A Bayesian Method to Infer Parameters in Power Flow Models Using Linear Sensitivities

Drew Séguin†, Xun Huan‡, Yu Christine Chen†
†The University of British Columbia, Vancouver, BC, Canada, {drewseg, chen}@ece.ubc.ca
‡University of Michigan, Ann Arbor, MI, United States, xhuan@umich.edu

***Abstract*—This paper presents a Bayesian method to infer parameters in distribution system power flow models from noisy measurements of voltage magnitudes and phase angles along with active- and reactive-power injections collected from a subset of buses with synchronized phasor measurement capability. The proposed method bypasses the large number of repeated nonlinear power flow solutions that would typically be required in sampling-based Bayesian inference. Instead, the proposed method iteratively and analytically linearizes the nonlinear power flow model, converging to the linearized model with the maximum probability of being (closest to) the actual nonlinear model that gave rise to the measurement data. The combination of the linear system, Gaussian parameter prior, and Gaussian measurement noise enables closed-form evaluation of the parameter posterior, model evidence, and their gradients. This can help to improve computational scalability for large-scale networks with potentially many unknown parameters to be inferred. We illustrate the effectiveness and key features of the proposed method with numerical case studies involving the IEEE 33-bus test system.**

***Index Terms*—Bayesian inference, Bayesian model selection, parameter estimation, power flow model, linear sensitivities**

I. Introduction

The transition to a low-carbon future calls for modern distribution grids to serve the ever-growing demand for electricity and newly electrified loads with renewable and distributed energy resources. The many expected new electric loads and energy sources can lead to unexpected or undesired changes to the voltage profile and power flows along a distribution feeder. Thus, there is growing need for utilities to pursue online monitoring and active control of distribution grids and devices therein via, e.g., state estimation, demand-side management, and optimal dispatch [1]–[3]. Such advanced monitoring and control techniques often rely on an offline distribution system power flow model constructed with accurate parameters for the loads, sources, network, and other components therein.

In this work, we compute approximate posterior probability density functions (PDFs) for parameters in the power flow model conditioned on noisy measurements of voltage phasors and complex-power injections collected from distribution-level phasor measurement units (D-PMUs) installed at a subset of buses in a distribution network. Our approach is rooted in Bayesian *inference*, which seeks the entire distribution of all plausible parameter values that give rise to the recorded measurements, circumventing practical challenges of satisfying the observability criterion, i.e., having sufficient sensor coverage, fraught in observer-based *estimation*. Moreover, inference techniques quantify the uncertainty in the resulting parameters affected by, e.g., the quantity and quality of measurement data, which is very useful to provide confidence measures that then inform risk-aware decision making [4]. However, uncertainty quantification is not featured in standard estimation-based approaches that tune parameter values to best match behaviour predicted by a given model with the measurement data [5].

Given the benefits on offer over estimation techniques, Bayesian methods have been applied to infer system state variables (i.e., bus voltages and injections) [6], [7]. Moreover, in anticipation of the need for distribution system model calibration, Bayesian methods have been used to infer locations of potential outages [8], parameters in composite load models [9], entries in the network admittance matrix [10], and changes in the network configuration [11]. Classical Bayesian inference uses Markov chain Monte Carlo (MCMC) algorithms that typically require thousands (or more) repeated simulations of the nonlinear system model under study. Our recent work in [4] to infer dynamic model parameters bypasses the large number of system model simulations typically required in sampling-based Bayesian inference. Instead, it relies on linearizations of the nonlinear system model, and the linear models are constructed to maximize the probability of being (closest to) the actual nonlinear model that gave rise to the measurement data.

This paper tailors the Bayesian inference method developed in [4] and offers extensions in several directions. First, we focus on inferring parameters pertinent to *static* power flow models more suitable for application in power distribution networks instead of the transmission-level *dynamic* models examined in [4]. Also, considering the expected heterogeneity in loads and sources in future distribution systems, we extend the power flow model typically featuring only constant-power loads to capture more realistic models of distribution-level nodal injections. Although the proposed method is general in the sense that it can infer parameters in any number of models for, e.g., electrified loads and distributed generation, numerical case studies involving the IEEE 33-bus system focus on parameters in the ZIP model for loads. The proposed method avoids high-volume nonlinear system simulations performed in [8], [9]. We further use measurements obtained at only a single snapshot in time instead of those collected over multiple time steps as in [8]–[10], over which the parameter values can potentially vary. Also, distinct from the discrete Bayesian model selection problem posed in [11] that relies on parallel calculations of posteriors for many candidate models, we consider a continuous spectrum of linearized models to arrive at the "best" one after examining only several candidates.

II. Preliminaries

In this section, we present the power flow model and pertinent linear sensitivities. We further describe the measurement statistical model and state the Bayesian parameter inference problem tackled in this paper.

A. Power Flow Model

Consider an AC distribution system with buses collected in the set $\mathcal{N} = \{1, 2, \ldots, n\}$. Over the interconnected distribution grid, a substation and potentially distributed generation serve nodal loads across the system. Let $x \in \mathbb{R}^{2n}$ collect voltage magnitudes and phase-angles of all buses in the power distribution network. Also let $u \in \mathbb{R}^{2n}$ collect net active- and reactive-power injections at all buses in the network. Further let $\lambda \in \mathbb{R}^p$ collect unknown parameters to be inferred. Then, the power flow equations can be expressed compactly as

$$0 = g(x, u; \lambda), \tag{1}$$

where, for a given λ, $g : \mathbb{R}^{2n} \times \mathbb{R}^{2n} \mapsto \mathbb{R}^{2n}$. In general, u may be modelled as a function of the nodal voltages and unknown parameters to be inferred, i.e.,

$$u = f(x; \lambda), \tag{2}$$

where, for a given λ, $f : \mathbb{R}^{2n} \mapsto \mathbb{R}^{2n}$. Also, system output $z \in \mathbb{R}^m$ can be mapped from the state variables, as follows:

$$z = h(x, u; \lambda). \tag{3}$$

where $f : \mathbb{R}^{2n} \times \mathbb{R}^{2n} \mapsto \mathbb{R}^m$, again, for a given λ. Note that entries of λ are distinct from state variables in x and nodal power injections in u. Next, via a remark, we detail the entries in x and u, potential parameters of interest in λ to be inferred, and possible quantities in z.

Remark 1 (Detailing system model in (1)–(3))**.** Let V_i and θ_i denote the voltage magnitude and phase angle at bus $i \in \mathcal{N}$, respectively. Also let P_i and Q_i denote the net active- and reactive-power injection at bus $i \in \mathcal{N}$. Power flow equations collected in (1) consist of nodal active- and reactive-power balance for each bus $i \in \mathcal{N}$ respectively expressed as

$$0 = V_i \sum_{k=1}^{n} V_k (G_{ik} \cos(\theta_i - \theta_k) + B_{ik} \sin(\theta_i - \theta_k)) - P_i, \tag{4}$$

$$0 = V_i \sum_{k=1}^{n} V_k (G_{ik} \sin(\theta_i - \theta_k) - B_{ik} \cos(\theta_i - \theta_k)) - Q_i, \tag{5}$$

where G_{ik} and B_{ik} are the real and imaginary parts of the (i, k) entry in the network admittance matrix, respectively. In (1), we let $x = [\theta_1, \ldots \theta_n, V_1, \ldots, V_n]^{\mathrm{T}}$ and $u = [P_1, \ldots, P_n, Q_1, \ldots, Q_n]^{\mathrm{T}}$. Unknown parameters of interest in λ can include G_{ik} and B_{ik} in (4)–(5), leading to inference of line impedance values. They may also include parameters needed to fully articulate nodal power injection models expressed in (2). An example of such parameters are coefficients in the ZIP model for a load at bus $i \in \mathcal{N}$, given by

$$P_i = c_i^{\mathrm{z}} V_i^2 + c_i^{\mathrm{i}} V_i + c_i^{\mathrm{p}}, \tag{6}$$

$$Q_i = d_i^{\mathrm{z}} V_i^2 + d_i^{\mathrm{i}} V_i + d_i^{\mathrm{p}}, \tag{7}$$

where c_i^{z}, c_i^{i}, c_i^{p}, d_i^{z}, d_i^{i}, and d_i^{p} can be inferred from online measurements to ensure the ZIP model remains representative of actual load characteristics potentially changing in real time. Additional examples include parameters pertinent to models of active- and reactive-power injections stemming from distributed generation or other DERs. Finally, system output z in (3) can comprise V_i, θ_i, P_i, and Q_i, $i \in \mathcal{N}$. It can also include line active- and reactive-power flows expressed as a function of bus voltage magnitudes and phase angles. ■

B. Linear Sensitivities

Suppose that for nominal parameter value $\lambda = \lambda^\star$, there exists power flow solution $(x^\star, u^\star; \lambda^\star)$. Denote by $x_\lambda^\star \in \mathbb{R}^{2n \times p}$ and $u_\lambda^\star \in \mathbb{R}^{2n \times p}$ the linear sensitivities of x and u, respectively, with respect to λ around the nominal operating point $(x^\star, u^\star; \lambda^\star)$. Applying the chain rule, differentiation of (1)–(2) with respect to λ yields

$$0 = g_x^\star x_\lambda^\star + g_u^\star u_\lambda^\star + g_\lambda^\star, \tag{8}$$

$$u_\lambda^\star = f_x^\star x_\lambda^\star + f_\lambda^\star, \tag{9}$$

where

$$g_x^\star = \frac{\partial g}{\partial x},\ g_u^\star = \frac{\partial g}{\partial u},\ g_\lambda^\star = \frac{\partial g}{\partial \lambda},\ f_x^\star = \frac{\partial f}{\partial x},\ f_\lambda^\star = \frac{\partial f}{\partial \lambda},$$

are suitably sized sensitivity matrices evaluated at the nominal operating point $(x^\star, u^\star; \lambda^\star)$. Similarly, differentiation of (3) with respect to λ around the nominal operating point yields sensitivities of output z, denoted by $z_\lambda^\star \in \mathbb{R}^{m \times p}$ and given by

$$z_\lambda^\star = h_x^\star x_\lambda^\star + h_u^\star u_\lambda^\star + h_\lambda^\star, \tag{10}$$

where we evaluate

$$h_x^\star = \frac{\partial h}{\partial x}, \quad h_u^\star = \frac{\partial h}{\partial u}, \quad h_\lambda^\star = \frac{\partial h}{\partial \lambda},$$

at $(x^\star, u^\star; \lambda^\star)$. Now substitute (9) into (8) and (10) to get

$$x_\lambda^\star = -(g_x^\star + g_u^\star f_x^\star)^{-1}(g_u^\star f_\lambda^\star + g_\lambda^\star), \tag{11}$$

$$z_\lambda^\star = (h_x^\star + h_u^\star f_x^\star) x_\lambda^\star + h_u^\star f_\lambda^\star + h_\lambda^\star, \tag{12}$$

where we have assumed that, at the nominal operating point, the power flow Jacobian matrix is invertible to arrive at (11). Finally, by substituting (11) into (12), we get the following expression describing how the system output sensitivities behave around the nominal operating point:

$$z_\lambda^\star = -(h_x^\star + h_u^\star f_x^\star)(g_x^\star + g_u^\star f_x^\star)^{-1}(g_u^\star f_\lambda^\star + g_\lambda^\star) + h_u^\star f_\lambda^\star + h_\lambda^\star. \tag{13}$$

Now, simultaneous solution of (1)–(3) and (13) yields output linear sensitivities in $z_\lambda^\star$ evaluated at the nominal operating point $(x^\star, u^\star; \lambda^\star)$ and the corresponding output $z^\star$.

We can use the linear sensitivities in (13) to approximate the output of a perturbed system resulting from variations in λ around $\lambda^\star$. To do this, let $z = z^\star + \Delta z$, where Δz results from $\Delta\lambda = \lambda - \lambda^\star$. Then, assuming that $\Delta\lambda$ is sufficiently small, we can approximate the perturbed output as

$$z \approx a(\lambda^\star)\lambda + b(\lambda^\star) =: \widetilde{z}(\lambda; \lambda^\star), \tag{14}$$

where $a(\lambda^\star) = z_\lambda^\star$ and $b(\lambda^\star) = z^\star - z_\lambda^\star \lambda^\star$ are parameterized by the choice of $\lambda^\star$.

C. Measurement Model and Problem Statement

Suppose D-PMUs provide synchronized phasor (or synchrophasor) measurements of nodal voltages and currents at (possibly a subset of) buses, and corresponding active- and reactive-power injections can be easily computed. Also available may be synchronized measurements of a subset of line active- and reactive-power flows. Let $\widehat{z}$ denote the measurement of system output at a particular snapshot in time. Considering noisy measurements, $\widehat{z}$ can be modelled as

$$\widehat{z} = z + \xi \approx a\lambda + b + \xi =: \widetilde{z} + \xi, \tag{15}$$

where z denote the *actual* system output, $\xi \in \mathbb{R}^m$ denote D-PMU measurement errors, and the approximation holds by substituting (14). In our simulations, z is supplied by an exact power flow solution furnished with the *true* parameter values. In (15), $\widetilde{z}$ is linear with respect to λ, and the coefficients a and b depend on the choice of $\lambda^\star$ as described in (14). Furthermore, entries of the error vector ξ are modelled to be independent and identically distributed under a joint Gaussian distribution with zero mean and covariance Σ_ξ, i.e., $\xi \sim \mathcal{N}(\mathbb{0}_m, \Sigma_\xi)$, where $\Sigma_\xi \in \mathbb{R}^{m\times m}$ is diagonal with the inverse of each diagonal entry reflecting the corresponding measurement precision.

Using the models established in this section and tailoring the method developed in [4], we tackle two interrelated but distinct problems. The first is to identify the best $\lambda^\star$ giving rise to the approximate linear model in (15) that most likely resembles the measurement-generating nonlinear system. Second, we infer the parameter λ by computing its entire distribution from measurement $\widehat{z}$, given a linearized model constructed around the nominal output resulting from a particular choice of $\lambda^\star$. We approach both problems under a Bayesian framework, where $\lambda^\star$ and λ are treated as random variables, as detailed next.

III. Bayesian Approach

We describe the strategy for the inference of λ given the model in (15) obtained by linearizing power flow equations around the nominal operating point induced by a *particular* $\lambda^\star$ choice, followed by the approach to find the *best* $\lambda^\star$ choice.

A. Inference on λ

We treat the unknown parameter λ in (15) as a random variable. It is endowed with a prior distribution representing the uncertainty in λ before making any observations, and it also has a corresponding posterior distribution representing the updated uncertainty after observing measurement data in $\widehat{z}$. Direct application of Bayes' theorem for conditional probability yields the following parameter-posterior PDF:

$$f(\lambda|\widehat{z}, \lambda^\star) = \frac{f(\widehat{z}|\lambda, \lambda^\star) f(\lambda)}{f(\widehat{z}|\lambda^\star)}, \tag{16}$$

where $f(\lambda)$ is the prior PDF for λ (assuming that it is independent of the point of linearization, i.e., $\lambda^\star$), $f(\widehat{z}|\lambda, \lambda^\star)$ is the likelihood, and $f(\widehat{z}|\lambda^\star)$ is the model evidence.

Specifically, we prescribe Gaussian prior $\lambda \sim \mathcal{N}(\mu_\circ, \Sigma_\circ)$ to represent the initial uncertainty in λ. The likelihood then follows from the linearized measurement model in (15):

$$f(\widehat{z}|\lambda, \lambda^\star) = (2\pi)^{-\frac{m}{2}} |\Sigma_\xi|^{-\frac{1}{2}} e^{-\frac{1}{2}(\widehat{z}-\widetilde{z})^{\mathrm{T}} \Sigma_\xi^{-1} (\widehat{z}-\widetilde{z})}, \tag{17}$$

where $\widetilde{z}$ is evaluated for a given parameter value λ with the linearized model in (14) constructed from a particular choice of $\lambda^\star$. In (17), we drop the dependence of $\widetilde{z}$ on λ and $\lambda^\star$ to contain notational burden. The combination of linear model together with Gaussian prior and likelihood leads to a conjugate system with Gaussian posterior $(\lambda|\widehat{z}, \lambda^\star) \sim \mathcal{N}(\mu_{\pi,\lambda^\star}, \Sigma_{\pi,\lambda^\star})$, where the mean and covariance are given in closed form by

$$\mu_{\pi,\lambda^\star} = \Sigma_{\pi,\lambda^\star} \Big(\Sigma_\circ^{-1} \mu_\circ + a^{\mathrm{T}} \Sigma_\xi^{-1} (\widehat{z} - b) \Big), \tag{18}$$

$$\Sigma_{\pi,\lambda^\star} = \Big(\Sigma_\circ^{-1} + a^{\mathrm{T}} \Sigma_\xi^{-1} a \Big)^{-1}, \tag{19}$$

respectively. Above, the subscript π indicates posterior and the subscript $\lambda^\star$ reminds us that the mean and covariance of the posterior depend on the choice of $\lambda^\star$ for linearization.

B. Choice of $\lambda^\star$

Considering that the linear model in (15) is constructed given a particular $\lambda^\star$, we also formulate and solve an optimal model selection problem to identify the best $\lambda^\star$ giving rise to the linear model that most resembles the measurement-generating nonlinear system. Treating $\lambda^\star$ as a random variable and persisting under the Bayesian framework, we adopt the methods of Bayesian model selection (or equivalently, Bayes factors) applied to a continuous spectrum of models parameterized by $\lambda^\star$ [12], [13]. Application of Bayes' theorem given a candidate $\lambda^\star$ (mapping to a corresponding linear model) yields the following model-posterior PDF:

$$f(\lambda^\star|\widehat{z}) = \frac{f(\widehat{z}|\lambda^\star) f(\lambda^\star)}{f(\widehat{z})}, \tag{20}$$

representing the "goodness" of the linear model induced by $\lambda^\star$ given observations made from $\widehat{z}$. The *best* candidate $\lambda^\star$ thus maximizes this quantity (equivalently, its logarithm) as

$$\lambda^\star_{\mathrm{opt}} = \arg\max_{\lambda^\star} \ \ln f(\lambda^\star|\widehat{z}) = \arg\max_{\lambda^\star} \ \ln f(\widehat{z}|\lambda^\star), \tag{21}$$

where the second equality holds by recognizing that $f(\widehat{z})$ is a constant normalization factor that does not depend on $\lambda^\star$ and by adopting a uniform model-prior that does not initially favour any particular region of $\lambda^\star$ (i.e., $f(\lambda^\star)$ remains constant regardless of the choice of $\lambda^\star$).

The key to solving (21) is recognizing that the model likelihood $f(\widehat{z}|\lambda^\star)$ is precisely the model evidence (i.e., the denominator) in (16). Here, we can obtain it in closed form owing to the analytical parameter-posterior Gaussian PDF with mean and covariance in (18)–(19). Taking the logarithm of (16) and rearranging the resultant, we get

$$\ln f(\widehat{z}|\lambda^\star) = \ln f(\widehat{z}|\lambda, \lambda^\star) + \ln f(\lambda) - \ln f(\lambda|\widehat{z}, \lambda^\star). \tag{22}$$

We then substitute into (22) the likelihood expression in (17) along with the closed-form expressions for the prior $\lambda \sim \mathcal{N}(\mu_\circ, \Sigma_\circ)$ and posterior $(\lambda|\widehat{z}, \lambda^\star) \sim \mathcal{N}(\mu_{\pi,\lambda^\star}, \Sigma_{\pi,\lambda^\star})$, and finally arrive at the following analytical closed-form expression for the log-evidence:

$$\begin{aligned} \ln f(\widehat{z}|\lambda^\star) = &-\frac{m}{2}\ln(2\pi) - \frac{1}{2}\ln|\Sigma_\xi| - \frac{1}{2}\ln|\Sigma_\circ| \\ &- \frac{1}{2}(\widehat{z}-\widetilde{z})^{\mathrm{T}} \Sigma_\xi^{-1} (\widehat{z}-\widetilde{z}) - \frac{1}{2}(\lambda-\mu_\circ)^{\mathrm{T}} \Sigma_\circ^{-1} (\lambda-\mu_\circ) \end{aligned}$$

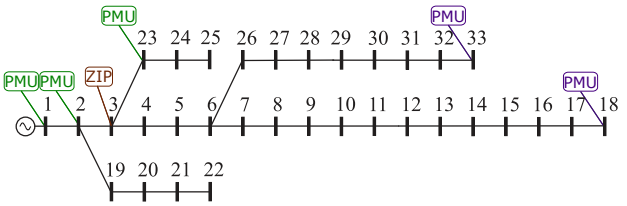

Fig. 1: One-line diagram for IEEE 33-bus test system. We infer ZIP-model parameters for the active-power load connected to bus 3. Synchrophasor measurements are available from buses 1, 2, and 23 (marked in green colour) for results presented in Fig. 3a, and at buses 18 and 33 (marked in purple colour) for those in Fig. 3b.

$$+\frac{1}{2}\ln|\Sigma_{\pi,\lambda^\star}|+\frac{1}{2}(\lambda-\mu_{\pi,\lambda^\star})^{\mathrm{T}}\Sigma_{\pi,\lambda^\star}^{-1}(\lambda-\mu_{\pi,\lambda^\star}), \quad (23)$$

where $\widetilde{z}$, $\mu_{\pi,\lambda^\star}$, and $\Sigma_{\pi,\lambda^\star}$ depend on the value of $\lambda^\star$.

Various optimization algorithms (e.g., gradient-based, quasi-Newton, and derivative-free methods) can be employed to iteratively update candidates for $\lambda^\star$ toward the optimizer $\lambda^\star_{\mathrm{opt}}$ of (21). For example, adopting gradient-ascent leads to the following update formula:

$$\lambda^\star_{(\ell+1)}=\lambda^\star_{(\ell)}+\gamma_{(\ell)}\nabla_{\lambda^\star}\ln f(\widehat{z}|\lambda^\star)|_{\lambda^\star_{(\ell)}}, \quad (24)$$

where $\gamma_{(\ell)}$ is a learning rate (gradient-ascent step size) and $\nabla_{\lambda^\star}\ln f(\widehat{z}|\lambda^\star)|_{\lambda^\star_{(\ell)}}$ is the gradient of the objective $\ln f(\widehat{z}|\lambda^\star)$ evaluated at $\lambda^\star_{(\ell)}$. A major advantage of the proposed framework is that the objective function in (21) and its gradient can be computed in closed form, potentially enabling greater scalability. Particularly, we completely bypass all numerical approximations of the gradient involving, e.g., finite differences, which may be computationally impractical for high-dimensional $\lambda^\star$. Interested readers may refer to [4] for details with respect to analytical computation of the gradient in (24).

IV. Numerical Case Studies

This section reports on simulations involving a modified IEEE 33-bus distribution test system for which the one-line diagram shown in Fig. 1, with a system power base of 10 MVA [14]. Simulations were performed using Python 3.10.12 & Jupyter Notebooks inside a Debian-based virtual machine on a personal computer with 32 GB RAM and Apple M1 Pro 10-core processor. Power flow solutions were obtained using `pandapower` [15], and custom Python code was used to implement the proposed method.

A. Simulation Setup

We infer unknown ZIP-model parameters c_3^{z}, c_3^{i}, and c_3^{p} for the active-power load connected to bus 3, i.e., $\lambda=[c_3^{\mathrm{z}},c_3^{\mathrm{i}},c_3^{\mathrm{p}}]^{\mathrm{T}}$. The true measurement-generating value for the unknown parameters is $\lambda_{\mathrm{true}}=[0.22,0.44,0.34]^{\mathrm{T}}$. Synchrophasor measurements of nodal active- and reactive-power injections as well as voltage magnitudes and phase angles collected in $\widehat{z}$ are subject to additive Gaussian noise with zero mean and 1% standard deviation, in accordance with the IEEE standard for PMU measurement error [16]. Particularly, we set the measurement noise $\xi\sim\mathcal{N}(\mathbb{0}_m,\Sigma_\xi)$, where $\Sigma_\xi=\mathrm{diag}(0.01z)^2$ is a diagonal matrix and z is the actual system output supplied by the exact nonlinear power flow solution using the true parameter values in λ_{true}.

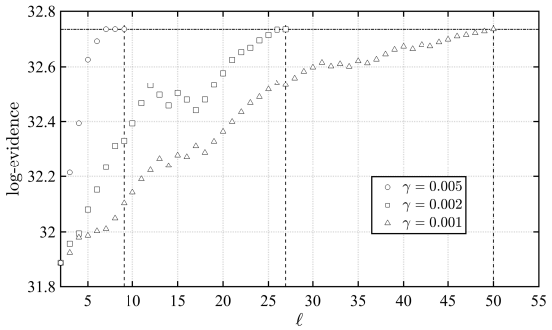

Fig. 2: Convergence of log-evidence $\ln f(\widehat{z}|\lambda^\star_{(\ell)})$ to $\ln f(\widehat{z}|\lambda^\star_{\mathrm{opt}})$ (marked as dash-dot trace) with learning rates of $\gamma_{(\ell)}=\gamma=0.005$, 0.002, and 0.001 for all ℓ. Dashed traces mark the iteration at which stopping criterion $||\lambda^\star_{(\ell+1)}-\lambda^\star_{(\ell)}||<10^{-6}$ is satisfied and we set $\lambda^\star_{\mathrm{opt}}=\lambda^\star_{(\ell+1)}$.

B. Simulation Results

We begin by considering measurement scenario i) with $z=[V_1,P_1,V_2,P_2,V_{23},P_{23}]^{\mathrm{T}}$. Measurements are obtained from buses near (but *not* at) the load with unknown parameters.

1) Choosing $\lambda^\star$: We set $\lambda^\star_{(0)}=[0.242,0.484,0.374]^{\mathrm{T}}$. In each iteration ℓ, we solve (1)–(3) and (13) with the updated candidate $\lambda^\star_{(\ell)}$, with which the approximate output $\widetilde{z}(\lambda,\lambda^\star_{(\ell)})$ is constructed from (14). We next evaluate the mean $\mu_{\pi,\lambda^\star_{(\ell)}}$ and covariance $\Sigma_{\pi,\lambda^\star_{(\ell)}}$ in closed form using (18) and (19), respectively. We also evaluate the log-evidence via (23) leveraging $\mu_{\pi,\lambda^\star_{(\ell)}}$, $\Sigma_{\pi,\lambda^\star_{(\ell)}}$, and $\widetilde{z}(\lambda,\lambda^\star_{(\ell)})$. We then use (24) to obtain the next update $\lambda^\star_{(\ell)}$. The iterative procedure continues until the stopping criterion $||\lambda^\star_{(\ell+1)}-\lambda^\star_{(\ell)}||<10^{-6}$, at which point we set $\lambda^\star_{\mathrm{opt}}=\lambda^\star_{(\ell+1)}$. Plotted in Fig. 2 are log-evidence trajectories for three learning rates. Based on trends observed therein, we set $\gamma_{(\ell)}=0.005$, for all ℓ, in results reported below.

2) Inferring λ: For all values of $\lambda^\star_{(\ell)}$, we prescribe Gaussian prior $\lambda\sim\mathcal{N}(\mu_\circ,\Sigma_\circ)$, where $\mu_\circ=[0.33,0.66,0.51]^{\mathrm{T}}$ and $\Sigma_\circ=\mathrm{diag}([0.11^2,0.22^2,0.17^2])$. In practice, the prior PDF can be informed by prior experience or by domain experts. We evaluate the posterior mean $\mu_{\pi,\lambda^\star}$ and covariance $\Sigma_{\pi,\lambda^\star}$ in closed form again via (18) and (19), respectively, with $\lambda^\star=\lambda^\star_{\mathrm{opt}}$ upon convergence of the $\lambda^\star$-updates in (24). Using green-coloured traces in Fig. 3a, we plot the posterior marginal PDFs of each inferred ZIP-model parameter. We observe that the posterior distributions are narrower than their corresponding priors, indicating the uncertainty has reduced after being informed by the noisy measurement data. We also note that the centres of the posterior marginal distributions are quite close to the corresponding true parameter values.

3) Model Outputs with λ Posterior: We sample the posterior PDF of $\lambda\sim\mathcal{N}(\mu_{\pi,\lambda^\star_{\mathrm{opt}}},\Sigma_{\pi,\lambda^\star_{\mathrm{opt}}})$ 75 times and, for each, solve the *nonlinear* power flow. Resulting output values are shown in Fig. 4 as box-and-whisker plots, representing the posterior-predictive distribution. Also plotted are model outputs under the true parameter values and the recorded noisy measurements used to infer parameters. All values plotted in Fig. 4 are normalized using their respective model outputs under the true parameter values. The sampled outputs from the posterior-predictive distribution match well to the model outputs, so the inferred parameters indeed could have induced the measurement data. Further, the posterior-predictive distri-

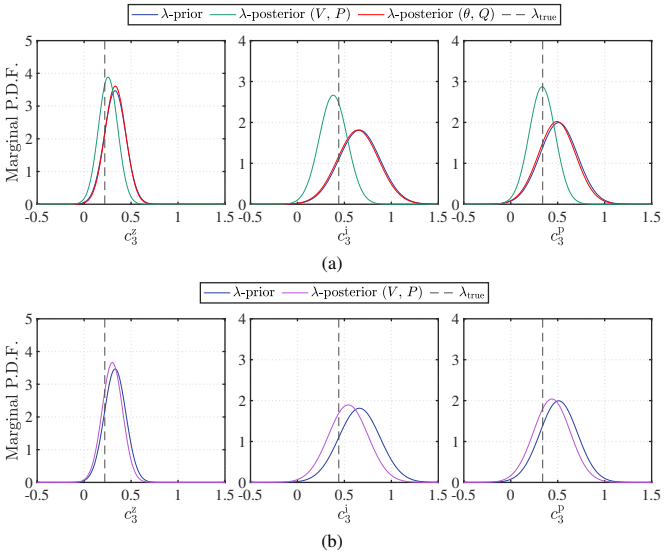

Fig. 3: Inference of ZIP-model parameters for active-power load at bus 3 using measurements obtained from (a) nearby buses 1, 2, and 23, and (b) far-away buses 18 and 33. Green-coloured traces correspond to measurement scenario i) with $z = [V_1, P_1, V_2, P_2, V_{23}, P_{23}]^{\mathrm{T}}$, red-coloured traces correspond to scenario ii) with $z = [\theta_1, Q_1, \theta_2, Q_2, \theta_{23}, Q_{23}]^{\mathrm{T}}$, purple-coloured traces correspond to scenario iii) with $z = [V_{18}, P_{18}, V_{33}, P_{33}]^{\mathrm{T}}$.

bution does not simply centre around the noisy measurements, but benefits from a regularizing effect from the Bayesian framework to navigate toward the true data-generating values.

Remark 2 (Effects of Measurement Type and Location)**.** We repeat the procedure described above, but considering scenario ii) with $z = [\theta_1, Q_1, \theta_2, Q_2, \theta_{23}, Q_{23}]^{\mathrm{T}}$, where measurements are obtained at the same locations as in scenario i) but they differ in *type*. We plot the posterior marginal PDFs of inferred ZIP-model parameters as red-coloured traces in Fig. 3a, which are nearly identical to those of the prior, indicating that measurements from scenario ii) do not serve to inform the uncertainty in unknown parameters. Finally, we consider scenario iii) with $z = [V_{18}, P_{18}, V_{33}, P_{33}]^{\mathrm{T}}$, where measurements are obtained at buses located far away from the load with unknown parameters. We plot the posterior marginal PDFs of the inferred parameters as purple-coloured traces in Fig. 3b. This set of measurements leads to very slight reduction in parameter uncertainty compared to the prior, while the centres of the distributions move closer to the true parameter values. ■

V. Concluding Remarks

We applied a Bayesian method to infer unknown parameters in a power flow model conditioned on noisy D-PMU measurements available from a subset of buses in the distribution network. The method is tailored from the one in [4] and inherits its main advantage of avoiding computationally expensive MCMC sampling of the posterior and instead computes an approximate posterior in closed form, promising for scalability to higher dimensional settings. Compelling directions for future work include assessing computational performance and scalability via simulations involving large-scale test systems with more unknown parameters, incorporating non-Gaussian measurement noise models into the proposed framework, and

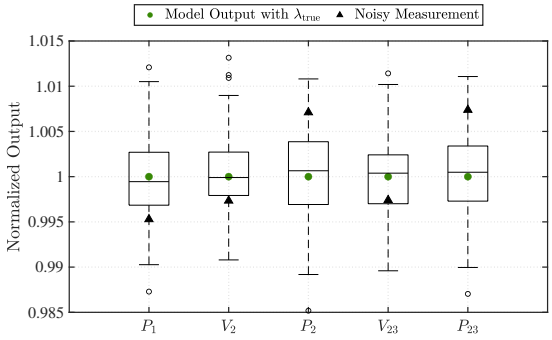

Fig. 4: Comparison of (normalized) outputs that are obtained from noisy measurements, power flow solution with true parameter values, and power flow solutions with parameter values sampled from the posterior PDF.

inferring parameters needed to model inter-temporal characteristics pertinent to, e.g., deferrable loads and energy storage.

References

[1] A. Primadianto and C.-N. Lu, "A Review on Distribution System State Estimation," *IEEE Trans. Power Syst.*, vol. 32, no. 5, pp. 3875–3883, 2017.

[2] S. Bahrami, Y. C. Chen, and V. W. S. Wong, "Deep Reinforcement Learning for Demand Response in Distribution Networks," *IEEE Trans. Smart Grid*, vol. 12, no. 2, pp. 1496–1506, 2021.

[3] E. Dall'Anese, S. S. Guggilam, A. Simonetto, Y. C. Chen, and S. V. Dhople, "Optimal Regulation of Virtual Power Plants," *IEEE Trans. Power Syst.*, vol. 33, no. 2, pp. 1868–1881, 2018.

[4] R. Nagi, X. Huan, and Y. C. Chen, "Bayesian Inference of Parameters in Power System Dynamic Models Using Trajectory Sensitivities," *IEEE Transa. on Power Syst.*, vol. 37, no. 2, pp. 1253–1263, 2022.

[5] K. Bollinger, H. Khalil, L. Li, and W. Norum, "A method for on-line identification of power system model parameters in the presence of noise," *IEEE Trans. Power App. Syst.*, no. 9, pp. 3105–3111, 1982.

[6] G. Ye, Y. Xiang, M. Nijhuis, V. Cuk, and J. F. G. Cobben, "Bayesian-Inference-Based Voltage Dip State Estimation," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 11, pp. 2977–2987, 2017.

[7] S. Dahale and B. Natarajan, "Bayesian Framework for Multi-Timescale State Estimation in Low-Observable Distribution Systems," *IEEE Trans. Power Syst.*, vol. 37, no. 6, pp. 4340–4351, 2022.

[8] J. A. D. Massignan, J. B. A. London, M. Bessani, C. D. Maciel, R. Z. Fannucchi, and V. Miranda, "Bayesian Inference Approach for Information Fusion in Distribution System State Estimation," *IEEE Trans. Smart Grid*, vol. 13, no. 1, pp. 526–540, 2022.

[9] C. Fu, Z. Yu, D. Shi, H. Li, C. Wang, Z. Wang, and J. Li, "Bayesian Estimation Based Parameter Estimation for Composite Load," in *IEEE Power & Energy Society General Meeting*, 2019, pp. 1–5.

[10] J.-S. Brouillon, E. Fabbiani, P. Nahata, K. Moffat, F. Dörfler, and G. Ferrari-Trecate, "Bayesian Error-in-Variables Models for the Identification of Distribution Grids," *IEEE Trans. Smart Grid*, vol. 14, no. 2, pp. 1289–1299, 2023.

[11] R. Singh, E. Manitsas, B. C. Pal, and G. Strbac, "A Recursive Bayesian Approach for Identification of Network Configuration Changes in Distribution System State Estimation," *IEEE Trans. Power Syst.*, vol. 25, no. 3, pp. 1329–1336, 2010.

[12] R. E. Kass and A. E. Raftery, "Bayes Factor," *Journal of American Statistical Association*, vol. 90, no. 430, pp. 773–795, 1995.

[13] L. Wasserman, "Bayesian model selection and model averaging," *Journal of Mathematical Psychology*, vol. 44, pp. 92–107, 2000.

[14] M. Baran and F. Wu, "Network reconfiguration in distribution systems for loss reduction and load balancing," *IEEE Trans. Power Del.*, vol. 4, no. 2, pp. 1401–1407, 1989.

[15] L. Thurner, A. Scheidler, F. Schäfer, J.-H. Menke, J. Dollichon, F. Meier, S. Meinecke, and M. Braun, "Pandapower—An Open-Source Python Tool for Convenient Modeling, Analysis, and Optimization of Electric Power Systems," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 6510–6521, 2018.

[16] "IEEE/IEC International Standard - Measuring relays and protection equipment - Part 118-1: Synchrophasor for power systems - Measurements," *IEC/IEEE 60255-118-1:2018*, pp. 1–78, 2018.