

IET Generation, Transmission & Distribution

Special issue

Call for Papers

**Be Seen. Be Cited.
Submit your work to a new
IET special issue**

Connect with researchers and
experts in your field and share
knowledge.

Be part of the latest research
trends, faster.

Read more



**The Institution of
Engineering and Technology**

ORIGINAL RESEARCH PAPER

Scalable design methods for online data-driven wide-area control of power systems

Jishnudeep Kar | Aranya Chakraborty

FREEDM Systems Center, North Carolina State University, 2816 Avent Ferry Road, Raleigh, NC 27606, USA (Email: achakra2@ncsu.edu)

Correspondence

Jishnudeep Kar, FREEDM Systems Center, North Carolina State University, 2816 Avent Ferry Road, Raleigh, NC 27606, USA.
Email: jkar@ncsu.edu

Funding information

National Science Foundation, Grant/Award Number: ECS-1711004

Abstract

A novel online system-identification-based control design for damping inter-area oscillations in power systems using supplementary wide-area excitation control is presented. The identification is based on a partly reduced-order and a partly full-order model of the grid. Assuming the grid to be divided into coherent areas with phasor measurement units located across the grid, the coherent states of the generators are averaged while the non-coherent states, such as the internal states of each power system stabilizer, are not. A hybrid state vector consisting of these averaged and non-averaged states is computed, and used for identifying a hybrid linear time-invariant state-space model. A linear quadratic regulator based on this hybrid model is, thereafter, designed. The reduced-order part of the model is found to save significant amount of online time for both learning and control, while the full-order part serves to enhance closed-loop stability. N4SID with randomized singular value decomposition (rSVD) is used for making the identification loop fast. The reduced-dimensional controller is finally implemented using a broadcast control strategy. The design is extended to non-linear models of power systems using Carleman bilinearization. Results for linear and bilinear control designs are validated using the IEEE 68-bus power system model.

1 | INTRODUCTION

Wide-area control of power systems using Synchrophasors has seen a suite of control algorithms developed so far [1–4]. The majority of these algorithms, however, are model based and are performed offline. With the growing levels of uncertainty in grid operations such as constantly changing loads, wind and solar penetration and changes in grid topology, system operators currently are inclining more towards model-free and online control approaches, but the greatest challenge standing in the way is the gigantic size and spatial scale of any real-world grid. As shown in [5], using online phasor measurement units (PMU) measurements to identify an entire grid model, and employing that model to take online control actions is practically impossible as the identification itself can take hours to complete even with the the most advanced computational facilities.

To circumvent this scalability issue of learning-based control for large and complex power systems, in this paper we propose to identify a severely reduced-dimensional grid model by mak-

ing use of time-scale separation property of the grid dynamics, using high-resolution and high-fidelity PMU measurements. This property is commonly seen in almost all practical grid models, either through coherency of the synchronous generators [6], or through the sparsity of the grid topology, or through controller time constants. Based on this identified reduced-order model, we design a reduced-dimensional linear quadratic regulator (LQR), which is implemented through an inverse projection-based broadcast control strategy. The challenge, however, arises from the fact that not all states of a generator exhibit time-scale separation, and, thus, cannot be reduced to an aggregate state. For example, from [6] we know that the electro-mechanical (swing) states and excitation system states exhibit coherency, and can be expressed as aggregate states by means of simple inertia-weighted averaging. The states of the power system stabilizer (PSS), however, do not exhibit this property. Thus, ideally speaking, the PSS states cannot be expressed as an aggregate state. In fact, forcing the reduced-order model to contain an aggregate PSS state can result in a controller which when

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *IET Generation, Transmission & Distribution* published by John Wiley & Sons Ltd on behalf of The Institution of Engineering and Technology

implemented via the broadcast rule can destabilize the original full-order grid model. In summary, designing a full-order controller based on a full-order grid model will result in optimal performance but will take an unacceptably long learning time, while on the other hand, designing a broadcast controller based on an *exact* reduced-order grid model will take significantly lesser amount of learning time but may cause closed-loop instability. We propose an approach that is midway between these two extreme scenarios. Assuming the grid to be divided into coherent areas, we average the coherent (swing and excitation) states of the generators, but do not average the non-coherent PSS states. A hybrid state vector consisting of a mixture of these averaged and non-averaged states is computed following any disturbance, and used for identifying a hybrid state-space model, followed by a LQR design based on this hybrid model. The reduced-order part of the model saves significant amount of online time for both learning and control, while retaining the full-order PSS states is found to enhance closed-loop stability. Singular value decomposition (SVD) and least squares are used in the internal steps of the identification. N4SID with randomized singular value decomposition (rSVD) is used for making the identification fast [7]. Results are validated using the 9-bus Kundur power system model [8] and the IEEE 68-bus power system model.

The second half of the paper extends this reduced-order learning-based control design to non-linear models of power systems. Identifying the linear model and implementing the linear state-feedback controller may not suffice when the system may have non-linear loads, or may be too close to voltage instability, for example, where it exhibits significant amount of non-linearity in its dynamics. Recent papers such as [9, 10] have considered control design for non-linear power system models using Koopman mode analysis, while [11] has designed approximate controllers using Carleman linearization. However, these are all model-based control designs, and do not consider any model reduction operation. In Section 6 of this paper we will show how the idea of learning a reduced-order model that we developed for the linear case can be extended to the Carleman linearized model, and subsequently develop a step-by-step iterative approach for designing a bilinear quadratic controller for the identified non-linear model containing both first- and second-order terms. The controller is validated using the IEEE 68-bus system with several non-conforming non-linear loads to show its superior performance compared to the linear controller.

While learning has been used in Synchrophasor applications on several occasions, majority of them are for open loop problems such as mode estimation [12], parameter estimation [13, 14], disturbance monitoring [15], etc. Very little work exists on how learning can be used for control, and that too taking the bottleneck of scalability into account. Recent works such as [16] have integrated model-based methods with data-driven methods for addressing frequency regulation, but not for LQR-based damping control. Other papers such as [17] have used direct adaptive control methods for learning real-time controllers using reinforcement learning. This approach, however, has two main drawbacks. First, during the *learning phase* the con-

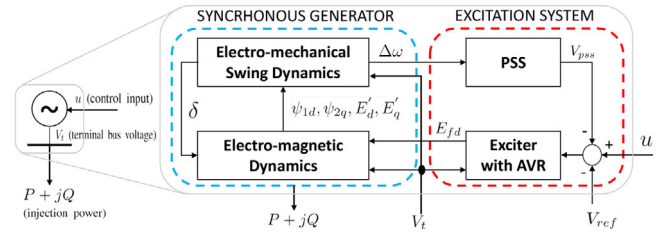


FIGURE 1 Synchronous generator with excitation system

troller may cause the system to behave erratically. Second, these designs rest on the idealistic assumption that a stabilizing controller can be found for the system model in order to start the iterative process for online control. Our proposed *indirect* approach does not have either of these drawbacks.

Our main contributions can be summarized as follows:

1. System identification of a *hybrid* model of a power system. The term ‘hybrid’ refers to the fact that the model is composed of both reduced-order and full-order dynamics.
2. Design of a reduced-dimensional LQR-based wide-area damping controller using the hybrid model, followed by a broadcasting technique for implementation.
3. Extension of the approach to non-linear Carleman models of power systems under high load stress.

The overall approach is termed as model free as we identify the reduced-order model only from measurement data.

The rest of the paper is organized as follows. Section 2 describes the problem formulation and identifying the cluster of generators for the reduced-order model. Section 3 presents the online algorithm for identification of the reduced-order linear time-invariant (LTI) system followed by its LQR control design in Section 4. Section 5 introduces in brief the concept of Carleman bilinearization for non-linear systems and design of quadratic controllers for such models. Section 6 implements the idea to power system taking hints from the actual model equations of a non-linear power system. Section 7 presents the simulations and Section 8 concludes the paper.

2 | PROBLEM FORMULATION

2.1 | Model description

Consider a power system with n synchronous generators. Every generator is modeled by a standard sixth-order dynamic model, as shown in Figure 1, with states as phase angle δ_i , rotor speed ω_i , transient emf due to field-flux linkage E_{qi} , transient emf due to field-flux linkage in q -axis E_{di} , sub-transient emf due to flux linkage in d -axis damper ψ_{1di} , and sub-transient emf due to flux linkage in q -axis damper ψ_{2qi} . The detailed non-linear dynamic equations can be found in [8]. For control applications, the stator transients are assumed to be much faster compared to swing dynamics, and follow from a standard set of algebraic equations. For the excitation system, we consider the

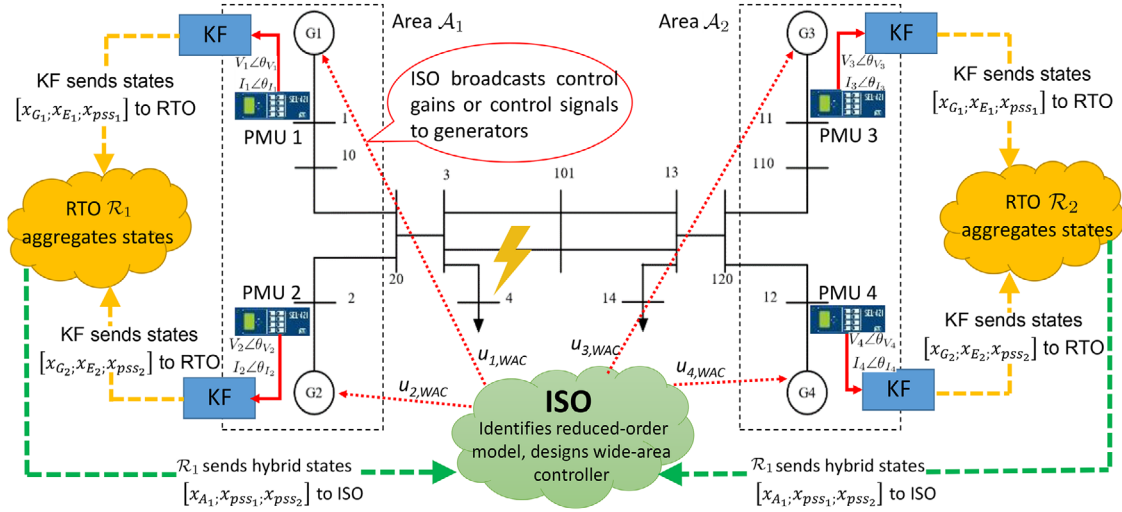


FIGURE 2 Implementation of learning and control scheme

IEEE-DC1A type exciter [18] with first-order dynamics of field voltage E_{fdi} , and a speed input PSS with third-order dynamics given by

$$V_{pss_i} = K_{pss_i} \frac{sT_w(1+sT_i)(1+sT_{3i})}{(1+sT_w)(1+sT_{2i})(1+sT_{4i})} \Delta\omega_i. \quad (1)$$

Details of these models and the corresponding notations can be found in [8], and are skipped here for brevity. The system model is linearized about the load flow solution.

Let us define the small-signal generator states as $x_{G_i} = [\Delta\delta_i \ \Delta\omega_i \ \Delta E_{qi} \ \Delta E_{di} \ \Delta\psi_{1di} \ \Delta\psi_{2qi}]^T \in \mathbb{R}^6$, the exciter state as $x_{E_i} = \Delta E_{fdi} \in \mathbb{R}$, and the three PSS states as $x_{pss_i} = [x_{pss_i}^{(1)} \ x_{pss_i}^{(2)} \ x_{pss_i}^{(3)}]^T \in \mathbb{R}^3$ for the i th generator, $i = 1, \dots, n$. The Kron-reduced model is written as

$$\dot{x}(t) = Ax(t) + Bu(t), \quad x(0) = x_0, \quad (2)$$

where $A \in \mathbb{R}^{(7n+3n_p) \times (7n+3n_p)}$, $B \in \mathbb{R}^{(7n+3n_p) \times n}$ are the state matrix and input matrix, and

$$x = [x_{G,E}; x_{pss}] \in \mathbb{R}^{7n+3n_p}, \quad (3)$$

$$x_{G,E} = [x_{G_1}; x_{E_1}; \dots x_{G_n}; x_{E_n}] \in \mathbb{R}^{7n}, \quad (4)$$

$$x_{pss} = [x_{pss_1}; \dots x_{pss_{n_p}}] \in \mathbb{R}^{3n_p}, \quad (5)$$

and $u(t) \in \mathbb{R}^n$ is the exciter input. n_p is the total number of generators with PSS (as not all generators may have PSS). The symbol $[\cdot]$ denotes a column vector of its arguments.

We assume that the matrices A and B are unknown, but the pair (A, B) is stabilizable. PMUs are assumed to be located at geometrically observable set of buses such that the voltage and

current phasors at every generator bus are computable from these measurements. Details of such PMU deployment can be found in [19]. A decentralized Kalman filter (KF) based on [20] is assumed to be present at every generator, which using these local bus phasors generates an estimate of all states of that generator, as shown in Figure 2. For implementing the KFs we will assume that the generator parameters are known individually to each generator owner, and that the uncertainty in the system model arises from variations in the grid-level parameters such as topology and loads, all of which collectively make A and B in (2) unknown. Since the KF estimates of $x(t)$ are highly accurate, we will ignore the error in the estimates throughout the paper, and implicitly assume that we can measure $x(t)$ that will be needed for both identification and control.

Problem statement

The goal is to design an LQR-based damping control input $u(t)$. For this, one would first need to identify A and B . Considering that the order of the system n can be in the hundreds, this identification, however, will take an unacceptably long amount of time. We, therefore, propose to identify a reduced-order model, as presented next.

2.2 | Model reduction using coherency

We assume that the power system is divided into r coherent areas. Two generators G_i and G_j are said to be coherent if for a certain disturbance the generators swing together [6], that is, their rotor angles δ satisfy $|\Delta\delta_i - \Delta\delta_j| < \epsilon$ for every time instant, where $0 < \epsilon \ll 1$ is a small specified tolerance. A smaller value of chosen tolerance gives a tighter coherent grouping. This definition can be used to impart a two time-scale separation in the system dynamics whereby (2) can be equivalently

written as

$$\epsilon \dot{x}_f = A_{11}x_f + A_{12}x_s + B_1u, \quad (6)$$

$$\dot{x}_s = A_{21}x_f + A_{22}x_s + B_2u, \quad (7)$$

following a similarity transformation $\text{col}(x_f, x_s) = Tx$, where T is an invertible transformation matrix (for its detailed construction please see [6]), x_f are the *fast* states, and x_s are the *slow* states. A_{11} through A_{22} , and B_1 and B_2 are matrices of appropriate dimensions. Assuming $\epsilon = 0$, the quasi-steady state of x_f can be approximated as $x_f \approx -A_{11}^{-1}A_{12}x_s - A_{11}^{-1}B_1u$, following which the slow dynamics in (7) can be written as

$$\dot{x}_s = \underbrace{(A_{22} - A_{21}A_{11}^{-1}A_{12})}_{A_{\text{red}}}x_s + \underbrace{(B_2 - A_{21}A_{11}^{-1}B_1)}_{B_{\text{red}}}u, \quad (8)$$

where A_{red} and B_{red} are reduced-order matrices.

Our approach will be to identify $(A_{\text{red}}, B_{\text{red}})$ using measurements of $x_s(t)$, assuming that the identity of the coherent areas is known, and thereafter design an LQR controller based on this tuple. The identity of the coherent groups usually do not change very frequently, and therefore a nominal knowledge base at the ISO can be used to identify the generator indices belonging to each coherent area. If the ISO intends to update its coherency map then several offline data-driven techniques such as spectral clustering reported in the recent literature [21] can be used.

2.2.1 | Aggregation of states

Consider the power system to be divided into $r < n$ coherent areas $\{\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_r\}$. The generator identities in each area are assumed to be known by the ISO. Following the coherency theory as in [6], we define the aggregate states as the area-wise average of the six generator states (x_{G_i}) and one exciter state (x_{E_i}). The seven aggregate *slow* state variables for \mathcal{A}_j are written as

$$\underbrace{x_{\mathcal{A}_j}^{1,\dots,6}}_{\text{Aggregate generator states}} = \frac{\sum_{i \in \mathcal{A}_j} M_i x_{G_i}}{\sum_{i \in \mathcal{A}_j} M_i} \in \mathbb{R}^6, \quad \underbrace{x_{\mathcal{A}_j}^7}_{\text{Aggregate exciter state}} = \frac{\sum_{i \in \mathcal{A}_j} x_{E_i}}{n_{\mathcal{A}_j}} \in \mathbb{R}^1, \quad (9)$$

where $n_{\mathcal{A}_j}$ is the number of generators in area j and M_i is the inertia of generator i . The symbol $i \in \mathcal{A}_j$ means generator i in the j th coherent area. Our strategy will be to identify a reduced-dimensional grid model based on these seven aggregate states obtained from each area. The identification will be done in a centralized way at the ISO. We will not consider aggregation of any of the PSS states because they are not physical states, and more importantly, they do not exhibit coherency. This is illustrated in Figure 3 where the first PSS state of generators 1 and 2, respectively, denoted as $x_{PSS1}^{(1)}$ and $x_{PSS2}^{(1)}$, for the IEEE 9-bus two-area power system model are clearly seen to be non-coherent. The $7r$ -dimensional state-space model which approximates the

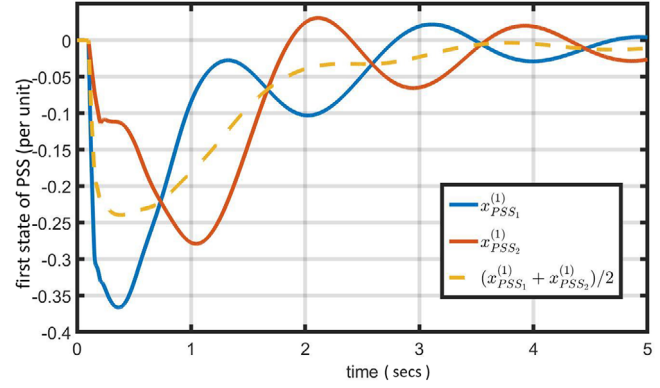


FIGURE 3 Non-coherent PSS states of IEEE 9-bus system

slow dynamics of (2) and is to be identified from the measurements of the $7r$ aggregate states following from (9) is written as

$$\dot{x}_s = A_{\text{red}}x_s + B'u, \quad x_s = \begin{bmatrix} x_{\mathcal{A}_1}^T & \dots & x_{\mathcal{A}_r}^T \end{bmatrix}^T, \quad (10)$$

where, $A_{\text{red}} \in \mathbb{R}^{7r \times 7r}$, and $B' \in \mathbb{R}^{7r \times n}$. We will design u as

$$u = P_B u_{\text{red}}, \quad (11)$$

where $u_{\text{red}} \in \mathbb{R}^r$ is a reduced-dimensional control input to be designed using LQR, and P_B is a broadcast matrix of the form

$$P_B(i, j) = \begin{cases} \frac{M_i}{\sum_{k \in \mathcal{A}_j} M_k}, & \text{if } i \in \mathcal{A}_j \\ 0, & \text{otherwise.} \end{cases} \quad (12)$$

The broadcast strategy indicates that all generators in area \mathcal{A}_j will receive the same control input (the j th element of u_{red}), but proportionally weighted according to its inertia relative to the total inertia of that area. We then write (10) as

$$\dot{x}_s = A_{\text{red}}x_s + B_{\text{red}}u_{\text{red}}, \quad B_{\text{red}} = B'P_B. \quad (13)$$

In the identification stage, as seen from (13), we ignore the effect of PSS states in the slow dynamics as they have a low participation factor on the slow poles. However, for the control stage, we will shortly see that ignoring the PSS states during control design may cause closed-loop instability. We will, therefore, develop an augmentation method to include the PSS model in A_{red} .

2.3 | Implementation architecture

Before proceeding to identification, it is important to discuss the implementation strategy of how the ISO may receive the aggregate states. As mentioned earlier, we assume that the

n generators are divided into r coherent areas. The set of coherent areas is denoted by $\mathcal{A} = \{\mathcal{A}_1, \dots, \mathcal{A}_r\}$. We also assume that the grid is divided into ℓ number of regional transmission operators (RTOs), denoted by the set $\mathcal{R} = \{\mathcal{R}_1, \dots, \mathcal{R}_\ell\}$. We discuss the following two cases.

Case 1: Consider the case when every coherent area \mathcal{A}_i is assumed to lie completely within a single RTO \mathcal{R}_j , that is, no coherent area belongs to more than one RTO. Let $\mathcal{A}^{(j)}$ denote the subset of \mathcal{A} that belongs to the region under RTO \mathcal{R}_j . In this case, the KF of each generator would send its states to the control center of the corresponding RTO. The aggregator or the phasor data concentrator (PDC) located at this control center would aggregate the states $x_{\mathcal{A}_i}, \forall \mathcal{A}_i \in \mathcal{A}^{(j)}$ following (9). Thereafter, the RTO \mathcal{R}_j would send all the aggregate states $x_{\mathcal{A}_i}$ (such that \mathcal{A}_i belongs to RTO \mathcal{R}_j) to the ISO. This case is illustrated in Figure 2.

Case 2: Consider the case when a coherent area \mathcal{A}_i may belong to more than one RTO. Denote the set of regions consisting of this coherent area \mathcal{A}_i as $\mathcal{R}^{(i)}$. Each RTO knows the identity of the generators belonging to every coherent area in its own region. However, it may not know the global coherency structure. Therefore, each RTO $\mathcal{R}_j \in \mathcal{R}^{(i)}$ sends the aggregate states for the generators belonging to its own coherent area as $x_{\mathcal{A}_i}^{\mathcal{R}_j}$ according to (9). The ISO, on the other hand, knows the global coherency structure. Thus, once the ISO receives the aggregate states for \mathcal{A}_i from each RTO in $\mathcal{R}^{(i)}$, it will compute the aggregate state as

$$x_{\mathcal{A}_i} = \frac{\sum_{\mathcal{R}_j \in \mathcal{R}^{(i)}} \left(\sum_{G_k \in \mathcal{A}_i \cap \mathcal{R}_j} M_k \right) x_{\mathcal{A}_i}^{\mathcal{R}_j}}{\sum_{G_k \in \mathcal{A}_i} M_k}. \quad (14)$$

Once the area-wise aggregate states $x_{\mathcal{A}_i}$ are calculated, the slow state vector x_s is constructed as in (10), and used for identification as explained next.

3 | REDUCED-ORDER MODEL IDENTIFICATION

3.1 | Identification of A_{red} and B_{red}

Following a small-signal disturbance, we first input a low-amplitude persistently exciting Gaussian noise to the exciter input u_i of every i th generator, $i = 1, \dots, n$. Generators belonging to the same coherent area are excited with the same noise input, that is,

$$\begin{cases} u_i = u_j = u_{\mathcal{A}_k} & i, j \in \mathcal{A}_k, \\ u_i \neq u_j & i \in \mathcal{A}_j, j \in \mathcal{A}_\ell, \end{cases} \quad (15)$$

from which we have $u_{\text{red},k} = u_{\mathcal{A}_k}$, $k = 1, 2, \dots, r$. Using this u_{red} and the estimated (or measured) aggregate states x_s as in (10), we apply the subspace identification method N4SID [22] to

identify A_{red} and B_{red} in (13). Traditional N4SID, however, suffers from a considerably large computational time complexity, which in our case would be $\mathcal{O}(r^3 N^3)$, where N is the number of measurement samples of x_s used for the identification. The computational time increases drastically with increasing number of samples and areas. Therefore, we instead propose the use of rSVD-based N4SID which has a computational complexity of $\mathcal{O}(r^3 N \log_2 N)$, and therefore reduces the ID time significantly. Traditional N4SID would involve the formation of Hankel matrix H_0 and H_1 as

$$H_0 = \begin{bmatrix} b_1 & b_2 & \dots & b_m \\ b_2 & b_3 & \dots & b_{m+1} \\ \vdots & \vdots & \ddots & \vdots \\ b_m & b_{m+1} & \dots & b_{2m-1} \end{bmatrix},$$

$$H_1 = \begin{bmatrix} b_2 & b_3 & \dots & b_{m+1} \\ b_3 & b_4 & \dots & b_{m+2} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m+1} & b_{m+2} & \dots & b_{2m} \end{bmatrix},$$

where b_i are the Hankel parameters and can be calculated from the measurement samples of x_s and u_{red} as in [22]. Since we are using all states as our output, the output matrix $C_{\text{red}} \in \mathbb{R}^{7r \times 7r}$ is the identity matrix. Once the Hankel matrices are computed, a rSVD is performed following [23] as

$$\text{rSVD}(H_0) = U \Sigma V^T = \underbrace{(U \Sigma^{1/2})}_{\mathcal{O}_r} \underbrace{(\Sigma^{1/2} V^T)}_{C_r}, \quad (16)$$

where \mathcal{O}_r and C_r are the extended observability and controllability matrices. H_1 can be further decomposed as

$$H_1 = \mathcal{O}_r A_r C_r, \quad (17)$$

where A_r is a similarity transformation of A_{red} as N4SID only preserves the input–output relationship and not the exact matrix structures. From (17) it follows that

$$\hat{A}_r = \mathcal{O}_r^{-1} H_1 C_r^{-1}. \quad (18)$$

Similarly, one can compute B_r and C_r as in [22]. Note that these matrices maybe in the transformed subspace. To get the estimates in the subspace we are interested in, the following transformation is performed based on the knowledge of the output matrix being the identity matrix:

$$\hat{A}_{\text{red}} = \hat{C}_r \hat{A}_r \hat{C}_r^{-1}, \quad \hat{B}_{\text{red}} = \hat{C}_r \hat{B}_r, \quad \hat{C}_{\text{red}} = I. \quad (19)$$

Results comparing the ID time between regular N4SID and randomized N4SID will be presented in Section 7. An important point to keep in mind is that the trade-off for using randomized N4SID is that the randomization brings in error in the

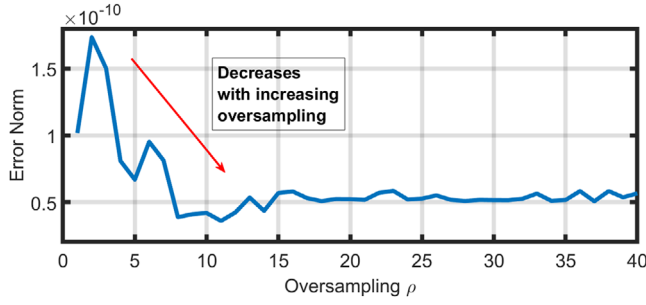


FIGURE 4 Error norm using r SVD for Kundur system

ALGORITHM 1 Reduced-order system identification

- 1: Disturbance (such as faults, etc.) occurs and is detected by ISO.
- 2: Low-amplitude, persistently-exciting Gaussian random noise is injected to the excitation input of each generator following (15).
- 3: Generator states in each area are transmitted in real time to its corresponding aggregator; the aggregator computes the aggregate state vector \mathbf{x}_j as in (9), and sends it to the ISO as shown in Figure 2 over a chosen interval of time.
- 4: ISO identifies $\hat{A}_r \in \mathbb{R}^{7r \times 7r}$, $\hat{B}_r \in \mathbb{R}^{7r \times r}$ and $\hat{C}_r \in \mathbb{R}^{7r \times 7r}$ using r SVD-N4SID.
- 5: ISO computes A_{red} and B_{red} following (19).

Outputs: $\hat{A}_{\text{red}}, \hat{B}_{\text{red}}$

computation of the SVD, which is given by

$$E\|H_0 - (U\Sigma V)_r\|_2 \leq \left[1 + \sqrt{\frac{r}{\rho-1}} + \frac{e\sqrt{r+\rho}}{\rho} \sqrt{N-r} \right]^{1/3} \sigma_{r+1}(\mathbf{H}), \quad (20)$$

where, $E\|H_0 - (U\Sigma V)_r\|_2$ is the expectation of the error norm over a randomly drawn Gaussian matrix, $(U\Sigma V)_r$ is the SVD using the randomized algorithm, $\sigma_{r+1}(H_0)$ denotes the largest discarded singular value of H_0 , e is Euler's number, and ρ is an oversampling parameter. We refer the reader to [23] for the details of these notations. In general, the error norm decreases by increasing the oversampling, as shown in Figure 4 for the 9-bus power system model.

The reduced-order system identification is summarized in Algorithm 1.

3.2 | Augmentation of PSS model

The next step is to augment the identified reduced-order model with the known PSS model. Ignoring the PSS states in the reduced-order model can easily lead to an unstable closed-loop system due to conflicting oscillatory modes [24], as will be seen shortly in Section 7.1. Structurally speaking, the PSS model can

be written as

$$\dot{\mathbf{x}}_{pss_i} = \underbrace{\begin{bmatrix} * & 0 & 0 \\ * & * & 0 \\ * & * & * \end{bmatrix}}_{A_{pss_i}} \begin{bmatrix} x_{pss_i}^{(1)} \\ x_{pss_i}^{(2)} \\ x_{pss_i}^{(3)} \end{bmatrix} + \underbrace{\begin{bmatrix} * \\ * \\ * \end{bmatrix}}_{B_{pss_i}} \Delta\omega_i, \quad (21)$$

where $*$ denotes the known parameters (gains/time constants) of the PSS. It is important to note that the PSS model depends only on its own internal states and the generator speed. Assuming that the PSS states evolve over a relatively slower time-scale by which the fast modes of the generator speed are well damped out, we approximate the speed of the i th generator by the aggregate speed $\Delta\omega_{A_j}$ of its area A_j . In that case, the i th PSS model can be approximated as

$$\dot{\mathbf{x}}_{pss_i} \approx A_{pss_i} \mathbf{x}_{pss_i} + B_{pss_i} \Delta\omega_{A_j}. \quad (22)$$

We assume A_{pss_i} and B_{pss_i} to be known for all the n_p PSSs. Also, note that only the exciter state (x_E) depends on the PSS states. Its dynamics can be structurally expressed as

$$\dot{\mathbf{x}}_{E_i} = \underbrace{\begin{bmatrix} * & * & * \end{bmatrix}}_{E_{pss_i}} \begin{bmatrix} x_{pss_i}^{(1)} \\ x_{pss_i}^{(2)} \\ x_{pss_i}^{(3)} \end{bmatrix} + f_i(x_{G_i}, x_{E_i}, u_i), \quad (23)$$

where $f_i(x_{G_i}, x_{E_i})$, following the model shown in [8], is a linear function of the exciter state and the corresponding generator states. For area j , the aggregate exciter dynamics ($\mathbf{x}_{A_j}^T$) can, therefore, analytically be written as

$$\dot{\mathbf{x}}_{A_j}^T = \frac{1}{n_{A_j}} \sum_{i \in A_j} E_{pss_i} \mathbf{x}_{pss_i} + f_{A_j}(x_{A_j}, u_{A_j}). \quad (24)$$

The contribution of $f_{A_j}(x_{A_j}, u_{A_j})$ is already accounted for in A_{red} and B_{red} , as identified in Section 3.1. To account for the first term in the RHS of (24) (i.e. the coupling of the exciter state with the PSS states), we define our augmented state vector $\mathbf{x}_{\text{aug}} \in \mathbb{R}^{(7r+3n_p)}$ as

$$\mathbf{x}_{\text{aug}} := [x_s^T \mid x_{pss_1}^T \dots x_{pss_{n_p}}^T]^T, \quad (25)$$

where x_s follows from (13) and n_p is the number of generators equipped with PSS. The augmented model is written as

$$\dot{\mathbf{x}}_{\text{aug}} = \hat{A}_{\text{aug}} \mathbf{x}_{\text{aug}} + \hat{B}_{\text{aug}} u_{\text{red}}, \quad (26)$$

where $\hat{A}_{\text{aug}} \in \mathbb{R}^{(7r+3n_p) \times (7r+3n_p)}$ can be constructed from the identified A_{red} and the known matrices A_{pss_i} , B_{pss_i} , and E_{pss_i} from (22) and (24). $\hat{B}_{\text{aug}} \in \mathbb{R}^{(7r+3n_p) \times r}$ is constructed as $\hat{B}_{\text{aug}} = [\hat{B}_{\text{red}}^T \ 0]^T$. For example, \hat{A}_{aug} for the 9-bus system, assuming equal generator inertias for simplicity of writing, can be written

where P_B follows from (12). Each element of $u(t)$ is broadcast to the corresponding generator at every sampling time instant t . This architecture is, thus, equivalent to a centralized controller.

Architecture 2: An alternate and perhaps more useful strategy can be that the ISO computes the full-dimensional feedback control matrix K as

$$K = P_B K_r V, \quad (33)$$

where V represents the averaging matrix such that $x_{avg}(t) = Vx(t)$, and does a one-time broadcast of the block-rows of K to the respective generators. Once each generator receives its control gain K_i (which is the i th block-row of K) it enables all-to-all communication with the other generators to implement its control input as $u_i(t) = -K_i x(t)$. This architecture pertains to a distributed implementation. The sparsity of K in this case will decide the structure of the wide-area communication network.

An important point in the implementation of the broadcast control is to account for the communication delays. For architecture 1, we assume that the ISO is moderately distant from all the generators, and therefore the communication delay for broadcasting the control input will be almost same for every generator. Referring to [29], in our simulations we will set the upper bound of this delay as $\tau_{max} = 300$ ms. For architecture 2, on the other hand, we consider two kinds of delays [29], namely, intra-area delay τ_s between generators belonging to the same coherent area whose typical value lies between 30 to 50 ms, and inter-area delay τ_b between generators across different coherent areas whose value usually lies between 200 to 300 ms.

5 | QUADRATIC CARLEMAN BILINEARIZATION FOR NON-LINEAR MODELS OF POWER SYSTEMS

5.1 | Motivation

So far we have identified small-signal models followed by linear control design. However, under certain operating conditions such as in the presence of non-linear non-conforming loads, or under high load stress, linear damping controllers may fail to guarantee closed-loop stability of the power system. For example, it was shown in [25] that when the system is stressed enough to be close to its voltage instability point, it becomes critical to consider higher-order models. The results in [26] showed that designing a linear controller for such stressed systems may potentially destabilize the system in case a contingency occurs. We cite a similar example for the IEEE 68-bus model. Figure 6 shows that when the load stress for this model is pushed to a limit where the power flow fails to converge, the reduced-order learning-based linear controller of Section 4 destabilizes the closed-loop system. Motivated by this fact, we, therefore, next propose to extend our control design to nonlinear dynamics. We use Carleman bilinearization [9, 10] for this purpose followed by an iterative technique to develop optimal quadratic controllers for bilinear approximation of a power system model.

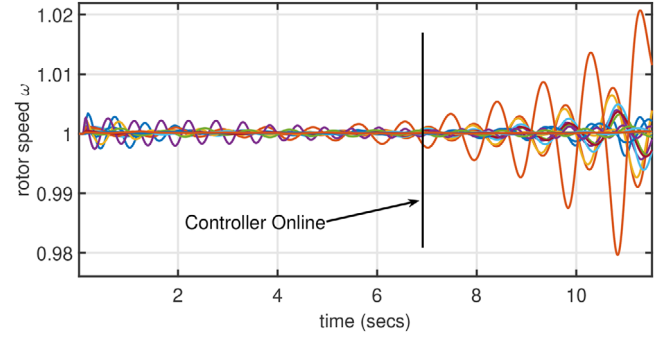


FIGURE 6 Linear controller fails for 68-bus system under high load stress

5.2 | About bilinearization technique

A non-linear system of the form

$$\dot{x} = f(x) + B_1 u, \quad x(0) = 0, \quad (34)$$

where $f(x) = \text{col}(f_1(x), f_2(x), \dots, f_n(x))$, $x \in \mathbb{R}^n$, can be approximated up to the second-order terms in the Taylor series expansion of the non-linear function as

$$\dot{x} = A_{11}x + A_{12}(x \otimes x) + B_1 u. \quad (35)$$

Here, n denotes the order of any general non-linear system, and is not related to the number of generators as used in Section 2. The state $x = x - x_0$ is the deviation of the actual state x from an equilibrium x_0 , $A_{11} \in \mathbb{R}^{n \times n}$ is the linear state matrix, and $A_{12} \in \mathbb{R}^{n \times k}$ is the quadratic state matrix, $k = n(n+1)/2$, and \otimes denotes the Kronecker product. Analytically, we have $A_{11}(i, j) = \frac{\partial f_i}{\partial x_j}$ and

$$A_{12} = \begin{bmatrix} \frac{\partial^2 f_1}{\partial x_1^2} & \frac{\partial^2 f_1}{\partial x_1 \partial x_2} & \dots & \frac{\partial^2 f_1}{\partial x_2^2} & \frac{\partial^2 f_1}{\partial x_2 \partial x_3} & \dots & \frac{\partial^2 f_1}{\partial x_n^2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial^2 f_n}{\partial x_1^2} & \frac{\partial^2 f_n}{\partial x_1 \partial x_2} & \dots & \frac{\partial^2 f_n}{\partial x_2^2} & \frac{\partial^2 f_n}{\partial x_2 \partial x_3} & \dots & \frac{\partial^2 f_n}{\partial x_n^2} \end{bmatrix}, \quad (36)$$

evaluated at the operating point x_0 . Carleman linearization of (35) yields the bilinear model

$$\dot{x}^\otimes = A^\otimes x^\otimes + B^\otimes x^\otimes u(t) + [B_1^T \ 0]^T u(t), \quad (37)$$

where $x^\otimes = [x^T \ (x \otimes x)^T]^T$ and

$$A^\otimes = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{21} \end{bmatrix}, B^\otimes = \begin{bmatrix} 0 & 0 \\ B_{21} & 0 \end{bmatrix}, \quad (38)$$

$$A_{21} = A_{12} \otimes I + I \otimes A_{12}, \quad B_{21} = B_1 \otimes I + I \otimes B_1. \quad (39)$$

We write (37) compactly as

$$\begin{bmatrix} \dot{x} \\ \dot{x}^2 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{21} \end{bmatrix} \begin{bmatrix} x \\ x^2 \end{bmatrix} + \underbrace{\begin{bmatrix} B_1 \\ B_{21}x \end{bmatrix}}_{B_c} u(t), \quad (40)$$

where, with a slight abuse of notation, x^2 denotes $x \otimes x$.

5.3 | Bilinear LQR control

The LQR objective function for the bilinear model (40) is given by

$$J = \frac{1}{2} \int_0^\infty ((x^\otimes)^T Q_2 x^\otimes + u^T R u) dt, \quad (41)$$

where Q_2 is defined as

$$Q_2 = \begin{bmatrix} Q_1 & 0 \\ 0 & 0 \end{bmatrix}, \quad (42)$$

and $R > 0$. The goal is to design $u = -K_{11}x - K_{12}x^2$ to minimize the objective function (41) for the closed-loop system

$$\dot{x}^\otimes = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{21} \end{bmatrix} \begin{bmatrix} x \\ x^2 \end{bmatrix} - \begin{bmatrix} B_1 K_{11} & B_1 K_{12} \\ 0 & K_{22} \end{bmatrix} \begin{bmatrix} x \\ x^2 \end{bmatrix}. \quad (43)$$

Since B_c is a function of the state x , we use an iterative method to converge to an optimal controller (modified from [27]). First, at iteration 0, the optimal input $u^{(0)} = -K_{11}^{(0)}x$ is computed for the linear system (A_{11}, B_1) using LQR weights (Q_1, R) . Once we have $K_{11}^{(0)}, K_{22}^{(1)}$ is computed as [27]

$$K_{22}^{(1)} = B_{21} \otimes K_{11}^{(0)}. \quad (44)$$

Next, we write the closed-loop system as

$$\dot{x}^\otimes = \underbrace{\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{21} - K_{22}^{(1)} \end{bmatrix}}_{A_d^{(1)}} x^\otimes + \underbrace{\begin{bmatrix} B_1 \\ 0 \end{bmatrix}}_{B_d} u(t). \quad (45)$$

The new control input $u^{(1)}$ is then designed as the LQR solution of the system $(A_d^{(1)}, B_d)$ with weights (Q_2, R) . In general, for any iteration index k , $A_d^{(k)}$ is calculated for $K_{22}^{(k)}$ given by

$$K_{22}^{(k)} = B_{21} \otimes K_{11}^{(k-1)}. \quad (46)$$

Following this, the new input $u^{(k)}$ is computed for $(A_d^{(k)}, B_d)$ till K_{11} converges, that is,

$$\|K_{11}^{(k)} - K_{11}^{(k-1)}\| < \gamma, \quad (47)$$

where $0 < \gamma \ll 1$ is a chosen threshold.

6 | IDENTIFICATION OF REDUCED-ORDER MODEL WITH CARLEMAN LINEARIZATION

6.1 | Deriving bilinear slow subsystem

Applying the transformation $T = \text{col}(G, F)$ as in Section 2.2 (for exact construction of G and F please see [6]), the fast and slow dynamics for (35) can be written as

$$\begin{bmatrix} \dot{x}_f \\ \dot{x}_s \end{bmatrix} = \begin{bmatrix} G \\ F \end{bmatrix} (A_{11}x + A_{12}(x \otimes x) + B_1 u), \quad (48)$$

from where we obtain

$$\begin{aligned} \dot{x}_f &= G A_1 T^{-1} \begin{bmatrix} x_f \\ x_s \end{bmatrix} \\ &\quad + G A_2 \left(T^{-1} \begin{bmatrix} x_f \\ x_s \end{bmatrix} \otimes T^{-1} \begin{bmatrix} x_f \\ x_s \end{bmatrix} \right) + G B_1 u, \\ \dot{x}_s &= E A_1 T^{-1} \begin{bmatrix} x_f \\ x_s \end{bmatrix} \\ &\quad + E A_2 \left(T^{-1} \begin{bmatrix} x_f \\ x_s \end{bmatrix} \otimes T^{-1} \begin{bmatrix} x_f \\ x_s \end{bmatrix} \right) + F B_1 u. \end{aligned} \quad (49)$$

From properties of Kronecker product, we can write

$$\begin{aligned} T^{-1} \begin{bmatrix} x_f \\ x_s \end{bmatrix} \otimes T^{-1} \begin{bmatrix} x_f \\ x_s \end{bmatrix} \\ = (T^{-1} \otimes T^{-1}) \left(\begin{bmatrix} x_f \\ x_s \end{bmatrix} \otimes \begin{bmatrix} x_f \\ x_s \end{bmatrix} \right), \end{aligned} \quad (50)$$

which transforms (49) to

$$\begin{aligned} \epsilon \dot{x}_f &= \Gamma_{1,f} x_f + \Gamma_{1,s} x_s + \Gamma_{2,ff} (x_f \otimes x_f) + \Gamma_{2,fs} (x_f \otimes x_s) \\ &\quad + \Gamma_{2,sf} (x_s \otimes x_f) + \Gamma_{2,ss} (x_s \otimes x_s) + G B_1 u, \\ \dot{x}_s &= Y_{1,f} x_f + Y_{1,s} x_s + Y_{2,ff} (x_f \otimes x_f) + Y_{2,fs} (x_f \otimes x_s) \\ &\quad + Y_{2,sf} (x_s \otimes x_f) + Y_{2,ss} (x_s \otimes x_s) + F B_1 u, \end{aligned} \quad (51)$$

where $\epsilon \ll 1$ is the coherency constant, and $\Gamma_{(\cdot)}$ and $Y_{(\cdot)}$ are appropriately sized matrices constructed from G, F, A_1, A_2, T , and B_1 . Substituting $\epsilon = 0$, the quasi-steady state of x_f follows as

$$\begin{aligned} x_f &= -\Gamma_{1,f}^{-1} (\Gamma_{1,s} x_s + \Gamma_{2,ff} (x_f \otimes x_f) + \Gamma_{2,fs} (x_f \otimes x_s) \\ &\quad + \Gamma_{2,sf} (x_s \otimes x_f) + \Gamma_{2,ss} (x_s \otimes x_s) + G B_1 u). \end{aligned} \quad (52)$$

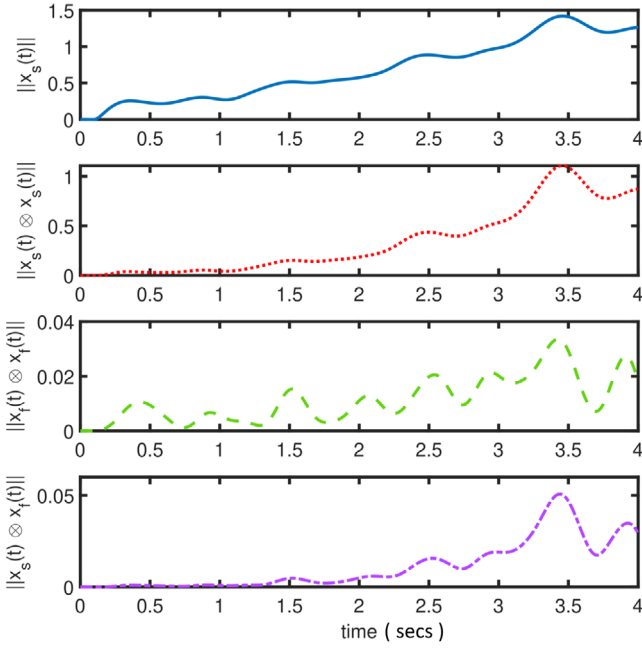


FIGURE 7 Norm comparison of $x_s, x_s \otimes x_s, x_f \otimes x_f$, and $x_s \otimes x_f$

Neglecting the third- and fourth-order terms, the slow subsystem can be written as

$$\dot{x}_s = \underbrace{\Psi_{1,s}x_s + B_\Psi u}_{\text{Linear part}} + \underbrace{\Psi_{2,ff}(x_f \otimes x_f) + \Psi_{2,fs}(x_f \otimes x_s) + \Psi_{2,sf}(x_s \otimes x_f) + \Psi_{2,ss}(x_s \otimes x_s)}_{\text{Second-order terms}} \quad (53)$$

where $\Psi_{(\cdot)}$ are appropriately sized matrices formed from Γ and Y . In practice, if the coherency is strong then the magnitude of the fast variable x_f is sufficiently small (after the fast transient is over), and so its Kronecker product with x_s or itself can be neglected compared to the most dominant term, which in this case is $(x_s \otimes x_s)$. This is numerically verified for the IEEE 68-bus model in Figure 7. This figure shows the Euclidean norm of $x_s(t) \in \mathbb{R}^r$, of $x_s(t) \otimes x_s(t)$, of $x_f(t) \otimes x_f(t)$, and of $x_s(t) \otimes x_f(t)$ using the measurements of the generator phase angles. Note that the Kronecker product is computed not over time, but over the variables; for example, the dimension of $x_s(t_i) \otimes x_s(t_i)$ at any given time instant t_i is $\mathbb{R}^{r(r+1)/2}$. The norm of this vector is plotted in Figure 7 for all time samples t_i till 4 s. The same applies for $x_f(t_i) \otimes x_f(t_i)$ and $x_s(t_i) \otimes x_f(t_i)$. It is clearly seen that $x_s(t) \otimes x_s(t)$ is the most dominant among the three Kronecker products. With this approximation and ignoring the terms involving the Kronecker product with the fast variable x_f , (53) is rewritten as

$$\dot{x}_s \approx \Psi_{1,s}x_s + B_\Psi u + \Psi_{2,ss}(x_s \otimes x_s). \quad (54)$$

We next describe how the matrices in the RHS of (54) can be identified using measurements of x_s .

6.2 | Carleman linearized system identification

For consistency of notations, we rewrite (54) as

$$\dot{x}_s = \mathcal{A}_{11,\text{red}}x_s + \mathcal{A}_{12,\text{red}}(x_s \otimes x_s) + B_{\text{red}}u_{\text{red}}, \quad (55)$$

where, $\mathcal{A}_{11,\text{red}} \in \mathbb{R}^{7r \times 7r}$ is the linear state matrix, $\mathcal{A}_{12,\text{red}} \in \mathbb{R}^{7r \times k}$, $k = 7r(7r+1)/2$. Furthermore, the power system model equations are non-linear functions of the phase angles δ_i (in terms of sine and cosine functions of the difference of phase angles $(\delta_i - \delta_j)$). The acceleration $\dot{\omega}(t)$ is also a quadratic function of the other states. However, as will be numerically verified shortly in the simulation section, the magnitudes of these quadratic states in $\dot{\omega}$ are negligible, as a result of which for all practical purposes it is sufficient to consider only the Kronecker products involving the aggregate rotor angles $\Delta\delta_{\mathcal{A}_i}$. Applying this approximation, we rewrite (55) as

$$\dot{x}_s = \mathcal{A}_{11,\text{red}}x_s + \mathcal{A}_{12,\text{red}}x_\delta^2 + B_{\text{red}}u_{\text{red}}, \quad (56)$$

where

$$x_\delta^2 := \text{col}(\Delta\delta_{\mathcal{A}_1}^2 (\Delta\delta_{\mathcal{A}_1} \Delta\delta_{\mathcal{A}_2}) \dots (\Delta\delta_{\mathcal{A}_1} \Delta\delta_{\mathcal{A}_r}) \Delta\delta_{\mathcal{A}_2}^2 (\Delta\delta_{\mathcal{A}_2} \Delta\delta_{\mathcal{A}_3}) \dots).$$

Treating x_δ^2 as a pseudo-input we write (56) as

$$\dot{x}_s(t) = \mathcal{A}_{11,\text{red}}x_s(t) + \underbrace{[\mathcal{A}_{12,\text{red}} \mid B_{\text{red}}]}_{B_\gamma} \underbrace{\begin{bmatrix} x_\delta^2(t) \\ u_{\text{red}}(t) \end{bmatrix}}_{u_\gamma}. \quad (57)$$

Using a similar approach as Algorithm 1, the ISO can employ measurements of x_s and x_δ^2 to estimate $\hat{\mathcal{A}}_{11,\text{red}}$ and \hat{B}_γ using the r SVD-based N4SID discussed in Section 3.1. $\hat{\mathcal{A}}_{12,\text{red}}$ and \hat{B}_{red} are extracted from \hat{B}_γ by partitioning it according to (57). Also, similar to (40), we can write the Carleman linearized equation for the slow subsystem as

$$\begin{bmatrix} \dot{x}_s \\ \dot{x}_\delta^2 \end{bmatrix} = \begin{bmatrix} \mathcal{A}_{11,\text{red}} & \mathcal{A}_{12,\text{red}} \\ 0 & \mathcal{A}_{21,\text{red}} \end{bmatrix} \begin{bmatrix} x_s \\ x_\delta^2 \end{bmatrix} + \begin{bmatrix} B_{\text{red}} \\ B_{2,\text{red}}x_s \end{bmatrix} u_{\text{red}}, \quad (58)$$

where the estimates of $\mathcal{A}_{21,\text{red}}$ and $B_{2,\text{red}}$ are calculated as

$$\begin{aligned} \hat{\mathcal{A}}_{21,\text{red}} &= \hat{\mathcal{A}}_{12,\text{red}} \otimes I + I \otimes \hat{\mathcal{A}}_{12,\text{red}} \\ \hat{B}_{2,\text{red}} &= \hat{B}_{\text{red}} \otimes I + I \otimes \hat{B}_{\text{red}}. \end{aligned} \quad (59)$$

The above steps are summarized in Algorithm 2.

6.3 | Augmentation of PSS model

After running Algorithm 2, we next augment the identified model with the known PSS model as in Section 3.2. The PSS model is augmented with the identified matrices $\hat{\mathcal{A}}_{11,\text{red}}$ and \hat{B}_{red} in a similar manner as it was augmented to $\hat{\mathcal{A}}_{\text{red}}$ and B_{red} . We

ALGORITHM 2 Reduced-order quadratic Carleman linearized system identification

- 1: Disturbance (such as faults, etc.) occurs and is detected by the ISO.
- 2: A low-amplitude, persistently exciting Gaussian random noise is injected to the excitation input of each generator following (15).
- 3: The aggregators compute their respective aggregate state vectors x_s and x_δ^2 , and sends them to the ISO, as shown in Figure 2.
- 4: Using x_s and x_δ^2 , ISO employs r SVD-based N4SID to estimate $\hat{A}_{11,\text{red}}, \hat{A}_{12,\text{red}}$ and \hat{B}_{red} .
- 5: $\hat{A}_{21,\text{red}}$ and $\hat{B}_{2,\text{red}}$ are computed using (59).

Outputs: $\hat{A}_{11,\text{red}}, \hat{A}_{12,\text{red}}, \hat{A}_{21,\text{red}}, \hat{B}_{\text{red}}, \hat{B}_{2,\text{red}}$

ALGORITHM 3 Bilinear quadratic controller

- 1: ISO identifies $\hat{A}_{\text{aug}}, \hat{A}_{12\text{r},\text{red}}, \hat{A}_{21,\text{red}}, \hat{B}_{\text{aug}}$ and $\hat{B}_{2,\text{red}}$ using Algorithm 2.
- 2: $K_{11}^{(0)}$ is designed for the linear system $\hat{A}_{\text{aug}}, \hat{B}_{\text{aug}}$ according to Section 5.3.
- 3: $K_{22}^{(1)}$ is designed using (46) applied to the first $r \times 7r$ partition of $K_{11}^{(0)}$, i.e., excluding the PSS states.
- 4: Compute the updated matrices $K_{11}^{(k)}$ and $K_{12}^{(k)}$ for the system (45).
- 5: Update $K_{22}^{(k+1)}$ according to (46).
- 6: If (47) is satisfied then stop, else go to step 4.

Outputs: K_{11}, K_{12}

refer to these augmented matrices as A_{aug} and B_{aug} . The matrix $\hat{A}_{12,\text{red}}$ is appropriately padded with extra zeroes (denote it by $\hat{A}_{12\text{r},\text{red}}$) to match the height of A_{aug} . The final augmented Carleman linearized model is written as

$$\begin{bmatrix} \dot{x}_s \\ \dot{x}_{\text{pss}} \\ \dot{x}_\delta^2 \end{bmatrix} = \begin{bmatrix} \hat{A}_{\text{aug}} & \hat{A}_{12\text{r},\text{red}} \\ 0 & \hat{A}_{21,\text{red}} \end{bmatrix} \begin{bmatrix} x_s \\ x_{\text{pss}} \\ x_\delta^2 \end{bmatrix} + \begin{bmatrix} \hat{B}_{\text{aug}} \\ \hat{B}_{2,\text{red}} x_s \end{bmatrix} u_{\text{red}}(t). \quad (60)$$

Note that (60) is not an LTI system as the input matrix is a function of $x_s(t)$ making it time variant. We next discuss how the iterative LQR control design presented in Section 5.3 can be used to account for the PSS augmentation.

6.4 | Reduced-order bilinear quadratic control design

We apply the iterative LQR design discussed in Section 5.3 to achieve our control objective for (60). The weight matrices Q_1 and R are chosen as in Section 4.1. Algorithm 3 summarizes the steps of this design.

The dimension of the reduced-order model is $7r + r(r + 1)/2$, and therefore the computational complexity of Algorithm 3 is $\mathcal{O}(r^6)$. For example, for the IEEE 68-bus 16-machine test system, Algorithm 3 is found to converge within two iterations and an average clock time of 0.11 s, making this algorithm feasible for online implementation. After the ISO calculates the linear and bilinear gains K_{11} and K_{22} , the reduced-order control

input is designed as

$$u_{\text{red}}(t) = -K_{11}x_s(t) - K_{12}x_\delta^2(t). \quad (61)$$

Finally, this reduced-order control input $u_{\text{red}} \in \mathbb{R}^r$ is broadcast to the excitation input of all generators $u \in \mathbb{R}^n$ using (32).

7 | SIMULATION RESULTS

We next test our algorithms on two IEEE prototype models. All simulations are carried out in a core i5 computer running Windows 10 operating system at 3.20 GHz, and with 16 GB RAM. The non-linear simulations for both are carried out using the Power System Toolbox (PST) [28] in MATLAB 2018a. The routine `s_simu` was used as the main non-linear simulation program for the complete simulation time. Suitable modifications were made to the exciter (actuator) routine `mexc_sig` to input the identification noise and state-feedback control as shown in Figure 8. The identification and broadcast control routines were executed in parallel and interfaced with `s_simu` to evaluate the performance of the controller.

Figure 8 shows the timing diagram for the whole process starting from injection of noise, to identification and control. T_0 is the time when a contingency occurs and is detected. Measurement data is captured from T_0 to T_1 , the system is identified from T_1 to T_2 . T_2 to T_3 is the time required for the LQR control design. T_4 is the settling time. In the figure, solid black arrows indicate flow of data, and the thin blue arrows indicate actuation signals (such as the exciter input).

7.1 | Two-Area, 4-machine Kundur system

We first illustrate the importance of augmenting the PSS model using the Kundur 9-bus model. The closed-loop eigenvalues for this model with the broadcast controller (13) are shown in Figure 9. It can be seen that if PSS states are ignored then the closed loop is unstable. Augmenting the PSS states, on the other hand, ensures closed-loop stability as well as the desired damping of the inter-area oscillation modes. Figure 10 shows how the reduced-order controller stabilizes the unstable open-loop system when a fault occurs on the line connecting buses 3 and 101 in Figure 2. T_0 is the time at which the fault occurs, T_1 is the time till which the measurement samples of x_s and u are collected (as defined in (10)), T_2 is the time at which the identification of A_{red} and B_{red} is complete. The reduced-dimensional wide-area controller is applied at T_3 via broadcasting, and the system is stabilized and the oscillations are damped by $T_4 = 8.4$ s. The entire exercise, starting from identification to control, is done using only data without knowing any model parameter other than the PSS gains.

7.2 | Linear controller on IEEE 68-bus, 5-area system

The 68-bus model is divided into five known coherent areas as marked in Figure 5 with PMUs placed on all generator buses.

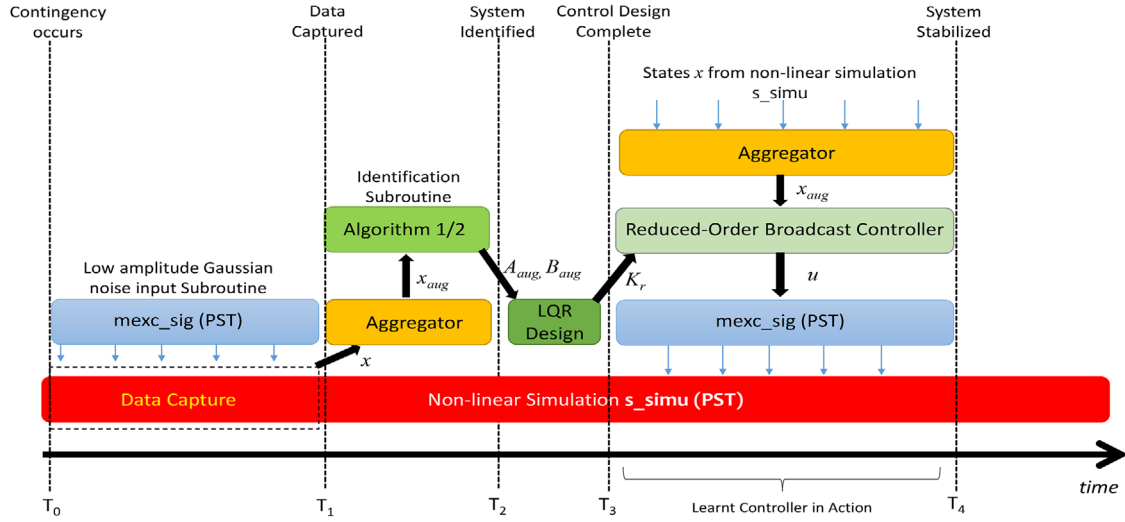


FIGURE 8 Execution flow of MATLAB-PST simulation

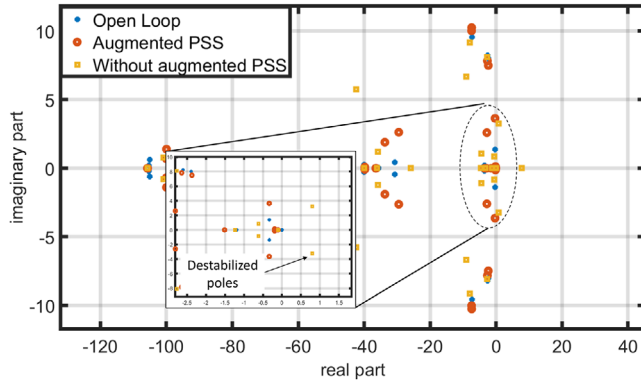


FIGURE 9 Unstable closed loop without augmenting PSS

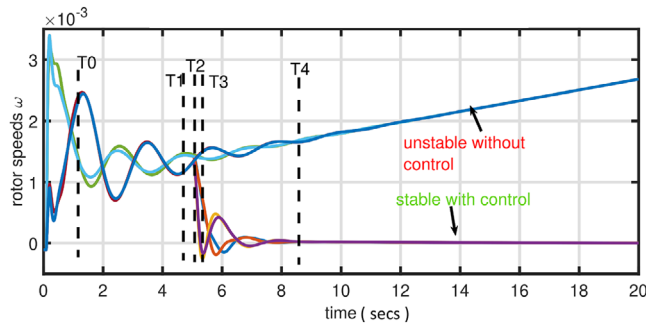


FIGURE 10 Non-linear fault simulation of 9-bus system

PSS is assumed to be present only at generators 1, 3, 8, 9, 13, and 14. The open-loop system is driven close to instability (Figure 11a) for a three-phase fault on the line connecting buses 55 and 56. Table 1 summarizes the result of the ID and control. The first three rows show that at least 4.5 s of data gathering time is needed to identify an accurate enough model that can produce a stable closed-loop system. Use of r -SVD drastically improves the ID time compared to standard N4SID, as seen from columns 2 through 4. Fidelity of the identification is quantified by a metric called final prediction error (FPE) following

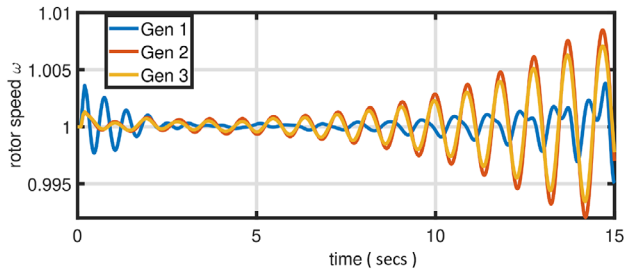
[7], which is defined as

$$\text{FPE} = \det \left(\frac{1}{N} \sum_{i=1}^N e(i) e(i)^T \right) \left(\frac{1 + \frac{d}{N}}{1 - \frac{d}{N}} \right), \quad (62)$$

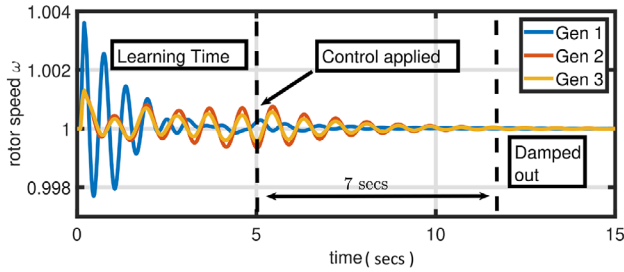
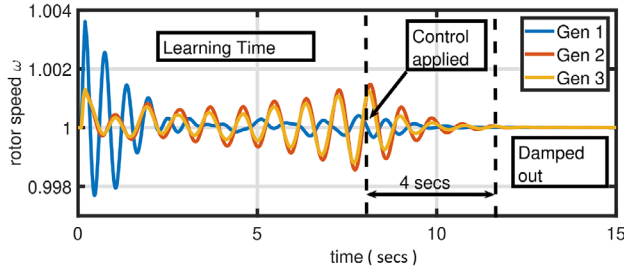
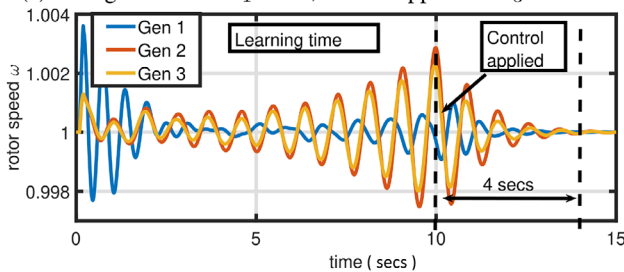
where $e(i) := x_{aug}(i) - \hat{x}_{aug}(i)$ is the estimation error vector for the i th sample, \hat{x}_{aug} being the value of the augmented state vector estimated from the identified model, $d = n_y + n_u$, n_y being the number of outputs and n_u the number of inputs. For this example, $n_y = 7r$ and $n_u = r$, and therefore $d = 8r$. A lower FPE indicates a more accurately estimated model. Increasing the initial amount of information for the ID by increasing the data gathering time improves the FPE, but after a certain threshold ($T_1 = 8$ s), as seen from the last two rows of the table, the initial information does not impact the FPE in any notable way. The closed-loop temporal responses of the first three generator frequencies are shown in Figure 11a–d. Figure 11b shows the case when data is gathered for 4.5 s. The control in this case is applied at approximately 5 s, and the oscillations are damped within 7 s starting from then. Figure 11c shows the case when data is gathered for 8 s, leading to a more accurate model and a more effective controller which damps the oscillations in only 4 s. The total time taken for damping (ID + control + damping time), however, for both Figure 11b,c, are almost same, which is 12 s starting from $t = 0$. Considering this fact, the former looks like a more feasible choice as it triggers the control faster than the latter, preventing the open-loop unstable oscillations to persist for longer. Figure 11d shows the case where adding more data samples or increasing the data gathering time before starting ID does not necessarily lead to a better controller. The right compartment of Table 1 shows the figures when the full-order model is identified. A least 8 s of initial data gathering time is needed, given the larger size of the parameter space that needs to be identified. The ID and control design takes 17.4 s in total. Our reduced-order controller, in contrast, takes only 8.5 s. This time-saving property becomes more beneficial as the

TABLE 1 Performance versus learning time

Data gathering time (T_1) (s)	Reduced order						Full Order				
	N4SID		rSVD-N4SID		Control		N4SID		rSVD-N4SID		Control LQR time ($T_3 - T_2$) (s)
	ID time ($T_2 - T_1$)(s)	FPE	ID Time ($T_2 - T_1$)(s)	FPE	LQR time ($T_3 - T_2$) (s)	Settling time (T_4) (s)	ID time ($T_2 - T_1$) (s)	FPE	ID time ($T_2 - T_1$) (s)	FPE	
2	2.5	10^{-4}	0.07	10^{-3}	0.09	Unstable	3.1	10^{-1}	0.07	10^{-1}	2.1 (unstable)
3	10.1	10^{-26}	0.13	10^{-26}	0.09	Unstable	23	10^{-7}	0.57	10^{-7}	2.1 (unstable)
4.5	93	10^{-56}	0.22	10^{-52}	0.09	12	398	10^{-34}	1.51	10^{-31}	2.1 (unstable)
8	236	10^{-59}	0.37	10^{-57}	0.09	12	1450	10^{-60}	7.3	10^{-51}	2.1
10	637	10^{-61}	0.53	10^{-57}	0.09	14	3360	10^{-62}	21	10^{-54}	2.1



(a) Open loop system is unstable

(b) Data gathered till $T_1 = 4s$, control applied at $T_3 = 5s$ (c) Data gathered till $T_1 = 8s$, control applied at $T_3 = 8.46s$ (d) Data gathered till $T_1 = 10s$, control applied at $T_3 = 10.6s$ **FIGURE 11** Non-linear fault simulation on 68-bus system

system size increases. One way to determine an appropriate value for T_1 in real time is to use neural networks (NN). For example, the ISO can train a NN for different operating conditions of the grid and use deep learning to correlate the corresponding values of T_1 for which the FPE remains below a chosen threshold. When the disturbance hits the power system, that specific operating point can then be fed to the NN to obtain the necessary value of T_1 . A second approach can be to repeat the ID multiple times for different values of T_1 , as in Table 1, and choose the model where the FPE starts plateauing.

We also make a comparison of the learning-based controller with the ideal LQR controller (i.e. the LQR controller if the system model was perfectly known) implemented just after the fault at T_0 in Figure 12. The figure shows the rotor speed of generator 2, indicating that the reduced-order controller performs sub-optimally compared to the ideal LQR controller. However, it stabilizes and damps the oscillations more effectively compared to when the LQR controller is designed based on a nominal system model.

7.2.1 | Effect of delays

Next, we study the effects of communication delays on the closed-loop response for the two architectures reported in Section 4.2. In the first architecture, that is, for centralized control, we assume the delay from the generators to the ISO, and ISO to the generators to be same. As shown in Figure 13a total delay of 200–700 ms is simulated until the system is destabilized. For a delay of 200 ms, the percentage degradation in the performance objective J is 1.55%, while increasing the delay to 500 ms increases the degradation to only 1.7%. The closed loop remains stable up to a delay of approximately 650 ms, which shows the resilience of our learning-based controller as in reality this delay will be much lesser.

For architecture 2, we consider two kinds of delays according to Section 4. For intra-area generators, we consider random delays varying between 30 and 50 ms. For inter-area generators, we consider delays between 200 and 300 ms. The closed-loop response of the generator speeds with and without delays is shown in Figure 14. The results testify the resilience of our

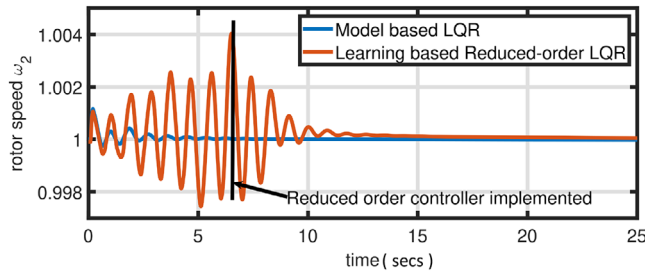


FIGURE 12 Performance of model-based versus learning-based controllers

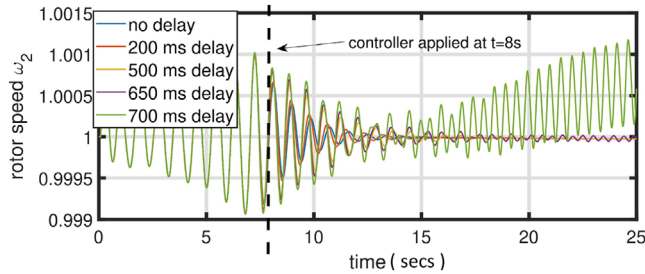


FIGURE 13 Effect of delay when input is broadcast through ISO

control strategy to the practical values of wide-area communication delays.

7.2.2 | Comparison with robust controller

We next show that the adaptive nature of our proposed controller is more advantageous in fighting against model uncertainty than a conventional model-based robust wide-area controller. For this, the nominal model of the 68-bus system is perturbed by stretching the loads at buses {18, 36, 39, 45, 52, 55, 67, 68} to their maximum limit. A robust H_∞ controller is designed assuming that the model can vary between the nominal model and this extremity. The controller is designed using the D-K iteration routine in the robust control toolbox in Matlab.

Figure 15 shows the simulation results comparing the H_∞ controller to the learning-based controller. Figure 15a shows that both controllers perform comparably well in damping the major oscillation lobes, but the learning-based controller

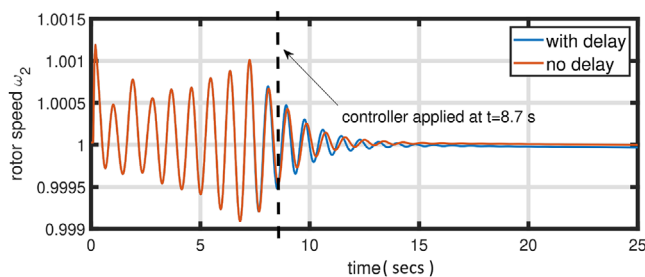


FIGURE 14 Effect of delay when control gain is broadcast

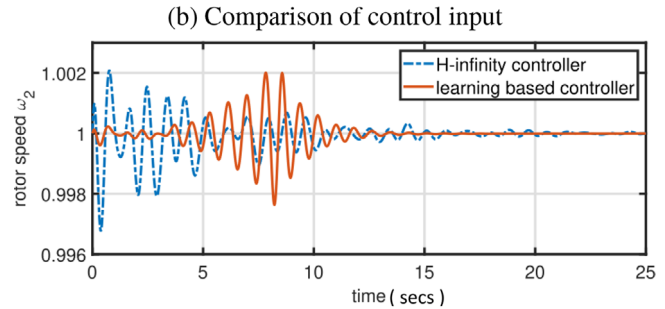
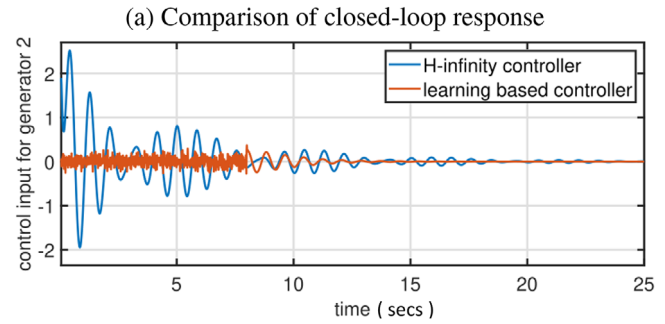


FIGURE 15 Comparison of H_∞ and learning-based controllers

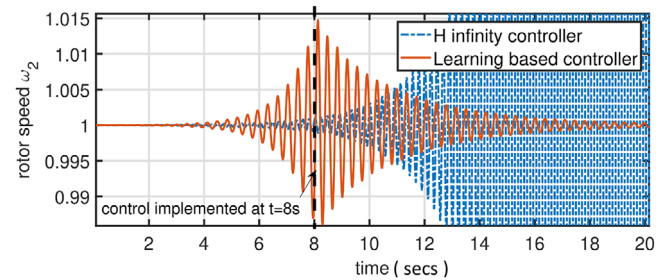
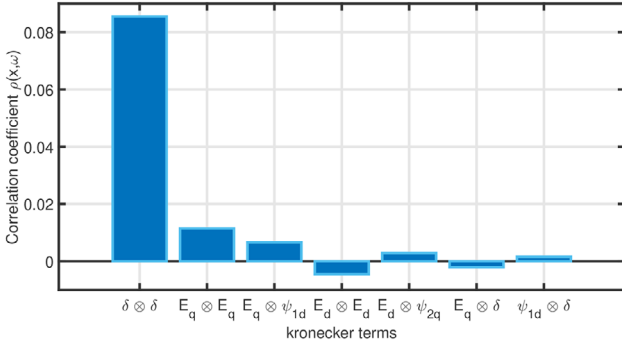
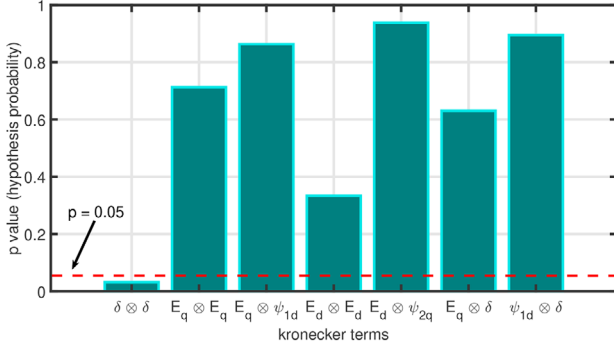


FIGURE 16 H_∞ versus learning-based control for major perturbations in the model

damps the residual oscillations faster. Figure 15b shows the time response of the control input to generator 2. The H_∞ controller comes into action immediately after the fault from $t = 0$, while the learning controller requires 8 s for system identification, and is actuated at $t = 8$ s. We can see that the H_∞ controller requires a significantly higher control effort as it is designed for the worst-case scenario while our controller learns the system as it is from the measured data and thereby requires a much smaller control effort. Figure 16 shows a case when the system is significantly deviated from the nominal model such that it no longer falls within the robustness margin of the H_∞ controller. We can see that in this case the H_∞ controller fails to stabilize the system as expected, whereas our learning-based controller is still able to damp the oscillations successfully.

7.3 | Results for bilinear quadratic control

We next test our reduced-order bilinear controller on the IEEE 68-bus, 16-machine test system. We add non-linear loads to

(a) Correlation coefficient ρ for ω w.r.t Kronecker terms

(b) p-value for correlation hypothesis testing

FIGURE 17 Correlation hypothesis testing

buses 25, 39, 42, 49, 62, and 64, and increase their consumption levels so that the load flow solution is close to its failing value. This is done using the `load_con` matrix in the 68-bus system data file of PST. For the second-order Kronecker terms, we only consider the interaction between ω and $\delta \otimes \delta$. From [8], we can see that $\dot{\omega}$ contains quadratic non-linearity with respect to the other states. However, we hypothesize that $\dot{\omega}$ is weakly correlated to the other cross-terms except for $\delta \otimes \delta$. This hypothesis is proved by performing a statistical analysis as shown in Figure 17. This includes calculating the correlation coefficient ρ of $\dot{\omega}$ with respect to the Kronecker terms, and finding the so-called the p -value that indicates the hypothesis probability of being uncorrelated. A p -value lesser than 0.05 indicates that the variables are correlated. Details of this hypothesis testing method and calculating them can be found in [30]. We can see from Figure 17 that the correlation coefficient of $\dot{\omega}$ with $\delta \otimes \delta$ given by $\rho(\delta \otimes \delta, \dot{\omega})$ is significantly higher compared to those with other relevant cross-terms. The hypothesis that $\dot{\omega}$ is only correlated to $\delta \otimes \delta$ is validated in Figure 17b, where the p -value is found to be only 0.018. The p -values for the other terms are higher, so we ignore their contribution while performing the Carleman linearized system identification. This reduces the dimension of $A_{12} \in \mathbb{R}^{7r \times k}$ from $k = \frac{7r(7r+1)}{2}$ to $k = \frac{r(r+1)}{2}$. For the 68-bus model, this means a reduction from a size of $k = 630$ to $k = 15$, which frees up significant computational time without noticeable loss of accuracy in the identification and the control.

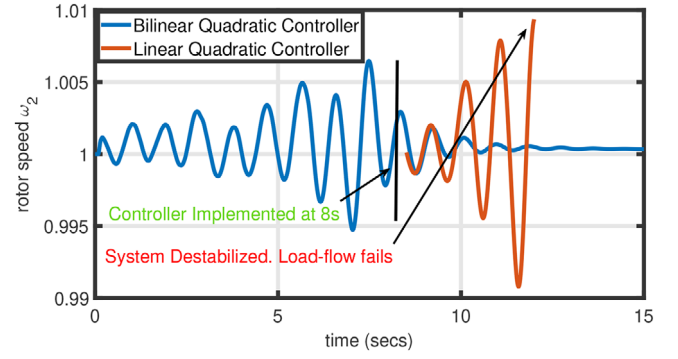


FIGURE 18 Performance of linear versus bilinear learning-based controllers

TABLE 2 Final prediction error and stability versus T_1

T_1	Algorithm 1		Algorithm 2	
	ID time (s)	FPE ₁	ID time (s)	FPE ₂
3s	0.15	10^{-5} (unstable)	0.19	10^{-3} (unstable)
5s	0.29	10^{-35} (unstable)	0.35	10^{-31} (unstable)
6s	0.31	10^{-39} (unstable)	0.41	10^{-49} (stable)
7s	0.38	10^{-38} (unstable)	0.52	10^{-54} (stable)

We simulate the same non-linear fault as shown in Figure 8. The bilinear quadratic wide-area controller designed in Section 6 is learned till 7.4 s and implemented at 7.5 s. From Figure 18, we can see that the bilinear controller successfully stabilizes and damps the system whereas the linear controller drives the system towards instability. Table 2 shows the FPEs using Algorithm 1 (FPE₁) for the linear system identification comparing it to Algorithm 2 (FPE₂) for the bilinear system. We can see that for the test system with the non-linear loads, Algorithm 1 fails to identify an accurate linear model compared to the bilinear system identification using Algorithm 2. As we saw in Section 6, a highly accurate model with low FPE is required to design a stabilizing quadratic controller. It becomes implausible to design a stabilizing controller using the identified linear model as the FPE cannot fit better than a value of 10^{-39} . For the bilinear system identification, it performs poorly for lesser number of samples due to underfitting. However, as T_1 increases, the FPE decreases significantly making it feasible for designing a stabilizing bilinear quadratic controller. From Figure 18 we can see that this controller successfully stabilizes and damps the oscillations within 15 s.

8 | CONCLUSION AND FUTURE WORK

In this paper, we presented a model-free learning-based broadcast control design for damping inter-area oscillations in a power systems without knowing its dynamic model. We used coherency theory and numerical methods such as randomized (r SVD) subspace identification (N4SID) for identifying a reduced-order power system model using real-time PMU

measurements. This reduced-order model is augmented with a known set of full-order PSS models to achieve closed-loop stability. We also extended the design to non-linear models where we identified a second-order Carleman linearized model, and designed a bilinear quadratic controller based on that. Results are illustrated by non-linear simulations of the IEEE 9-bus and 68-bus power system models. Robustness of the controller to wide-area communication delays and model uncertainties are also demonstrated through simulations. Our future work will be to evaluate how the proposed approximation-based identification and control extend to non-linear renewable-integrated grid models.

REFERENCES

- Chakraborty, A., Khargonekar, P.P.: Introduction to wide-area control of power systems. American Control Conference, Washington, DC, pp. 6758–6770, USA (2013)
- Yogarathinam, A., Chaudhuri, N.R.: Wide-area damping control using reduced copy under intermittent observation: A novel performance measure. *IEEE Trans. Power Syst.* 27(1), 434–442 (2019)
- Dörfler, F., et al.: Sparsity-promoting optimal wide-area control of power networks. *IEEE Trans. Power Syst.* 29(5), 2281–2291 (2014)
- Zima, M., et al.: Design aspects for wide-area monitoring and control systems. *Proc. IEEE* 93(5), 980–996 (2005)
- Kamwa, I., Gerin-Lajoie, L.: State-space system identification-toward MIMO models for modal analysis and optimization of bulk power systems. *IEEE Trans. Power Syst.* 15(1), 326–335 (2000)
- Chow, J.H.: *Power System Coherency and Model Reduction*. Springer, Berlin (2013)
- Ljung, L.: *System Identification: Theory for the User*. Prentice-Hall, Upper Saddle River (1999)
- Kundur, P.: *Power System Stability and Control*. McGraw-Hill, New York (1994)
- Sanchez-Gasca J.J., et al.: Inclusion of higher-order terms for small-signal (modal) analysis. *IEEE Trans. Power Syst.* 20(4), 1886–1904 (2005)
- Tian, T., et al.: An accurate third-order normal form approximation for power system nonlinear analysis. *IEEE Trans. Power Syst.* 33(2), 2128–2139 (2018)
- Rajkumar, V., et al.: A bilinear self-tuning controller for multimachine transient stability. *IEEE Trans. Power Syst.* 9(3), 1379–1384 (1994)
- Zhou, N., et al.: Robust RLS methods for online estimation of power system electro-mechanical modes. *IEEE Trans. Power Syst.* 22(3), 1240–1249 (2007)
- Dehghani, M., Nikraves, S.K.Y.: State-space model parameter identification in large-scale power systems. *IEEE Trans. Power Syst.* 23(3), 1449–1457 (2008)
- Schiffer, J., Aristidou, P., Ortega, R.: Online estimation of power system inertia using dynamic regressor extension and mixing. *IEEE Trans. Power Syst.* 34(6), 1–9 (2019)
- O'Shea, P.: A high resolution. Algorithm for power system disturbance monitoring. *IEEE Trans. Power Syst.* 17(3), 676–680 (2002)
- Wang, Q., et al.: Integrating model-driven and data-driven methods for power system frequency stability assessment and control. *IEEE Trans. Power Syst.* 34(6), 4557–4568 (2019)
- Mukherjee, S., Bai, H., Chakraborty, A.: Reduced-dimensional reinforcement learning control using singular perturbation approximations. *Automatica* 126, 109451, (2021)
- IEEE recommended practice for excitation system models for power system stability studies. *IEEE Std 421.5-2016* (Revision of IEEE Std 421.5-2005), 1–96 (2016)
- Jain, A., Chakraborty, A., Biyik, E.: Distributed wide-area control of power system oscillations under communication and actuation constraints. *Control Eng. Pract.* 74, 132–143 (2018)
- Singh, A.K., Pal, B.C.: Decentralized dynamic state estimation in power systems using unscented transformation. *IEEE Trans. Power Syst.* 29(2), 794–804 (2014)
- Lin, Z., et al.: Data-driven coherency identification for generators based on spectral clustering. *IEEE Trans. Ind. Inf.* 14(3), 1275–1285 (2018)
- Van Overschee, P., de Moor, B.: N4SID: Numerical algorithms for state space subspace system identification. *IFAC Proceedings Volumes* 26(2), 5–58 (1993)
- Liberty, E., et al.: Randomized algorithms for the low-rank approximation of matrices. *Proc. Natl. Acad. Sci.* 104, 20167–20172 (2007)
- Hung, S.T., Kokotovic, P.V., Winkelman, J.R.: Self tuning of conflicting oscillatory modes: A case study. *IEEE Trans. Autom. Control* 34(2), 250–253 (1989)
- Pariz, N., Shانهchi, H.M., Vaahedi, E.: Explaining and validating stressed power systems behavior using modal series. *IEEE Trans. Power Syst.* 18(2), 778–785 (2003)
- Lomei, H., et al.: A new approach to reduce the nonlinear characteristics of a stressed power system by using the normal form technique in the control design of the excitation system. *IEEE Trans. Ind. Appl.* 53(1), 492–500 (2017)
- Aganovic, Z., Gajic, Z.: The successive approximation procedure for finite-time optimal control of bilinear systems. *IEEE Trans. Autom. Control* 39(9), 1932–1935 (1994)
- Chow, J.H., Cheung, K.W.: A toolbox for power system dynamics and control engineering education and research. *IEEE Trans. Power Syst.* 7(4), 1559–1564 (1992)
- Dibaji, S.M., et al.: Delay-aware control designs of wide-area power networks. *IFAC-PapersOnLine* 50(1), 79–84 (2017)
- Casella G, G., Berger, R.L.: *Statistical Inference*. Duxbury Press, New York (2001)

How to cite this article: Kar J, Chakraborty A. Scalable design methods for online data-driven wide-area control of power systems. *IET Gener. Transm. Distrib.* 2021;15:2085–2100.
<https://doi.org/10.1049/gtd2.12159>