

Inferring Power System Frequency Oscillations using Gaussian Processes

Mana Jalali, Vassilis Kekatos, Siddharth Bhela, and Hao Zhu

Abstract—Synchronized data provide unprecedented opportunities for inferring voltage frequencies and rates of change of frequencies (ROCOFs) across the buses of a power system. Aligned to this goal, this work puts forth a novel framework for learning dynamics after small-signal disturbances by leveraging the tool of Gaussian processes (GPs). We extend results on inferring the input and output of a linear time-invariant system using GPs to the multi-input multi-output setup by exploiting power system swing dynamics. This physics-aware learning technique captures time derivatives in continuous time, accommodates data streams sampled potentially at different rates, and can cope with missing data and heterogeneous levels of accuracy. While Kalman filter-based approaches require knowing all system inputs, the proposed framework handles readings of system inputs, outputs, their derivatives, and combinations thereof on an arbitrary subset of buses. Relying on minimal system information, it further provides uncertainty quantification in addition to point estimates for dynamic grid signals. The required spatiotemporal covariances are obtained by exploring the statistical properties of approximate swing dynamics driven by ambient disturbances. Numerical tests verify that this technique can infer frequencies and ROCOFs at non-metered buses under (non)-ambient disturbances for a linearized dynamic model of the IEEE 300-bus benchmark.

I. INTRODUCTION

Maintaining the stability of power systems requires knowing frequency oscillations across the grid. Due to the high penetration of distributed and renewable energy resources, power systems may experience increased oscillations. On the other hand, modern devices (such as the frequency disturbance recorders (FDRs) comprising the frequency monitoring network (FNET) [1], [2]) can collect time-synchronized power grid measurements (frequency, voltage magnitude, voltage phase angle) across the grid and relay them over the Internet for processing on a centralized server. Since FDRs are installed at the distribution level, their voltage phase readings are susceptible to local injection fluctuations and arbitrary shifts caused by transformers, yet frequency readings could help predict and control system instabilities [3]. Reliable estimates of frequency excursions and their accelerations are also instrumental towards realizing load frequency control [4]. Nonetheless, frequency measurements are not available at all buses, while the available data streams may be sampled at different rates or even include missing entries due to communication failures, and computing ROCOFs as time differences is sensitive to measurement noise.

M. Jalali and V. Kekatos are with the Bradley Dept. of ECE, Virginia Tech, VA, USA. Emails: {manaj2, kekatos}@vt.edu. S. Bhela is with Siemens Technology, Princeton, NJ, USA. Email: siddharth.bhela@siemens.com. H. Zhu is with the Dept. of ECE, the University of Texas at Austin, Austin, TX 78712, USA. Email: haozhu@utexas.edu. The work was supported by the US NSF grants 1751085 and 1923221.

The existing literature on inferring power system dynamics can be broadly classified into data- and model-based methods. Data-based methods typically engage synchrophasor measurements to learn the system’s dynamic states. For example, missing entries from PMU data streams can be recovered via the low-rank plus sparse decomposition of the PMU data matrix pursued in [5]. If all PMU data is lost for one or more consecutive times, a robust matrix completion approach stacking data in a Hankel matrix shows promise to estimate the lost signal [6], [7]. Arranging synchrophasor readings in higher-order tensors rather than matrices could potentially impute data over prolonged communication failures via tensor factorization [8]. Nonetheless, data-based techniques cannot extrapolate on non-metered buses and ignore any information on the dynamic system model.

Dynamic state estimation (DSE) aims at inferring the power system states using both a system model and measurements processed typically through a Kalman filter (KF); see [9] for a comprehensive review. Plain KFs are optimal estimators that adopt a linear system model, while nonlinear power system dynamics can be handled through KF variants, such as the extended [10], [11]; the unscented [12], [13]; and ensemble KFs [14], [15]. By and large, KF-based DSE solutions operate on a localized fashion assuming each bus is connected to an infinite bus, and thus ignore the dynamic correlation among generators and loads [16], [17]. KFs operate on data collected over uniformly sampled time intervals, which renders them vulnerable to missing data. DSE approaches approximate continuous differential equations with discrete finite differences. Reference [18] suggests a function basis expansion to predict the local frequency under again a localized single-bus state-space model using expectation maximization and prediction error minimization algorithms. The work in [19] uses physics-informed deep neural networks to find frequencies for a single-machine infinite bus model.

Synchrophasor readings can be used to infer more coarse dynamical system information, such as locating the sources of oscillations. The latter task can be accomplished by comparing the arrival time of traveling waves [20]; by measuring the dissipated energy of power flows [21], [22]; or via robust principal component analysis [23]. Yet these methods require all buses to be monitored.

In a nutshell, existing methods for inferring power system oscillations cannot extrapolate to other buses, presume uniform sampling rates, and cannot naturally deal with missing data. In this paper, we aim at inferring the voltage frequencies and acceleration across buses and time instances using the powerful toolbox of GPs. We consider a centralized esti-

mation scheme in which system operators collect frequency and/or ROCOF measurements on a subset of buses and times possibly at different sampling rates. Depending on which dynamic grid quantities are measured and which are wanted, one can envision various application scenarios, such as *a1*) Estimating system frequencies and/or ROCOFs at non-metered buses to ensure stability; *a2*) Predicting frequencies and ROCOFs at monitored buses; or *a3*) Imputing missing values from an FDR data stream.

The contribution of this work is twofold. First, introduce GP modeling to learning power system frequencies and ROCOFs. This provides a comprehensive learning paradigm for processing data streams with different sampling rates and missing data, extrapolating to non-metered buses, obviating the need for approximating ROCOFs from finite differences, and providing a complete pdf description in addition to point estimates. Second, extend the idea of GP modeling linear dynamical systems to the MIMO setting by exploring an approximate homogeneous model for swing dynamics driven by ambient disturbances and utilizing its statistical properties to build physically meaningful covariances for the GP model.

Notation: Column vectors (matrices) are denoted by lower-(upper-) case boldface letters. Operator $\text{dg}(\mathbf{x})$ returns a diagonal matrix with \mathbf{x} on its main diagonal. Symbol $(\cdot)^\top$ stands for transposition, matrix \mathbf{I}_N is the $N \times N$ identity matrix, $\dot{x} = \frac{dx}{dt}$ denotes time differentiation, and \mathbb{E} is the expectation operator. Notation $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ means \mathbf{x} is drawn from a multivariate Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$.

II. PROBLEM STATEMENT AND PRELIMINARIES

Consider a power system with N buses comprising set \mathcal{N} . The dynamic behavior of the power system is captured by a set of nonlinear differential-algebraic equations. Under small-signal analysis and certain simplifications, power system dynamics can be approximated by a second-order multi-input multi-output (MIMO) linear time-invariant (LTI) system described by the *swing equation* [24, Ch. 3]

$$\mathbf{M}\dot{\boldsymbol{\omega}} + \mathbf{D}\boldsymbol{\omega} + \mathbf{L}\boldsymbol{\theta} = \mathbf{p} \quad (1)$$

where vectors $\boldsymbol{\theta} := [\theta_1 \dots \theta_N]^\top$, $\boldsymbol{\omega} = \dot{\boldsymbol{\theta}}$, and $\mathbf{p} := [p_1 \dots p_N]^\top$ collect respectively the deviations of nodal voltage angles, voltage frequencies, and active power injections from their nominal steady-state values. We henceforth drop the term *deviations* for brevity. Vector $\dot{\boldsymbol{\omega}}$ collects the related ROCOFs. Matrices $\mathbf{M} := \text{dg}(\{M_n\}_{n=1}^N)$ and $\mathbf{D} := \text{dg}(\{D_n\}_{n=1}^N)$ carry the per-generator inertia and damping coefficients. Matrix \mathbf{L} is the (reduced) negative Jacobian matrix of active power injections with respect to voltage angles evaluated at the current operating point. Within standard approximations, matrix \mathbf{L} is assumed symmetric and positive semi-definite (psd); see [25]. Note that the dynamic model in (1) involves only buses hosting generators as load buses have been eliminated via Kron reduction [26].

The task of inferring oscillations dealt with in this work can be abstractly posed as: Knowing model (1) and given measurements of frequencies and/or ROCOFs at different

buses and times, estimate frequencies and ROCOFs at non-metered buses and times. The proposed learning framework can accommodate different power systems monitoring tasks. For example, one can infer oscillations at non-metered buses; whereas for metered buses, one can impute missing frequency data, predict future values, or provide more reliable estimates of ROCOFs from frequency measurements.

The proposed framework is built upon a physics-aware statistical model for power system oscillations using Gaussian processes (GPs). Gaussian processes have previously been adopted for learning in single-input single-output (SISO) LTI systems [27], or differential equations evolving across time and/or continuous spatial dimensions [28]. Leveraging the particular properties of power systems oscillations, we put forth a GP-based Bayesian model for learning under the MIMO LTI system setup of (1). Before presenting our model, let us briefly review GPs.

A GP is a random process with the additional property that any collection of a finite number of its samples forms a Gaussian random vector [29, Ch.1]. Take for instance a time series $z(t)$ and two sets of time indices \mathcal{T}_1 and \mathcal{T}_2 . If $z(t)$ is a GP, the two vectors \mathbf{z}_1 and \mathbf{z}_2 collecting the samples over \mathcal{T}_1 and \mathcal{T}_2 are jointly Gaussian or

$$\mathbf{z} = \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{21}^\top \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix} \right). \quad (2)$$

Now suppose one may wish to estimate \mathbf{z}_2 given an implementation of \mathbf{z}_1 . From (2), the conditional pdf of \mathbf{z}_2 given \mathbf{z}_1 is Gaussian with mean and covariance [30, Ch. 6.4]

$$\mathbb{E}[\mathbf{z}_2 | \mathbf{z}_1] = \boldsymbol{\mu}_2 + \boldsymbol{\Sigma}_{21} \boldsymbol{\Sigma}_{11}^{-1} (\mathbf{z}_1 - \boldsymbol{\mu}_1) \quad (3a)$$

$$\text{Cov}[\mathbf{z}_2 | \mathbf{z}_1] = \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21} \boldsymbol{\Sigma}_{11}^{-1} \boldsymbol{\Sigma}_{21}^\top. \quad (3b)$$

Equation (3a) provides the minimum mean square error (MMSE) estimate of \mathbf{z}_2 given \mathbf{z}_1 , while the covariance in (3b) quantifies its uncertainty. Judiciously selecting the parameters for the Gaussian distribution of \mathbf{z} is obviously important. The mean vector is usually set to zero, while the covariance matrix takes the parametric form of a *kernel function* as $\mathbb{E}[z(t)z(t')] = k(t, t')$; see [30]. The Gaussian kernel $k(t, t') = e^{-\beta(t-t')^2}$ for $\beta > 0$ is a popular choice for learning processes $z(t)$ that are known to be smooth.

Granted a kernel function, applying GP learning to power system oscillations is straight-forward: Stack all frequency/ROCOF measurements in \mathbf{z}_1 , gather the unknown quantities in \mathbf{z}_2 , and apply (3) to obtain point estimates and confidence metrics. These steps are depicted in Figure 1. The key challenge here is in selecting meaningful covariances or equivalently kernel functions. In doing so, two observations are in order: First, we seek spatiotemporal kernels as the processes at hand evolve both across time and the spatial dimension of finitely many buses. Second, one should exploit the fact that the involved processes are related through time-differentiation and the swing dynamics. We pursue these two objectives by considering an approximate model for ambient dynamics as detailed next.

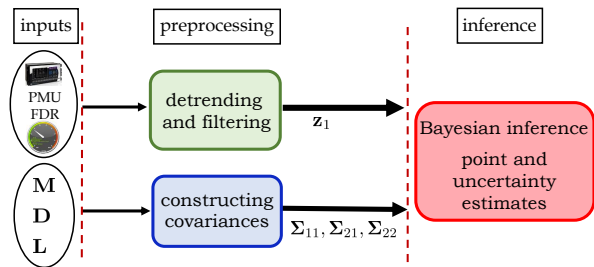


Fig. 1. Workflow for inferring frequencies/ROCOFs.

III. DESIGNING COVARIANCES UNDER HOMOGENEOUS AMBIENT DYNAMICS

To design the covariances required in modeling frequency oscillations as GPs, we resort to the particular setting of swing dynamics driven by ambient disturbances and over buses with a homogeneous inertia-damping model. We commence with the latter by postulating the next assumption.

Assumption 1. *The inertia and damping coefficients are homogeneous across buses, i.e., $\mathbf{D} = \gamma\mathbf{M}$ for a given $\gamma > 0$.*

This approximation is reasonable since such coefficients scale roughly proportionally to the power rating of each generator [31]. Assumption 1 has been frequently adopted for studying power system dynamics [32], [33]. From Assumption 1, the MIMO LTI system of (1) decomposes into a system of N separable SISO LTI eigen-systems. The novel idea here is to first model each SISO eigen-system using GPs, and then linearly combine the eigen-states to arrive at a statistical model for the original MIMO system of (1). Starting with the decomposition, we paraphrase the next claim from [25] and [33] for completeness.

Proposition 1. *Consider matrix $\mathbf{L}_M := \mathbf{M}^{-1/2}\mathbf{L}\mathbf{M}^{-1/2}$ and its eigenvalue decomposition $\mathbf{L}_M = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^\top$. Under Assumption 1, the swing dynamics of (1) are equivalent to*

$$\ddot{\mathbf{y}} + \gamma\dot{\mathbf{y}} + \mathbf{\Lambda}\mathbf{y} = \mathbf{x} \quad (4)$$

under the state and input transformations

$$\mathbf{y} = \mathbf{V}^\top \mathbf{M}^{1/2} \boldsymbol{\theta} \quad \text{and} \quad \mathbf{x} = \mathbf{V}^\top \mathbf{M}^{-1/2} \mathbf{p}. \quad (5)$$

Proof. Matrix \mathbf{L}_M is symmetric psd since \mathbf{L} is symmetric psd. Therefore, its eigenvectors are orthonormal or $\mathbf{V}^\top \mathbf{V} = \mathbf{V}\mathbf{V}^\top = \mathbf{I}_N$. Moreover, its eigenvalues $\{\lambda_i\}_{i=1}^N$, placed in the main diagonal of $\mathbf{\Lambda}$, satisfy $0 = \lambda_1 < \lambda_2 \leq \dots \leq \lambda_N$. This is because \mathbf{L} is known to be a weighted Laplacian of the power grid graph [26]. If the grid is connected, then \mathbf{L} has a single zero eigenvalue and the remaining eigenvalues are positive [32]. The claim carries over to \mathbf{L}_M due to the congruence transformation $\mathbf{L}_M = \mathbf{M}^{-1/2}\mathbf{L}\mathbf{M}^{-1/2}$. If $\mathbf{D} = \gamma\mathbf{M}$, the equivalence of (1) with (4) follows by substituting $\boldsymbol{\theta} = \mathbf{M}^{-1/2}\mathbf{V}\mathbf{y}$ and $\mathbf{p} = \mathbf{M}^{1/2}\mathbf{V}\mathbf{x}$ provided by (5) in (1), and premultiplying by $\mathbf{V}^\top \mathbf{M}^{-1/2}$. \square

Because $\mathbf{\Lambda}$ is diagonal, the MIMO LTI system of (4) is decomposed into the N SISO eigen-systems

$$\ddot{y}_i + \gamma\dot{y}_i + \lambda_i y_i = x_i, \quad i = 1, \dots, N \quad (6)$$

where $\lambda_i \geq 0$ is the i -th eigenvalue of \mathbf{L}_M . We term \dot{y}_i an eigen-frequency or eigen-state and x_i an eigen-input. As discussed in [25], [34], and [33], the transfer function of eigen-system i with respect to state $\dot{y}_i(t)$ is

$$H_i(s) = \frac{s}{s^2 + \gamma s + \lambda_i}. \quad (7)$$

The first eigen-system corresponding to $\lambda_1 = 0$ is first-order, and the remaining ones are second order. It is not easy to verify they are all stable. Heed that if rather than \dot{y}_i , we select y_i as the state, the first eigen-system becomes marginally unstable. This is an additional reason we chose (eigen)-frequencies as the system states.

To apply a GP model on eigen-states, we need to describe the spatiotemporal covariance $\mathbb{E}[\dot{y}_i(t+\tau)\dot{y}_j(t)]$ for all i and $j \in \mathcal{N}$, and $t+\tau$ and $t \in \mathcal{T}$. To do so, we focus on ambient oscillations occurring when the swing equation is continuously driven by small-magnitude random variations in generation and demand at all buses. Ambient oscillations occur under normal conditions and thus correspond to the majority of observed frequency data. They are typically attributed a white noise profile across time and buses with variances scaling proportionally to generator ratings as described next; see [25] and references therein.

Assumption 2. *Ambient disturbances are zero-mean random processes with covariance $\mathbb{E}[\mathbf{p}(t+\tau)\mathbf{p}^\top(t)] = \epsilon\mathbf{M}\delta(\tau)$, where $\delta(\tau)$ is the Dirac delta function and $\epsilon > 0$ is a given constant.*

Under this statistical model for ambient disturbances, we next compute the spatiotemporal covariances of eigen-states by adopting a result from [25]. It is worth noticing that Assumption 2 can be waived as demonstrated in [35], which generalizes the current work. Excluding $\lambda_1 = 0$, we assumed $4\lambda_i > \gamma$ for all $i = 2, \dots, N$, although over-damped eigen-systems can be handled similarly.

Proposition 2. *Under Assumption 2, the covariance of eigen-states is*

$$\mathbb{E}[\dot{\mathbf{y}}(t+\tau)\dot{\mathbf{y}}^\top(t)] = \text{dg}(\{k_i(\tau)\}) \quad (8)$$

where $k_1(\tau) = \frac{1}{2\gamma}e^{-\gamma|\tau|}$ and for $i = 2, \dots, N$

$$k_i(\tau) = B_i e^{-\gamma|\tau|/2} \cos(2\pi f_i |\tau| + \psi_i) \quad (9)$$

with $f_i = \sqrt{|4\lambda_i - \gamma^2|}/(4\pi)$; $B_i := \frac{\epsilon}{2} \sqrt{\frac{1}{(4\pi f_i)^2} + \frac{1}{\gamma^2}}$; and $\psi_i = \arctan\left(\frac{\gamma}{4\pi f_i}\right)$.

Proof. Under Assumption 2, the covariance of the eigen-inputs can be easily derived as

$$\begin{aligned} \mathbb{E}[\mathbf{x}(t+\tau)\mathbf{x}^\top(t)] &= \mathbf{V}^\top \mathbf{M}^{-1/2} \mathbb{E}[\mathbf{p}(t+\tau)\mathbf{p}^\top(t)] \mathbf{M}^{-1/2} \mathbf{V} \\ &= \epsilon \mathbf{V}^\top \mathbf{M}^{-1/2} \mathbf{M} \mathbf{M}^{-1/2} \mathbf{V} \delta(\tau) \\ &= \epsilon \mathbf{I} \delta(\tau). \end{aligned} \quad (10)$$

Because the eigen-inputs are uncorrelated and the eigen-systems are decoupled, the eigen-states are uncorrelated with each other. Moreover, since each eigen-system i is driven

by zero-mean white noise, its output $y_i(t)$ is wide-sense stationary with mean zero and covariance $k_i(\tau) = \mathbb{E}[\dot{y}_i(t + \tau)\dot{y}_i(t)]$ that can be computed as the convolution

$$k_i(\tau) = \epsilon h_i(\tau) * h_i(-\tau) * \delta(\tau).$$

Here $h_i(\tau)$ is the impulse response of eigen-system i computed as the inverse Laplace transform of $H_i(s)$ in (7)

$$h_i(t) = A_i e^{-\gamma t/2} \cos(2\pi f_i t + \psi_i) u(t)$$

where $A_i := \sqrt{1 + \frac{\gamma^2}{|4\lambda_i - \gamma^2|}}$ and $u(t)$ being the unit step function. Interestingly, the convolution $\epsilon h_i(\tau) * h_i(-\tau)$ yields (9) as shown in [25], [34]. \square

Having captured the first- and second-order statistics of eigen-states, we can now transition to the original space of power system frequencies via the transformation $\boldsymbol{\omega} = \mathbf{M}^{-1/2} \mathbf{V} \dot{\mathbf{y}}$. Frequency oscillations are apparently zero-mean and their covariance can be easily computed from (8) as

$$\mathbb{E}[\boldsymbol{\omega}(t + \tau) \boldsymbol{\omega}^\top(t)] = \mathbf{M}^{-1/2} \mathbf{V} \text{dg}(\{k_i(\tau)\}) \mathbf{V}^\top \mathbf{M}^{-1/2}. \quad (11)$$

This covariance provides an explicit form for the kernel function $k((n, t_1), (m, t_2)) = \mathbb{E}[\omega_n(t_1) \omega_m(t_2)]$, which can be evaluated as the (m, n) -th entry of the matrix on the right-hand side (RHS) of (11) upon setting $\tau = t_1 - t_2$. Note that the kernel function k_i exhibits shift-invariance in the sense that $k_i(t_1, t_2) = k_i(t_1 - t_2)$. This is why we have oftentimes abused notation and denoted k_i as taking a single-argument. The covariance of (11) applies to noiseless readings of system frequencies. In practice, the collected frequency data are corrupted by zero-mean Gaussian measurement noise with variance σ_n^2 per bus. Then, the RHS of (11) should be appended a $\text{dg}(\{\sigma_n^2 \mathbf{1}_{n=1}^N\}) \delta(\tau)$ term.

It should be emphasized that if ambient inputs are Gaussian distributed, eigen-states and hence original states (frequencies) are Gaussian as well. Then GPs become the natural choice. A GP model however can still be used when swing dynamics are driven by non-Gaussian ambient disturbances or non-ambient disturbances (e.g., deterministic inputs of impulse- or step-type caused by generator or line trips).¹ We elaborate on this in the next section, which reviews how GPs extend to the time derivatives of random processes.

IV. COVARIANCES UPON TIME DIFFERENTIATION

So far, we have presented a statistical model assuming that both vector \mathbf{z}_1 of observations and vector \mathbf{z}_2 of wanted quantities in (2)–(3) involve only frequencies. An appealing property of GPs is that time differentiation of a GP yields a GP. Therefore, our GP model for frequencies can be easily modified to yield a GP model for ROCOFs. This is important as it enables including available ROCOF measurements in \mathbf{z}_1 and/or estimating ROCOFs by having them in \mathbf{z}_2 .

To explain this key feature of GPs, we adopt the alternative interpretation of GPs provided in [27]. One can postulate

¹Our numerical tests considered ambient and non-ambient disturbances, both under a non-homogeneous inertia/damping swing model.

that each eigenstate $\dot{y}_i(t)$ can be expressed as a linear combination of K fixed basis functions $\{\phi_i^k(t)\}_{k=1}^K$ as [30]

$$\dot{y}_i(t) = \sum_{k=1}^K w_i^k \phi_i^k(t) = \mathbf{w}_i^\top \boldsymbol{\phi}_i(t), \quad i = 1, \dots, N \quad (12)$$

where $\boldsymbol{\phi}_i(t) := [\phi_i^1(t) \dots \phi_i^K(t)]^\top$ is a mapping from t to \mathbb{R}^K and \mathbf{w}_i is a vector of unknown coefficients. Following a Bayesian approach, we further postulate that \mathbf{w}_i is random with prior distribution $\mathbf{w}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_K)$. We are interested in the joint probability density function (pdf) for samples of $\dot{y}_i(t)$ collected at times $\mathcal{T} := \{t_1, \dots, t_T\}$, and stacked in $\dot{\mathbf{y}}_i := [\dot{y}_i(t_1) \dots \dot{y}_i(t_T)]^\top$. From (12), it follows that $\dot{\mathbf{y}}_i = \boldsymbol{\Phi}_i \mathbf{w}_i$ for the $T \times K$ matrix $\boldsymbol{\Phi}_i := [\boldsymbol{\phi}_i(t_1) \dots \boldsymbol{\phi}_i(t_T)]^\top$. Because \mathbf{w}_i is Gaussian, vector $\dot{\mathbf{y}}_i$ becomes Gaussian too. It is not hard to verify that $\dot{\mathbf{y}}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{K}_i)$ where $\mathbf{K}_i = \boldsymbol{\Phi}_i \boldsymbol{\Phi}_i^\top \succeq 0$. Since the latter holds for any collection of sampling instances \mathcal{T} , signal $\dot{y}_i(t)$ is a GP. If we assume $\mathbb{E}[\mathbf{w}_i \mathbf{w}_j^\top] = \mathbf{0}$ for $j \neq i$ and that matrix \mathbf{K}_i is Toeplitz with entries dictated by (9), then (12) agrees with the modeling of Section III. Observe also that its (ℓ, k) -th is

$$[\mathbf{K}_i]_{\ell, k} = k_i(t_\ell - t_k) = \phi_i^\top(t_\ell) \phi_i(t_k) \quad (13)$$

with the kernel function $k_i(t_\ell - t_k)$ defined in (9).

If $y_i(t)$ is provided by (12), its time derivative is $\ddot{y}_i(t) = \dot{\mathbf{w}}_i^\top \dot{\boldsymbol{\phi}}_i(t)$. And because \mathbf{w}_i is Gaussian, the random process $\ddot{y}_i(t)$ is also a GP. If vector $\ddot{\mathbf{y}}_i$ collects the samples of $\ddot{y}_i(t)$ over \mathcal{T} , its covariance matrix is

$$\mathbb{E}[\ddot{\mathbf{y}}_i \ddot{\mathbf{y}}_i^\top] = \dot{\boldsymbol{\Phi}}_i \mathbb{E}[\mathbf{w}_i \mathbf{w}_i^\top] \dot{\boldsymbol{\Phi}}_i^\top = \dot{\boldsymbol{\Phi}}_i \dot{\boldsymbol{\Phi}}_i^\top = \ddot{\mathbf{K}}_i \quad (14)$$

where $\dot{\boldsymbol{\Phi}}_i$ is the element-wise time derivative of $\boldsymbol{\Phi}_i$. Interestingly, matrix $\ddot{\mathbf{K}}_i := \dot{\boldsymbol{\Phi}}_i \dot{\boldsymbol{\Phi}}_i^\top$ relates to \mathbf{K}_i as

$$\begin{aligned} [\ddot{\mathbf{K}}_i]_{\ell, k} &= \dot{\phi}_i^\top(t_\ell) \dot{\phi}_i(t_k) \\ &= \frac{\partial^2}{\partial t_\ell \partial t_k} \phi_i^\top(t_\ell) \phi_i(t_k) \\ &= \frac{\partial^2 k_i(t_\ell, t_k)}{\partial t_\ell \partial t_k}. \end{aligned}$$

The first equality stems from the definition of $\ddot{\mathbf{K}}_i$, the second from the definition of time differentiation, and the third one because $k_i(t_\ell, t_k) = \phi_i^\top(t_\ell) \phi_i(t_k)$. The cross-covariances $\mathbb{E}[\ddot{\mathbf{y}}_i \dot{\mathbf{y}}_i^\top]$ can be computed likewise as $\mathbb{E}[\ddot{\mathbf{y}}_i \dot{\mathbf{y}}_i^\top] = \dot{\boldsymbol{\Phi}}_i \boldsymbol{\Phi}_i^\top = \ddot{\mathbf{K}}_i$, whose (ℓ, k) -th entry would be $\frac{\partial k_i(t_\ell, t_k)}{\partial t_\ell}$.

In a nutshell, eigen-ROCOFs can be modeled as zero-mean GPs with covariances obtained by differentiating the related kernel functions. Based on these GP models, it is straightforward to arrive at a GP model for the original ROCOFs $\dot{\boldsymbol{\omega}} = \mathbf{M}^{-1/2} \mathbf{V} \ddot{\mathbf{y}}$ similar to (11). Having modeled frequencies and ROCOFs as GPs, we can utilize the Bayesian inference of (3) to estimate frequencies and ROCOFs at non-metered buses and times.

In some applications, one might be interested in a particular frequency band of power system frequencies. Recall that the term power system frequencies $\{\omega_n(t)\}_{n=1}^N$ refers in fact to the time-domain signal of interest here. For example, inter-area oscillations typically correspond to the frequency

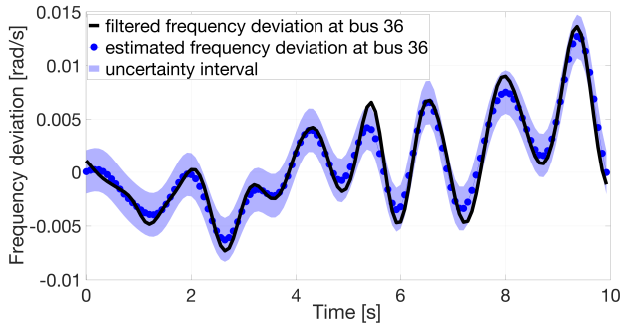


Fig. 2. Frequency estimate at bus 36 under ambient disturbances.

content of $\omega_n(t)$ in the range of 0.1-0.8Hz [24, Ch. 12]. As indicated by (7), each eigen-system i exhibits either a low-pass or band-pass behavior around its resonance frequency f_i . If the focus lies on a particular band, the eigen-systems with out-of-band resonant frequencies can be ignored and the observed signals can be passed through bandpass filters. In other words, filtered frequency measurements can be approximated as $\tilde{\omega} = \mathbf{M}^{-1/2} \mathbf{V}_D \mathbf{y}_D + \mathbf{n}$, where \mathbf{n} is measurement and modeling noise; \mathbf{y}_D is the subvector of \mathbf{y} collecting the eigen-states with in-band resonant frequencies; and matrix \mathbf{V}_D samples the associated eigenvectors (columns of \mathbf{V}). To reduce the computational complexity of the proposed learning scheme, matrix \mathbf{V} should be replaced by \mathbf{V}_D in (11). Another advantage of such filtering is that having to infer fewer eigen-states means that learning can be conducted using fewer metered buses.

V. NUMERICAL TESTS

The novel GP-based learning framework was evaluated using synthesized data on the IEEE 300-bus power system benchmark, which was Kron-reduced to a 69-bus system upon eliminating load buses [26]. Generator parameters were obtained from [36], so that Assumption 1 on homogeneity was actually not met. To build the postulated covariances, we need to know $(\mathbf{M}, \gamma, \mathbf{L})$. We chose $\gamma = \frac{\mathbf{1}^\top \mathbf{D} \mathbf{1}}{\mathbf{1}^\top \mathbf{M} \mathbf{1}}$ and \mathbf{L} was set to the B-matrix obtained from MATPOWER's function `makeJac`. Although covariances were derived presuming homogeneous inertia-damping coefficients, power systems dynamics data were generated using the non-homogeneous (\mathbf{M}, \mathbf{D}) parameters from the benchmark. We converted the swing equation to its state-space representation using MATLAB's `ss` function, and solved given input $\mathbf{p}(t)$ using the `ode45` command. Dynamics were simulated at a time resolution of 1 ms, but sampled and processed at the reporting rate of 15 samples per second according to the IEEE C37.118.1-2011 standard [37]. Synthesized frequency readings were corrupted by zero-mean additive white Gaussian noise (AWGN) with standard deviation of 0.01 rad/s to comply with the maximum frequency errors allowed by [37]. Benchmark data and the developed code can be found at <https://github.com/manajalali/GP4GridDynamics>.

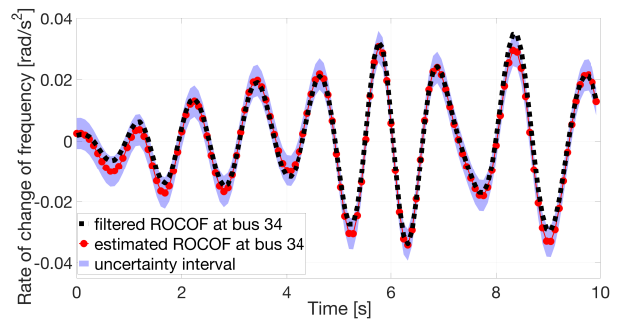


Fig. 3. ROCOF estimate at bus 34 under ambient disturbances.

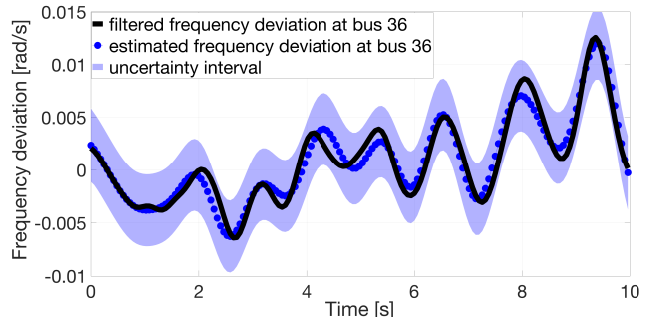


Fig. 4. Frequency estimate for bus 36 under ambient disturbances with metered buses selected at random.

We first evaluated the proposed GP framework under ambient disturbances with covariance $\mathbb{E}[\mathbf{p}(t)\mathbf{p}^\top(t)] = 0.01 \cdot \mathbf{M}$. Focusing on inter-area oscillations, we filtered frequency measurements using MATLAB's `filtfilt` function. The filter was designed to maintain the $K = 5$ eigen-states with resonant frequencies f_i 's falling within 0.1-0.8Hz. We assumed frequencies were being collected at $M = 20$ out of the 69 buses. Metered buses were selected based on the placement scheme proposed in [38], which aims at maximizing the minimum eigenvalue of the posterior covariance matrix when estimating \mathbf{y}_D by row-sampling the linear-Gaussian model $\tilde{\omega} = \mathbf{M}^{-1/2} \mathbf{V}_D \mathbf{y}_D + \mathbf{n}$. The non-metered frequency deviations and ROCOFs were estimated using (3). The $\pm\sigma$ uncertainty interval was computed by taking the square root of the diagonal entries of (3b). Figure 2 and 3 depict the estimation results respectively for the frequency and ROCOF at a non-metered bus. The results confirm that the GP framework can successfully infer oscillations under ambient disturbances even when Assumption 1 on inertia/damping homogeneity is waived. To study the effect of meter placement, we repeated the previous test by selecting $M = 20$ metered buses at random. Figure 4 shows the frequency estimation outcome. The results demonstrate that the sensor placement does not significantly affect the point estimation of the GP paradigm. However, the uncertainty interval was increased.

We also analyzed the capability of the GP paradigm to predict system states. Unlike the previous tests, we did not focus on inter-area oscillations but used the raw frequency

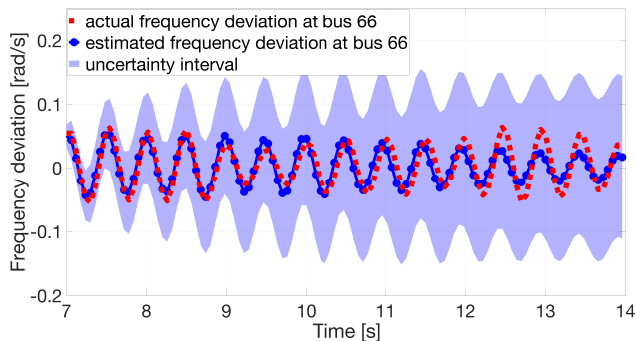


Fig. 5. Frequency prediction for bus 66 under ambient disturbance with randomly sampled measurements.

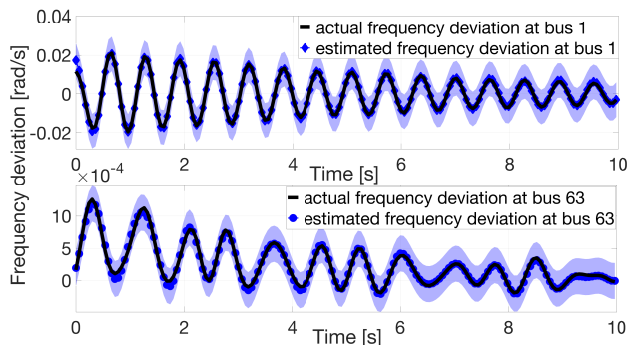


Fig. 6. Estimated frequencies at buses 1 and 63 of the 69-bus system under a non-ambient disturbance (a fault near bus 1).

measurements from 60 randomly selected buses across the grid. We used the frequencies observed from these buses during 7 seconds to predict the frequencies at non-metered buses for the next 7 seconds. Figure 5 demonstrates how uncertainty increases as we advance further in the future.

In this final test, we evaluated the proposed scheme under non-ambient dynamics, thus waiving both assumptions under which covariances were derived. To generate a non-ambient setup, we simulated a generator trip on bus 1 using an impulse function. Further, we assumed that bus 1 at which the generator trip has occurred is not metered. Figure 6 depicts the performance of the GP paradigm for estimating frequencies at buses 1 and 63. The results confirm that although the proposed GP framework was formulated for ambient disturbances, it applies also to other settings.

VI. CONCLUSIONS AND FUTURE WORK

A novel method for learning power system frequencies and ROCOFs using GPs has been put forth. The required spatiotemporal covariances have been derived upon leveraging the statistical properties of ambient dynamics under an approximate rendition of the swing equation. The proposed framework can accommodate different sampling rates, missing data, extrapolate dynamics at non-metered buses, and estimates ROCOFs in a parametric fashion without resorting to finite differences. The point estimates are endowed with a complete pdf characterization, which can be instrumental in detecting and identifying bad, erroneous, and/or attacked

data. Numerical tests on a non-homogeneous power system benchmark corroborated the efficacy of the proposed scheme in extrapolating and predicting frequencies and ROCOFs under ambient and non-ambient disturbances. This work sets the foundations for interesting research directions, including real-time implementations, parameter estimation, and modal analysis. Utilizing the GP estimation paradigm to infer voltage angles and power deviations would be useful for fault identification. Finally, exploring more detailed generator models and using other measurements, such as field voltages, could improve observability and enhance frequency estimation.

REFERENCES

- [1] Y. Zhang, P. Markham, T. Xia, L. Chen, Y. Ye, Z. Wu, Z. Yuan, L. Wang, J. Bank, J. Burgett, R. W. Conners, and Y. Liu, "Wide-area frequency monitoring network (FNET) architecture and applications," *IEEE Trans. Smart Grid*, vol. 1, no. 2, pp. 159–167, Sep. 2010.
- [2] Z. Zhong, C. Xu, B. J. Billian, L. Zhang, S. S. Tsai, R. W. Conners, V. A. Centeno, A. G. Phadke, and Y. Liu, "Power system frequency monitoring network (FNET) implementation," *IEEE Trans. Power Syst.*, vol. 20, no. 4, pp. 1914–1921, Nov. 2005.
- [3] W. Yu, W. Yao, and Y. Liu, "Definition of system angle reference for distribution level synchronized angle measurement applications," *IEEE Trans. Power Syst.*, vol. 34, no. 1, pp. 818–820, Oct. 2019.
- [4] C. Zhao and S. H. Low, "Optimal decentralized primary frequency control in power networks," in *Proc. IEEE Conf. on Decision and Control*, Los Angeles, CA, Dec. 2014, pp. 2467–2473.
- [5] P. Gao, M. Wang, S. G. Ghiocel, and J. H. Chow, "Modeless reconstruction of missing synchrophasor measurements," in *Proc. IEEE Power & Energy Society General Meeting*, National Harbor, MD, Jul. 2014, pp. 1–5.
- [6] S. Zhang and M. Wang, "Correction of corrupted columns through fast robust Hankel matrix completion," *IEEE Trans. Signal Process.*, vol. 67, no. 10, pp. 2580–2594, May 2019.
- [7] S. Zhang, Y. Hao, M. Wang, and J. H. Chow, "Multichannel Hankel matrix completion through nonconvex optimization," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 4, pp. 617–632, Aug. 2018.
- [8] D. Osipov and J. H. Chow, "PMU missing data recovery using tensor decomposition," *IEEE Trans. Power Syst.*, vol. 35, no. 6, pp. 4554–4563, May 2020.
- [9] J. Zhao, A. Gomez-Exposito, M. Netto, L. Mili, A. Abur, V. Terzija, I. Kamwa, B. Pal, A. K. Singh, J. Qi, Z. Huang, and A. P. S. Meliopoulos, "Power system dynamic state estimation: Motivations, definitions, methodologies, and future work," *IEEE Trans. Power Syst.*, vol. 34, no. 4, pp. 3188–3198, Jul. 2019.
- [10] Zhenyu Huang, K. Schneider, and J. Nieplocha, "Feasibility studies of applying Kalman filter techniques to power system dynamic state estimation," in *Intl. Power Engineering Conf.*, Singapore, Dec. 2007, pp. 376–382.
- [11] Kuang-Rong Shih and Shyh-Jier Huang, "Application of a robust algorithm for dynamic state estimation of a power system," *IEEE Trans. Power Syst.*, vol. 17, no. 1, pp. 141–147, Aug. 2002.
- [12] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proc. IEEE*, vol. 92, no. 3, pp. 401–422, Nov. 2004.
- [13] S. Wang, W. Gao, and A. P. S. Meliopoulos, "An alternative method for power system dynamic state estimation based on unscented transform," *IEEE Trans. Power Syst.*, vol. 27, no. 2, pp. 942–950, May 2012.
- [14] W. S. Rosenthal, A. M. Tartakovsky, and Z. Huang, "Ensemble Kalman filter for dynamic state estimation of power grids stochastically driven by time-correlated mechanical input power," *IEEE Trans. Power Syst.*, vol. 33, no. 4, pp. 3701–3710, Jul. 2018.
- [15] N. Zhou, D. Meng, and S. Lu, "Estimation of the dynamic states of synchronous machines using an extended particle filter," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4152–4161, Nov. 2013.
- [16] A. K. Singh and B. C. Pal, "Decentralized dynamic state estimation in power systems using unscented transformation," *IEEE Trans. Power Syst.*, vol. 29, no. 2, pp. 794–804, Mar. 2014.
- [17] J. Zhao and L. Mili, "Power system robust decentralized dynamic state estimation based on multiple hypothesis testing," *IEEE Trans. Power Syst.*, vol. 33, no. 4, pp. 4553–4562, Jul. 2018.

- [18] J. Dong, X. Ma, S. M. Djouadi, H. Li, and Y. Liu, "Frequency prediction of power systems in FNET based on state-space approach and uncertain basis functions," *IEEE Trans. Power Syst.*, vol. 29, no. 6, pp. 2602–2612, May 2014.
- [19] G. S. Misyris, A. Venzke, and S. Chatzivasileiadis, "Physics-informed neural networks for power systems," in *Proc. IEEE Power & Energy Society General Meeting*, Montreal, Canada, Aug. 2020, pp. 1–5.
- [20] P. N. Markham and Y. Liu, "Electromechanical speed map development using FNET/GridEye frequency measurements," *IEEE Power & Energy Society General Meeting*, National Harbor, MD, Jul. 2014, pp. 1–5.
- [21] X. Deng, D. Bian, D. Shi, W. Yao, Z. Jiang, and Y. Liu, "Line outage detection and localization via synchrophasor measurement," in *Proc. IEEE Conf. on Innovative Smart Grid Technologies*, Chengdu, China, May 2019, pp. 3373–3378.
- [22] S. N. Nuthalapati, *Power System Grid Operation Using Synchrophasor Technology*. Cham, Switzerland: Springer, 2019.
- [23] T. Huang, N. M. Freris, P. R. Kumar, and L. Xie, "A synchrophasor data-driven method for forced oscillation localization under resonance conditions," *IEEE Trans. Power Syst.*, pp. 1–1, Mar. 2020.
- [24] P. Kundur, *Power System Stability and Control*. New York, NY: McGraw-Hill, 1994.
- [25] P. Huynh, H. Zhu, Q. Chen, and A. E. Elbanna, "Data-driven estimation of frequency response from ambient synchrophasor measurements," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 6590–6599, Nov. 2018.
- [26] T. Ishizaki, A. Chakraborty, and J. Imura, "Graph-theoretic analysis of power systems," *Proc. IEEE*, vol. 106, no. 5, pp. 931–952, May 2018.
- [27] T. Graepel, "Solving noisy linear operator equations by Gaussian processes: Application to ordinary and partial differential equations," in *Intl. Conf. on Machine Learning*, Washington, DC, Aug. 2003, p. 234–241.
- [28] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Inferring solutions of differential equations using noisy multi-fidelity data," *Journal of Computational Physics*, vol. 335, pp. 736–746, Apr. 2017.
- [29] C. Rasmussen and C. Williams, *Gaussian Processes for Machine Learning*. Cambridge, MA: MIT Press, 2006.
- [30] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY: Springer, 2006.
- [31] B. K. Poolla, S. Bolognani, and F. Dorfler, "Optimal placement of virtual inertia in power grids," *IEEE Trans. Autom. Contr.*, vol. 62, no. 12, pp. 6209–6220, Dec. 2017.
- [32] L. Guo, C. Zhao, and S. H. Low, "Graph Laplacian spectrum and primary frequency regulation," in *Proc. IEEE Conf. on Decision and Control*, Miami Beach, FL, Dec. 2018, pp. 158–165.
- [33] F. Paganini and E. Mallada, "Global analysis of synchronization performance for power systems: Bridging the theory-practice gap," *IEEE Trans. Autom. Contr.*, vol. 65, no. 7, pp. 3007–3022, Sep. 2020.
- [34] S. Liu, H. Zhu, and V. Kekatos, "A dynamic response recovery framework using ambient synchrophasor data," *IEEE Trans. Power Syst.*, Jul. 2021, (submitted). [Online]. Available: <https://arxiv.org/abs/2104.05614>
- [35] M. Jalali, V. Kekatos, S. Bhela, H. Zhu, and V. Centeno, "Inferring power system dynamics from synchrophasor data using Gaussian processes," *IEEE Trans. Power Syst.*, May 2021, (submitted). [Online]. Available: <https://arxiv.org/abs/2105.12608>
- [36] P. Demetriou, M. Asprou, J. Quiros-Tortos, and E. Kyriakides, "Dynamic IEEE test systems for transient analysis," *IEEE Systems Journal*, vol. 11, no. 4, pp. 2108–2117, Jul. 2017.
- [37] *IEEE C37.118.1-2011 Std. for Synchrophasor Measurements for Power Systems*, IEEE Std., 2011.
- [38] 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.