Studying the Electromechanical Oscillations using Ambient Synchrophasor Data

Phuc Huynh[†], Qianli Chen[*], Ahmed Elbanna[*], Hao Zhu[†]
[†] Dept. of ECE, Univ. of Illinois, Urbana, IL, USA
[*] Dept. of CEE, Univ. of Illinois, Urbana, IL, USA

***Abstract*—We study the propagation of electromechanical oscillation waves across an interconnected power system using ambient synchrophasor measurements. Leveraging recent advances in seismology and earth sciences, we propose to estimate the system impulse response from one location to another by cross-correlating the corresponding ambient frequency data. We have presented initial analytical results to demonstrate the validity of this approach for undamped stable systems, corroborated by numerical tests on small (and even damped) systems. Additional validations have been performed using two dynamic test cases representing realistic power systems. By comparing the cross-correlation results to the model-based solutions, we have identified several cases of successful recovery, and also a few cases of visible mismatches that need further investigations.**

I. INTRODUCTION

Electromechanical oscillations in interconnected power system are complex in nature and can cause severe problems in power system operations. The root causes of these oscillations can be attributed to fast excitation systems and to weak tie lines between different areas [1, Ch. 1]. As long as well damped, oscillations are generally acceptable in the interconnected grids. However, existence of poorly damped oscillations are not uncommon. Even worse, for some poorly damped oscillation modes, even a small disturbance event would trigger growing level of oscillations and system instability issues. This was exactly the cause of the August 1996 blackout at the US/Canada Western Interconnection, with thousands of megawatts of load and generation outaged [2].

Studying the oscillation phenomena has traditionally relied on analyzing the corresponding system dynamic system models. Building these dynamic models, typically nonlinear, require the complete and accurate information regarding synchronous generators, fast/slow exciters, and network information; see e.g., [3]–[5]. By linearizing the dynamic models, one can perform the *small-signal analysis* to study the oscillation mode frequency/damping and identify coherent generator groups. However, this model-based framework would experience issues such as outdated model information, frequent topology changes, or high computational complexity. All of these issues could challenge the performance fidelity of the model-based stability analysis results.

Over the past decades, high-resolution synchronized measurements provided by phasor measurement units (PMUs) have been advocated for a measurement-based framework for power system analysis, including the oscillation related dynamics analysis. Synchronized frequency/angle and voltage data at high sampling rates could reveal important characteristics of the power system dynamics and require little information of the underlying models. A classical problem is to estimate the oscillation modes and mode shapes using either ambient data, ring-down signals, or probing responses; see e.g., [6]–[9]. More recently, it has been shown that the statistical information in ambient synchrophasor data can be used to estimate the system Jacobian matrix [10] or analyze the sustained oscillations [11].

The present work aims to develop a data-driven framework to estimate the generator's frequency response due to any input disturbance to better understand the oscillation behaviors of an interconnected system. This impulse response can be obtained by the small-signal analysis of the linearized dynamic model. Albeit focused on the linearize system regime, it can provide valuable information regarding inherent dynamic characteristics such as system stability and generator coherency. In addition, with the time-scale separation in power system dynamics [4][Ch. 1], the initial behavior of generator response can well approximate the very fast system response due to a disturbance.

The frequency deviation at different locations across an interconnected power system after occurrence of a big disturbance can be closely related to the propagation of electromechanical (EM) waves over continuum medium; see e.g., [12], [13]. Under this framework, the angular variation across an interconnected system can be modeled as a two-dimensional planar wave. Although the perspective of EM wave propagation fails to capture the full spectrum of actual grid behaviors, it offers the macroscopic insights into disturbance response over a large-scale interconnection. For example, the propagation speed has been approximately computed to be at $\sim$ 1000 km/sec [12]. Interestingly, the EM wave model for grid disturbance propagation has inspired the adoption of data processing techniques that have been developed for other types of waves in natural systems. Specifically, earth scientists and seismologists have discovered that some of the wave propagation properties can be extracted by cross-correlating high-resolution ambient measurements at any two locations. The relation between the cross-correlation of time-history responses of two record points and the impulse response, or the *Green's function* has been established for uncorrelated diffusive field [14]. The condition of this relation was later

extended in the view of scattered waves [15], energy flux [16], and reciprocity theorem [17]. More detailed discussions can be also found in [18]–[20] regarding why seismic ambient noise may give good estimate of Green's function even when the conditions are not strictly satisfied. This cross-correlation method has been successfully validated by seismic coda waves [21], long time Rayleigh waves [22], and crustal body waves [23]. It has been proposed to divide the data processing scheme into multiple phases as summarized in [24]. Most recently, this method has also been used to predict the strong ground motion under "virtual earthquakes" [25].

We propose to leverage these advancements of wave studies in seismology to estimate the grid frequency response by cross-correlating ambient PMU measurements. Such an approach has been pioneered in [26], however, the estimated Green's function therein corresponds to a disturbance in terms of step function change in frequency. Moreover, although [26] has presented preliminary results on real data validation, it has not provided any insights to justify the adoption of the cross-correlation approach based on power system dynamic modeling. Compared to [26], we present some initial analytical results to better understand the applicability and potential limitations of the cross-correlation approach. Specifically, we have shown the equivalence relation between the impulse response and the cross-correlation of frequency data for undamped stable systems under random white noise inputs. Moreover, we have tested the cross-correlation approach on two power system dynamic cases: the Western System Coordinating Council (WSCC) 9-bus system [27] and the Northeast Power Coordinating Council (NPCC) 48-machine system [28]. With the benchmark impulse response obtained by the model-based method, we have identified several success cases where the cross-correlation can accurately recover the disturbance response. Our numerical simulation results also point out the potential to generalize the validation of the proposed cross-correlation method to certain types of damped systems.

II. Power System Dynamic Modeling

Dynamics of a power system can be modeled by a set of differential and algebraic equations (DAEs). To this end, consider the following set of DAEs that can be used to describe a general non-linear dynamical system

$$\begin{cases} \dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}, \mathbf{y}, \mathbf{u}) \\ \mathbf{0} &= \mathbf{g}(\mathbf{x}, \mathbf{y}) \end{cases} \tag{1}$$

with dynamics of the state vector $\mathbf{x}$ depending on the algebraic variables in $\mathbf{y}$ and the input variables in $\mathbf{u}$. In the context of power systems, vector $\mathbf{x}$ contains the internal angle and speed of generators and status of their connected equipments such as exciters and governors. Also, vector $\mathbf{y}$ constitutes of the voltage magnitudes and angles at all buses to establish the power flow equations, while the input $\mathbf{u}$ typically includes the mechanical power injected to all generators. The set of DAEs in (1) can be linearized around an operating point, i.e., a solution to the steady-state power flow equations, in order to perform the small-signal stability analysis. Furthermore, one can eliminate the vector $\mathbf{y}$ from the resultant linearized model to obtain a set of ordinary differential equations (ODEs) for the state vector $\mathbf{x}$.

We adopt the classical model for synchronous generators [4, Sec. 6.6] to make the dynamic model in (1) more specific. Under this simplified generator model, the system states only involve the angle and speed of all the synchronous generators. For a system of n generators with a given operating point, one can partition $\mathbf{x}$ into the angle deviation vector $\boldsymbol{\delta} = \left[\delta_1, \delta_2, \cdots, \delta_n\right]^\mathsf{T}$ and the speed deviation one $\boldsymbol{\omega} = \left[\omega_1, \omega_2, \cdots, \omega_n\right]^\mathsf{T}$. Meanwhile, the vector $\mathbf{u} = \left[u_1, u_2, \cdots, u_n\right]^\mathsf{T}$ stands for deviation from the nominal mechanical power input values at all the generators. After linearization and elimination of algebraic variables, the system dynamical model is given by the so-called swing equations [4, Sec. 6.6]:

$$\begin{cases} \dot{\boldsymbol{\delta}} &= \boldsymbol{\omega} \\ \dot{\boldsymbol{\omega}} &= -\mathbf{M}^{-1}\mathbf{K}\boldsymbol{\delta} - \mathbf{M}^{-1}\mathbf{D}\boldsymbol{\omega} + \mathbf{M}^{-1}\mathbf{u} \end{cases} \tag{2}$$

where $\mathbf{M} = \text{diag}\{M_1, M_2, \cdots, M_n\}$ and $\mathbf{D} = \text{diag}\{D_1, D_2, \cdots, D_n\}$ are the positive diagonal matrices containing the generator moment of inertia and damping coefficients, respectively. Also, matrix $\mathbf{K} := \frac{\partial \mathbf{P}^e}{\partial \boldsymbol{\delta}}$ is the partial derivative for the generator electrical power output $\mathbf{P}^e = \left[P_1^e, P_2^e, \cdots, P_n^e\right]^\mathsf{T}$ with respect to $\boldsymbol{\delta}$. The relation of the two is based on the power flow equations. Clearly, one can construct the linearized model (2) with given generator parameters and network information. There also exist several simulation tools that can automate this process such as the Power System Toolbox [28] and Power System Analysis Toolbox (PSAT) [27].

The rest of this paper will develop a model-free approach for estimating the impulse response $T_{k\ell}$ from any input u_k to any speed state ω_ℓ by using only ambient synchrophasor frequency data collected at the two nodes. The ensuing section will present this cross-correlation based approach using the swing equation model in (2).

III. Impulse Response Estimation using Cross-correlation

We propose to estimate the impulse response by cross-correlating the ambient frequency measurements under small perturbation inputs. It can be shown that the impulse response between any two locations is related to the cross-correlation of corresponding ambient generator speed responses. For simplicity, here we establish this relation for an *undamped and stable system*, resulting in an orthogonal transformation to the oscillation modes. Extension to the more realistic damped systems is possible under orthogonal transformation condition, which will be explored in future.

To this end, the system in (2) can be represented as a linear combination of its n modes associated with the $2n$ eigenvalues of the state transition matrix. For undamped system with $\mathbf{D} = \mathbf{0}$, each eigenvalue λ_i can be found by solving the generalized eigenvalue problem $\mathbf{K}\mathbf{c}_i = -\lambda_i^2\mathbf{M}\mathbf{c}_i$ for given

$\mathbf{M}$ and $\mathbf{K}$, with $\mathbf{c}_i$ denoting the corresponding eigenvector. To ensure stability, we can assume that the matrices $\mathbf{M}$ and $\mathbf{K}$ are symmetric and positive definite. This way, we have $\lambda_i^2 \leq 0$ and every eigenvalue is purely imaginary; i.e., $\lambda_i = \pm \mathtt{j}\beta_i$ as characterized by the oscillation frequency β_i. Accordingly, the normalized eigenvectors are also orthogonal to each other, as given by the following conditions [29, Chap. 15]

$$\mathbf{c}_i^{\mathsf{T}} \mathbf{M} \mathbf{c}_j = \begin{cases} 0, & \text{if } i \neq j \\ 1, & \text{if } i = j \end{cases}. \tag{3}$$

The orthogonality of the modes can simplify the analysis of the full system (2) by projecting it to each mode, denoted by z_i. The response of each mode is determined by the second-order dynamics:

$$\ddot{z}_i = \lambda_i^2 z_i + \mathbf{c}_i^{\mathsf{T}} \mathbf{u}. \tag{4}$$

Under zero initial condition at time $t = 0$, the time derivative of z_i's trajectory is given by

$$\dot{z}_i(t) = \int_0^t \cos(\beta_i \tau) \mathbf{c}_i^{\mathsf{T}} \mathbf{u}(t-\tau) d\tau, \quad t > 0. \tag{5}$$

Inverting the projection, the trajectory of ω_ℓ is recovered as

$$\begin{aligned} \omega_\ell(t) &= \sum_{i=1}^n c_{\ell i} \dot{z}_i(t) \\ &= \sum_{i=1}^n c_{\ell i} \int_0^t \cos(\beta_i \tau) \mathbf{c}_i^{\mathsf{T}} \mathbf{u}(t-\tau) d\tau. \end{aligned} \tag{6}$$

where $c_{\ell i}$ is the ℓ-th entry of eigenvector $\mathbf{c}_i$. With a slight abuse of notation, use $\delta(t)$ to denote the Dirac delta function. Under an impulse input $u_k = \delta(t)$ with the other inputs kept as zero, the impulse response at ω_ℓ for any $k, \ell \in \{1, \cdots, n\}$ becomes

$$T_{k\ell}(t) = \sum_{i=1}^n c_{ki} c_{\ell i} \cos(\beta_i t), \quad t \geq 0. \tag{7}$$

To generate the ambient response data, assume there exist small system perturbations due to e.g., random load variations and thus the input $\mathbf{u}$ is modeled by zero-mean white-noise process with covariance matrix equals to $\mathbf{\Sigma}$. Using (6), the normalized cross-correlation between $\omega_k(t)$ and $\omega_\ell(t)$ over a duration of T seconds becomes

$$\begin{aligned} C_{k\ell}(\tau) &= \frac{1}{T} \int_{T_1}^{T_1+T} \omega_k(t) \omega_\ell(t-\tau) dt \\ &= \frac{1}{T} \int_{T_1}^{T_1+T} \Big\{ \sum_{i=1}^n c_{ki} \int_0^t \cos(\beta_i \tau_1) \mathbf{c}_i^{\mathsf{T}} \mathbf{u}(t-\tau_1) d\tau_1 \\ &\quad \sum_{j=1}^n c_{\ell j} \int_0^{t-\tau} \cos(\beta_j \tau_2) \mathbf{c}_j^{\mathsf{T}} \mathbf{u}(t-\tau-\tau_2) d\tau_2 \Big\} dt. \end{aligned} \tag{8}$$

where $T_1 \gg 0$ is a sufficiently large time after zero initialization. Under the white-noise input condition, we have

$$\begin{aligned} &\frac{1}{T} \int_{T_1}^{T_1+T} \{\mathbf{c}_i^{\mathsf{T}} \mathbf{u}(t-\tau_1)\} \{\mathbf{c}_j^{\mathsf{T}} \mathbf{u}(t-\tau-\tau_2)\}^{\mathsf{T}} dt \\ &\qquad \cong \ \mathbf{c}_i^{\mathsf{T}} \mathbf{\Sigma} \mathbf{c}_j \delta(\tau_1 - \tau - \tau_2), \end{aligned} \tag{9}$$

recalling that $\delta(t)$ is the Dirac delta function. To capitalize the orthogonal condition in (3), we assume the variance of input perturbation per node scales proportionally with the inertia of the generator; i.e., $\mathbf{\Sigma} = \mu \mathbf{M}$. This assumption can be justified as higher inertia value would have been allocated to areas at larger disturbance level at the system planning phase. Interestingly, this assumption simplifies the cross-correlation in (9) as

$$\mathbf{c}_i^{\mathsf{T}} \mathbf{\Sigma} \mathbf{c}_j \delta(\tau_1 - \tau - \tau_2) = \mu \mathbf{c}_i^{\mathsf{T}} \mathbf{M} \mathbf{c}_j \delta(\tau_1 - \tau - \tau_2), \tag{10}$$

which is only non-zero when $i = j$ [recalling (3)]. Under the orthogonality of (10) and sufficiently long system response ($T_1 \gg 0$), the cross-correlation in (8) can be further simplified as

$$\begin{aligned} C_{k\ell}(\tau) &\cong \sum_{i=1}^n \mu c_{ki} c_{\ell i} \int_0^{T_1} \cos(\beta_i \tau_1) \cos\left(\beta_i (\tau_1 - \tau)\right) d\tau_1 \\ &= \sum_{i=1}^n \mu c_{ki} c_{\ell i} \left[\frac{T_1}{2} \cos(\beta_i \tau) + \frac{1}{2} \int_0^{T_1} \cos(2\beta_i \tau_1 - \beta_i \tau) d\tau_1 \right]. \end{aligned} \tag{11}$$

The second summand of (11) can be easily bounded by 1 and becomes negligible compared to the first one as T_1 grows, and thus the cross-correlation becomes

$$C_{k\ell}(\tau) \cong \frac{\mu T_1}{2} T_{k\ell}(\tau). \tag{12}$$

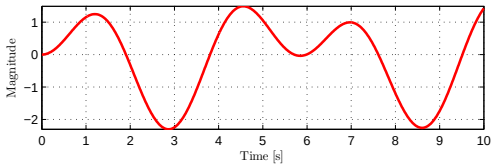
This establishes the equivalence between the cross-correlation of ambient data and the system impulse response, up to a scaling difference. Several key assumptions have been made, especially regarding the perturbation inputs. Although these assumptions like the covariance condition in (10) may seem restrictive, they are equivalent to having homogeneous excitation inputs across the system, which should hold for large-scale interconnected power grid.

To quickly verify our analysis, we will test it on a two-node toy undamped case. Consider the system (2) with the following parameters:

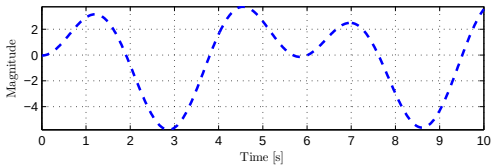
$$\mathbf{M} = \begin{bmatrix} 0.25 & 0 \\ 0 & 0.5 \end{bmatrix}, \quad \mathbf{K} = \begin{bmatrix} 1 & -0.5 \\ -0.5 & 1 \end{bmatrix}, \quad \mathbf{D} = \mathbf{0}. \tag{13}$$

The generalized eigenvalues are given by $-\lambda_i^2 = \{4.73, 1.27\}$. Hence, the system is stable with oscillation frequency at $\beta_i = \{2.18, 1.13\}$. We compute the cross-correlation using the response of $\boldsymbol{\omega}$ to white-noise input. The system response is generated with $T_1 = 500$ seconds from the zero initial condition at time step of $dt = 0.01$ second.

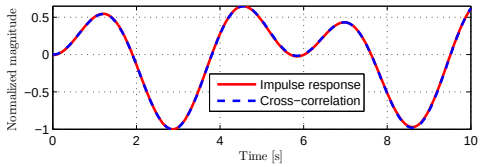
Fig. 1 plots the comparison between the cross-correlation result and the impulse response for input-output pair T_{21}. Based on (12), the two are different by a scaling factor at



(a) Impulse response

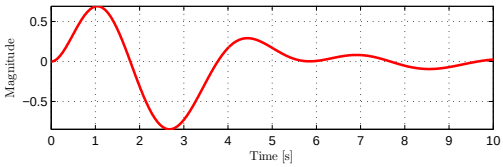


(b) Cross-correlation

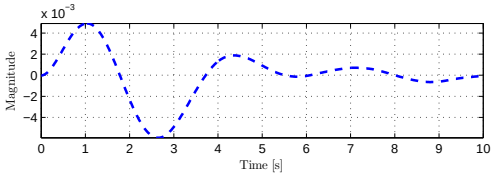


(c) Normalized comparisons

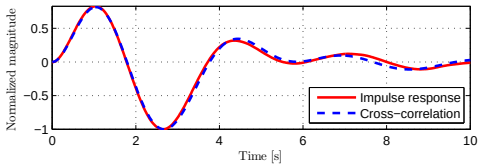
Fig. 1: Impulse response versus cross-correlation from u_2 to ω_1 for the two-node undamped case.



(a) Impulse response



(b) Cross-correlation



(c) Normalized comparisons

Fig. 2: Impulse response versus cross-correlation from u_1 to ω_2 for the two-node damped case.

TABLE I: Dissimilarity scores for the two-node undamped case.

S_{kl}	u_1	u_2
ω_1	0.008	0.010
ω_2	0.010	0.015

TABLE II: Dissimilarity scores for the two-node damped case.

S_{kl}	u_1	u_2
ω_1	0.045	0.052
ω_2	0.083	0.051

$\frac{\mu T_1}{2}$. To account for the scaling difference, we perform the normalization using the maximum absolute value and compare the two using the following dissimilarity score

$$S_{k\ell} = \frac{\|\bar{T}_{k\ell} - \bar{C}_{k\ell}\|_2}{\|\bar{T}_{k\ell}\|_2} \tag{14}$$

where $\bar{T}_{k\ell}$ ($\bar{C}_{k\ell}$) is the normalized impulse response (cross-correlation) from input u_k to speed state ω_ℓ, while $\|\cdot\|_2$ denotes the L^2 norm of the vector corresponding to the discrete-time response. Hence, perfect similarity would correspond to a zero score. The similarity scores for each input-output pair are listed in Table I. The extremely small values verify that the two time series are very similar up to a scaling difference.

As mentioned earlier, the equivalence analysis can be potentially extended to damped system. To demonstrate this possibility, we have slightly modified the system in (13) by changing the damping matrix to

$$\mathbf{D} = 0.75\mathbf{M}. \tag{15}$$

For this modified damped case, the dissimilarity scores are updated in Table II, and the comparison corresponding to the pair (u_1, ω_2) with the highest dissimilarity score is plotted in Fig. 2. Again, the updated dissimilarity scores are close to zero, suggesting that the cross-correlation method works in estimating the impulse response for a system that has damping matrix $\mathbf{D}$ proportional to the mass matrix $\mathbf{M}$.

The analysis and the examples so far shows for undamped stable systems and certain damped systems, the impulse response between any input-output pair can be recovered by the cross-correlation of the generator speed responses under ambient conditions. Lacking of direct measurements of generator internal speed in practical implementations, one can use the synchrophasor frequency data at the closest bus as the surrogate.

IV. Case Studies

This section will test the proposed cross-correlation based approach the WSCC 9-bus system (modified from the one provided in [27]) and the NPCC 48-machine system (as given in [28]). Although our analysis in Section III is limited to several simplification assumptions such as undamped and symmetric systems, we demonstrate the effectiveness of the cross-correlation approach for practical systems with damping and even governor/exciter controls. Specifically, for the WSCC 9-bus case, we compare the model-based impulse response and the estimated one for every pair of generators. As for the NPCC 48-machine case, we have identified some local or inter-area oscillation scenarios where the proposed method is more effective. Other less successful scenarios are also presented to better understand the applicability of the proposed method.

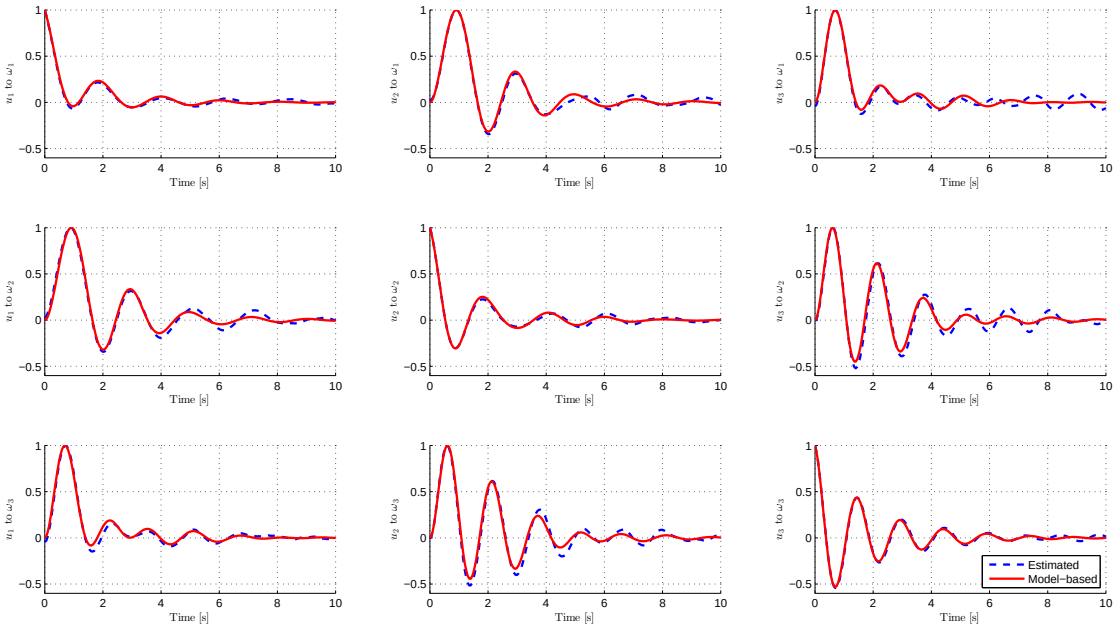


Fig. 3: WSCC 9-bus system: model-based impulse responses (solid red lines) versus the estimated ones using cross-correlation (dash blue lines) for each input-generator pair.

A. *WSCC 9-bus system*

We first test on the WSCC 9-bus system, popular for power system dynamics analysis. Following from the approach in [10], we modify the WSCC 9-bus case in [27] by only consider the classical synchronous generator models without exciters or governors. The parameter values are as follow: $M_1 = 0.63, M_2 = 0.34, M_3 = 0.16$, and $D_1 = 0.63, D_2 = 0.34, D_3 = 0.16$. Under these parameter values, the $\mathbf{K}$ matrix in synchronous reference frame is given by

$$\mathbf{K} = \begin{bmatrix} 2.819 & -1.523 & -1.294 \\ -1.611 & 2.723 & -1.112 \\ -1.338 & -1.108 & 2.447 \end{bmatrix}. \qquad (16)$$

Clearly, matrix $\mathbf{K}$ is not perfectly symmetric for this system, which may affect the validity of the condition on vectors $\{\mathbf{c}_i\}$ in (3). After obtaining the linearized model (2), we can directly compute the model-based impulse response. To obtain the ambient frequency measurements, we generate random white Gaussian processes of inertia proportional variance as the input to the linearized system. The speed responses are cross-correlated to estimate the corresponding impulse response.

Figure 3 plots the modeled-based impulse responses along with the estimated counterparts for every input to every generator output. To tackle the scaling ambiguity between the two, we have normalized each time series to be of unit peak value. High consistency has been observed for each input-generator pair, demonstrating the effectiveness of the proposed approach. We also calculate the dissimilarity scores to quantify the similarity using equation (14). The numerical values are listed in Table III.

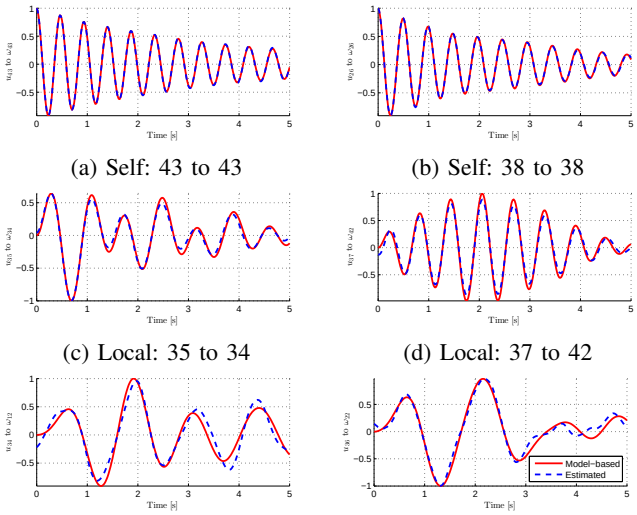
TABLE III: Dissimilarity scores between the model-based and estimated impulse responses for WSCC 9-bus case.

$S_{k\ell}$	u_1	u_2	u_3
ω_1	0.095	0.107	0.182
ω_2	0.123	0.103	0.188
ω_3	0.120	0.166	0.100

Overall speaking, the reconstructed impulse responses match well with the actual ones for the WSCC 9-bus case. The best match turns out to be from u_1 to ω_1, with dissimilarity score at 0.095. The least similar one appears to be from u_3 to ω_2, with dissimilarly score of 0.188. Even at the highest dissimilarity, the two time series are still shown to be very close, as illustrated in the right middle plot in Figure 3. The comparisons again verify the effectiveness of the proposed data-driven approach in estimating the actual system impulse responses.

B. *NPCC 48-machine system*

We further test the proposed method on the NPCC 48-machine system with 140 buses, the detailed case information of which is available from [28]. Generators in this case are equipped with governors and exciters. They could be partitioned into 17 areas, as discussed in [30]. Using the area partition information therein, we can categorize the local or inter-area oscillations for this system. We use the same techniques to generate the ambient frequency measurements

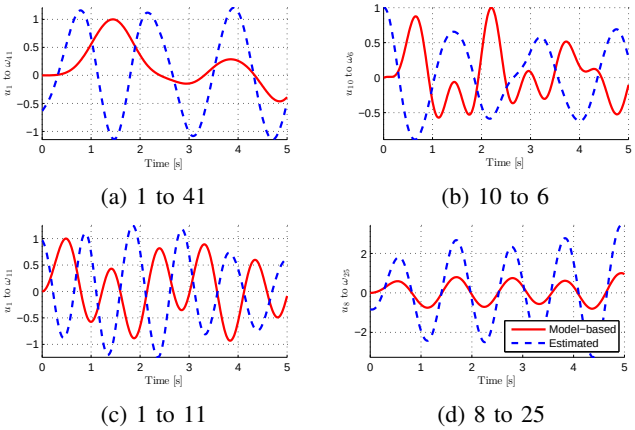


(a) Self: 43 to 43 (b) Self: 38 to 38

(c) Local: 35 to 34 (d) Local: 37 to 42

(e) Inter: 34 to 12 (f) Inter: 36 to 22

Fig. 4: NPCC 48-machine system: model-based and estimated impulse responses for selected input-generator pairs under local or inter-area oscillations.



(a) 1 to 41 (b) 10 to 6

(c) 1 to 11 (d) 8 to 25

Fig. 5: NPCC 48-machine system: model-based and estimated impulse responses for example input-generator pairs at high level of mismatch.

everywhere in the system for performing the cross-correlation analysis.

We compare some selected pairs of model-based and estimated impulse responses in Figures 4 and 5. Figures 4(a)-(b) correspond to those from the input at generator 43 or 38 to itself, respectively. These two self-impulse responses are perfectly reconstructed. Figures 4(c)-(d) plot two selected pairs corresponding to local area oscillation at area 11 and area 9, respectively. Both also show good estimation accuracy. Figures 4(e)-(f) illustrate two inter-area oscillations, one from area 11 to area 4 and the other from area 12 to area 6. For the selected pairs, the impulse responses are well estimated by the cross-correlation.

Although the proposed approach has been successfully validated for a majority of input-generator pairs, there also exists some high level of mismatch for several pairs. Some of these example cases are plotted in Figure 5. Since our analysis has been focused on very simple system models using swing equations, we have provided certain characteristics of real systems that may have led to the observed mismatches. One factor could be the fact that in real power systems not all lines are purely inductive, as some lines have resistance even comparable to its inductance. Therefore, the from-end and to-end power flows do not equal to each other, affecting the symmetric property of matrix $\mathbf{K}$. This asymmetry has already been observed by the WSCC 9-bus system, as illustrated in (16). For the example pairs with high level of mismatch in Fig. 5, we notice that most of the generator buses are directly connected to highly-resistive transmission lines. For example, generator 11 at bus 47 is connected to bus 46 by a line whose resistance value is at more than half of the inductance value.

Another factor that could cause the mismatch is the existence of controllers (exciters, turbine governors and power system stabilizers) for improved voltage or frequency stability. The additional controllers affect the validity of the swing equations based dynamic model. From the simulation results, we observe that cross-correlation usually works better if the associated generators are equipped with less complicated controllers . For example in Fig. 4(a)-(d), the generators 34, 35, 37, 38, 42, 43 neither have exciter or turbine governor only. If both types of controllers exist, the cross-correlation results are typically different from the model-based ones. As in Fig. 5(b)-(c), the generators 1, 6, 10, 11 are all equipped with both exciter and turbine governor. We suspect the higher-order dynamics due to these control actions have affected the performance of the proposed data-based method.

V. Concluding Remarks

This paper studied the electromechanical disturbance propagation phenomenon, a long-standing problem for bulk interconnected power grids, from a new data-driven perspective. Our approach leverages the increasing volume of high-resolution synchrophasor data available to understand the impact of an input disturbance at one location to another location in the interconnected power system. To estimate this impulse response, we proposed to cross-correlate ambient frequency measurements at the two locations without requiring any model information. The estimated impulse responses could be used to further determine the modes and mode shapes, as well the generator participation factor for a particular mode. Albeit successfully validated by the WSCC 9-bus case, the proposed data-driven method turns out to be at visibly high level of mismatch for selected input-generator pairs for the NPCC 48-machine system. Certain factors have been identified, such as the existence of highly-resistive transmission lines and generator control actions. It is crucial to perform real data based validations and to further understand the applicability of the proposed cross-correlation approach for large-scale interconnected systems.

VI. Acknowledgment

This work was supported by the Siebel Energy Institute and the National Science Foundation (NSF) under Award Number ECCS-1653706. QC acknowledges travel support from South California Earthquake Center (SCEC) to attend 2016 Summer School on Earthquake Science that covered frontiers of ambient noise correlation techniques.

References

[1] G. Rogers, *Power system oscillations*. Springer Science & Business Media, 2012.

[2] C. W. Taylor, "Improving grid behaviour," *IEEE Spectrum*, vol. 36, no. 6, pp. 40–45, Jun 1999.

[3] M. Klein, G. J. Rogers, and P. Kundur, "A fundamental study of inter-area oscillations in power systems," *IEEE Trans. Power Syst.*, vol. 6, no. 3, pp. 914–921, Aug 1991.

[4] M. A. P. Peter W. Sauer, *Power System Dynamics and Stability*. Stipes Publishing L.L.C., 1997.

[5] P. Kundur, N. J. Balu, and M. G. Lauby, *Power system stability and control*. McGraw-hill New York, 1994, vol. 7.

[6] J. W. Pierre, D. J. Trudnowski, and M. K. Donnelly, "Initial results in electromechanical mode identification from ambient data," *IEEE Trans. Power Syst.*, vol. 12, no. 3, pp. 1245–1251, Aug 1997.

[7] N. Zhou, J. W. Pierre, D. J. Trudnowski, and R. T. Guttromson, "Robust rls methods for online estimation of power system electromechanical modes," *IEEE Trans. Power Syst.*, vol. 22, no. 3, pp. 1240–1249, Aug 2007.

[8] N. Zhou, D. J. Trudnowski, J. W. Pierre, and W. A. Mittelstadt, "Electromechanical mode online estimation using regularized robust rls methods," *IEEE Trans. Power Syst.*, vol. 23, no. 4, pp. 1670–1680, Nov 2008.

[9] N. Zhou, Z. Huang, L. Dosiek, D. Trudnowski, and J. W. Pierre, "Electromechanical mode shape estimation based on transfer function identification using pmu measurements," in *2009 IEEE Power Energy Society General Meeting*, July 2009, pp. 1–7.

[10] X. Wang and K. Turitsyn, "Pmu-based estimation of dynamic state jacobian matrix," *arXiv preprint arXiv:1510.07603*, 2015.

[11] X. Wang and K. Turitsyn, "Data-driven diagnostics of mechanism and source of sustained oscillations," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 4036–4046, Sept 2016.

[12] J. S. Thorp, C. E. Seyler, and A. G. Phadke, "Electromechanical wave propagation in large electric power systems," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl. (until 2003)*, vol. 45, no. 6, pp. 614–622, Jun 1998.

[13] M. Parashar, J. S. Thorp, and C. E. Seyler, "Continuum modeling of electromechanical dynamics in large-scale power systems," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 51, no. 9, pp. 1848–1858, Sept 2004.

[14] O. I. Lobkis and R. L. Weaver, "On the emergence of the greenŠs function in the correlations of a diffuse field," *The Journal of the Acoustical Society of America*, vol. 110, no. 6, pp. 3011–3017, 2001.

[15] R. Snieder, "Extracting the greenŠs function from the correlation of coda waves: A derivation based on stationary phase," *Physical Review E*, vol. 69, no. 4, p. 046610, 2004.

[16] R. Snieder, K. Wapenaar, and U. Wegler, "Unified greenŠs function retrieval by cross-correlation; connection with energy principles," *Physical Review E*, vol. 75, no. 3, p. 036103, 2007.

[17] K. Wapenaar, "Retrieving the elastodynamic green's function of an arbitrary inhomogeneous medium by cross correlation," *Physical review letters*, vol. 93, no. 25, p. 254301, 2004.

[18] V. C. Tsai, "On establishing the accuracy of noise tomography travel-time measurements in a realistic medium," *Geophysical Journal International*, vol. 178, no. 3, pp. 1555–1564, 2009.

[19] V. C. Tsai, "The relationship between noise correlation and the green's function in the presence of degeneracy and the absence of equipartition," *Geophysical Journal International*, vol. 182, no. 3, pp. 1509–1514, 2010.

[20] V. C. Tsai, "Understanding the amplitudes of noise correlation measurements," *Journal of Geophysical Research: Solid Earth*, vol. 116, no. B9, 2011.

[21] M. Campillo and A. Paul, "Long-range correlations in the diffuse seismic coda," *Science*, vol. 299, no. 5606, pp. 547–549, 2003.

[22] N. M. Shapiro and M. Campillo, "Emergence of broadband rayleigh waves from correlations of the ambient seismic noise," *Geophysical Research Letters*, vol. 31, no. 7, 2004.

[23] P. Roux, K. G. Sabra, P. Gerstoft, W. Kuperman, and M. C. Fehler, "P-waves from cross-correlation of seismic noise," *Geophysical Research Letters*, vol. 32, no. 19, 2005.

[24] G. Bensen, M. Ritzwoller, M. Barmin, A. Levshin, F. Lin, M. Moschetti, N. Shapiro, and Y. Yang, "Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements," *Geophysical Journal International*, vol. 169, no. 3, pp. 1239–1260, 2007.

[25] M. Denolle, E. Dunham, G. Prieto, and G. Beroza, "Strong ground motion prediction using virtual earthquakes," *Science*, vol. 343, no. 6169, pp. 399–403, 2014.

[26] S. Backhaus and Y. Liu, "Electromechanical wave green's function estimation from ambient electrical grid frequency noise," in *2012 45th Hawaii International Conference on System Sciences*, Jan 2012, pp. 2054–2061.

[27] F. Milano, "An open source power system analysis toolbox," *IEEE Trans. Power Syst.*, vol. 20, no. 3, pp. 1199–1206, Aug 2005.

[28] J. H. Chow and K. W. Cheung, "A toolbox for power system dynamics and control engineering education and research," *IEEE Trans. Power Syst.*, vol. 7, no. 4, pp. 1559–1564, Nov 1992.

[29] B. N. Parlett, *The symmetric eigenvalue problem*. SIAM, 1998.

[30] J. H. Chow, R. Galarza, P. Accari, and W. W. Price, "Inertial and slow coherency aggregation algorithms for power system dynamic model reduction," *IEEE Trans. Power Syst.*, vol. 10, no. 2, pp. 680–685, May 1995.