Statistical guarantees in the sparse recovery of an electrical network

Alberto González–Sanz*[1] and Álvaro Samperio♭,[2]

(*) Department of Statistics, Columbia University, New York, USA,
(♭) Departamento de Álgebra, Análisis Matemático, Geometría y Topología and IMUVA, Universidad de Valladolid, Valladolid, Spain.

1 Introduction

Several authors have studied different problems of recovering some unknown information within an electrical network using measured data such as voltage, power, or current. Some of these problems arise in the context of real networks (see e.g., [1, 5, 6]), where the information about the network is often incorrect or outdated. This information is crucial for monitoring and operating the network, making it worthwhile to develop estimation methods based on measured data. This is especially pertinent given the recent general improvement in the precision and availability of data in real networks. Other network recovery problems have emerged as discrete approximations to Calderón's problem (see [2, 4] and references therein).

In this talk, we concentrate on the sparse network recovery problem as introduced in [10], which involves reconstructing a sparse network topology and its associated cable parameters using power and voltage data collected from all vertices. In Section 4 (Ibidem), an algorithm designed to address this problem is introduced. First, we will briefly introduce the notation used in [10] to clearly formulate the main problem.

We denote by $\Gamma = (N, E, w)$ an electrical network, where (N, E) is an undirected, finite and simple graph (called the *"network topology"*) with vertex set N and edge set E, and w is a vector of weights at the edges E (the cable parameters of the network). In the case of an Direct Current (DC) network, each edge is characterized by a real non-negative parameter, so $w \in \mathbb{R}_{\geq 0}^{|E|}$, and in the case of an Alternating Current (AC) network, each edge can be characterized by two real non-negative parameters, so $w \in \mathbb{R}_{\geq 0}^{2|E|}$.

We will denote as *"state"* of an AC network Γ a vector in $\mathbb{R}^{4n}$ containing the real and imaginary parts of the values of voltage and power injected at each vertex at a given time. In the case of DC networks, voltage and power are real, so a state is a vector in $\mathbb{R}^{2n}$. Any state must satisfy the *"power flow equations"*, derived from Kirchhoff's and Ohm's laws. They are quadratic in the state variables and linear in the cable parameters of Γ.

Given a data set of states Ω which may have some errors, we define the fitting error $\mathsf{rms}(\Gamma, \Omega)$ of any network Γ as the root mean square of the vector of evaluations of all the power flow equations of Γ at every element of Ω.

[1] ag4855@columbia.edu
[2] alvaro.samperio@uva.es

Problem 1. *[Main Problem] Given a set of nodes N, a data set of states Ω which may have some errors, and a tolerance $\mathsf{tol} \in \mathbb{R}$, we seek to determine a network $\Gamma = (N, E, w)$ such that $\mathsf{rms}(\Gamma, \Omega) \leq \mathsf{tol}$, and Γ is "minimal" in the following sense. Given any electrical network $\Gamma' = (N, E', w')$, we have that:*

1. *If $E' = E$, then $\mathsf{rms}(\Gamma', \Omega) \geq \mathsf{rms}(\Gamma, \Omega)$.*
2. *If $E' \subsetneq E$, then $\mathsf{rms}(\Gamma', \Omega) > \mathsf{tol}$.*

It is shown in [10] that in many cases the problem of simultaneously recovering the topology and cable parameters of an electrical network from power and voltage data at all vertices is ill-posed. Even if the data comes from a sparse network without error, there may be dense networks with zero fitting error. The formulation of the main problem is of significant interest from an applied perspective because recovering a sparse network topology facilitates the efficient resolution of operational challenges within electrical networks, such as power flow optimization.

In this presentation, we begin by reviewing the key theoretical results from [10], which provide the foundation for the algorithm proposed to solve the main problem. We then delve into additional findings related to this algorithm. Our discussion includes the consistency of the estimator for cable parameters, assuming a fixed network topology and specific data distribution assumptions.

The algorithm in [10] employs a random iterative process, during which several network edges may be eliminated in some iterations. In this talk, we demonstrate that as the number of iterations approaches infinity, the algorithm is guaranteed to identify a minimal network with probability one. Furthermore, we propose a stopping criterion for the algorithm, which is derived from the analysis of a random walk on a graph. This criterion helps in determining the optimal point to terminate the algorithm, ensuring efficiency while maintaining accuracy.

2 Methods

2.1 Algorithm of sparse network recovery

In this section, we will provide an overview of the primary techniques utilized in the algorithm proposed in [10] for solving the main problem of sparse network recovery.

When considering a dataset of states Ω and a fixed network topology (N, E), the task of determining a parameter vector w that minimizes the fitting error is equivalent to solving a Non-Negative Least Squares (NNLS) problem, as outlined in more detail in [10, Section 3]. NNLS problems are convex optimization problems [7], ensuring that every local minimum is also a global minimum. Solutions to this problem can be obtained using interior point methods.

Adding or removing edges to a network whose parameters are all equal to zero does not change the power flow equations, so the algorithm starts by finding a vector of parameters w with minimum error with the complete graph as fixed topology. As the problem of parameter estimation with this topology is usually ill-posed, w is usually dense, so no edges can be trivially removed from the nertwork.

In order to remove edges of a DC network, we use the random procedure in [3] for spectral graph sparsification, which is a notion introduced by Spielman and Teng in [9]. Given an electrical network Γ and $\varepsilon > 0$, the procedure has a certain probability of obtaining an ε-approximation of Γ, which is defined as follows.

Definition 1. For $\varepsilon \geq 0$, we say that a DC network Γ', with Laplacian matrix $L_{\Gamma'}$, is an ε-approximation [8] of a DC network Γ, with Laplacian matrix L_Γ, if they have the same vertex set and for all $z \in \mathbb{R}^n$:

$$\frac{1}{1+\varepsilon} z^T L_\Gamma z \leq z^T L_{\Gamma'} z \leq (1+\varepsilon) z^T L_\Gamma z. \tag{1}$$

When addressing an AC network Γ, wherein each edge entails two parameters, we can depict the network through two real weighted graphs, with each graph corresponding to one of the parameters. By independently applying the spectral sparsification procedure to each graph, there exists a positive probability of obtaining an *"ε-approximation"* AC electrical network of Γ. This approximation satisfies a property analogous to the one described in Definition 1. For a comprehensive explanation, please refer to [10, Section 4].

The core theoretical results upon which the sparse network recovery algorithm is founded are [10, Theorems 2 and 3]. These theorems assert that for a given dataset Ω, an electrical network Γ, and a positive value ε, if Γ' represents an ε-approximation of Γ, then the fitting error in Γ', denoted as $\mathsf{rms}(\Gamma', \Omega)$, does not surpass the previous error $\mathsf{rms}(\Gamma, \Omega)$ by more than ε times a fixed quantity, which solely depends on Γ and Ω.

The algorithm of sparse network recovery is [10, Algorithm 2]. Starting from a data set Ω, a tolerance tol and an initial value of ε and a value $\psi > 1$, the algorithm starts by recovering a network Γ with the complete graph topology. Then, it performs successive iterations. Each iteration begins using the procedure in [3] for obtaining an ε-approximation of Γ. In the case the procedure does not remove any edges, we increase the value of ε multiplying it by ψ and we finish the iteration. By increasing the value of ε it becomes more likely that in the next iteration the procedure in [3] removes some edges.

In the scenario where certain edges have been removed, we proceed to estimate the parameters of the network based on the updated topology. If the error in the new network is either lower or equal to the tolerance threshold tol, we then substitute Γ with this new network and conclude the iteration. Conversely, if the error in the new network surpasses tol, we discard the network, reduce the value of ε by dividing it by ψ, and conclude the iteration. By diminishing the value of ε, the upper bound for the error, as provided by [10, Theorems 2 and 3], will subsequently decrease in the subsequent ε-approximation.

2.2 Statistical consistency under fixed topology

From a statistical standpoint, the data obtained by the practitioner may be subject of some randomness (*e.g.,* measurement errors). In this talk, we also address this scenario. That is, assume the observation of a sample $(\mathbf{X}, \mathbf{Y})^{[m]} = (\mathbf{X}_1, \mathbf{Y}_1) \dots (\mathbf{X}_m, \mathbf{Y}_m)$ of i.i.d. states

$$\mathbf{X}_l = \left(e_1^l, f_1^l, ..., e_n^l, f_n^l\right))^T \text{ and } \mathbf{Y}_l = \left(P_1^l, Q_1^l, ..., P_n^l, Q_n^l\right))^T \quad l = 1, \dots, m.$$

The empirical network, namely $\hat{\mathbf{w}}$, is obtained as the solution the non-negative least-mean-square minimization problem

$$\begin{aligned} \hat{\mathbf{w}} \in \operatorname{argmin} \quad & \frac{1}{m} \sum_{l=1}^{m} ||M(\mathbf{X}_l)\mathbf{w} - \mathbf{Y}_l||^2 \\ \text{s.t.} \quad & \kappa \geq w_k \geq 0, \quad \text{for all} \quad k = 1, \dots, 2|E|. \end{aligned} \tag{2}$$

where $M(\mathbf{X}_l)$ is the multivariate Vandermonde matrix of the data point $\mathbf{X}_l$ (see Section 3 in [10] for further details). We will show that as the number of samples increases, $\hat{\mathbf{w}}$ converges almost surely to a solution of the population problem

$$\begin{aligned} \text{minimize} \quad & \mathbb{E}\left(||M(\mathbf{X})\mathbf{w} - \mathbf{Y}||^2\right) \\ \text{s.t.} \quad & \kappa \geq w_k \geq 0, \quad \text{for all} \quad k = 1, \dots, 2|E|. \end{aligned} \tag{3}$$

This provides statistical guarantees on the consistency of the empirical network towards the population one as the number of observations increases.

2.3 Stopping criterion for the algorithm

Consider any iteration of the network recovery algorithm, in which the current network is $\Gamma = (N, E, w)$, with $\mathsf{rms}(\Gamma, \Omega) < \mathsf{tol}$. The randomness of the spectral sparsification procedure in [3] assures that every subgraph of (N, E) has a non-zero probability of being selected. Thus, when the number of iterations of the algorithm increases, if there is any network $\Gamma' = (N, E', w')$ with the property that $E' \subsetneq E$ and $\mathsf{rms}(\Gamma', \Omega) \leq \mathsf{tol}$, then the algorithm is guaranteed to converge to a network with this property. This guarantees that when the number of iterations of the algorithm goes to infinity, the algorithm converges to a minimal network.

Despite in several experiments the algorithm of sparse network recovery gives a minimal network in few iterations, a stopping criteria for the algorithm has not been established before. Let ε_0 be the initial value of ε. The following points summarize the main steps of the algorithm that will enable us to derive such a stopping criterion:

1. Throughout the algorithm ε can take values $\varepsilon_k := \varepsilon_0 \psi^k$, with $k \in \mathbb{Z}$.

2. The probability of increasing the value of ε in the algorithm is equal to the probability of having zero empty cells in a multinomial distribution of $t = 8|N| \cdot log(|N|)/\varepsilon^2$ samples with $|E|$ possible outcomes.

3. There exists an integer $k_{\max} \in \mathbb{Z}$ such that for any integer $k \geq k_{\max}$, the probability of increasing the value of ε for ε_k is zero. This occurs because $t < |E|$, ensuring that the sparsification procedure is guaranteed to remove edges.

4. There exists an integer $k_{\min} \in \mathbb{Z}$ such that for any integer $k \leq k_{\min} + 1$, according to [10, Theorems 2 and 3], we are guaranteed that $\mathsf{rms}(\Gamma', \Omega) \leq \mathsf{tol}$ for any Γ' that is an ε-approximation of Γ. If during an iteration the value of ε is $\varepsilon_{k_{\min}}$ and it must decrease, it indicates that the sparsification procedure failed to produce an ε-approximation of Γ. This means that this value is sufficiently low to ensure that we will obtain a network with an acceptably low error if we obtain an ε-approximation of Γ with fewer edges than Γ in a subsequent iteration. Hence, the algorithm is adjusted to increase the value of ε at the next iteration.

Therefore, every iteration in the algorithm, until it reaches a new network with fewer edges than the previous one and an error lower than the tolerance, can be viewed as a step in a random walk within a finite directed graph. The set of nodes in this graph is $\{1, ..., s+1\}$, where each of the first s nodes represents one of the s distinct values that ε can assume, sorted in ascending order. The last node, $s+1$, symbolizes the success in discovering a new network with the desired property.

For any j, we denote by π_j the probability of increasing the value of ε, and by δ_j the probability of obtaining a new network when ε is equal to the j-th possible value it can assume, respectively.

The random walk represents a Markov chain without memory. The probability of transitioning from state i to state j is determined by the (j, i)-th entry of the matrix:

$$\left(\begin{array}{ccccccc|c}
0 & 1-\pi_2-\delta_2 & & & & & & 0 \\
1-\delta_1 & 0 & 1-\pi_3-\delta_3 & & & & & \vdots \\
 & \pi_2 & 0 & \ddots & & & & \vdots \\
 & & \pi_3 & \ddots & 1-\pi_{s-1}-\delta_{s-1} & & & \vdots \\
 & & & \ddots & 0 & 1-\delta_s & & \vdots \\
 & & & & \pi_{s-1} & 0 & & 0 \\
\hline
\delta_1 & \delta_2 & \delta_3 & \dots & \delta_{s-1} & \delta_s & & 1
\end{array}\right)$$

In the presentation, it will be observed that by monitoring the values taken by ε and under the assumption that any of the δ_j is sufficiently large, a stopping criterion for the algorithm can be established. This criterion determines a specific number of iterations beyond which there is a low probability of encountering a new network that is sparser than the current one.

Acknowledgments

Á. Samperio was supported by a FPI grant of the Research Project PGC2018-096446-BC21 (with the help of the FEDER Program), and partially supported by grant PID2022-138906NB-C21 funded by MICIU/AEI/ 10.13039/501100011033 and by ERDF/EU.

References

[1] Abur, A., Gómez, A., Power System State Estimation: Theory and Implementation. 0-8247-5570-7, Marcel Dekker, Inc., 2004.

[2] Araúz, C., Carmona, Á., Encinas, A. M., Mitjana, M., Recovering the conductances on grids: a theoretical justification. In A panorama of mathematics: pure and applied. 978-1-4704-1668-3, Amer. Math. Soc., Providence, RI, 2016, pp. 149–166.

[3] Batson, J., Spielman, D., Srivastava, N., Teng, S., Spectral Sparsification of Graphs: Theory and Algorithms *Communications of the ACM*, 56: 87–94, 2013.

[4] Carmona, Á., Encinas, A. M., Jiménez, M. J., Samperio, Á., Stable recovery of piecewise constant conductance on spider networks *Arxiv preprint*, 2023.

[5] Deka, D., Backhaus, S., Chertkov, M., Structure Learning in Power Distribution Networks *IEEE Transactions on Control of Network Systems*, PP: 1061–1074, 2015.

[6] Khan, A., Chakrabarti, S., Sharma, A., Alam, Mahamad N., Parameter and Topology Estimation for Electrical Power Distribution System. In 2019 8th International Conference on Power Systems (ICPS 2019). 9781728141046, Institute of Electrical and Electronics Engineers (IEEE), 2019, pp. 1–5.

[7] Slawski, M., Hein, M., Non-negative least squares for high-dimensional linear models: Consistency and sparse recovery without regularization *Electronic Journal of Statistics*, 7: 3004—3056, 2013.

[8] Spielman, D., Graphs, Vectors, and Matrices *Bulletin of the American Mathematical Society*, 54: 45–61, 2017.

[9] Spielman, D. and Srivastava, N., Graph Sparsification by Effective Resistances *Proc. 40th STOC*, 40: 563–568, 2008.

[10] Samperio, Á., Sparse recovery of an electrical network *Arxiv preprint*, 2023.