we introduce the s-th *restricted isometry constant (RIC)* of the sensing matrix $\mathbf{B} \in \mathbb{R}^{N^2\times N^2}$, denoted by $\varsigma_s(\mathbf{B}) \in [0,1]$, as the smallest $\varsigma \geq 0$ that ensures $(1-\varsigma)||\mathbf{u}||_2^2 \leq \|\mathbf{B}\mathbf{u}\|_2^2 \leq (1+\varsigma)||\mathbf{u}||_2^2$ for all s-sparse vectors $\mathbf{u} \in \mathbb{R}^{N^2}$. In situations where the context clearly refers to $\mathbf{B}$, we will omit the argument $\mathbf{B}$ and use ς_s to represent its RIC.

Proposition 2. Suppose the following conditions hold:

i. ρ in (16) satisfies $\rho > 0$.
ii. Assumption 1 holds with $\|\text{vec}(\mathcal{H}(\mathbf{L}^{(1)}))\|_0 \leq s$. Furthermore, $\|\text{vec}(\mathcal{H}(\mathbf{L}^{(1)}))\|_1 \leq B$ for some constant $B < \infty$, i.e., the true graph filter has a finite objective value to the problem in (25a).
iii. The RIC value $\varsigma_{2s}(\mathbf{B}) < 0.475$.
iv. Let $\delta = \arg\min_{k\in[K]} \min\{\lambda_k - \lambda_{k+1}, \lambda_{k-1} - \lambda_k\}$ be the minimum spectral gap of the first K eigenvalues of $\mathcal{H}(\mathbf{L}^{(1)})$. Define an auxiliary constant $\bar{\delta}$ as follows:

$$\bar{\delta} \triangleq \min\left\{\frac{1}{2}, \frac{\sqrt{2}\rho}{8\sqrt{K}}\right\}(0.19-0.4\varsigma_{2s})\delta - \frac{B(0.15\varsigma_{2s}+0.45)}{\sqrt{s}}. \tag{28}$$

The value of $\bar{\delta}$ is required to be bounded below by the noise variance as

$$\nu^2 \leq \frac{\bar{\delta}^2}{2N^2}. \tag{29}$$

Then, there exists some $\eta > 0$ such that the solution to (25) satisfies that

$$\Pr\left(\widehat{\sigma}(n) = \sigma^\star(n), \forall n \in \mathcal{AS}(\mathcal{G}_2)\right) \geq (1 - e^{-\frac{\bar{\delta}^2 - 2N^2\nu^2}{3\nu^2}})^{K+1}. \tag{30}$$

Additionally, suppose the following conditions also hold:

v. $\tilde{\mathbf{z}}$ in (24) is uniformly bounded above almost surely.
vi. $M \geq N$.
vii. K is chosen such that K/δ^2 remains constant.

Then, we further have

$$\Pr\left(\widehat{\sigma}(n) = \sigma^\star(n), \forall n \in \mathcal{AS}(\mathcal{G}_2)\right) \geq 1 - Ne^{-Const\cdot N^2(1-\varsigma_{2s})}, \tag{31}$$

Type	Sparsity	# of samples for a diminishing error
Sparse filter	$s = \mathcal{O}(N)$	$M \propto \log N$
Semi-sparse filter	$s = \mathcal{O}(N \log N)$	$M \propto (\log N)^2$
Dense filter	$s = \mathcal{O}(N^2)$	$M \propto N$

TABLE I: Comparison of the sufficient number of signal samples for achieving a diminishing error probability for different graph filters.

where $Const$ denotes an absolute constant.

Proof. See Appendix C. □

We note that the critical RIC value ς_{2s} in (31) depends on the excitation signal matrix $\mathbf{X}$, and thus on the sampling size M. To better understand the scaling order of the de-anonymization error probability in relation to these factors, we present the following result on ς_{2s}.

Proposition 3. Suppose $\{\mathbf{x}_m\}_{m=1}^M$ are i.i.d. sub-Gaussian random vectors satisfying $\mathbb{E}[\mathbf{x}_m] = \mathbf{0}$ and $\mathbb{E}[\mathbf{x}_m\mathbf{x}_m^T] = \mathbf{I}_N$ for $\forall m$. For any s such that $2s$ is divisible by N, we have $\varsigma_{2s}(\mathbf{B}) \leq t$ holds with probability at least $1 - e^{-Mt^2/(2C')}$, provided that

$$t^2 \geq \frac{4C's}{MN}\ln(eN^2/(2s)). \tag{32}$$

where C' is a constant defined in (55).

Proof. See Appendix D. □

Proposition 3 requires i.i.d. excitation signals. In practice, this condition can be approximately met by pre-whitening the graph signals. This lemma shows that the value of ς_{2s} decreases with the sample size M with high probability, and thus Conditions iii and iv in Proposition 2 can be met with a large M.

C. Impacts of Graph Filter Sparsity

Building upon the above analysis, we examine how filter sparsity impacts the accuracy of graph filter inference, which in turn affects the de-anonymization error probability. In Table I, we explore various scenarios where the graph filters—defining interactions among nodes—possess sparse, semi-sparse, and dense structures. These configurations are analogous to real-world network structures like trees, small-world networks, and complete graphs, respectively. The results indicate that leveraging the sparsity of the filter matrix enhances the accuracy of filter inference and consequently reduces the required signal sample size to achieve a diminishing error probability. Moreover, it is observed that the sample size required in these instances is considerably lower than that needed for the covariance-based method discussed in Section V, highlighting the advantage of incorporating knowledge of excitation signal values.

On the other hand, it is crucial to emphasize that the sparsity of the filter matrix and the sparsity of its eigenvectors $\mathbf{V}^{(1)}$ in (7) are related yet distinct concepts. A matrix with sparse entries often leads to sparse principal eigenvector(s).

Moreover, it has been proved in [42] that the presence of symmetric nodes results in the existence of sparse eigenvectors in $\widetilde{\mathbf{V}}^{(1)}$. We emphasize that while sparsity in the filter matrix, as utilized in the filter inference (25), can improve the estimation accuracy of its eigenmodes, eigenvector sparsity, especially in principal eigenvectors of the graph filter, adversely affects the accuracy of graph matching in the GS2GM problem (13). This negative impact is observed even when the graph filter and its eigenvectors are perfectly estimated. The reason is that sparse eigenvectors imply weak edge connectivity and/or graph symmetry and thus a small value of ρ. This, in turn, reduces the de-anonymizability of the corresponding matching problem.

Remark 3. While blind graph matching and graph filter inference have been explored separately in existing literature, our work extends beyond simply merging these methodologies. Specifically, our analysis provides a unified analytical framework that allows for a consistent comparison of different GS2GM approaches under the same metric of matching error probability. The results offer critical insights into: 1) the impact of graph filter estimation errors on de-anonymization accuracy, 2) the number of signal samples required to achieve accurate de-anonymization, and 3) the impact of the graph filter sparsity and spectral gaps on the de-anonymization performance.

VII. Experimental Results

In this section, we present experimental results to verify the performance of the proposed approach.

A. Accuracy of Symmetry Detection

We begin by analyzing the symmetric node detection mechanism discussed in Section IV-A. We consider the *Facebook* dataset from [43], which comprises ten real-world *ego-networks* with a collective total of $4,039$ users and includes ten principal users referred to as egos. In Table II, we recursively execute symmetric detection by using (14) to pinpoint all symmetric nodes subject to single or multiple swaps within each ego-network. Owing to the inherent small-world characteristic of these social networks, all symmetric nodes across the examined graphs are efficiently identifiable via single-swap permutations as articulated in (14).

B. Social Network De-anonymization

We explore the use of the proposed GS2GM method for de-anonymizing social network users with the observed graph signals, specifically under the conditions of Problem 1. Referring to Example 1, the diffused opinions in social networks can be represented as filtered graph signals, with the initial opinions (i.e., excitation signals) typically being unknown to the attacker. Our performance evaluation focuses on graphs derived from the following models:

- ER random graphs comprising $N = 50$ nodes and an edge connection probability of 0.3. As shown by Tran and Ahn [34], large ER graphs are asymmetric with high probability.
- The first ego network extracted from the *Facebook* dataset [43], which includes 334 users and $2,519$ edges representing user friendships. As demonstrated in Table II, this graph contains 45 symmetric nodes, and our de-anonymization accuracy assessment focuses on the remaining 289 nodes.

# of nodes	# of single-swap symmetric nodes	# of other symmetric nodes
334	45	0
1,035	15	0
225	9	0
151	11	0
169	5	0
62	11	0
787	35	0
748	10	0
535	33	0
53	12	0

TABLE II: Symmetric nodes in the *Facebook* ego-networks.

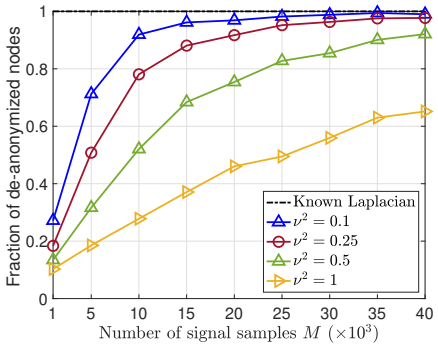

Fig. 3: De-anonymization accuracy versus the sample size M for the ER graph with uncorrelated excitation signals.

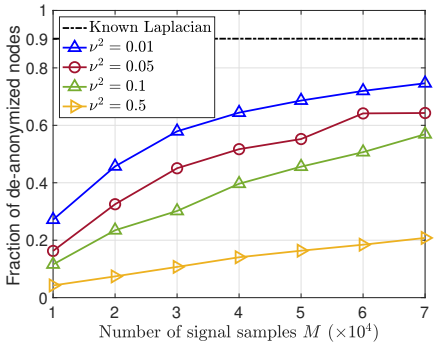

Fig. 4: De-anonymization of the *Facebook* ego-network with uncorrelated excitations.

The filtered graph signals are generated by (3) over the unknown graph $\mathcal{G}_1$ by employing a low-pass opinion-dynamic graph filter [12], as defined by $\mathcal{H}(\mathbf{L}^{(1)}) = (\mathbf{I}_N + 0.1\mathbf{L}^{(1)})^{-1}$. The excitation signals are formulated as per (20), with their covariance matrix $\mathbf{C}_x$ delineated later. We construct the reference graph $\mathcal{G}_2$ by randomly shuffling the nodes of $\mathcal{G}_1$. The linear assignment problem in (13) is solved by the Hungarian method [38].

We first consider the model with uncorrelated excitation signals, where each signal $\mathbf{x}_m$ is independently drawn from

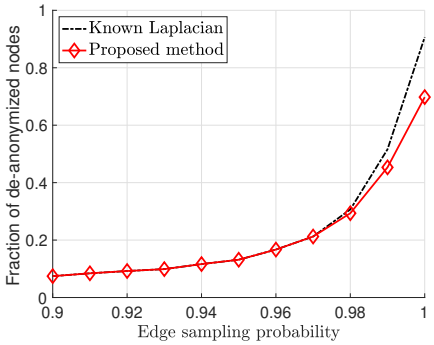

Fig. 5: De-anonymization accuracy versus edge sampling probability for graph matching with the non-isomorphic reference.

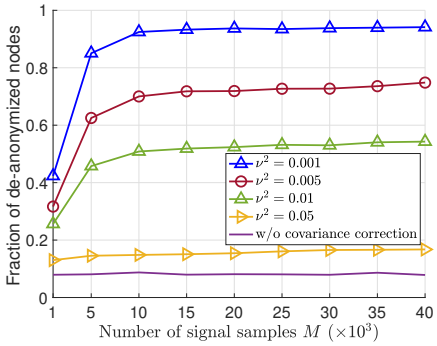

Fig. 7: De-anonymization of the ER graph with correlated excitations.

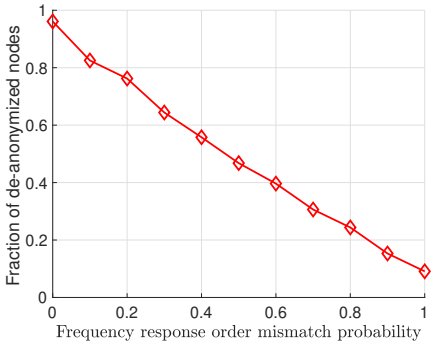

Fig. 6: Network de-anonymization accuracy versus the mismatch probability in the ordering of the filter frequency responses.

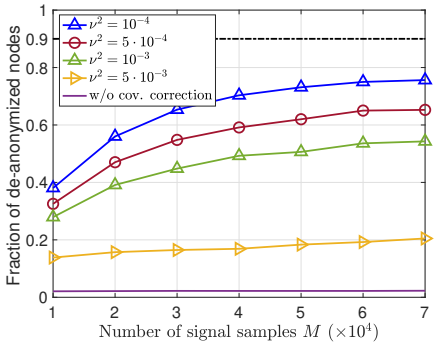

Fig. 8: De-anonymization of the *Facebook* ego-network with correlated excitations.

a standard normal distribution $\mathcal{N}(\mathbf{0}, \mathbf{I}_N)$. Figures 3 and 4 display the proportion of accurately de-anonymized nodes, within a range of $[0, 1]$, for the proposed approach in (18), applied to ER graphs and the *Facebook* ego-network. In these figures, we adjust the variance of the measurement noise ν^2 to demonstrate the impact of signal noise. An ideal baseline, assuming complete knowledge of the graph Laplacian $\mathbf{L}^{(1)}$ and employing the graph matching algorithm from [37], is included for comparison. This baseline delineates the optimal performance achievable with our method under perfect eigenvector estimation. The findings indicate that the error probability declines exponentially with an increase in the signal sample size M. Additionally, larger signal noise contributes to an increased matching error, aligning with the analytical result in Proposition 1. A critical insight from Figure 4 reveals that perfect de-anonymization is unattainable in the *Facebook* network, even with an error-free Laplacian matrix $\mathbf{L}^{(1)}$. This limitation arises due to the symmetric structures in the network, which results in the de-anonymizability metric ρ in (16) of approximately -0.45, thus not satisfying the condition of $\rho > 0$ in Theorem 1. It shows that our proposed method successfully de-anonymizes above 75% of the asymmetric nodes even though this network exhibits symmetry.

Next, we investigate a more challenging scenario where the reference graph $\mathcal{G}_2$ possesses only a subset of the information from $\mathcal{G}_1$, resulting in a non-isomorphic GS2GM challenge. We constructed $\mathcal{G}_1$ as an ER random graph and $\mathcal{G}_2$ through independent edge sampling from $\mathcal{G}_1$. Figure 5 plots the matching accuracy against the edge sampling probability, with $M = 5{\cdot}10^4$ and $\nu^2 = 0.01$. A pronounced decrease in de-anonymization accuracy is observed as the reference graph $\mathcal{G}_2$ presents reduced edge information. This highlights the critical impact of reference graph isomorphism on matching accuracy.

Recall from (9) that our method requires to know the correct order to align the eigenvectors corresponding to $\{\lambda_k\}_{k=1}^N$ and $\{\tilde{h}_k\}_{k=1}^N$. Here, we examine the sensitivity of the de-anonymization accuracy when the correct order is partially known. We simulate a scenario where the frequency responses may be incorrectly ordered among neighboring values. This incorrect ordering could occur when we have only a general understanding of the graph filter's global tendency (e.g., low-pass or high-pass), but local tendencies may not align. Specifically, we assume that for any $k = 1, 3, 5, \cdots$, each pair of the adjacent index ordering values $\text{ord}(k)$ and $\text{ord}(k+1)$ in (9) has a probability of $\frac{p_m}{2}$ of being swapped. Consequently, on average, a fraction p_m of the frequency response orders are incorrect. Figure 6 illustrates the de-anonymization accuracy of the proposed method as p_m varies for the ER random graphs. The other simulation parameters are the same as those in Figure 5. The results indicate that achieving accurate de-anonymization becomes challenging with significant order mismatches, highlighting the critical role of knowing the frequency response ordering.

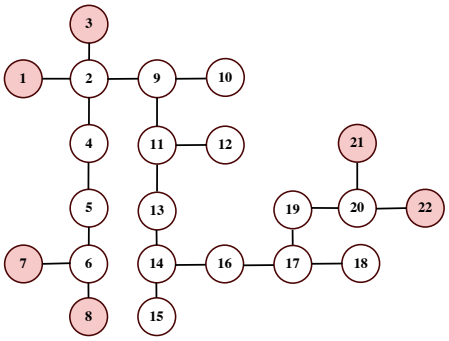

Fig. 9: Topology of the IEEE 22-bus system, where the numbers represent the labels of the nodes and the pink circles represent the symmetric nodes in the graph.

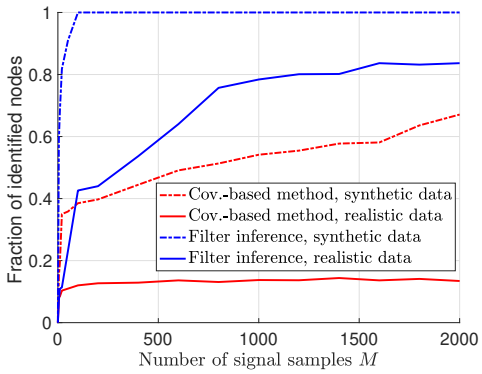

Fig. 10: Node de-anonymization accuracy in the power grid.

Finally, we study a model with *correlated* excitation signals by following [29]. Here, the excitation covariance is generated according to $\mathbf{C}_x = \mathbf{V}_x\mathbf{\Lambda}_x\mathbf{V}_x^T$, where the eigenmatrix $\mathbf{V}_x$ is uniformly drawn from the group of $N \times N$ orthogonal matrices and the eigenvalues in $\mathbf{\Lambda}_x$ are set to the absolute values of i.i.d. samples drawn from $\mathcal{N}(0,1)$. Figures 7 and 8 show the de-anonymization accuracy of the proposed method with the correlation adjustment in (23). For comparison, we also present results from the method in (18), which treats the excitations as uncorrelated signals. The comparison between these two methodologies underscores the superior capability of the proposed strategy in (23) to effectively address spatial correlations within the signals.

C. Node Identification in Power Grids

In this section, we apply GS2GM for node label identification in power grid networks, specifically addressing Problem 2. We consider the IEEE 22-bus radial distribution system with $N = 22$ nodes depicted in Fig. 9. The current and voltage phasors in the network are characterized in (5). By employing the model in (3), we can conceptualize $\boldsymbol{i}_m$ and $\boldsymbol{v}_m$ as the filtered and excitation graph signals, respectively, with $\mathbf{Y}$ acting as the associated high-pass graph filter; see (5). The value of $\mathbf{Y}$ can be found in the Supplementary Material of this paper. Here, we adopt two types of measurement data:

- *Synthetic uncorrelated data*: Each voltage phasor $\boldsymbol{v}_m$ is independently drawn from $\mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$, and $\boldsymbol{i}_m$ is generated by (5) with $\mathbf{w}_m \sim \mathcal{CN}(\mathbf{0}, \nu^2\mathbf{I}_N)$ and $\nu^2 = 0.2$.
- *Real-world demand data*: Realistic demand profiles sourced from New York State are used to calculate voltage and current phasor measurements by Matpower [44]. The resultant voltage measurements exhibit high correlation both spatially and temporally.

For M voltage phasor measurements $\{\boldsymbol{v}_m \in \mathbb{C}^{N\times 1}\}_{m=1}^M$, we assess the strength of correlations using the average lag-1 auto-correlation. Denote the n-th entry of $\boldsymbol{v}_m$ by $v_m[n]$. The auto-correlation is calculated as follows:

$$R_{\text{avg}} = \frac{1}{N}\sum_{n=1}^{N}\left|\frac{\sum_{m=2}^{M}(v_m[n]-\bar{v}_n)(v_{m-1}[n]-\bar{v}_n)^\star}{\sum_{m=1}^{M}|v_m[n]-\bar{v}_n|^2}\right|,$$

where $(\cdot)^\star$ denotes the conjugate, and $\bar{v}_n = \frac{1}{M}\sum_{m=1}^{M} v_m[n]$ denotes the sample mean over the n-th dimension. This lag-1 auto-correlation intuitively measures the extent to which each sample correlates on average with its immediate predecessor. A higher value of R_{avg} indicates a stronger correlation within the dataset. For the real-world dataset, comprising a total of 1.7×10^4 samples, the corresponding average auto-correlation equals to 0.7962. In contrast, the synthetic i.i.d. Gaussian data, expectedly uncorrelated, shows an average auto-correlation of 0.0072.

Given the current and voltage measurements $\{\boldsymbol{i}_m, \boldsymbol{v}_m\}_{m=1}^M$ collected from the phasor measurement units, our goal is to identify the node labels of the anonymous network with the aid of a reference graph $\mathcal{G}_2$. Observing the sparsity of $\mathbf{Y}$ as confirmed in Figure 9, we employ the filter estimation method in (25) to determine $\mathbf{Y}$ as (cf. [33, Eqs. (47)-(49)])

$$\min_{\mathbf{Y}\in\mathbb{C}^{N\times N}} \ \|\text{vec}(\text{Re}(\mathbf{Y}))\|_1 + \|\text{vec}(\text{Im}(\mathbf{Y}))\|_1 \tag{33a}$$

$$\text{subject to} \ \sum_{m=1}^{M}\|\boldsymbol{i}_m - \mathbf{Y}\boldsymbol{v}_m\|_2^2 \le \eta^2, \tag{33b}$$

$$\mathbf{Y}^T = \mathbf{Y}, \mathbf{Y}\mathbf{1} = \mathbf{0}, \tag{33c}$$

$$\text{Re}(\text{tr}(\mathbf{Y})) = \alpha N, \text{Im}(\text{tr}(\mathbf{Y})) = \beta N, \tag{33d}$$

where $\text{Re}(\mathbf{Y})$ and $\text{Im}(\mathbf{Y})$ are the real and imaginary parts of $\mathbf{Y}$, respectively, and η, α, and β are adjustable parameters. Moreover, we examine the covariance-based GS2GM method in (18), which predominantly relies on current phasor data for node matching. Figure 10 plots the proportion of nodes correctly identified using these techniques. The filter-inference-based method outperforms the method in (18) in terms of de-anonymization accuracy, benefiting from the incorporation of voltage measurements. This observation aligns with the analysis in Section VI, showing that exploiting the knowledge of excitations reduces the required number of signal samples. Notably, under an equivalent number of signal samples, the node identification accuracy with realistic demand data is significantly lower compared to that of the synthetic uncorrelated data, attributed to the high temporal correlation in the real-world dataset that results in a reduced number of independent signal samples.

VIII. Conclusions

This work studied network de-anonymization attacks by matching asymmetric nodes from graph signals to a refer-

ence graph. We developed a methodology to estimate the eigenvectors of the graph filter using either the signal sample covariance matrix or the inferred graph filter and subsequently compute node matching by solving a linear assignment problem. We conducted theoretical analysis on both scenarios and characterized the signal sample requisites to achieve perfect de-anonymization. Empirical results validate our analysis and demonstrate the efficiency of the proposed algorithm. The analysis reveals that graph-signal-based de-anonymization can significantly expose private information concerning node identities, given an adequate amount of signal samples. However, when graph signals are highly correlated or available in limited quantities, they afford a robust level of privacy, safeguarding against de-anonymization attacks. Developing more sophisticated methods to de-anonymize highly correlated graph signals is a promising direction for future research.

APPENDIX A
PROOF OF THEOREM 1

For notational convenience, we define the following events:

$$\mathcal{A} = [\exists n \in \mathcal{AS}(\mathcal{G}_2) \text{ s.t. } \widehat{\sigma}(n) \neq \sigma^\star(n)]; \tag{34a}$$

$$\mathcal{B} = \left[|\hat{\lambda}_k^2 - \lambda_k^2| < \frac{\delta_k}{2} \text{ for } \forall 1 \leq k \leq K\right]; \tag{34b}$$

$$\mathcal{C}_k = \left[|\hat{\lambda}_k^2 - \lambda_k^2| \geq \frac{\delta_k}{2}\right], \forall 1 \leq k \leq K. \tag{34c}$$

It follows from the definitions of $\mathcal{B}$ and $\mathcal{C}_k$ that $\Pr(\mathcal{B}) = \prod_{k=1}^K (1 - \Pr(\mathcal{C}_k))$. Applying the law of total probability, we have

$$\begin{aligned}
&\Pr\left(\widehat{\sigma}(n) = \sigma^\star(n), \forall n \in \mathcal{AS}(\mathcal{G}_2)\right) = 1 - \Pr(\mathcal{A}) \\
&= 1 - \Pr(\mathcal{A}|\mathcal{B})\Pr(\mathcal{B}) - \Pr(\mathcal{A}|\mathcal{B}^c)\Pr(\mathcal{B}^c) \\
&\overset{(a)}{\geq} 1 - \Pr(\mathcal{A}|\mathcal{B})\Pr(\mathcal{B}) - \Pr(\mathcal{B}^c) \\
&= (1 - \Pr(\mathcal{A}|\mathcal{B})) \prod_{k=1}^K (1 - \Pr(\mathcal{C}_k)).
\end{aligned} \tag{35}$$

where $\mathcal{B}^c$ is the complementary event of $\mathcal{B}$, and (a) follows from $\Pr(\mathcal{A}|\mathcal{B}^c) \leq 1$. To prove Theorem 1, it suffices to bound $\Pr(\mathcal{A}|\mathcal{B})$. Following [28, Eqs. (35)-(39)], with $\rho > 0$ we have $\Pr(\mathcal{A}|\mathcal{B}) \leq \Pr(\|\mathbf{E}\|_{\max} \geq \rho/2)$, where $\mathbf{E} = (\overline{\mathbf{U}}_K - \overline{\widetilde{\mathbf{V}}}_K^{(1)})(\overline{\widetilde{\mathbf{V}}}_K^{(2)})^T$ represents the perturbation in the cost matrix due to the eigenvector estimation error, and $\|\cdot\|_{\max}$ is the max norm. Applying the result in [28, Eq. (51)], we have

$$\begin{aligned}
\|\mathbf{E}\|_{\max} &\leq \|(\overline{\mathbf{U}}_K - \overline{\widetilde{\mathbf{V}}}_K^{(1)})(\overline{\widetilde{\mathbf{V}}}_K^{(2)})^T\|_F \leq \|\overline{\mathbf{U}}_K - \overline{\widetilde{\mathbf{V}}}_K^{(1)}\|_F \\
&= \sqrt{2\sum_{k=1}^K (1 - |\mathbf{u}_k^T \widetilde{\mathbf{v}}_k^{(1)}|)} \leq \sqrt{2\sum_{k=1}^K (1 - (\mathbf{u}_k^T \widetilde{\mathbf{v}}_k^{(1)})^2)},
\end{aligned} \tag{36}$$

where the last inequality follows from $0 \leq |\mathbf{u}_k^T \widetilde{\mathbf{v}}_k^{(1)}| \leq 1$.

Combining (35) and (36) completes the proof.

APPENDIX B
PROOF OF PROPOSITION 1

According to [28, Propositon 1], under the condition of bounded and i.i.d. filtered graph signals $\{\mathbf{z}_m\}$, it follows that

$$\Pr\left(|\hat{\lambda}_k^2 - \lambda_k^2| \geq \frac{\delta_k}{2}\right) \leq \frac{4\kappa_k}{M\delta_k^2} \tag{37}$$

where δ_k is the k-th spectral gap with respect to the square of the k-th eigenvalue λ_k^2, and κ_k is a constant given by $\mathbb{E}_{\mathbf{z}_m}[\|\mathbf{z}_m \mathbf{z}_m^T \mathbf{v}_k^{(1)}\|_2^2]$. Applying the Cauchy–Schwarz inequality and the boundedness of $\|\mathbf{z}_m\|_2$, we have $\kappa_k \leq \mathbb{E}_{\mathbf{z}_m}[\|\mathbf{z}_m\|_2^4] \leq Z^4$. Therefore, it follows that

$$\prod_{k=1}^K \left(1 - \Pr\left(|\hat{\lambda}_k^2 - \lambda_k^2| \geq \frac{\delta_k}{2}\right)\right) \geq \prod_{k=1}^K (1 - \frac{4Z^4}{M\delta_k^2}) \geq 1 - \frac{4KZ^4}{M\delta^2}, \tag{38}$$

where $\delta = \min_{k=1}^K \delta_k$ denotes the minimum spectral gap.

Next, applying the Davis-Khan theorem [45], we have

$$1 - (\mathbf{u}_k^T \widetilde{\mathbf{v}}_k^{(1)})^2 \leq \frac{4\|\widehat{\mathbf{C}}_z - \mathbf{C}_z\|_2^2}{\delta_k^2}. \tag{39}$$

According to [28, Lemma 2], it follows with probability at least $1 - 2t$ for sufficiently large M and any $t > 0$ that

$$\|\widehat{\mathbf{C}}_z - \mathbf{C}_z\|_2 \leq C(\nu^2 + Z)\sqrt{\frac{N \ln(N/t)}{M\delta^2}}, \tag{40}$$

where C is an absolute constant. Equivalently, for any $t' > 0$, we have $\Pr(\|\widehat{\mathbf{C}}_z - \mathbf{C}_z\|_2 \geq t') \leq 2Ne^{-\frac{M(t')^2}{C^2 N(\nu^2+Z)}}$. Setting $t' = \frac{\delta\rho}{4\sqrt{2K}}$, we have

$$\begin{aligned}
&\Pr\left(\rho \leq 2\sqrt{2\sum_{k=1}^K (1 - (\mathbf{u}_k^T \widetilde{\mathbf{v}}_k^{(1)})^2)}\right) \\
&\leq \Pr\left(\frac{\delta\rho}{4\sqrt{2K}} \leq \|\widehat{\mathbf{C}}_z - \mathbf{C}_z\|_2\right) \leq 2Ne^{-\frac{M\rho^2\delta^2}{32C^2(\nu^2+Z)KN}}.
\end{aligned} \tag{41}$$

Substituting (38) and (41) to (15) leads to (19).

APPENDIX C
PROOF OF PROPOSITION 2

We extend the recovery error bound from the compressed sensing literature by calculating an upper bound that scales linearly with the RIC value ς_{2s}. The coefficient of this bound is computed numerically to ensure the tightest fit, leading directly to the precise formulas presented in Conditions iii and iv of Proposition 2. We first state the following two lemmas.

Lemma 1. Let the conditions in Proposition 2 hold. We have

$$\|\widehat{\mathcal{H}}(\mathbf{L}^{(1)}) - \mathcal{H}(\mathbf{L}^{(1)})\|_F \leq \frac{BC_1}{\sqrt{s}} + C_2\eta, \tag{42}$$

where the constant B is defined in Condition ii of Proposition

2, and C_1 and C_2 are given by

$$C_1 = \frac{4}{1-C_0}\sqrt{\frac{2(2-\varsigma_{2s})}{(1-\varsigma_{2s})(32-25\varsigma_{2s})}}, \tag{43}$$

$$C_2 = \frac{2}{\sqrt{1-\varsigma_{2s}}}\left(1+\frac{C_1}{\sqrt{2}}\right), \tag{44}$$

with $C_0 = 2\sqrt{\frac{1+5\varsigma_{2s}-4\varsigma_{2s}^2}{(1-\varsigma_{2s})(32-25\varsigma_{2s})}} < 1$.

Proof. When $\varsigma_{2s} < 0.4931$, (42) follows from a simple combination of the results in [41, Theorem 6.11] and [46, Theorem 4.6], and [41, Proposition 2.3]. □

Lemma 2. For the parameters C_1 and C_2 defined in (43) and (44) and any $\varsigma_{2s} \in (0, 0.475)$, it holds that

$$\frac{C_1}{C_2} \le 0.15\varsigma_{2s} + 0.45, \quad -\frac{1}{C_2} \le 0.4\varsigma_{2s} - 0.19.$$

Proof: To prove the result, it is equivalent to verify the functions $f(\varsigma_{2s}) = \frac{C_1}{C_2} - 0.15\varsigma_{2s} - 0.45$ and $g(\varsigma_{2s}) = -\frac{1}{C_2} - 0.4\varsigma_{2s} + 0.19$ are non-positive in the domain of $\varsigma_{2s} \in (0, 0.475)$. To find the maximum values of these one-dimensional functions f and g, we adopt the numerical optimization method in [47], implemented by the *fminbnd* function in *MATLAB*. As a result, we have $\max_{\varsigma_{2s}} f(\varsigma_{2s}) = f(0.2445) \approx -0.015 < 0$ and $\max_{\varsigma_{2s}} g(\varsigma_{2s}) \le g(0) \approx -0.006 < 0$ for $\varsigma_{2s} \in (0, 0.5)$. This completes the proof. Note that the result holds with the following condition:

$$0.4\varsigma_{2s} - 0.19 < 0 \Leftrightarrow \varsigma_{2s} < 0.475. \tag{45}$$

■

We are ready to prove Proposition 2 using the above results. In particular, we pick $\eta = \|\tilde{\mathbf{y}} - \mathbf{B}\text{vec}(\mathcal{H}(\mathbf{L}^{(1)}))\|_2 = \|\tilde{\mathbf{w}}\|_2^2$ in the following proof. By the Bauer-Fike theorem [48], it follows that

$$\begin{aligned}
&\Pr\left(|\hat{\lambda}_k - \lambda_k| \ge \frac{\delta_k}{2}\right) \\
&\le \Pr\left(\|\widehat{\mathcal{H}}(\mathbf{L}^{(1)}) - \mathcal{H}(\mathbf{L}^{(1)})\|_2 \ge \frac{\delta_k}{2}\right) \\
&\overset{\text{Lemma 1}}{\le} \Pr\left(\frac{BC_1}{\sqrt{s}} + C_2\eta \ge \frac{\delta_k}{2}\right) \\
&\overset{\text{Lemma 2}}{\le} \Pr\left(\eta \ge \frac{0.19 - 0.4\varsigma_{2s}}{2}\delta - \frac{B(0.15\varsigma_{2s}+0.45)}{\sqrt{s}}\right) \\
&\overset{(28)}{\le} \Pr\left(\eta \ge \bar{\delta}\right) = \Pr\left(\frac{\|\tilde{\mathbf{w}}\|_2^2}{\nu^2} \ge \frac{\bar{\delta}^2}{\nu^2}\right).
\end{aligned} \tag{46}$$

Next, applying the Davis-Khan theorem [45], we have

$$1 - (\mathbf{u}_k^T\widetilde{\mathbf{v}}_k^{(1)})^2 \le 4\frac{\|\widehat{\mathcal{H}}(\mathbf{L}^{(1)}) - \mathcal{H}(\mathbf{L}^{(1)})\|_F^2}{\delta_k^2}. \tag{47}$$

Therefore,

$$\begin{aligned}
&\Pr\left(\rho \le 2\sqrt{2\sum_{k=1}^{K}(1-(\mathbf{u}_k^T\widetilde{\mathbf{v}}_k^{(1)})^2)}\right) \\
&\overset{(47)}{\le} \Pr\left(\rho \le 4\sqrt{2\sum_k 1/\delta_k^2}\cdot\|\widehat{\mathcal{H}}(\mathbf{L}^{(1)}) - \mathcal{H}(\mathbf{L}^{(1)})\|_F\right) \\
&\overset{\text{Lemma 1}}{\le} \Pr\left(\frac{BC_1}{\sqrt{s}} + C_2\eta \ge \frac{\sqrt{2}\rho\delta}{8\sqrt{K}}\right) \\
&\overset{\text{Lemma 2}}{\le} \Pr\left(\eta \ge \frac{0.19 - 0.4\varsigma_{2s}}{2}\frac{\sqrt{2}\rho\delta}{8\sqrt{K}} - \frac{B(0.15\varsigma_{2s}+0.45)}{\sqrt{s}}\right) \\
&\overset{(28)}{\le} \Pr\left(\eta \ge \bar{\delta}\right) = \Pr\left(\frac{\|\tilde{\mathbf{w}}\|_2^2}{\nu^2} \ge \frac{\bar{\delta}^2}{\nu^2}\right).
\end{aligned} \tag{48}$$

Substituting (46) and (48) into (15), we have

$$\Pr\left(\widehat{\sigma}(n) = \sigma^\star(n), \forall n\right) \le \left(1 - \Pr\left(\frac{\|\tilde{\mathbf{w}}\|_2^2}{\nu^2} \ge \frac{\bar{\delta}^2}{\nu^2}\right)\right)^{K+1}. \tag{49}$$

The final step is to bound the chi-squared random variable $\|\tilde{\mathbf{w}}\|_2^2$. To this end, we apply the following concentration bound.

Lemma 3. Let Q be a chi-squared random variable with d degrees of freedom. For any $t \ge 2d$, we have $\Pr(Q \ge t) \le e^{-\frac{t-2d}{3}}$.

Proof. The result in [49, Eq. (4.3)] shows that, for any $x > 0$, $\Pr(Q \ge d + 2\sqrt{dx} + 2x) \le e^{-x}$. Note that $d + 2\sqrt{dx} + 2x \le 2d + 3x$. Setting $t = 2d + 3x$ leads to the result. □

Applying Lemma 3, when $\nu^2 \le \frac{\bar{\delta}^2}{2N^2}$, we have

$$\Pr\left(\frac{\|\tilde{\mathbf{w}}\|_2^2}{\nu^2} \ge \frac{\bar{\delta}^2}{\nu^2}\right) \le e^{-\frac{\bar{\delta}^2/\nu^2 - 2N^2}{3}}. \tag{50}$$

Substituting (50) to (49) results in (30). Finally, when $\tilde{\mathbf{z}}$ in (24) has a bounded ℓ_2 norm, it follows from (24) that $\nu^2 = \mathcal{O}(1/(MN))$. When K is chosen such that $\delta/\sqrt{K}$ remains constant, $\bar{\delta}$ in (28) becomes a linearly decreasing function of ς_{2s}. Substituting the above results to (50) as well as taking into account $K \le N$ and (29), a simple calculation yields

$$1 - (1 - e^{-\frac{\bar{\delta}^2/\nu^2 - 2N^2}{3}})^{K+1} \le Ne^{-N^2(c_1 - c_2\varsigma_{2s})}, \tag{51}$$

where c_1 and c_2 are some constants. This completes the proof.

Appendix D
Proof of Proposition 3

We first consider the RIC value $\varsigma_r(\mathbf{B})$ for any r that is divisible by N. Consider an arbitrary r-sparse vector $\mathbf{u}$ in $\mathbb{R}^{N^2}$. It can be verified that $\mathbf{u}$ can be partitioned into N sub-vectors in $\mathbb{R}^N$, each of sparsity $r' = r/N \in \mathbb{Z}$. We denote each sub-vector as $\mathbf{u}_i, 1 \le i \le N$, with $\mathbf{u}_i \in \mathbb{R}^N$ and $\|\mathbf{u}_i\|_0 = r'$. We note that the r'-th RIC of $\mathbf{X}^T$, denoted by $\varsigma_{r'}(\mathbf{X}^T)$, satisfies that

$$(1 - \varsigma_{r'}(\mathbf{X}^T))\|\mathbf{u}_i\|^2 \le \|\mathbf{X}^T\mathbf{u}_i\|^2 \le (1 + \varsigma_{r'}(\mathbf{X}^T))\|\mathbf{u}_i\|^2, \forall i. \tag{52}$$

Summing up all the N inequalities, we have

$$(1-\varsigma_{r'}(\mathbf{X}^T))\sum_i\|\mathbf{u}_i\|^2\leq\sum_i\|\mathbf{X}^T\mathbf{u}_i\|^2\leq(1+\varsigma_{r'}(\mathbf{X}^T))\sum_i\|\mathbf{u}_i\|^2$$
$$\Leftrightarrow(1-\varsigma_{r'}(\mathbf{X}^T))\|\mathbf{u}\|^2\leq\|\mathbf{B}\mathbf{u}\|^2\leq(1+\varsigma_{r'}(\mathbf{X}^T))\|\mathbf{u}\|^2. \tag{53}$$

From the definition of the RIC, we know that $\varsigma_r(\mathbf{B})$ is the smallest value satisfies that

$$(1-\varsigma_r(\mathbf{B}))\|\mathbf{u}\|^2\leq\|\mathbf{B}\mathbf{u}\|^2\leq(1+\varsigma_r(\mathbf{B}))\|\mathbf{u}\|^2.$$

Consequently, it follows from (53) that

$$\varsigma_{Nr'}(\mathbf{B})\leq\varsigma_{r'}(\mathbf{X}^T),\forall r'\in\mathbb{Z}. \tag{54}$$

According to [41, Theorem 9.2], when $\mathbf{X}$ (and thus $\mathbf{X}^T$) is an i.i.d. zero-mean unit-variance sub-Gaussian matrix, $\varsigma_r(\mathbf{X}^T)\leq t$ holds for any $r',t>0$ with probability at least $1-e^{-t^2M/(2C')}$, provided that

$$M\geq\frac{2C'}{t^2}r'\ln(eN/r'), \tag{55}$$

where C' is a constant depending on the distribution of $\mathbf{X}$. Substituting (55) to (54) with $r'=2s/N$, we have

$$\Pr(\varsigma_{2s}(\mathbf{B})\leq t)\geq1-e^{-t^2M/(2C')}, \tag{56}$$

provided that (32) holds.

REFERENCES

[1] H. Liu, A. Scaglione, and S. Peisert, "Privacy leakage in graph signal to graph matching problems," in *IEEE Int. Conf. Acoustics Speech Signal Process. (ICASSP)*, Apr. 2024, pp. 9371–9375.

[2] A. Narayanan and V. Shmatikov, "Robust de-anonymization of large sparse datasets," in *IEEE Symposium on Security and Privacy*, 2008, pp. 111–125.

[3] ——, "De-anonymizing social networks," in *IEEE Symposium on Security and Privacy*, 2009, pp. 173–187.

[4] R. Pang, M. Allman, V. Paxson, and J. Lee, "The devil and packet trace anonymization," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 36, no. 1, p. 29–38, Jan. 2006.

[5] V. Lyzinski, D. E. Fishkind, and C. E. Priebe, "Seeded graph matching for correlated Erdös-Rényi graphs," *J. Mach. Learn. Res.*, vol. 15, no. 1, p. 3513–3540, Jan. 2014.

[6] P. Pedarsani and M. Grossglauser, "On the privacy of anonymized networks," in *Proc. ACM SIGKDD*, 2011, p. 1235–1243.

[7] E. Onaran, S. Garg, and E. Erkip, "Optimal de-anonymization in random graphs with community structure," in *Proc. 50th Asilomar Conf. Signals Syst. Comput.*, 2016, pp. 709–713.

[8] J. Zhang, S. Qu, Q. Li, H. Kang, L. Fu, H. Zhang, X. Wang, and G. Chen, "On social network de-anonymization with communities: A maximum a posteriori perspective," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 3, pp. 2859–2874, Mar. 2023.

[9] B. Miao, S. Wang, L. Fu, and X. Lin, "De-Anonymizability of social network: Through the lens of symmetry," in *Proc. ACM Mobihoc*, Oct. 2020, p. 71–80.

[10] G. Mateos, S. Segarra, A. G. Marques, and A. Ribeiro, "Connecting the dots: Identifying network structure via graph signal processing," *IEEE Signal Process. Mag.*, vol. 36, no. 3, pp. 16–43, May 2019.

[11] A. Ortega, P. Frossard, J. Kovačević, J. M. F. Moura, and P. Vandergheynst, "Graph signal processing: Overview, challenges, and applications," *Proc. IEEE*, vol. 106, no. 5, pp. 808–828, May 2018.

[12] 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 Process. Mag.*, vol. 37, no. 6, pp. 74–85, Nov. 2020.

[13] C. Huang, M. Li, F. Cao, H. Fujita, Z. Li, and X. Wu, "Are graph convolutional networks with random weights feasible?" *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 3, pp. 2751–2768, 2023.

[14] L. Bai, L. Cui, Y. Wang, M. Li, J. Li, P. S. Yu, and E. R. Hancock, "Haqjsk: Hierarchical-aligned quantum jensen-shannon kernels for graph classification," *IEEE Trans. Knowl. Data Eng.*, pp. 1–14, 2024.

[15] M. Li, A. Micheli, Y. G. Wang, S. Pan, P. Lió, G. S. Gnecco, and M. Sanguineti, "Guest editorial: Deep neural networks for graphs: Theory, models, algorithms, and applications," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 35, no. 4, pp. 4367–4372, 2024.

[16] K. Huang, Y. G. Wang, M. Li, , and P. Liò, "How universal polynomial bases enhance spectral graph neural networks: Heterophily, over-smoothing, and over-squashing," *ArXiv Preprint: 2405.12474*, 2024.

[17] S. Segarra, A. G. Marques, G. Mateos, and A. Ribeiro, "Network topology inference from spectral templates," *IEEE Trans. Signal Inform. Process. Netw.*, vol. 3, no. 3, pp. 467–483, Sept. 2017.

[18] S. Grotas, Y. Yakoby, I. Gera, and T. Routtenberg, "Power systems topology and state estimation by graph blind source separation," *IEEE Trans. Signal Process.*, vol. 67, no. 8, pp. 2036–2051, 2019.

[19] Z. Xiang, K. Huang, W. Deng, and C. Yang, "Blind topology identification for smart grid based on dictionary learning," in *IEEE Symp. Ser. Comput. Intell. (SSCI)*, 2019, pp. 1319–1326.

[20] D. Deka, V. Kekatos, and G. Cavraro, "Learning distribution grid topologies: A tutorial," *IEEE Trans. Smart Grid*, vol. 15, no. 1, pp. 999–1013, 2024.

[21] M. Halihal, T. Routtenberg, and H. V. Poor, "Estimation of complex-valued laplacian matrices for topology identification in power systems," *IEEE Trans. Signal Process.*, pp. 1–15, 2024.

[22] 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 Trans. Signal Process.*, vol. 68, pp. 436–451, Dec. 2020.

[23] C. Ye, R. Shafipour, and G. Mateos, "Blind identification of invertible graph filters with multiple sparse inputs," in *European Signal Process. Conf. (EUSIPCO)*, 2018, pp. 121–125.

[24] B. Rathnayake *et al.*, "Graph-based blind hyperspectral unmixing via nonnegative matrix factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 9, pp. 6391–6409, 2020.

[25] J. Miettinen, E. Nitzan, S. A. Vorobyov, and E. Ollila, "Graph signal processing meets blind source separation," *IEEE Trans. Signal Process.*, vol. 69, pp. 2585–2599, 2021.

[26] A. Einizade and S. H. Sardouie, "Joint graph learning and blind separation of smooth graph signals using minimization of mutual information and laplacian quadratic forms," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 9, pp. 35–47, 2023.

[27] M.-H. A. Yarandi and M. Babaie-Zadeh, "A new approach for graph signal separation based on smoothness," *IEEE Trans. Signal Process.*, vol. 72, pp. 972–981, 2024.

[28] H. Liu, A. Scaglione, and H.-T. Wai, "Blind graph matching using graph signals," *IEEE Trans. Signal Process.*, vol. 72, pp. 1766–1781, Apr. 2024.

[29] R. Shafipour, S. Segarra, A. G. Marques, and G. Mateos, "Identifying the topology of undirected networks from diffused non-stationary graph signals," *IEEE Open J. Signal Process.*, vol. 2, pp. 171–189, 2021.

[30] M. H. DeGroot, "Reaching a consensus," *J. Amer. Stat. Assoc.*, vol. 69, no. 345, pp. 118–121, 1974.

[31] M. Billio, M. Getmansky, A. W. Lo, and L. Pelizzon, "Econometric measures of connectedness and systemic risk in the finance and insurance sectors," *J. Financ. Econ.*, vol. 104, no. 3, pp. 535–559, 2012.

[32] T. Ishizaki, A. Chakrabortty, and J.-I. Imura, "Graph-theoretic analysis of power systems," *Proc. IEEE*, vol. 106, no. 5, pp. 931–952, 2018.

[33] R. Ramakrishna and A. Scaglione, "Grid-graph signal processing (Grid-GSP): A graph signal processing framework for the power grid," *IEEE Trans. Signal Process.*, vol. 69, pp. 2725–2739, 2021.

[34] Q. Van Tran and H.-S. Ahn, "Further analysis on structure and spectral properties of symmetric graphs," *arXiv preprint arXiv:2203.01408*, 2022. [Online]. Available: https://arxiv.org/abs/2203.01408

[35] P. J. Cameron and Q. Mary, "Automorphisms of graphs," *Topics in Algebraic Graph Theory*, vol. 102, pp. 137–155, 2004.

[36] L. Fu, J. Zhang, S. Qu, H. Kang, X. Wang, and G. Chen, "Measuring social network de-anonymizability by means of morphism property," *IEEE/ACM Trans. Netw.*, vol. 30, no. 6, pp. 2744–2759, Dec. 2022.

[37] S. Umeyama, "An eigendecomposition approach to weighted graph matching problems," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 10, no. 5, pp. 695–703, Sept. 1988.

[38] J. Munkres, "Algorithms for the assignment and transportation problems," *J. Soc. Ind. Appl. Math.*, vol. 5, no. 1, pp. 32–38, Mar. 1957.

[39] Z. Fan, C. Mao, Y. Wu, and J. Xu, "Spectral graph matching and regularized quadratic relaxations: Algorithm and theory," in *Proc. Int. Conf. Mach. Learn. (ICML)*, vol. 119, July 2020, pp. 2985–2995.

[40] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[41] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*. Springer, 2012.
[42] O. Teke and P. P. Vaidyanathan, "Sparse eigenvectors of graphs," in *IEEE Int. Conf. Acoustics Speech Signal Process. (ICASSP)*, 2017, pp. 3904–3908.
[43] J. J. McAuley and J. Leskovec, "Learning to discover social circles in ego networks," in *Proc. Int. Conf. Neural Inf. Process. Syst. (NeurIPS)*, 2012, pp. 548–556.
[44] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "Matpower: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 12–19, Feb. 2011.
[45] Y. Yu, T. Wang, and R. J. Samworth, "A useful variant of the Davis-Kahan theorem for statisticians," *Biometrika*, vol. 102, no. 2, pp. 315–323, 2014.
[46] Q. Mo and S. Li, "New bounds on the restricted isometry constant δ_{2k}," *Appl. Comput. Harmon. Anal.*, vol. 31, no. 3, pp. 460–468, 2011.
[47] G. E. Forsythe, M. A. Malcolm, and C. B. Moler, *Computer Methods for Mathematical Computations*. Englewood Cliffs, NJ: Prentice Hall, 1976.
[48] F. L. Bauer and C. T. Fike, "Norms and exclusion theorems," *Numer. Math.*, vol. 2, pp. 137–141, 1960.
[49] B. Laurent and P. Massart, "Adaptive estimation of a quadratic functional by model selection," *Ann. Statist.*, vol. 28, no. 5, p. 1302–1338, 2000.