Error Bounds for Radial Network Topology Learning from Quantized Measurements

Samuel Talkington[★], Aditya Rangarajan[♯], Pedro A. Alcântara[†], Line Roald[♯], Daniel K. Molzahn[★], Daniel R. Fuhrmann[†]

Abstract—We probabilistically bound the error of a solution to a radial network topology learning problem where both connectivity and line parameters are estimated. In our model, data errors are introduced by the precision of the sensors, i.e., quantization. This produces a nonlinear measurement model that embeds the *operation of the sensor communication network* into the learning problem, expanding beyond the additive noise models typically seen in power system estimation algorithms. We show that the error of a learned radial network topology is proportional to the width of the quantization bin and grows sublinearly in the number of nodes, provided that the number of samples per node is logarithmic in the number of nodes.

I. Introduction

Efficiently allocating *limited smart meter bandwidth* is an emerging challenge at the intersection of power and communication engineering [1], [2]. As smart meters proliferate, power engineers will increasingly rely on analog-to-digital *quantization* methods to improve computational and communication efficiency by mapping continuous measurements to discrete intervals [3]; see Fig. 1 (left). While coarse quantization accelerates computations [4], it introduces *nonlinear*, typically *non-Gaussian*, measurement noise [5], [6], contrasting with the additive Gaussian assumptions common in power system estimation.

To help address these challenges, we present a new *sample complexity* analysis method for distribution topology learning, a well-studied inference task in power engineering; see [7] for a comprehensive review. Sample complexity refers to the *number of measurements needed* to ensure that the *error of an estimated parameter is bounded* by a chosen tolerance [8]. In particular, this work provides such topology learning error bounds under the effects of practically relevant communication non-idealities.

To the knowledge of the authors, these contributions are the first of their kind and expand upon recent deterministic studies of measurement requirements for topology learning [9]. Concretely, we address the following question:

> Consider an *n*-node distribution network with unknown topology (connectivity and parameters). Suppose that we collect measurements at every node with a uniform quantization bin width $\Delta > 0$. How many samples per node $s = m/n$ are needed to recover the topology of the network from the $m = sn$ quantized measurements, up to a desired error tolerance?

This research question is illustrated in Fig. 1 (right), outlined in Section II. We give a precise answer in Section III by providing a bound on the error of a topology estimate given a prescribed number of measurements.
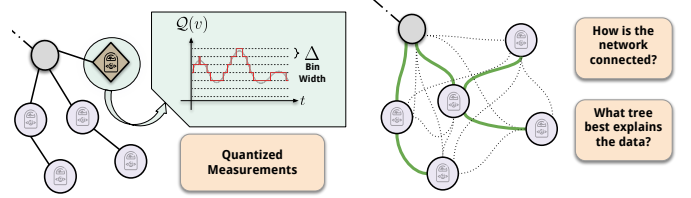
Fig. 1. Conceptual illustration of the proposed problem. Left: Distribution network smart meter measurements have realistic quantization noise corresponding to *bin width* $\Delta > 0$. Right: Estimate the topology *and* parameters of the distribution network (green, solid lines).

II. Problem formulation

A. Communication model

We wish to recover an unknown vector of line parameters $\boldsymbol{w}_\star \in \mathbb{R}^d$ from $m \geq d$ *quantized measurements* of the form

$$p_i \triangleq \mathcal{Q}(\langle \boldsymbol{a}_i, \boldsymbol{w}_\star \rangle), \quad i = 1, \ldots, m, \qquad (1)$$

where $\mathcal{Q} : \mathbb{R} \rightarrow \mathbb{R}$ is a nonlinear *quantization function* and $(p_i, \boldsymbol{a}_i)$, $i = 1, \ldots, m$ are measurements collected from smart meters (both spatially *and* temporally). Vectors $\{\boldsymbol{a}_i\}$ are rows of a particular sensing matrix $\boldsymbol{A} \in \mathbb{R}^{m \times d}$ and contain pairwise voltage differences measured co-temporally with active power injection p_i. We will define matrix $\boldsymbol{A}$ explicitly in Section III-A as a consequence of linearizing the power flow equations. In the same way, we will argue that the sparsity pattern of $\boldsymbol{w}_\star$ represents the network topology (cf. [7, Sec. IV-B]).

Concretely, we focus on the setting in which coarsely quantized measurements (1) are generated from a uniformly dithered quantization function (cf. [3], [6]) with bin width $\Delta > 0$. These measurements take the form

$$p_i = \Delta \cdot \left(\left\lfloor \frac{\langle \boldsymbol{a}_i, \boldsymbol{w}_\star \rangle + \tau_i}{\Delta} \right\rfloor + \frac{1}{2} \right) \quad i = 1, \ldots, m, \quad (2)$$

where $\tau_i \sim \text{Uniform}\left(-\frac{\Delta}{2}, \frac{\Delta}{2}\right)$, and $\mathsf{E}[p_i] = \langle \boldsymbol{a}_i, \boldsymbol{w}_\star \rangle$. The *dither* τ_i is a purposely applied random noise component that is generated by a sensor and added to an input signal prior to quantization. This technique is rather well-established and commonly used by sensors in practice, and by statisticians in theory; see [10], [3].

B. Statistical tools

Suppose that $\boldsymbol{w}_\star$ lies in a convex constraint set $\mathcal{K} \subseteq \mathbb{R}^d$ that encodes some sort of *known structure* of the parameter $\boldsymbol{w}_\star$, e.g., knowledge of where certain lines are located (but not their switching status), or whether the network is operated radially. We can solve for a constrained estimate $\hat{\boldsymbol{w}} \in \mathbb{R}^d$ that obeys this structure via the program (3), known as the *generalized LASSO*:

$$\hat{\boldsymbol{w}} = \arg\min_{\boldsymbol{w} \in \mathcal{K}} \frac{1}{2m} \sum_{i=1}^{m} \left(\langle \boldsymbol{a}_i, \boldsymbol{w} \rangle - p_i \right)^2. \qquad (3)$$

The quality of the estimate produced by the program (3) can be quantified with a statistical tool known as the *Gaussian width*. We introduce this tool briefly and refer the reader to [11], [12, Ch. 7.5] and [5], [6] for further technical information. Formally, the Gaussian width $\omega(\mathcal{T})$ of a set $\mathcal{T} \subseteq \mathbb{R}^d$ is defined as

$$\omega(\mathcal{T}) \triangleq \mathsf{E}\left[\sup_{\boldsymbol{u} \in \mathcal{T}} \langle \boldsymbol{u}, \boldsymbol{g}\rangle\right], \tag{4}$$

where $\boldsymbol{g} \sim \mathcal{N}(0, \mathbf{I})$ is a vector of i.i.d. standard Gaussians. The Gaussian width ω is a useful tool for predicting the behavior of structured convex problems with random input data [11]. It represents the "size" of a standard Gaussian process over $\mathcal{T}$, which we can use as a stochastic *comparison* for our structured problem. Two classic examples of structured constraint sets with well-behaved Gaussian widths are sparse and low-rank constraint sets [11].

The squared Gaussian width, $\omega^2(\mathcal{T})$, provides a measure of the effective dimension of the set $\mathcal{T}$; informally, this represents the minimum number of orthogonal directions required to capture the variation or structure within $\mathcal{T}$. We will also need another object for our analysis—the *tangent cone* of a set $\mathcal{K} \subset \mathbb{R}^d$ at $\boldsymbol{x} \in \mathbb{R}^d$, which is defined as

$$\mathcal{D}(\mathcal{K}, \boldsymbol{x}) \triangleq \{\lambda \boldsymbol{u} \ : \ \lambda \geq 0, \ \boldsymbol{u} \in \mathcal{K} - \boldsymbol{x}\}, \tag{5}$$

where $\mathcal{K} - \boldsymbol{x}$ is the set $\mathcal{K}$ translated by $\boldsymbol{x}$. The tangent cone (5) generalizes the idea of the tangent space to a nonlinear surface that may have non-differentiable points.

C. Network and power flow model

The abstract quantities in (1) and (2) can be applied to the Linear Coupled Power Flow (LCPF) model [7], [13]. As the focus of this work is on sample complexity, we make the following simplifying assumption on the reactive power.

Assumption 1 (Fixed power factor). *Let $\boldsymbol{p} \in \mathbb{R}^n$ be active power injections. Assume that reactive power injections satisfy $\boldsymbol{q} = \kappa \boldsymbol{p}$, where $\kappa \in \mathbb{R}$ is a known constant.*

Assumption 1 is equivalent to defining a fixed power factor $\phi \in (0, 1]$ such that $\kappa = \pm \phi^{-1}\sqrt{1 - \phi^2}$, where $\mathrm{sign}(\kappa) = 1$ if the injections are inductive and $\mathrm{sign}(\kappa) = -1$ if the injections are capacitive. For more information on the consequences of this assumption, see [14], [15].

1) Linear Coupled Power Flow (LCPF) model: For a radial network with $n + 1$ nodes, let $\boldsymbol{C} \in \{-1, 0, 1\}^{n \times n}$ be the invertible, reduced, branch-to-node incidence matrix of a tree graph with the column corresponding to the slack node removed. Let $\left\{(\boldsymbol{v}_t, \boldsymbol{p}_t)\right\}_{t=1}^m$ be a sequence of nodal voltage magnitude and active power measurements. Under Assumption 1, the voltage magnitudes satisfy [7], [13]:

$$\boldsymbol{v}_t - \mathbf{1} = (\boldsymbol{R} + \boldsymbol{X} \cdot \kappa) \boldsymbol{p}_t, \quad t = 1, \ldots, m \tag{6}$$

where $\boldsymbol{R} + \boldsymbol{X}\kappa \triangleq \boldsymbol{Z}$ is the (reduced) *equivalent impedance matrix*. We can write $\boldsymbol{Z}$ as the inverse of a *real-valued equivalent admittance matrix* $\boldsymbol{Y} \in \mathbb{R}^{n \times n}$, where

$$\boldsymbol{Z} = \boldsymbol{C}^{-1} \mathrm{diag}\left(\boldsymbol{r} + \kappa \boldsymbol{x}\right) \boldsymbol{C}^{-\top} \triangleq \boldsymbol{Y}^{-1}, \tag{7}$$

where $\boldsymbol{z} \triangleq \boldsymbol{r} + \kappa \boldsymbol{x} \in \mathbb{R}^n$ are the line impedances (scaled). The matrix $\boldsymbol{C}^{-1} \in \{-1, 0\}^{n \times n}$ is a lower triangular matrix where the non-zero entries of each column j represent the descendants of node j in the tree. In particular, over the indices $i, j = 1, \ldots, n$, we have $\left(\boldsymbol{C}^{-1}\right)_{ij} = -1$ if $i = j$ or i is a descendant of j, and $\left(\boldsymbol{C}^{-1}\right)_{ij} = 0$ otherwise.

2) Lifting to the complete graph: Echoing [9], we can view the problem of recovering $\boldsymbol{w}_\star$ as *sparsifying* a complete undirected graph K_{n+1} with $n + 1$ nodes and $\binom{n+1}{2}$ lines. Accordingly, let $\tilde{\boldsymbol{Y}}(\cdot)$ be the *admittance matrix operator* that sends line parameters $\boldsymbol{w} \in \mathbb{R}^{\binom{n+1}{2}}$ to a corresponding (non-reduced) $(n + 1) \times (n + 1)$ admittance matrix:

$$\tilde{\boldsymbol{Y}}(\boldsymbol{w}) = \tilde{\boldsymbol{C}}^\top \mathrm{diag}(\boldsymbol{w})\tilde{\boldsymbol{C}} = \sum_{i=1}^n \sum_{j=i+1}^n \boldsymbol{E}_{ij} w_{ij}. \tag{8}$$

Here, $\boldsymbol{E}_{ij} \triangleq \left(\boldsymbol{e}_i - \boldsymbol{e}_j\right)\left(\boldsymbol{e}_i - \boldsymbol{e}_j\right)^\top$ is an *elementary Laplacian matrix*, where $\boldsymbol{e}_i$ is the i-th standard basis vector in $\mathbb{R}^{n+1}$. The matrix $\tilde{\boldsymbol{C}} \in \{-1, 0, 1\}^{\binom{n+1}{2} \times n+1}$ is the incidence matrix that corresponds to a complete graph. We can now represent the set of all radial networks as the set of all line parameter vectors $\boldsymbol{w}$ that correspond to a connected radial network with $n + 1$ nodes. This set takes the form

$$\mathcal{R} = \left\{\boldsymbol{w} \in \mathbb{R}^{\binom{n+1}{2}} \ : \ ||\boldsymbol{w}||_0 = n, \ \ \lambda_2(\tilde{\boldsymbol{Y}}(\boldsymbol{w})) > 0\right\}. \tag{9}$$

The sparsity condition $||\boldsymbol{w}||_0 = n$ ensures n lines, and the condition on the second smallest eigenvalue $\lambda_2(\tilde{\boldsymbol{Y}}(\boldsymbol{w})) > 0$ means the graph is connected, i.e., $\boldsymbol{x}^\top \tilde{\boldsymbol{Y}}(\boldsymbol{w})\boldsymbol{x} > 0$ for all $\boldsymbol{x} \perp \mathbf{1}$. The constraint set (9) is intractable; however, in Section III-A we provide a relaxation that performs well both experimentally and theoretically. Hereafter, we implicitly remove the slack node column of $\tilde{\boldsymbol{C}}$ in our calculations, and refer to the network as having n nodes.

III. MAIN RESULT

A. Topology learning problem

Let $\boldsymbol{P}, \boldsymbol{V} \in \mathbb{R}^{n \times s}$ be data matrices whose columns are s samples of active power and voltage magnitude measurement vectors, respectively, across all n nodes. We seek to recover an n-sparse vector $\boldsymbol{w}_\star$ of line parameters, where $w_{\star,i} = 1/z_{\star,i}$ if $z_{\star,i} \neq 0$, and 0 otherwise. This problem can be written as a sparse recovery problem over the *complete graph* by defining the measurement system

$$\boldsymbol{p} = \mathcal{Q}\left(\mathrm{vec}\left(\boldsymbol{P}\right)\right) = \mathcal{Q}\left(\boldsymbol{A}\boldsymbol{w}_\star\right), \quad \boldsymbol{p} \in \mathbb{R}^{sn}. \tag{10}$$

The $sn \times \binom{n+1}{2}$ sensing matrix is $\boldsymbol{A} \triangleq \boldsymbol{V}^\top \tilde{\boldsymbol{C}}^\top \odot \tilde{\boldsymbol{C}}^\top$, where $\odot$ is the Khatri-Rao matrix product.

B. Bounding the error of the parameter estimate

Taking advantage of the sparsity, we can characterize the Gaussian width of the tree set (9).

Lemma 1. *Define the set $\mathcal{T}_{\mathcal{R}, \boldsymbol{w}_\star} \triangleq \mathcal{D}(\mathcal{R}, \boldsymbol{w}_\star) \cap \mathbb{S}^{d-1}$ as the intersection of the tangent cone (5) of constraint set $\mathcal{R}$ with the unit shell $\mathbb{S}^{d-1} \triangleq \left\{\boldsymbol{u} \in \mathbb{R}^d \ : \ ||\boldsymbol{u}||_2 = 1\right\}$, $d = \binom{n+1}{2}$. The squared Gaussian width of the set of radial networks (9) satisfies $\omega^2\left(\mathcal{T}_{\mathcal{R}, \boldsymbol{w}_\star}\right) \leq 2n \log\left(\frac{n+1}{2}\right) + \frac{3}{2}n$.*

Proof. The ground truth $\boldsymbol{w}_\star$ corresponds to a spanning tree of the complete graph with $n+1$ nodes. Since this tree must have n edges, vector $\boldsymbol{w}_\star$ must be n-sparse, i.e., $||\boldsymbol{w}_\star||_0 = n$. Dropping connectivity, define the relaxation $\mathcal{K} \supseteq \mathcal{R}$ as

$$\mathcal{K} \triangleq \left\{\boldsymbol{w} \in \mathbb{R}^{\binom{n+1}{2}} : ||\boldsymbol{w}||_1 \leq ||\boldsymbol{w}_\star||_1\right\}, \tag{11}$$

i.e., the ℓ_1-norm ball of radius $||\boldsymbol{w}_\star||_1$. It is well known (cf. [11], [6]) that, as $\boldsymbol{w}_\star$ is n-sparse, we can conclude that $\omega^2(\mathcal{T}_{\mathcal{R}, \boldsymbol{w}_\star}) \leq \omega^2(\mathcal{T}_{\mathcal{K}, \boldsymbol{w}_\star}) \leq 2n \log(\frac{n+1}{2}) + \frac{3}{2}n$. $\qquad \square$
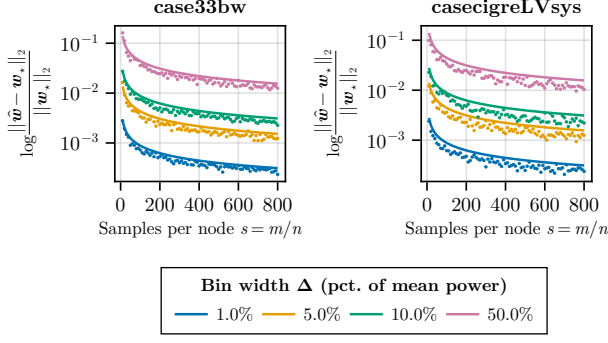
Fig. 2. Relative error vs. number of samples for repeated runs of the topology learning program (3) on the Baran & Wu (left) and CIGRE low-voltage (right) test cases. The scatter markers show the experimental error obtained for one run of (3), where each color corresponds to a bin width Δ. The solid lines show the error prescribed by (12).

Using Lemma 1, the error of the topology learned by (3) with constraint set (11) and bin width $\Delta > 0$ is bounded and can be quantified.

Theorem 1. *Suppose that $s = \frac{m}{n}$ samples with bin width $\Delta > 0$ are collected throughout an n-node distribution network, and suppose that the $\{\boldsymbol{a}_i\}$ are any i.i.d. sub-Gaussian random vectors. Then, there exist constants $C, c_1, c_2 > 0$ such that the error of the topology estimate from solving (3), with probability at least 0.99 (i.e., 99%) is bounded as*

$$||\hat{\boldsymbol{w}} - \boldsymbol{w}_\star||_2 \le C\Delta \cdot \sqrt{\frac{2\log\left(\frac{n+1}{2}\right) + 3/2}{s}}, \qquad (12)$$

provided the number of samples $s \ge \mathcal{O}(\log n)$.

Proof. Combining Lemma 1 with (13) and applying [6, Thm. III.1], there exists a constant $C > 0$ such that the minimizer $\hat{\boldsymbol{w}}$ of the program (3) satisfies, with probability at least 0.99,

$$||\hat{\boldsymbol{w}} - \boldsymbol{w}_\star||_2 \le C\Delta \cdot \frac{\omega\left(\mathcal{T}_{\mathcal{K},\boldsymbol{w}_\star}\right)}{\sqrt{m}}, \qquad (13a)$$

$$\le C\Delta \cdot \sqrt{\frac{2\mathcal{H}\log\left(\frac{n+1}{2}\right) + \frac{3}{2}\mathcal{H}}{s\mathcal{H}}}, \qquad (13b)$$

provided the total number of measurements $m = sn \ge c_1\omega^2\left(\mathcal{T}_{\mathcal{K},\boldsymbol{w}_\star}\right) + c_2$ for some absolute constants $c_1, c_2 > 0$. Due to Lemma 1, if the number of samples per node $s \ge c_1\left(2\log\left(\frac{n+1}{2}\right) + 3/2\right) + c_2 = \mathcal{O}(\log n)$ we satisfy this measurement requirement. $\square$

C. Numerical results

Fig. 2 compares the error bound (12) with the parameter error obtained by numerically solving many instances of the program (3). The errors are plotted against the number of samples per node, varied over 100 discrete uniformly spaced points $10 \le s \le 800$. For each instance of the program (3), a random matrix $\boldsymbol{V} \in \mathbb{R}^{n \times s}$ of independent Gaussian voltages is generated, where the mean of each column is the AC power flow solution of the feeder, and the variance of each column entry is 10% thereof. Active power measurements are then generated as in (10). The quantization bin width Δ is varied as a percentage of the sample mean of the absolute active power injections

$\frac{1}{m}||\boldsymbol{A}\boldsymbol{w}_\star||_1$. The solid curves in Fig. 2 depict the error bounds given by (12), for each bin width Δ.

Following [6], the constant C is not specified by the theorem. Therefore, we numerically select C for each feeder by computing the maximum ratio of empirical errors and (12) over all s, Δ. We found $C \approx 13$ for case33bw and $C \approx 10$ for casecigreLV, which we use for the solid curves in Fig. 2. The scaling $\mathcal{O}(\Delta\sqrt{\log n/s})$ predicted by (12) matches the experimental errors across all Δ.

IV. CONCLUSION

We provided a framework to *predict the error of a learned distribution network topology—before solving any optimization problem*. By exploiting the radial structure, we achieved an error *sublinear* in the number of nodes, up to constant factors determined by the precision of the sensor and the underlying probability distributions. Formally, we showed that the topology *and* line parameters of a radial network can be recovered up to a relative error of $\mathcal{O}(\Delta\sqrt{\log n/s})$, given a collection of $\mathcal{O}(\log n)$ samples from every node with quantization bin width $\Delta > 0$. Thus, before collecting data, one can size quantization and sampling budgets to hit a target topology error. Note that (13) is determined by the bin width Δ. Deriving bounds involving communication bit rates would require assumptions on the variance of the underlying measurement being quantized.

REFERENCES

[1] C. Huang *et al.*, "Smart Meter Pinging and Reading Through AMI Two-Way Communication Networks to Monitor Grid Edge Devices and DERs," *IEEE Trans. Smart Grid*, pp. 4144–4153, 2022.
[2] A. Rangarajan, R. K. Gupta, D. K. Molzahn, and L. A. Roald, "Forecast-Aided State Estimation in Unbalanced Distribution Networks Using Smart Meter Data under Limited Communication Bandwidth," to appear in *Sustainable Energy, Grids and Networks*, June 2025.
[3] R. Gray and D. Neuhoff, "Quantization," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2325–2383, 1998.
[4] A. Gholami, S. Kim, Z. Dong, Z. Yao, M. W. Mahoney, and K. Keutzer, "A Survey of Quantization Methods for Efficient Neural Network Inference," in *Low-Power Computer Vision*, pp. 291–326, Chapman and Hall/CRC, 2022.
[5] Y. Plan and R. Vershynin, "The Generalized LASSO with Non-Linear Observations," *IEEE Trans. Inf. Theory*, vol. 62, no. 3, pp. 1528–1537, 2016.
[6] C. Thrampoulidis and A. S. Rawat, "The Generalized LASSO for Sub-Gaussian Measurements With Dithered Quantization," *IEEE Trans. Inf. Theory*, vol. 66, no. 4, pp. 2487–2500, 2020.
[7] 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.
[8] V. Vapnik, *The Nature of Statistical Learning Theory*. Springer Science & Business Media, 2013.
[9] N. Rin, I. Shames, I. R. Petersen, and E. L. Ratnam, "Electric Grid Topology and Admittance Estimation: Quantifying Phasor-based Measurement Requirements." to appear in the *American Control Conference (ACC)*, 2025.
[10] R. Gray and T. Stockham, "Dithered Quantizers," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 805–812, 1993.
[11] V. Chandrasekaran, B. Recht, P. A. Parrilo, and A. S. Willsky, "The Convex Geometry of Linear Inverse Problems," *Foundations of Computational Mathematics*, vol. 12, no. 6, pp. 805–849, 2012.
[12] R. Vershynin, *High-Dimensional Probability: An Introduction with Applications in Data Science*. Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, 2025.
[13] S. Bolognani and F. Dörfler, "Fast Power System Analysis via Implicit Linearization of the Power Flow Manifold," in *53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 402–409, 2015.
[14] O. Stanojev, L. Werner, S. Low, and G. Hug, "Tractable Identification of Electric Distribution Networks," in *62nd IEEE Conference on Decision and Control (CDC)*, pp. 549–555, 2023.
[15] S. Bolognani and K. Moffat, "Chapter Six - Grid Topology Identification via Distributed Statistical Hypothesis Testing," in *Big Data Application in Power Systems*, pp. 109–134, Elsevier, 2nd ed., 2024.