

Multigrid for model reduction of power grid networks

B. Lee 

Department of Mathematics, Southern Methodist University, Dallas, Texas,

Correspondence

B. Lee, Department of Mathematics,
Southern Methodist University, Dallas,
TX 75275.
Email: barryl@smu.edu

AMS Classification: 65M55; 65M70;
65N55; 65R20; 65Z05

Summary

This paper presents a method for determining the relevant buses for reduced models of power grid networks described by systems of differential-algebraic equations and for constructing the coarse-grain dynamical power grid systems. To determine these buses, time integration of differential equations is not needed, but rather, a stationary system is analyzed. However, unlike stationary-system approaches that determine only coarse generator buses by approximating the coherency of the generators, the proposed method analyzes the graph Laplacian associated with the admittance matrix. The buses for the reduced model are chosen to ensure that the graph Laplacian of the reduced model is an accurate approximation to the graph Laplacian of the full system. Both load and generator buses can be selected by this procedure since the Laplacian is defined on all the buses. The basis of this proposed approach lies in the close relationship between the synchrony of the system and the spectral properties of this Laplacian, that is, conditions on the spectrum of this Laplacian that almost surely guarantee the synchrony of the system. Thus, assuming that the full system is in synchrony, our strategy is to coarsen the full-system Laplacian such that the coarse Laplacian possesses good approximation to these spectral conditions. Accurate approximation to these conditions then can better lead to synchronous reduced models. The coarsened Laplacian is defined on coarse degrees of freedom (DOFs), which are associated with the relevant buses to include in the reduced model. To realize this coarse DOF selection, we use multigrid coarsening techniques based on compatible relaxation. Multigrid is the natural choice since it has been extensively used to coarsen Laplacians arising from discretizations of elliptic partial differential equations and is actively being extended to graph Laplacians. With the selection of the buses for the reduced model, the reduced model is completed by constructing the coarse admittance matrix values and other physical parameters using standard power grid techniques or by using the intergrid operators constructed in the coarse DOFs selection process. Unfortunately, the selection of the coarse buses and the coarsening of the admittance matrix and physical parameters are not sufficient by themselves to produce a stable reduced system. To achieve a stable system, system structures of the fine-grain model must be preserved in the reduced model. We analyze this to develop a multigrid

methodology for constructing stable reduced models of power grid systems. Numerical examples are presented to validate this methodology.

KEYWORDS

coarse-grid coarsening, compatible relaxation, differential-algebraic systems, graph Laplacians, multigrid, power grid, reduced models, synchrony

1 | INTRODUCTION

The emergent power grid is fundamentally changing the structure of the power grid network. Because of the incorporation of new technologies and new market policies, analysis of the grid will require new techniques and tools. These new technologies and policies can arise from a shift to renewable energy sources and from deregulation of the electric markets. Implementation of these in turn will affect the dynamics of the grid and, thus, affect the stability of the system. Fortunately, the incorporation of new technologies and policies can be included in the mathematical formulations of the grid. For example, modeling renewable energy leads to stochastic models, and modeling deregulation leads to decentralized or multicomponent systems. In this paper, we consider a computational method to handle one aspect of power grid analysis of the emerging grid. This is the construction of coarse-grain reduced models for large parts of the grid. This aspect can arise from the need to conduct quick real-time analysis of the grid, in order to take advantage of new system devices. It can also arise from decentralization of the full network where different utility companies manage sections of the network but with limited data on the state of the full grid. Constrained by limited data, it is argued that reduced models can be used for sections of the network that the companies do not manage.

Model reduction is not new in the power grid community. Probably, the most commonly used method is to approximate the coherency of the generator dynamics when selected disturbances to the system are introduced (e.g., the works of Chow et al.,¹ Kundur,² and Machowski et al.³). To obtain these approximations, the system of differential-algebraic equations is integrated for a sufficiently long time for each disturbance scenario in order to determine the dynamical response of the system. Time integration makes this method very expensive and it has been observed that different coherent groups/patterns can be produced for different disturbance scenarios. Reduced model techniques have also been developed from stationary systems. These methods are often based on aggregation of nodes and may have electric circuit analogies. Two of these methods are Dimo's aggregation, which requires introducing fictitious branches possibly with negative admittance weights (see the work of Dimo⁴; this is summarized in the work of Machowski et al.³) and Zhukov's aggregation, which requires that the voltage and currents at the retained nodes be unchanged and that the injected power at each aggregated node equals the sum of the injected power at the nodes making up the aggregate (see the work of Zhukov⁵; this is summarized in the work of Machowski et al.³). One stationary system approach that is rather different and of interest is Machowski's work.³ This method approximates generator coherency using measures on coefficients that are closely related to the graph Laplacian. We will give a brief description of this method later. Finally, in recent years, there has been some development in SVD/eigenvalue or proper orthogonal decomposition techniques for model reduction of power systems (e.g., the work of Barocio et al.⁶). These approaches can be viewed as projection methods that approximate the full system using the "dominant" eigenvectors of a sample covariance matrix from time snapshots of the system dynamics. A disadvantage of these approaches is that they do not determine the most relevant buses of the system, which has practical relevance in power grid analysis. A related disadvantage is that they do not address many of the fundamental challenges in developing good coarse-grain models. One goal of this paper is to address some of these challenges.

In this paper, we develop a fast computational technique that also avoids time integration. The method is based on some recent mathematical development on the synchrony of the grid.^{7,8} Instead of determining the groups of coherent generators, we determine the buses that lead to good approximation of the synchrony of the full grid. To be precise, we determine the buses, load, and generator buses that are most relevant for synchrony. The key behind this strategy is that synchrony can be determined from the graph Laplacian associated with the admittance matrix. The selection of the buses is then based on ensuring that the graph Laplacian of the reduced model is an accurate approximation to the graph Laplacian of the full system, which in turn is based on accurately approximating spectral properties of the full-system Laplacian. In particular, assuming that the full system is in synchrony, its graph Laplacian will satisfy certain spectral properties. Construction of the reduced model will be governed by ensuring that Laplacian of the reduced model accurately satisfies these spectral properties. Now, to achieve this, it is important that both the appropriate buses are retained and the values of the reduced Laplacian matrix are appropriately formed. We will achieve these using multigrid

coarsening techniques. However, it is unclear if the multigrid techniques designed for discretized scalar partial differential equations (PDEs) will be sufficient. The Laplacians for the power grid networks have more unstructured sparsity patterns. Thus, we will use some of the recently developed multigrid techniques for graph Laplacians.^{9–11}

We mention beforehand that the numerical experiments will involve power systems with network patterns given by discretizations of PDEs, although the graph weights are random. One reason for this is the ease in constructing stable large-scale power systems using these connection patterns. Another is that the goal of this paper is to demonstrate that multigrid techniques can be used to construct reduced models for these dynamical systems. This is not obvious because of the nonlinearity and nonelliptic characteristics of these systems. Furthermore, it is not obvious if a multigrid-coarsened Laplacian is all that is needed to produce a stable reduced model. Thus, the first step is to determine the necessary multigrid and system components that must be satisfied to ensure stable and accurate reduced models. Hence, it is sufficient to use these power networks for the numerical experiments. Employing the recently developed multigrid techniques for graph Laplacians will ensure that all the parts of the new model reduction methodology are in place to handle real power systems, which we are currently in discussion with power grid engineers for obtaining. Application of the new methods to these real systems will be described in a future paper.

Two of these necessary components in constructing stable and accurate reduced models, besides the selection of coarse buses and formation of the coarse Laplacian, are carefully coarsening the physical parameters (e.g., inertial and damping coefficients, and mechanical powers) and preserving system structures of the fine-grain model. We will see that the physical parameters can be coarsened using the intergrid operators produced in the coarsening procedure. We will also see that coarse dynamical system should not introduce variable coupling other than through the coarsened Laplacian.

This paper progresses as follows. In Section 2, the classical power grid problem is reviewed and a fully differential model to replace the system of differential-algebraic equations is described. We will assume that the classical model is used for the whole grid, although extensions to higher-order models will be explored in the future. In Section 3, synchrony and the graph Laplacian synchrony conditions will be presented. In Section 4, multigrid methods for scalar PDEs are reviewed, particularly the strength of connection between nodes in PDE-based Laplacians and compatible relaxation to determine a generalized strength for general graph Laplacians. Furthermore, the formulas for the intergrid operators are presented since these operators will be needed to form the coarse-grain Laplacian operator. In Section 5, we combine the synchrony condition, the multigrid techniques, and techniques for coarsening the power grid parameters to produce the reduced models. We then analyze these models to determine principles that must hold to produce an accurate and stable reduced model. Lastly, Section 6 will present some numerical results.

2 | MODEL REDUCTION OF THE POWER NETWORK

We assume that the full power grid network is composed of m generators and n buses of which $(n - m)$ are load buses and the remaining m buses are generator buses. Each bus is represented as a node on a graph, and the transmission lines connecting the buses are the edges of the graph. With E denoting the set of edges, we use $\mathcal{G}(n, E)$ to denote this graph. In addition, let N_g and N_l represent the set of nodes corresponding to the generator and load buses, respectively. The most general mathematical formulation that describes the dynamics of the power grid is the differential-algebraic system defined on the graph as follows:

$$\dot{\mathbf{x}} = F(\mathbf{x}, \mathbf{y}) \quad (1)$$

$$\mathbf{0} = G(\mathbf{x}, \mathbf{y}), \quad (2)$$

where $\mathbf{x}$ contains the differential variables describing the dynamics of the generators and $\mathbf{y}$ contains the algebraic variables governed by algebraic constraints, for example, the power-flow equations given by algebraic relations on the voltages and currents through Kirchhoff's law.

The most basic model for (1), or the lowest-order model in the sense that only the rudimentary details of the physics are included, is the swing equations or classical model. It has the form

$$H_i \ddot{\delta}_i + D_i \dot{\delta}_i + P_{e,i} - P_{m,i} = 0 \quad i \in N_g, \quad (3)$$

where, for generator i , unknown δ_i is the phase angle, H_i and D_i are respectively the inertial and damping parameters, and $P_{e,i}$ and $P_{m,i}$ are the electric power output and mechanical power. Introducing the speed/frequency $\omega_i = \dot{\delta}_i$, the above equation can be converted into the first-order system (1). Knowing the states δ_i and ω_i at any time, we can assess the state

of the power grid system at any time. Better assessment can be gained by using higher-order models that involve more dynamic variables for some of the generators. These models will not be considered in this paper.

Turning to the algebraic constraints, the power-flow equations are based on the nodal network equations

$$\mathbf{I} = \mathbf{YV}, \quad (4)$$

where $\mathbf{I}$ and $\mathbf{V}$ are the complex-valued currents and voltages at the n nodes and $\mathbf{Y}$ is the $(n \times n)$ complex-valued admittance matrix describing the conductances and susceptances of the network. The power injected at node i is given by

$$V_i I_i^* := P_i + \sqrt{-1} Q_i, \quad (5)$$

where the asterisk denotes the complex conjugate, and the real (P_i) and imaginary (Q_i) components of the power are called the active and reactive powers, respectively. Separating the admittance matrix into its real and imaginary components, that is

$$\mathbf{Y}_{ij} = \mathbf{G}_{ij} + \sqrt{-1} \mathbf{B}_{ij},$$

the power-flow equations can be expressed as

$$P_i = \hat{V}_i^2 \mathbf{G}_{ii} + \sum_{j=1, j \neq i}^n \hat{V}_i \hat{V}_j [\mathbf{B}_{ij} \sin(\delta_{ij}) + \mathbf{G}_{ij} \cos(\delta_{ij})] \quad (6)$$

$$Q_i = -\hat{V}_i^2 \mathbf{B}_{ii} + \sum_{j=1, j \neq i}^n \hat{V}_i \hat{V}_j [\mathbf{G}_{ij} \sin(\delta_{ij}) - \mathbf{B}_{ij} \cos(\delta_{ij})], \quad (7)$$

where $\hat{V}_j$ is now just the voltage magnitude at the j th bus, $\mathbf{G}_{ij}$ and $\mathbf{B}_{ij}$ are the conductance and susceptance of the transmission line connecting buses i and j given in the symmetric matrices $\mathbf{G}$ and $\mathbf{B}$, and $\delta_{ij} = (\delta_i - \delta_j)$. The four types of quantities in these equations are the active and reactive power, the voltage magnitude, and the phase angle. The buses are categorized according to the specified quantities at the bus: a PV bus (usually a generator bus) has its active power and voltage magnitude specified, a PQ bus has its active and reactive power specified, and a slack bus has its voltage magnitude and phase angle specified. The unspecified quantities are determined by solving the power-flow equations.

The algebraic power-flow equations can be approximated with a differential system. One way to do this is to view the algebraic variables as fast time-scale variables, which then allows us to express the algebraic equations as

$$\epsilon \dot{\mathbf{y}} = G(\mathbf{x}, \mathbf{y}),$$

for small $\epsilon > 0$. Another way that leads to a more explicit form is to take a frequency-dependent model for the power drawn at the load buses. Assuming that the conductances are negligible, the electric power is given by the formula

$$P_{e,i} = \sum_{j=1}^n \hat{V}_i \hat{V}_j \mathbf{B}_{ij} \sin(\delta_{ij}). \quad (8)$$

The electric power drawn by load bus i can be expressed in a similar form. Now, assuming that the drawn load is frequency dependent, for small frequency variations about a stable operating point that has drawn power $P_{d,i}^o$, we have

$$P_{d,i} = P_{d,i}^o + D_i \dot{\delta}_i, \quad D_i > 0. \quad (9)$$

Equating $P_{d,i}$ with its series expansion, we have

$$-\sum_{j=1}^n \hat{V}_i \hat{V}_j \mathbf{B}_{ij} \sin(\delta_{ij}) = P_{d,i}^o + D_i \dot{\delta}_i \quad (10)$$

(the negative sign arises because power is drawn out). Collecting (3), (8), and (10) gives the system

$$H_i \ddot{\delta}_i + D_i \dot{\delta}_i + \sum_{j=1}^n a_{ij} \sin(\delta_{ij}) = P_{m,i} \quad i \in N_g \quad (11)$$

$$D_i \dot{\delta}_i + \sum_{j=1}^n a_{ij} \sin(\delta_{ij}) = -P_{d,i}^o \quad i \in N_l, \quad (12)$$

where $a_{ij} = \hat{V}_i \hat{V}_j \mathbf{B}_{ij}$. This system is known as the structure-preserving model for power grid networks since it involves the complete generator and load structure in the model, as opposed to the network-reduced model that applies a Schur complement reduction to eliminate the load nodes from the system. (This reduction is known as the Kron reduction in the power grid and electrical engineering communities.) With the Schur complement reduction, the network-reduced model

involves only a system of differential equations with a denser Schur-complement admittance matrix. Not only does the denser matrix destroy the structure of the original network, but also the loads do not participate in the network-reduced model. One feature of the emerging power grid is the stochastic behavior of some loads. Thus, it may be beneficial to keep some of the loads in the network.

Even for the lowest-order model, the structure-preserving model is too complex for analyzing realistic power grid systems. For example, because of the deregulation of energy markets, only part of the grid may be accurately known and monitored. With this limitation, it is often argued by practitioners that a model reduction is adequate. For the deregulation scenario, the reduction involves partitioning the grid into an internal subsystem that the utility company manages, an external subsystem that it does not manage, and an interface subsystem. A higher-order model in fact can be used in the internal subsystem. However, for the external system, a reduced classical model can be sufficient given the limited data for this subsystem. A common methodology used in the power grid community for formulating the reduced model is first to eliminate the load buses using a Kron reduction and then retaining only selective external generators. The retained generators can be determined by grouping the generators into phase coherent sets, where phase angles δ_i and δ_j are said to be phase coherent if

$$|\delta_i(t) - \delta_j(t)|_{\text{mod } 1} \leq \epsilon \quad t \geq t_0 \quad (13)$$

for some ϵ and t_0 , where “mod 1” means that this relation holds within a constant determined by the initial conditions of δ_i and δ_j (i.e., the two trajectories are almost parallel). By selecting one representative generator from each coherent set, the set of external generators is effectively *coarsened*.

However, determining the coherent phases is nontrivial. As mentioned earlier, approaches for determining them generally require numerically integrating the dynamical system under different scenarios to observe the dynamical responses of the generators. One method that avoids time integration is described in the work of Machowski et al.³ This method determines the coherent phases by perturbing the angle at one of the boundary interface generators and approximating the change in the power output at the external generators. Simplifications of the system are assumed: the angles at the other nodes are constant, all the voltage magnitudes are constant, and the mutual inductance $\mathbf{G}_{ij}$ and damping D_i are negligible. The argument follows. Let δ_k be the phase angle of an interface generator that is perturbed by $\Delta\delta_k$. The active power at an external generator i changes to

$$P_i(\Delta\delta_k) = \hat{V}_i^2 \mathbf{G}_{ii} + \hat{V}_i \hat{V}_k \mathbf{B}_{ik} \sin(\delta_{ik} - \Delta\delta_k) + \sum_{j=1, j \neq i, j \neq k}^n \hat{V}_i \hat{V}_j \mathbf{B}_{ij} \sin(\delta_{ij}), \quad (14)$$

so that, using (6) and the simplifications, the power change at generator i is

$$\begin{aligned} \Delta P_i(\Delta\delta_k) &= \hat{V}_i \hat{V}_k \mathbf{B}_{ik} [\sin(\delta_{ik}) - \sin(\delta_{ik} - \Delta\delta_k)] \\ &\approx \hat{V}_i \hat{V}_k \mathbf{B}_{ik} \cos(\delta_{ik}) \Delta\delta_k \\ &= M_{ik} \Delta\delta_k, \end{aligned} \quad (15)$$

where $M_{ik} := \hat{V}_i \hat{V}_k \mathbf{B}_{ik} \cos(\delta_{ik})$. Since the damping is assumed negligible, the change in the active power leads to the following change in the acceleration of the rotor:

$$\Delta \ddot{\delta}_i \approx \frac{\Delta P_i(\Delta\delta_k)}{H_i} \approx \frac{M_{ik}}{H_i} \Delta\delta_k. \quad (16)$$

Notice, however, that this change in the acceleration generally leads to an instantaneous change in δ_i , which would mean that the angles at other nodes are not constant, which then implies that the active power $P_i(\Delta\delta_k)$ has a more complicated form than (14). Nevertheless, assuming that this method is sufficiently accurate, two external generators are coherent if the change in the rotor accelerations are similar: Generators i and j are coherent if

$$\frac{M_{ik}}{H_i} = \frac{M_{jk}}{H_j}. \quad (17)$$

For practical purposes, this coherency condition only has to approximately hold. However, creating the coherent groups of generators is still challenging. One way to partition the generators is to consider all generators i and j to be in the same group S if

$$\left[\max_{i \in S} \frac{M_{ik}}{H_i} - \min_{j \in S} \frac{M_{jk}}{H_j} \right] \leq \rho_k, \quad (18)$$

for a small $\rho_k > 0$. In the work of Machowski et al.,³ this measure is normalized as

$$\frac{\max_{i \in S} \frac{M_{ik}}{H_i} - \min_{j \in S} \frac{M_{jk}}{H_j}}{d_S} \leq \rho_k, \quad (19)$$

where

$$d_S = \min_{(i,j) \in T} d_{ij} \quad d_{ij} = \min_{i,j} \left\{ \frac{M_{ij}}{H_i}, \frac{M_{ji}}{H_j} \right\}$$

and T is a tree in S . A tree of edges is required because two generators can be strongly connected through a sequence of edges rather than being directly connected. Notice that a group is adaptively constructed in the sense that, as new generators are included in the group, measure (19) must be modified. This makes the algorithm nontrivial. Moreover, notice that the groups are strongly dependent on the interface generator k . Thus, selecting a different interface generator can dramatically change the partitioning. The work of Machowski et al.³ does not describe the sensitivity of this interface dependence and does not present a method that constructs one partition using all the interface generators. Moreover, even with a partitioning, a systematic procedure for constructing the admittance matrix for the reduced model can be difficult, particularly because a Kron reduction is performed before the partitioning, making physics-based constructions of this admittance matrix difficult. Nevertheless, because the M_{ik} 's depend only on the susceptances and initial voltage magnitudes, the algorithm does not require time-integration simulations.

To complete the construction of the reduced model, the inertia, damping, and mechanical power for the retained generators must be formed. In the power grid community, the inertial and mechanical power are constructed with the aggregation formulas

$$H_S = \sum_{i \in S} H_i \quad P_{m,S} = \sum_{i \in S} P_{m,i}. \quad (20)$$

2.1 | Comment

In this paper, we will assume that the voltage magnitudes are held constant in structure-preserving system (11)–(12). In practice, this is reasonable since the algebraic power-flow equations first can be solved for the voltage magnitudes and angles, with the resulting angles taken as the initial conditions for the structure-preserving differential equations and with the voltage magnitudes held constant in these equations. These differential equations then are integrated up to a fixed time, after which the power-flow equations are again solved for an updated set of voltage magnitudes and angles and the integration process repeated. Because of the required small timesteps, the integration procedure can be more computationally intensive. In this paper, the coarse-grain technique is specifically developed to improve this procedure.

3 | SYNCHRONY OF THE POWER GRID

The solution δ of (11)–(12) is said to be phase cohesive if there exists a $\gamma \in [0, \pi)$ such that

$$|\delta_i(t) - \delta_j(t)| \leq \gamma \quad \forall (i, j) \in E, t \geq 0.$$

The solution is also said to be frequency synchronized if all the frequencies $\dot{\delta}_i$ converge to a common frequency as $t \rightarrow \infty$. Following the work of Dörfler et al.,⁸ a power grid system is said to be synchronized if it is phase cohesive and frequency synchronized with $\gamma < \frac{\pi}{2}$. One can see that synchrony is closely related to phase coherency.

Moreover, synchronization is related to the stability of the power grid. Since the frequency of electricity must be the same over the whole network (e.g., in North America, the common house supply of electricity is 60 Hz), all the generators must oscillate at a synchronized frequency for the system to be stable. Now, it is clear from (11) and (12) that the bus angles couple only through the network term $\sum_{j=1}^n a_{ij} \sin(\delta_{ij})$. Without this term, each generator oscillates at its own natural frequency, implying that the generators are not in synchrony and so the system is unstable. The network coupling needs to be sufficiently strong for a stable system to arise. Quantification of this coupling strength is revealed through the graph Laplacian of the system, that is, the matrix $\mathbf{L}_f$ with elements

$$\mathbf{L}_{f,ij} = \begin{cases} -a_{ij} & i \neq j \\ \sum_{j \neq i} a_{ij} & i = j, \end{cases} \quad (21)$$

where we have used subscript f to distinguish the Laplacian of the full system from the Laplacian of the reduced system. Note that, because each bus is connected to only a small number of buses in a power grid, this graph Laplacian is sparse.

Although properties of $\mathbf{L}_f$ have been known to reveal the synchrony of similar but simpler systems of differential equations such as the Kuramoto model, which can be used to describe chemical and biological oscillators,^{12,13} only recently have extensions been derived for power grid networks.^{7,8} Let

$$P_i = \begin{cases} P_{m,i}, & i \in N_g \\ -P_{d,i}^o, & i \in N_l \end{cases}$$

and consider the system

$$\sum_{j=1}^n a_{ij} \sin \delta_{ij} = P_i \quad i \in N_g \cup N_l. \quad (22)$$

The solutions of this system are the equilibrium points of the structure-preserving system. Stable equilibrium points, which describe the stable solutions of the dynamical system, are possible if the system satisfies a synchrony condition. To describe this condition, let $\mathbf{L}_f^\dagger$ denote the pseudoinverse of $\mathbf{L}_f$. If

$$\|\mathbf{L}_f^\dagger \mathbf{p}\|_{E,\infty} \leq \sin \gamma, \quad (23)$$

where $\mathbf{p} = (P_i)_{i=1}^n$, $\|\mathbf{x}\|_{E,\infty} = \max_{(i,j) \in E} |x_i - x_j|$, and $\gamma < \frac{\pi}{2}$, then for almost all power grid networks, there exists a solution δ_{ij} of (22) that satisfies

$$|\delta_{ij}| \leq \gamma \quad \forall (i,j) \in E$$

(see the work of Dörfler et al.⁸). This is a stable solution. Thus, condition (23) is the synchrony condition we will use.

Additional information can be extracted from condition (23). Consider the eigenvalue/eigenvector decomposition of $\mathbf{L}_f$,

$$\mathbf{L}_f = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^t.$$

From (21), we see that $\mathbf{L}_f$ has zero row sum meaning that zero is an eigenvalue with the constant vector being the corresponding eigenvector. Furthermore, vector $\mathbf{L}_f^\dagger \mathbf{p}$ can be decomposed into an eigenvector expansion, with its dominant components being the projections onto the eigenvectors corresponding to the smaller nonzero eigenvalues. For a general $\mathbf{p}$, to accurately approximate $\mathbf{L}_f^\dagger \mathbf{p}$ in a reduced model, we require that the smaller eigenvalues/eigenvectors of the coarse-grain Laplacian be accurate approximations to the smaller eigenvalues/eigenvectors of $\mathbf{L}_f$. Denote such a coarse-grain Laplacian by $\mathbf{L}_c$. Because $\mathbf{L}_c$ approximates $\mathbf{L}_f$ for the dominant modes and because the dominant components of $\mathbf{L}_f^\dagger \mathbf{p}$ consist of these modes, the synchrony of the full system ideally should be captured by the reduced model. Hence, we have a strategy for choosing the buses for the reduced model; choose the buses that best ensures that the resulting $\mathbf{L}_c$ accurately approximates $\mathbf{L}_f$, particularly for the smaller eigenvalues/eigenvectors. Unlike other reduced model methods, we will select both generator and load buses because both are relevant for accurately representing the full power grid (e.g., the loads can affect the stability of the grid, especially in the modern grid). This coarse bus/node selection problem has been extensively studied in another context, that is, for developing fast solvers for discretizations of elliptic PDEs. For these Laplacians, multigrid techniques for selecting the coarse degrees of freedom (DOFs) and constructing the coarse-grain Laplacian lead to a good approximation to the fine-grain Laplacian. Moreover, these techniques allow the construction of not just one coarse Laplacian but also a hierarchy of coarser Laplacians, which can be useful for observing the scale at which difficulties can arise. For Laplacians coming from more general graphs, multigrid techniques are being actively investigated. We will employ some of these recently developed methods to determine the coarse buses.

4 | MULTIGRID

One of the most efficient method for solving systems of equations described by weighted Laplacians coming from scalar elliptic PDEs is multigrid. Multigrid achieves its efficiency by using a hierarchy of grids, where the computation on the coarser grids costs only a fraction of the effort in computing on the original grid. Moreover, by carefully designing the grid-level computation to handle only solution–error components on the scale of the level, the goal of the grid-level computation is to resolve only grid-scale features. The solution/error components are then handled levelwise.

The two complementary components for a robust and efficient multigrid solver are the smoothing and coarse-grid correction. The purpose of smoothing (e.g., a few sweeps of Gauss–Seidel or weighted Jacobi iteration) is to “smooth” out the error in an approximate solution. For Laplacians arising from structured grid discretizations of scalar elliptic PDE, smooth errors actually correspond to geometrically smooth errors. However, for general Laplacians, these errors

correspond to the algebraic near-null space of the Laplacian. As for the complementary coarse-grid correction, this process refers to the updating of the fine-grid approximation with the solution of the coarse-grid problem. The ideal situation is to have the coarse-grid correction resolve error components that are not well handled by the smoothing procedure. This is the essence of the so-called complementary smoothing/coarse-grid principle in constructing a good multigrid method.

One of the challenges in developing a multigrid method that satisfies the complementary principle is constructing an accurate coarse-grid problem. This, in turn, requires defining the coarse DOFs and the intergrid operators between the coarse and fine levels. The coarse DOFs must be chosen to permit formation of accurate coarse-grid problems, or viewed differently, weighted sums of the error at selected coarse DOFs must be able to accurately represent the error at fine DOFs that are not sufficiently handled by the smoother. Hence, the selection of the coarse DOFs is intimately connected to the effectiveness of the smoother, which is directly related to the Laplace operator. To select these coarse DOFs, a measure based on the operator is needed. For discretized PDEs on unstructured grids, this measure is given by the strength of connections between the DOFs and is determined from the coefficients of the Laplacian. For example, in a classical Ruge–Stuben coarsening approach¹⁴ for discretized PDEs, the following strength of connection measure is used: DOF x_j is said to strongly influence DOF x_i if

$$|\mathbf{L}_{f,ij}| \geq \theta \max_{k \neq i} |\mathbf{L}_{f,ik}| \quad \text{for } 0 < \theta \leq 1. \quad (24)$$

The coarse nodes are essentially selected to be the DOFs that strongly influence the most number of fine DOFs (i.e., DOFs that have not been selected to be coarse nodes) and form a maximal independent set. Since this selection is based on a measure explicitly involving the coefficients of $\mathbf{L}_f$, these coarse DOFs, together with other procedures in the Ruge–Stuben multigrid method, can accurately represent the eigenvectors with smaller eigenvalues when the Laplacian comes from a discretization of a scalar elliptic PDE.^{14,15}

For the graph Laplacians arising in power grids, other measures may be needed since the Ruge–Stuben coarsening might not be able to detect the appropriate nodes for these matrices. An alternative method is based on compatible relaxation,^{9–11,16} which allows the selection of the coarse DOFs using approximations to the algebraically smooth eigenvectors themselves. However, this selection procedure requires a more general “affinity” measure between the DOFs.

To describe this selection procedure, we use the fact that the coarse DOFs can be implicitly exposed through the performance of the smoother itself. From the complementary principle, algebraically smooth error components that are poorly handled by the smoother must first be *representable* on the coarser level. Thus, what remains after smoothing can themselves expose candidates for the coarse DOFs, particularly small sets of them that can accurately represent the poorly handled error. Specifically, consider solving a linear system $\mathbf{L}_f \mathbf{x} = \mathbf{b}$ using a smoother/relaxation scheme like Gauss–Seidel. After a few sweeps of the smoother, the dominant component in the error of the approximate solution consists of the desired “smooth” eigenvectors. Compatible relaxation exploits this fact. It applies the smoother to the homogeneous problem

$$\mathbf{L}_f \mathbf{x} = \mathbf{0} \quad (25)$$

with exact solution $\mathbf{x} = \mathbf{0}$ (choosing a particular $\mathbf{b}$ for the right-hand side would make $\mathbf{x}$ dependent on $\mathbf{b}$, but the goal is to approximate the eigenvectors with smaller eigenvalues). Starting with a random initial guess, after several smoothing sweeps, the resulting approximation, which is also the error, will expose these eigenvectors. By repeating this process for a collection of random initial guesses that are orthogonal to previously smoothed vectors, a good approximation to these eigenvectors can be obtained. (To be precise, this procedure is not exactly compatible relaxation. In compatible relaxation, relaxation is applied only to subsets of DOFs to determine the ones that lead to slow convergence in the relaxation scheme. By repeating this process, all the DOFs that lead to slow convergence are progressively determined. For power grid applications, this meticulous method may be an overkill. Thus, we only consider relaxation on all the DOFs and do not monitor the convergence of each or subsets of DOFs.)

From this collection of smoothed vectors, the coarse DOFs can be selected based on an appropriate measure. Two measures that are of particular interest for power grids are given in the works of Ron et al.⁹ and Livne et al.¹⁰: Let $\{\mathbf{x}^l\}_{l=1}^K$ be the set of smoothed vectors, let α and β denote two arbitrary components of the vector $\mathbf{x}^l$ and define the node inner product

$$(\mathbf{x}_\alpha, \mathbf{x}_\beta) = \sum_{l=1}^K \mathbf{x}_\alpha^l \mathbf{x}_\beta^l.$$

The measures are the distance and correlation-related quantities

$$d_{\alpha,\beta} = \frac{1}{\max_{l=1,\dots,K} |\mathbf{x}_\alpha^l - \mathbf{x}_\beta^l|} \quad (26)$$

$$c_{\alpha\beta} = \frac{|\langle \mathbf{x}_\alpha, \mathbf{x}_\beta \rangle|^2}{(\mathbf{x}_\alpha, \mathbf{x}_\alpha)(\mathbf{x}_\beta, \mathbf{x}_\beta)}. \quad (27)$$

Both have physical relevance in power systems (distance and correlation between buses can reveal coherency behavior). For a given threshold θ , nodes α and β are said to be of *close* affinity if the selected measure is greater than this threshold, for example, for $c_{\alpha\beta}$, we can choose $0.5 \leq \theta < 1$. Nodes that are close have a better chance of being “aggregated” with similar coarse nodes, which are referred to as *seeds*. As the strength of connection is used to determine the coarse nodes in the classical Ruge–Stuben method, the affinity measure is used in an analogous way. In particular, as in the Ruge–Stuben coarsening, a coarse node and all the fine nodes that it strongly influences are “fractionally” aggregated in the sense that these fine nodes may be aggregated with several coarse nodes; the affinity measure will be used to fractionally aggregate the fine nodes with the seeds. This will lead to an overlapping partitioning of the nodes in the sense that a nonseed node can be part of several aggregates.

As developed in the work of Ron et al.,⁹ to determine which nodes become the seeds, the affinity measure is used to construct the projected volume of the resulting aggregate if a node was selected as a seed. To describe this, for each node i , let $N(i)$ be its neighborhood (i.e., all $j \in N_g \cup N_l$ such that $ij \in E$) and let v_i be its volume, which is initially 1. With z denoting either measure, the projected volume v_i for node i is defined as

$$v_i = v_i + \sum_{ij \in E} v_j \frac{z_{ij}}{\sum_{jk \in E} z_{jk}}.$$

All nodes that have projected volumes greater than a factor of the average projected volume over the whole grid become seeds. This is the first pass for determining the set of coarse nodes, which we denote by C . The second pass can convert a nonseed node into a seed if its affinity measures to other nonseed nodes are relatively large or if its weighted graph connections in $\mathbf{L}_f$ with nonseed nodes are relatively large: that is, for some threshold Q and for a nonseed i , if

$$\frac{\sum_{j \in (C \cap N(i))} z_{ij}}{\sum_{j \in N(i)} z_{ij}} \leq Q \quad \text{or} \quad \frac{\sum_{j \in (C \cap N(i))} \mathbf{L}_{f,ij}}{\sum_{j \in N(i)} \mathbf{L}_{f,ij}} \leq Q, \quad (28)$$

then move i to C . Interpreting, the first inequality holds if the z_{ij} 's to its neighboring nonseed nodes, $N(i)/(C \cap N(i))$, are large compared to the affinity measures to its neighboring C nodes, $(C \cap N(i))$. The second inequality has a similar interpretation but in terms of the graph connections. Notice that the selected seeds are nodes themselves, and thus, they physically represent buses.

After this second pass, all nonseeds form the F nodes. For $i \in F$, i can be associated with several aggregates/seeds. The maximal number of seeds that each nonseed is associated with is referred to as the caliber of interpolation. Given caliber l , the seeds that i is associated with are determined by the maximal $z_{ij}, j \in C$, affinity to seed nodes, and maximal graph connections $\mathbf{L}_{f,ij}, j \in C$ with seeds.⁹ Finally, denoting the seeds that an F node i is associated with by $N^{C_l}(i)$, the caliber l interpolation operator $\mathcal{I}$ is computed as

$$\mathcal{I}_{ij} = \begin{cases} \mathbf{L}_{f,ij} / \sum_{k \in N^{C_l}(i)} \mathbf{L}_{f,ik}, & i \in F, j \in N^{C_l}(i) \\ 1, & j = i, i \in C \\ 0, & \text{elsewhere.} \end{cases} \quad (29)$$

With $\mathcal{I}$, the coarse-grain Laplacian is

$$\mathbf{L}_c = \mathcal{I}^t \mathbf{L}_f \mathcal{I}, \quad (30)$$

which is often referred to as the RAP (the triple matrix product of the restriction R , operator A , and interpolation P) or Galerkin coarsening of $\mathbf{L}_f$.

4.1 | Comments

- a. An immediate conclusion for power grid applications can be deduced from (27). Assume that some of the rows of graph Laplacian are strongly diagonally dominant. Then, relaxing (25) leads to quick convergence to zero for the DOFs corresponding to these rows. Now, assuming that the probability distribution of $\mathbf{x}_\alpha$ has mean zero, then (27) is an estimate of the correlation between nodes α and β . For power grid systems, if α and β correspond to generators,

then diagonal dominance means that these generators have weak correlation with other generators, that is, these generators swing independently. On the other hand, strong correlation between groups of α 's and β 's reflects strong coupling between these generators. Since the smoothed vectors are randomly selected, this sample correlation measure will give a good estimate of the physical coherent groups.

- b. Compared to other techniques for selecting the coarse generators, one of the advantages of the relaxation-based method is that it uses the diffusivity of information to determine the important generators (see the work of Coifman et al.¹⁷, which describes diffusion for feature extractions in large data sets). That is, since the smoothing sweeps reveal how information propagates through the graph, selecting the coarse generators using compatible relaxation reveals the diffusion of electricity through the graph. Using several random initial vectors, we can capture the statistical behavior of the electricity diffusion essentially with random power flows. This can lead to more accurate assessments of the power grid network than traditional techniques. For example, unlike a graph node-degree approach (e.g., the work of Wang et al.¹⁸) that selects the relevant buses to be the ones that have the most number of graph links, the physical conductance/susceptance coefficients of the admittance matrix together with the power flow behavior are used in the relaxation-based method.
- c. Assuming that the mean of the set $\{\mathbf{x}^l\}_{l=1}^K$ at each node α is zero, $c_{\alpha\beta}$ is the observed correlation of the phases at α and β . Hence, in computing this measure, we also get an important practical monitor on the sensitivity of the buses to each other.

5 | MULTIGRID FOR MODEL REDUCTION OF THE POWER GRID

The above multigrid procedure will be used to construct the reduced models. Although only a two-grid method is needed for model reduction, incorporating more levels can expose the spatial multiscale nature of the synchronization. However, this depends on the accuracy of the coarse graph Laplacians generated by the multilevel procedure and the construction of the coarse damping and inertial coefficients. It further heavily depends on preserving the structure of the first-order system. To see all of this, we first consider only a system of swing equations that may correspond to a Kron reduction of the structure-preserving system. Linearizing the nonlinear components in (11)–(12), we have the matrix system

$$\begin{pmatrix} \dot{\delta}_f \\ \dot{\omega}_f \end{pmatrix} = \begin{pmatrix} \mathbf{0} & \mathbf{I}_f \\ -\mathbf{H}_f^{-1}\mathbf{L}_f & -\mathbf{H}_f^{-1}\mathbf{D}_f \end{pmatrix} \begin{pmatrix} \delta_f \\ \omega_f \end{pmatrix} + \begin{pmatrix} \mathbf{0} \\ \mathbf{H}_f^{-1}\mathbf{p}_f \end{pmatrix}, \quad (31)$$

where subscript f is used to denote the nontrivial fine-grain components that have to be coarsened and $\mathbf{H}_f$ and $\mathbf{D}_f$ are diagonal matrices containing the inertial and damping coefficients, respectively. Since $\mathbf{H}_f$ and $\mathbf{D}_f$ are diagonal, the bus coupling occurs only from the $(2, 1)$ row-scaled Laplacian term of (31). This is relevant because synchrony in similar dynamical systems is known to occur if the coupling occurs only through the weighted graph Laplacian with these couplings being sufficiently strong.¹² If coupling arises through other terms in the differential equations or through a more complicated network coupling that is not reflective of a graph Laplacian, obtaining synchrony can be very difficult. These system structure facts will be used in constructing our coarse-grain models.

5.1 | Simplified system

For simplicity, assume that $\mathbf{H}_f = \mathbf{D}_f = \mathbf{I}_f$. The solution of (31) can be expressed in terms of the eigenvalues and eigenvectors of

$$\begin{pmatrix} \mathbf{0} & \mathbf{I}_f \\ -\mathbf{L}_f & -\mathbf{I}_f \end{pmatrix}. \quad (32)$$

Denoting the eigenvalue/eigenvector pairs of $\mathbf{L}_f$ by $\{\mu_i, \mathbf{e}_i\}$, the eigenvalues/vectors of (32) are

$$\lambda_{i,n+i} = \frac{-1 \pm \sqrt{1 - 4\mu_i}}{2} \quad \text{with corresponding eigenvectors} \quad \begin{pmatrix} \mathbf{e}_i \\ \lambda_{i,n+i}\mathbf{e}_i \end{pmatrix}.$$

The close relationship between the eigenvalues/eigenvectors of $\mathbf{L}_f$ and matrix (32) shows that the coarse-grain Laplacian must sufficiently approximate $\mathbf{L}_f$ for the reduced model to be effective. Let $\mathcal{I}$ denote an accurate interpolation operator constructed from $\mathbf{L}_f$ to interpolate the phases, and let $\mathcal{J}$ be the interpolation operator for the frequencies. Applying a RAP

coarsening procedure to (32) using the block diagonal interpolation operator

$$\begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{J} \end{pmatrix}$$

gives the coarse-grid operator

$$\begin{pmatrix} \mathbf{0} & \mathbf{I}^t \mathbf{J} \\ -\mathbf{J}^t \mathbf{L}_f \mathbf{I} & -\mathbf{J}^t \mathbf{J} \end{pmatrix}. \quad (33)$$

Since a good coarse-grid $\mathbf{L}_c$ is $\mathbf{I}^t \mathbf{L}_f \mathbf{I}$, we should take $\mathbf{J} = \mathbf{I}$. This is also reasonable since the phase and frequency should require the same “spatial” interpolation as the frequency is the *temporal* derivative of the phase. With this choice, one coarse-grain model that follows is

$$\begin{pmatrix} \dot{\delta}_c \\ \dot{\omega}_c \end{pmatrix} = \begin{pmatrix} \mathbf{0} & \mathbf{I}^t \mathbf{I} \\ -\mathbf{I}^t \mathbf{L}_f \mathbf{I} & -\mathbf{I}^t \mathbf{I} \end{pmatrix} \begin{pmatrix} \delta_c \\ \omega_c \end{pmatrix} + \begin{pmatrix} \mathbf{0} \\ \mathbf{I}^t \mathbf{p}_f \end{pmatrix}. \quad (34)$$

However, (34) does not reflect the system structure of (31). In particular, since generally $\mathbf{I}^t \mathbf{I} \neq \mathbf{I}_c$, the equations for $\dot{\delta}_c$ do not reflect the phase–frequency relation $\omega = \dot{\delta}$. Furthermore, a nondiagonal $\mathbf{I}^t \mathbf{I}$ implies that the frequencies and phases couple in more ways than through the Laplacian, which as mentioned earlier, can lead to more complicated synchronization or even instability in general. We thus introduce a coarse-grain model that preserves more of the structure of (31) but is not a RAP coarsening of operator (32). First, for a given vector $\mathbf{x}$, let $\text{diag}(\mathbf{x})$ be the diagonal matrix with the values of $\mathbf{x}$ along the diagonal, and vice-versa for a diagonal matrix $\mathbf{E}$, we represent it as a vector $\mathbf{e}$ containing the diagonal components. To keep the phase–frequency relation and to restrict variable coupling to occur only through the Laplacian, we consider the coarse-grain model

$$\begin{pmatrix} \dot{\delta}_c \\ \dot{\omega}_c \end{pmatrix} = \begin{pmatrix} \mathbf{0} & \mathbf{I}_c \\ -[\text{diag}(\mathbf{I}^t \mathbf{1}_f)]^{-1} \mathbf{I}^t \mathbf{L}_f \mathbf{I} & -[\text{diag}(\mathbf{I}^t \mathbf{1}_f)]^{-1} \text{diag}(\mathbf{I}^t \mathbf{1}_f) \end{pmatrix} \begin{pmatrix} \delta_c \\ \omega_c \end{pmatrix} + \begin{pmatrix} \mathbf{0} \\ [\text{diag}(\mathbf{I}^t \mathbf{1}_f)]^{-1} \mathbf{I}^t \mathbf{p}_f \end{pmatrix}, \quad (35)$$

where $\mathbf{1}_f$ is the vector representation of $\mathbf{I}_f$. To see the intuition behind the construction of this model, the (1, 2) block is defined to ensure the phase–frequency relation $\omega_i = \dot{\delta}_i$. For the second row, consider the second row of (31). First, since we need an accurate coarsening of $\mathbf{L}_f$, we take $\mathbf{I}^t \mathbf{L}_f \mathbf{I}$ since $\mathbf{I}$ is constructed from $\mathbf{L}_f$. To get a good coarsening of $\mathbf{H}_f = \mathbf{I}_f$ and $\mathbf{D}_f = \mathbf{I}_f$ while preserving their diagonal structure to prevent variable coupling, we recall that $\mathbf{H}_f$ and $\mathbf{D}_f$ are respectively the inertial and damping coefficients at the buses. A good way to coarsen the coefficients is to take a weighted average of them in the aggregates. This can be achieved using $\mathbf{I}^t$. Converting the diagonal matrices to $\mathbf{H}_f = \mathbf{D}_f = \mathbf{I}_f$ to the vector representation $\mathbf{1}_f$, the coarse inertial and damping coefficients for this simplified system is $\mathbf{I}^t \mathbf{1}_f$. Finally, taking the matrix representation of these vectors, we have the coarse diagonal matrix $\text{diag}(\mathbf{I}^t \mathbf{1}_f)$. Clearly this coarse-grain construction is applicable to more general diagonal $\mathbf{H}_f$ and $\mathbf{D}_f$, as we will see later (models (44) and (45)).

To illustrate the performance of these two models, we consider a system where the Laplacian is obtained from a bilinear finite-element discretization of the Laplace operator but modified so that the weights of the graph are random. We apply a six-level multigrid coarsening for constructing $\mathbf{L}_c$ using compatible relaxation with affinity measure (27) (measure (26) produces similar results) and caliber-4 interpolation. For the compatible relaxation, 20 weighted-Jacobi smoothing steps and 20 test vectors are used with $Q = 0.5$ in (28), the threshold in the second pass of the seed selection. This coarsening reduces the total number of nodes from 324 to roughly 30. The external force $\mathbf{p}_f$ is generated to have tenfold jumps at time intervals $0.5 \leq t \leq 0.6$ and $4.5 \leq t \leq 4.6$ to simulate faults. That is, we take $\mathbf{p}_f$ to be a spatial sinusoidal wave ($\sin \pi x \sin \pi y$) with tenfold jumps at the fault intervals $0.5 \leq t \leq 0.6$ and $4.5 \leq t \leq 4.6$. Figure 1 is a plot of the fine-grain model and coarse-grain model (34) for the coarse phases. Although the reduced model captures the location of the “faults” and the eventual synchrony of the system, details of the dynamics are not captured. Figure 2 shows the results for model (35), from which we see that the details of the dynamics are captured. Using fewer levels leading to more coarse-grain DOFs gives more accuracy.

These results demonstrate that the system structures should be preserved as much as possible, which a RAP coarsening to operator (32) generally does not achieve. One more coarse-grain model to illustrate the importance of structure preserving is

$$\begin{pmatrix} \dot{\delta}_c \\ \dot{\omega}_c \end{pmatrix} = \begin{pmatrix} \mathbf{0} & \mathbf{I}_c \\ -[\mathbf{I}^t \mathbf{I}]^{-1} \mathbf{I}^t \mathbf{L}_f \mathbf{I} & -\mathbf{I}_c \end{pmatrix} \begin{pmatrix} \delta_c \\ \omega_c \end{pmatrix} + \begin{pmatrix} \mathbf{0} \\ [\mathbf{I}^t \mathbf{I}]^{-1} \mathbf{I}^t \mathbf{p}_f \end{pmatrix}. \quad (36)$$

This is obtained by applying the RAP coarsening now to the whole linearized dynamical system, where the fine-grid unknowns are assumed to have the form

$$\begin{pmatrix} \delta_f \\ \omega_f \end{pmatrix} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \delta_c \\ \omega_c \end{pmatrix}.$$

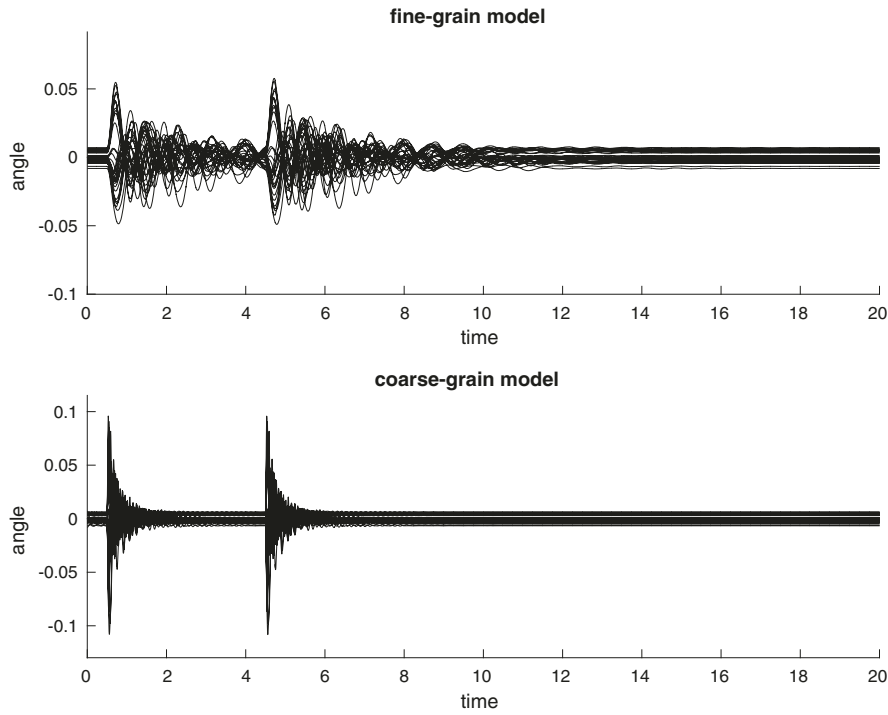


FIGURE 1 Tracking of coarse trajectories using six multigrid levels with caliber-4 interpolation with measure $c_{\alpha,\beta}$ to construct coarse-grain model (34). The coarse-grain model contains only 32 buses, a reduction from 324 buses. Poor but stable tracking is obtained

This gives the system

$$\begin{pmatrix} I^t I \ddot{\delta}_c \\ I^t I \dot{\omega}_c \end{pmatrix} = \begin{pmatrix} \mathbf{0} & I^t I \\ -I^t \mathbf{L}_f I & -I^t I \end{pmatrix} \begin{pmatrix} \delta_c \\ \omega_c \end{pmatrix} + \begin{pmatrix} \mathbf{0} \\ I^t \mathbf{p}_f \end{pmatrix} \quad (37)$$

from which we obtain (36). Now the phase–frequency relation $\omega = \dot{\delta}$ is maintained and the coupling occurs only through the 21 block. However, since $I^t I \neq \mathbf{I}_c$, the coupling pattern of $[I^t I]^{-1} I^t \mathbf{L}_f I$ is not the coupling pattern of a row-scaled

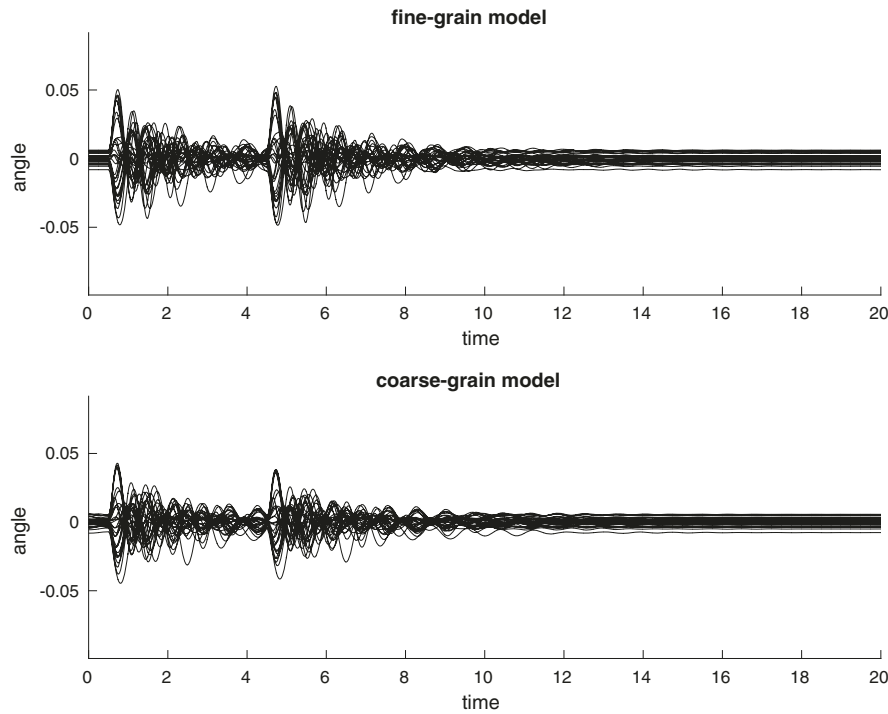


FIGURE 2 Tracking of coarse trajectories using six multigrid levels with caliber-4 interpolation with measure $c_{\alpha,\beta}$ to construct coarse-grain model (35). The coarse-grain model contains only 31 buses, a reduction from 324 buses. We see detailed tracking

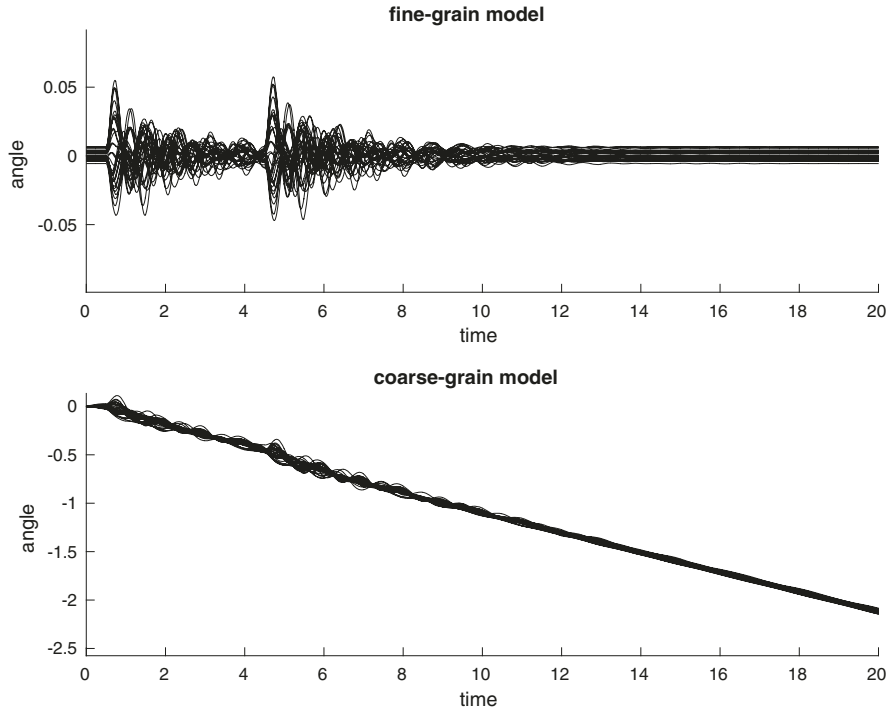


FIGURE 3 Tracking of coarse trajectories using six multigrid levels with caliber-4 interpolation with measure $c_{\alpha,\beta}$ to construct coarse-grain model (36). The coarse-grain model contains only 32 buses, a reduction from 324 buses. We see poor tracking, in fact, instability

graph Laplacian. As mentioned earlier, this generally models a different dynamical system. Figure 3 shows that this coarse-grain model is not even stable, that is, synchrony is not preserved.

5.1.1 | Comment

In principle, with regard to coarse model (34), one can choose a restriction $\mathcal{R}$ such that $\mathcal{R}\mathbf{I} = \mathbf{I}_c$ as in, for example, the work of Falgout et al.¹⁹ However, as mentioned in the work of Falgout et al.,¹⁹ $\mathcal{R}$ is generally not the restriction $\mathbf{I}^t$, implying a nonsymmetric coarse graph Laplacian would result from a RAP coarsening of the dynamical system.

5.2 | More complex system

The above technique can be used for more general inertial and damping coefficients, that is, applied to (31). However, unlike in the simplified case, there is now not an easy relationship between the eigenvalues/vectors of $\mathbf{L}_f$ and the matrix operator of (31). Thus, it is unclear without analyzing the matrix operator to determine if multigrid techniques applied to $\mathbf{L}_f$ will select coarse DOFs for good tracking of the trajectories. Rather than an explicit spectral analysis, we will examine the iteration matrix for a time integration of the system to observe what components of the operator must be accurately approximated. This will also expose scaled Laplacians that must be approximated well. First, consider a simplified homogeneous system with no damping, uniform inertia (i.e., $\mathbf{H}_f = \mathbf{I}_f$), and integrated with an explicit one-step method as follows:

$$\begin{aligned}
 \begin{pmatrix} \delta_f^j \\ \omega_f^j \end{pmatrix} &= \begin{pmatrix} \mathbf{I}_f & \Delta t \mathbf{I}_f \\ -\Delta t \mathbf{L}_f & \mathbf{I}_f \end{pmatrix} \begin{pmatrix} \delta_f^{j-1} \\ \omega_f^{j-1} \end{pmatrix} \\
 &= \left[\begin{pmatrix} \mathbf{0} & \Delta t \mathbf{I}_f \\ -\Delta t \mathbf{L}_f & \mathbf{0} \end{pmatrix} + \begin{pmatrix} \mathbf{I}_f & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_f \end{pmatrix} \right]^j \begin{pmatrix} \delta_f^0 \\ \omega_f^0 \end{pmatrix} \\
 &:= [\mathcal{M} + \mathcal{N}]^j \begin{pmatrix} \delta_f^0 \\ \omega_f^0 \end{pmatrix},
 \end{aligned} \tag{38}$$

where $\mathcal{M}$, $\mathcal{N}$ are the first and second matrices in (38), respectively. Because $\mathcal{M}$ and $\mathcal{N}$ commute, using a matrix binomial expansion, we have

$$[\mathcal{M} + \mathcal{N}]^j = \sum_{l=0}^j \binom{j}{l} \mathcal{M}^{j-l} \quad (39)$$

with

$$\mathcal{M}^r = \begin{cases} (-1)^m (\Delta t)^{2m} \begin{pmatrix} \mathbf{L}_f^m & \mathbf{0} \\ \mathbf{0} & \mathbf{L}_f^m \end{pmatrix} & r = 2m \\ (-1)^m (\Delta t)^{2m+1} \begin{pmatrix} \mathbf{0} & \mathbf{L}_f^m \\ -\mathbf{L}_f^{m+1} & \mathbf{0} \end{pmatrix} & r = 2m + 1. \end{cases} \quad (40)$$

Expression (40) shows the relevance of powers of the Laplacian, which expectedly implies that the coarse Laplacian must sufficiently approximate the fine Laplacian. However, this relevance does not mean that power iterations should be used instead of relaxation sweeps for determining the coarse nodes. Since the coarse spatial grid can represent only algebraically smooth functions, the coarse Laplacian should be constructed to ensure accurate approximation to smooth functions. Hence, relaxation sweeps that expose the smooth modes rather than power iterations that amplify the oscillatory modes should be used to measure the bus affinities for selecting the coarse nodes.

Turning to system (31), the explicit one-step integrator modifies (38) with

$$\mathcal{M} = \begin{pmatrix} \mathbf{0} & \Delta t \mathbf{I}_f \\ -\Delta t \mathbf{H}_f^{-1} \mathbf{L}_f & \mathbf{0} \end{pmatrix} \quad \mathcal{N} = \begin{pmatrix} \mathbf{I}_f & \mathbf{0} \\ \mathbf{0} & (\mathbf{I}_f - \Delta t \mathbf{H}_f^{-1} \mathbf{D}_f) \end{pmatrix}.$$

Since these two matrices do not commute, a matrix binomial expansion cannot be used to simplify the iteration matrix. $[\mathcal{M} + \mathcal{N}]^j$ will involve expressions containing powers of both $\mathcal{M}$ and $\mathcal{N}$. Hence, both of these block matrices must be coarsened well, of which $\mathcal{M}$ is more difficult since $\mathcal{N}$ is a diagonal matrix. We have

$$\mathcal{M}^r = \begin{cases} (-1)^m (\Delta t)^{2m} \begin{pmatrix} (\mathbf{H}_f^{-1} \mathbf{L}_f)^m & \mathbf{0} \\ \mathbf{0} & (\mathbf{H}_f^{-1} \mathbf{L}_f)^m \end{pmatrix} & r = 2m \\ (-1)^m (\Delta t)^{2m+1} \begin{pmatrix} \mathbf{0} & (\mathbf{H}_f^{-1} \mathbf{L}_f)^m \\ -(\mathbf{H}_f^{-1} \mathbf{L}_f)^{m+1} & \mathbf{0} \end{pmatrix} & r = 2m + 1. \end{cases} \quad (41)$$

showing that powers of a nonsymmetric row-scaled Laplacian arise. This then raises the question of whether the multigrid techniques applied to $\mathbf{L}_f$ would detect the appropriate coarse nodes for a good representation of $\mathbf{H}_f^{-1} \mathbf{L}_f$. In turn, this leads to questions on compatible relaxation for nonsymmetric matrices. We will address the nonsymmetry by scaling the dynamical system and will see that reasonably coarsened $\mathbf{H}_f^{-1}$ and $\mathbf{L}_f$ for coarse-grain systems that preserve structures of the fine-grain system is sufficient enough.

First, consider the new variable $\tilde{\omega}_f = \mathbf{H}_f^{\frac{1}{2}} \omega_f$. The linearized dynamical system becomes

$$\begin{pmatrix} \dot{\delta}_f \\ \dot{\tilde{\omega}}_f \end{pmatrix} = \begin{pmatrix} \mathbf{0} & \mathbf{H}_f^{-\frac{1}{2}} \\ -\mathbf{H}_f^{-\frac{1}{2}} \mathbf{L}_f & -\mathbf{H}_f^{-\frac{1}{2}} \mathbf{D}_f \mathbf{H}_f^{-\frac{1}{2}} \end{pmatrix} \begin{pmatrix} \delta_f \\ \tilde{\omega}_f \end{pmatrix} + \begin{pmatrix} \mathbf{0} \\ \mathbf{H}_f^{-\frac{1}{2}} \mathbf{p}_f \end{pmatrix} \quad (42)$$

and

$$\mathcal{M}^r = \begin{cases} (-1)^m (\Delta t)^{2m} \begin{pmatrix} (\mathbf{H}_f^{-1} \mathbf{L}_f)^m & \mathbf{0} \\ \mathbf{0} & (\mathbf{H}_f^{-\frac{1}{2}} \mathbf{L}_f \mathbf{H}_f^{-\frac{1}{2}})^m \end{pmatrix} \\ (-1)^m (\Delta t)^{2m+1} \begin{pmatrix} \mathbf{0} & \mathbf{H}_f^{-\frac{1}{2}} (\mathbf{H}_f^{-\frac{1}{2}} \mathbf{L}_f \mathbf{H}_f^{-\frac{1}{2}})^m \\ -\mathbf{H}_f^{-\frac{1}{2}} \mathbf{L}_f (\mathbf{H}_f^{-1} \mathbf{L}_f)^m & \mathbf{0} \end{pmatrix} \end{cases}$$

$$= \begin{cases} (-1)^m (\Delta t)^{2m} \begin{pmatrix} \mathbf{H}_f^{-\frac{1}{2}} \left(\mathbf{H}_f^{-\frac{1}{2}} \mathbf{L}_f \mathbf{H}_f^{-\frac{1}{2}} \right)^m \mathbf{H}_f^{\frac{1}{2}} & \mathbf{0} \\ \mathbf{0} & \left(\mathbf{H}_f^{-\frac{1}{2}} \mathbf{L}_f \mathbf{H}_f^{-\frac{1}{2}} \right)^m \end{pmatrix} \\ (-1)^m (\Delta t)^{2m+1} \begin{pmatrix} \mathbf{0} & \mathbf{H}_f^{-\frac{1}{2}} \left(\mathbf{H}_f^{-\frac{1}{2}} \mathbf{L}_f \mathbf{H}_f^{-\frac{1}{2}} \right)^m \\ - \left(\mathbf{H}_f^{-\frac{1}{2}} \mathbf{L}_f \mathbf{H}_f^{-\frac{1}{2}} \right)^{m+1} \mathbf{H}_f^{\frac{1}{2}} & \mathbf{0} \end{pmatrix}, \end{cases} \quad (43)$$

The powers now occur in a symmetrically scaled Laplacian. This means compatible relaxation with basic iteration schemes can be used for determining the affinity between buses for this system, albeit applied to the matrix $\mathbf{H}_f^{-\frac{1}{2}} \mathbf{L}_f \mathbf{H}_f^{-\frac{1}{2}}$.

(A similar expression involving $(\mathbf{H}_f^{-\frac{1}{2}} \mathbf{L}_f \mathbf{H}_f^{-\frac{1}{2}})^m$ can be used in (41), but scaling the dynamical system introduces other formulations that may be useful. Nevertheless, it implies that relaxation can be used to determine the bus affinities in (41).)

Second, given that relaxation can be used to detect the coarse nodes for general system (31), we introduce the following reduced models for (31) and its scaled form (42), which are derived in a similar fashion as in reduced model (35):

$$\begin{pmatrix} \dot{\delta}_c \\ \dot{\omega}_c \end{pmatrix} = \begin{pmatrix} \mathbf{0} & \mathbf{I}_c \\ -[\text{diag}(\mathbf{I}^t \mathbf{h}_f)]^{-1} \mathbf{I}^t \mathbf{L}_f \mathbf{I} - [\text{diag}(\mathbf{I}^t \mathbf{h}_f)]^{-1} \text{diag}(\mathbf{I}^t \mathbf{d}_f) & \mathbf{0} \end{pmatrix} \begin{pmatrix} \delta_c \\ \omega_c \end{pmatrix} + \begin{pmatrix} \mathbf{0} \\ [\text{diag}(\mathbf{I}^t \mathbf{h}_f)]^{-1} \mathbf{I}^t \mathbf{p}_f \end{pmatrix}, \quad (44)$$

and

$$\begin{pmatrix} \dot{\delta}_c \\ \dot{\omega}_c \end{pmatrix} = \begin{pmatrix} \mathbf{0} & [\text{diag}(\mathbf{I}^t \mathbf{h}_f)]^{-\frac{1}{2}} \\ -[\text{diag}(\mathbf{I}^t \mathbf{h}_f)]^{-\frac{1}{2}} \mathbf{I}^t \mathbf{L}_f \mathbf{I} - [\text{diag}(\mathbf{I}^t \mathbf{h}_f)]^{-\frac{1}{2}} \text{diag}(\mathbf{I}^t \mathbf{d}_f) [\text{diag}(\mathbf{I}^t \mathbf{h}_f)]^{-\frac{1}{2}} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \delta_c \\ \omega_c \end{pmatrix} + \begin{pmatrix} \mathbf{0} \\ [\text{diag}(\mathbf{I}^t \mathbf{h}_f)]^{-\frac{1}{2}} \mathbf{I}^t \mathbf{p}_f \end{pmatrix}. \quad (45)$$

Here, to determine the bus affinities, relaxation can be applied to $\mathbf{H}_f^{-\frac{1}{2}} \mathbf{L}_f \mathbf{H}_f^{-\frac{1}{2}}$, but the weights in interpolation $\mathbf{I}$ are determined using only the values of $\mathbf{L}_f$. This is a compromise between constructing a coarse-grain model with an ideal multigrid method that determines the coarse DOFs and interpolation weights based on the matrix (i.e., based on the complementary principle) and constructing a practical method that only approximates an ideal multigrid method but preserves the structure of the dynamical system. In addition, note that, to coarsen the inertial and damping coefficients (important in both $\mathcal{M}$ and $\mathcal{N}$), a delicate weighted averaging procedure is needed. Standard arithmetic, harmonic, and geometric averaging of these coefficients over the coarse aggregates is not as effective as the averaging effected through $\mathbf{I}^t$ constructed from $\mathbf{L}_f$. That is, $\mathbf{I}$ achieves the delicate ratio of how each coarse variable contributes to a fine variable. Equivalently, $\mathbf{I}^t$ achieves the appropriate weighted averaging of fine-grid values to produce a coarse-grid value. Hence, $\mathbf{I}^t \mathbf{h}_f$ and $\mathbf{I}^t \mathbf{d}_f$ can be used for the coarse inertia and coarse damping.

Using the first block of equations in (44) and (45), the dynamical systems corresponding to these systems are respectively

$$\begin{aligned} \mathbf{0} &= \dot{\omega}_c + [\text{diag}(\mathbf{I}^t \mathbf{h}_f)]^{-1} \{ \mathbf{I}^t \mathbf{L}_f \mathbf{I} \delta_c + \text{diag}(\mathbf{I}^t \mathbf{d}_f) \dot{\delta}_c - \mathbf{I}^t \mathbf{p}_f \} \\ &:= \dot{\omega}_c + \mathbf{H}_c^{-1} \{ \mathbf{L}_c \delta_c + \mathbf{D}_c \dot{\delta}_c - \mathbf{p}_c \} \end{aligned} \quad (46)$$

$$\begin{aligned} \mathbf{0} &= \dot{\delta}_c + [\text{diag}(\mathbf{I}^t \mathbf{h}_f)]^{-\frac{1}{2}} \{ \mathbf{I}^t \mathbf{L}_f \mathbf{I} \delta_c + \text{diag}(\mathbf{I}^t \mathbf{d}_f) \dot{\delta}_c - \mathbf{I}^t \mathbf{p}_f \} \\ &:= \dot{\delta}_c + \mathbf{H}_c^{-\frac{1}{2}} \{ \mathbf{L}_c \delta_c + \mathbf{D}_c \dot{\delta}_c - \mathbf{p}_c \}. \end{aligned} \quad (47)$$

Comparing these to the fine-grain system

$$\dot{\omega}_f + \mathbf{H}_f^{-1} \{ \mathbf{L}_f \delta_f + \mathbf{D}_f \dot{\delta}_f - \mathbf{p}_f \} = \mathbf{0}, \quad (48)$$

we see that (46) and (47) are structurally similar to (48). For comparison, consider the scaled system

$$\begin{aligned} \begin{pmatrix} \dot{\delta}_c \\ \dot{\omega}_c \end{pmatrix} &= \begin{pmatrix} \mathbf{0} & \mathbf{I}_c \\ -\mathbf{I}_1^t \left[\mathbf{H}_f^{-\frac{1}{2}} \mathbf{L}_f \mathbf{H}_f^{-\frac{1}{2}} \right] \mathbf{I}_1 - [\text{diag}(\mathbf{I}_2^t \mathbf{h}_f)]^{-\frac{1}{2}} \text{diag}(\mathbf{I}_2^t \mathbf{d}_f) [\text{diag}(\mathbf{I}_2^t \mathbf{h}_f)]^{-\frac{1}{2}} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \delta_c \\ \omega_c \end{pmatrix} \\ &+ \begin{pmatrix} \mathbf{0} \\ [\text{diag}(\mathbf{I}_2^t \mathbf{h}_f)]^{-\frac{1}{2}} \mathbf{I}_1^t \mathbf{p}_f \end{pmatrix} \end{aligned} \quad (49)$$

with $\tilde{\delta} = \mathbf{H}_f^{\frac{1}{2}} \delta$, $\tilde{\omega} = \mathbf{H}_f^{\frac{1}{2}} \omega$, $\mathcal{I}_1$ generated from $\mathbf{H}_f^{\frac{1}{2}} \mathbf{L}_f \mathbf{H}_f^{\frac{1}{2}}$, and $\mathcal{I}_2$ based on some type of weighted averaging. Its corresponding dynamical system is

$$\dot{\tilde{\omega}}_c + \mathcal{I}_1^t \left[\mathbf{H}_f^{-\frac{1}{2}} \mathbf{L}_f \mathbf{H}_f^{-\frac{1}{2}} \right] \mathcal{I}_1 \tilde{\delta}_c + [\text{diag}(\mathcal{I}_2^t \mathbf{h}_f)]^{-\frac{1}{2}} \left\{ \text{diag}(\mathcal{I}_2^t \mathbf{d}_f) [\text{diag}(\mathcal{I}_2^t \mathbf{h}_f)]^{-\frac{1}{2}} \tilde{\delta}_c - \mathcal{I}_1^t \mathbf{p}_f \right\} = \mathbf{0}$$

from which we see that the terms cannot be factored as in (48). Further note that scaled system (49) is valid only when $\mathbf{H}_f = \mathbf{I}_f$ since the nonlinearity in the dynamical system occurs in δ , for which the scaling $\tilde{\delta} = \mathbf{H}_f^{\frac{1}{2}} \delta$ is generally invalid.

5.3 | Systems with generators and loads

The above model reduction technique can be applied to systems with generators and loads, that is, to system (11)–(12). Distinguishing the generator and load values with subscripts g and l , the linearized system is

$$\begin{pmatrix} \dot{\delta}_{f,g} \\ \dot{\delta}_{f,l} \\ \dot{\omega}_{f,g} \end{pmatrix} = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{I}_{f,g} \\ -\mathbf{D}_{f,l}^{-1} \mathbf{L}_{f,lg} & -\mathbf{D}_{f,l}^{-1} \mathbf{L}_{f,ll} & \mathbf{0} \\ -\mathbf{H}_{f,g}^{-1} \mathbf{L}_{f,gg} & -\mathbf{H}_{f,g}^{-1} \mathbf{L}_{f,gl} & -\mathbf{H}_{f,g}^{-1} \mathbf{D}_{f,g} \end{pmatrix} \begin{pmatrix} \delta_{f,g} \\ \delta_{f,l} \\ \omega_{f,g} \end{pmatrix} + \begin{pmatrix} \mathbf{0} \\ -\mathbf{D}_{f,l}^{-1} \mathbf{p}_{f,d}^o \\ \mathbf{H}_{f,g}^{-1} \mathbf{p}_{f,m}^o \end{pmatrix}. \quad (50)$$

We will assume that most of the buses are generators, which can represent the scenario where most of the loads have been removed through network reduction while the major ones and the ones with sources of uncertainty have been retained. Because the drawn load is a small frequency perturbation of $\mathbf{p}_{f,d}^o$, we will assume that the damping at the loads is very large to ensure a small perturbation. The reasoning behind this can be seen from (12):

$$\dot{\delta}_i = \frac{-\sum_{j=1}^n a_{ij} \sin(\delta_{ij}) - P_{d,i}^o}{D_i} \quad i \in N_l$$

to ensure that a small frequency perturbation D_i should be large. This means that the oscillations in these phases damp down quickly. The large dampings also pose a problem in constructing the coarse damping coefficients for coarse generator buses that have been “aggregated” with load buses. Because the damping coefficient for a coarse bus is formed by weighted averaging the values of the buses that have been aggregated to form this coarse bus, the large damping from the loads will overdamp a coarse generator bus. The simple remedy we take in this paper is to omit the damping contributions from the loads when forming the damping for the coarse generators. This seems to be sufficient for the examples of this paper, but a more meticulous approach will be examined in the future.

6 | NUMERICAL EXPERIMENTS

Constructing large stable dynamical systems can be difficult. One way to construct stable systems is through semidiscretizations of time-dependent elliptic PDEs. This is the approach we take. In particular, the graph patterns in the admittance matrix will be taken from discretizations of scalar PDEs and systems of PDEs. However, assuming that conductances are negligible, the values in the admittance matrix will be random and positive. Depending on whether the graph pattern is derived from a scalar system of PDEs, each bus will have between three to nine connections (scalar case) or six to 18 connections (system case). To complete the construction of the dynamical systems, we randomly draw the inertial and damping coefficients from uniform distributions defined over positive subintervals. Because of the different random features in the setup, each experiment will represent a different scenario, although all the scenarios should result in qualitatively similar trajectories. Furthermore, the compatible relaxation procedure will involve 20 weighted-Jacobi smoothing steps and 20–30 test vectors with $Q = 0.5$ in (28) in the second pass of the seed selection. The simulations will be time integrated up to 20 seconds with $\Delta t = 0.01$ using an explicit fourth-order Runge–Kutta scheme, with the same forcing used in the experiments in Section 5.1 with faults occurring at intervals $0.5 \leq t \leq 0.6$ and $4.5 \leq t \leq 4.6$. Finally, for comparison, only the trajectories of the coarse buses generated from the full-system dynamics and generated from the reduced-system dynamics will be plotted.

We also mention that the sizes of the fine-grain models are small compared to the steady-state elliptic problems multigrid methods are often applied to. However, the goal of these experiments is to show that a carefully constructed multigrid can be applied to time-dependent dynamical systems. Coarse-graining complex networks has been an active area of research in the physics community during the past decade. Furthermore, with respect to practical power grids,

large networks are of the order 100,000 buses, and hence, one may question whether coarse-graining is really needed in the power grid community. However, the time steps must be very small to ensure that the dynamics are captured. Thus, reducing the size of the network can greatly reduce the simulation time. This is particular true for more detailed models, which will involve more variables of interest at the nodes, and for real-time analysis and parametric studies (uncertainty quantification) in decentralized systems and deregulated electric grids.

Only Generator Buses with Nonuniform Inertia and Damping. Our first set of experiments involves systems consisting of only generators but with nonuniform inertial and damping coefficients. These coefficients are drawn from uniform distributions over $[0.1, 50.1]$ and $[0.1, 10.1]$, respectively. (Larger intervals can be chosen, but the stability of the constructed dynamical system must be guaranteed.) We first consider experiments using affinity measures $c_{\alpha,\beta}$ and $d_{\alpha,\beta}$ of (26)–(27) for reduced model (44). Caliber-4 interpolation and seven multigrid levels are used to reduce the number of buses by a factor of 15–20, and the coarse inertial and damping coefficients are computed using the restriction operators generated by the multigrid procedure. Figures 4 and 5 show the results. We see that good and rather similar tracking is achieved by both measures, so that both measures are effective. More accurate tracking can be achieved by using less multigrid levels but at the expense of retaining more buses in the reduced model.

The next set of experiments involves scaled model (45). We also consider the coarse DOFs selection using relaxation on $\mathbf{H}_f^{-\frac{1}{2}} \mathbf{L}_f \mathbf{H}_f^{-\frac{1}{2}}$ in the affinity measure, although the interpolation weights in $\mathcal{I}$ involve only the values of $\mathbf{L}_f$. Because the affinity measures $c_{\alpha,\beta}$ and $d_{\alpha,\beta}$ give similar results, only results for $c_{\alpha,\beta}$ are given. Again, the coarse inertial and damping coefficients are constructed using the restriction operators generated by the multigrid coarsening. Figure 6 shows the tracking of the phases using relaxation on $\mathbf{L}_f$ to determine the coarse DOFs, and Figure 7 shows the tracking using relaxation on $\mathbf{H}_f^{-\frac{1}{2}} \mathbf{L}_f \mathbf{H}_f^{-\frac{1}{2}}$. Good results are obtained for both approaches, even for a coarsening from 1,568 to 53 buses.

Now, consider reduced model (49), which does not preserve some of the structures of system (31) where scaling $\tilde{\delta} = \mathbf{H}_f^{\frac{1}{2}} \delta$ is invalid. The coarse inertial and damping coefficients are constructed using weighted volume averaging. Figure 8 illustrates the poor coarse-grain tracking and, in fact, unstable coarse-grain system. Moreover, the fine-grain trajectories are dissimilar to the trajectories in the other reduced models, which is due to the invalid scaling.

Finally, to show that the coarsening of the inertial and damping coefficients is more appropriately computed using the multigrid restriction operators, we consider model (45) but with these coefficients constructed with volume averaging. Figure 9 illustrates the results, which are very different from the plot in Figure 6 where these coefficients were constructed with the restriction operators. Although the tracking is poor, the system does appear to be slowly stabilizing as it evolves.

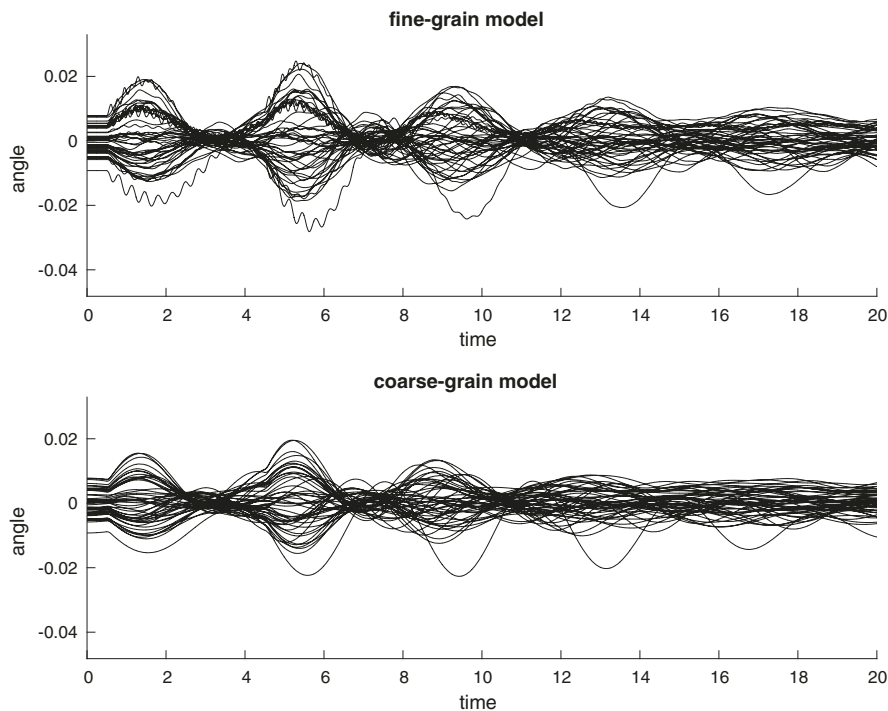


FIGURE 4 Tracking of coarse trajectories using seven multigrid levels with caliber-4 interpolation and $c_{\alpha,\beta}$ measure (27) to construct coarse-grain model (44). The coarse-grain model contains only 49 buses, a reduction from 784 buses. Detailed tracking is obtained

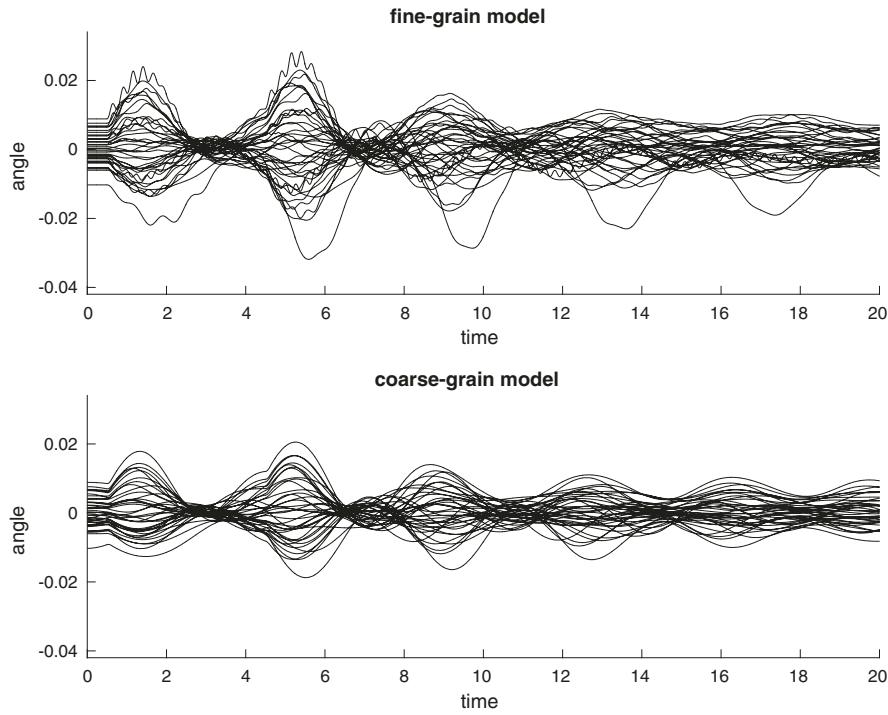


FIGURE 5 Tracking of coarse trajectories using seven multigrid levels with caliber-4 interpolation and $d_{\alpha,\beta}$ measure (26) to construct coarse-grain model (44). The coarse-grain model contains only 39 buses, a reduction from 784 buses. Detailed tracking is obtained

Generator and Load Buses with Nonuniform Inertia and Damping. The next set of experiments involves both generator and load buses. The inertial and damping coefficients for the generators are drawn again from uniform distributions over $[0.1, 50.1]$ and $[0.1, 10.1]$, respectively, but the damping in the loads are drawn from $[10, 1010]$. The first

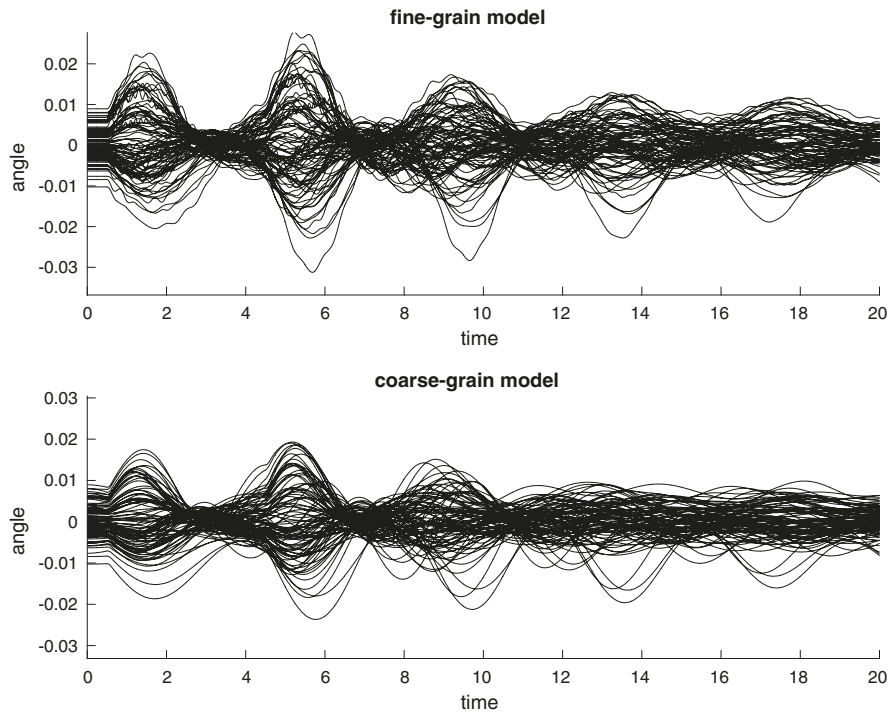


FIGURE 6 Tracking of coarse trajectories using seven multigrid levels with caliber-4 interpolation and $c_{\alpha,\beta}$ measure (27) to construct coarse-grain model (45). The coarse-grain model contains only 43 buses, a reduction from 784 buses. Detailed tracking is obtained

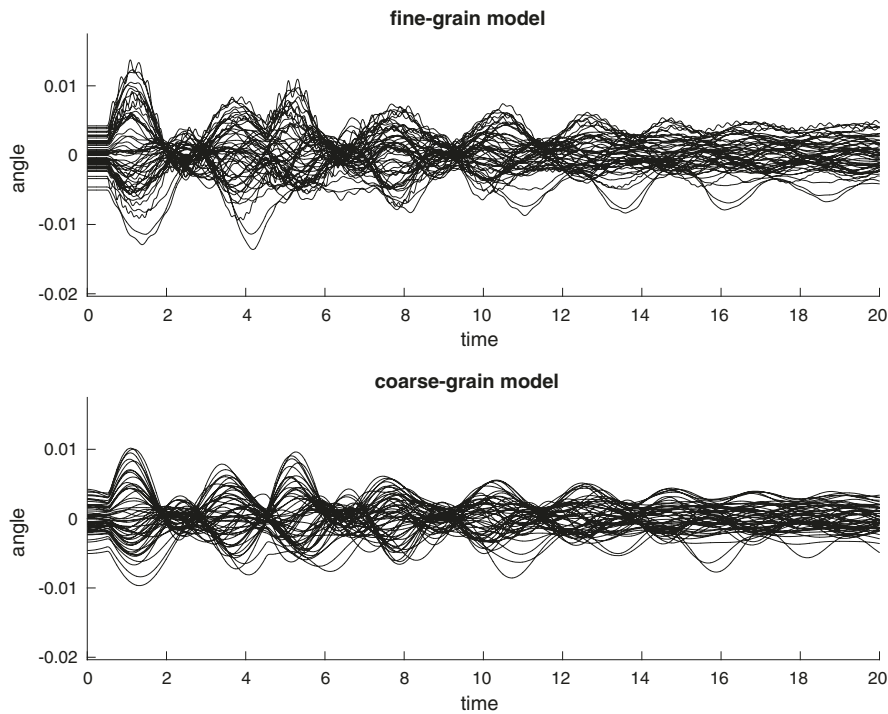


FIGURE 7 Tracking of coarse trajectories using eight multigrid levels with caliber-4 interpolation and $c_{\alpha,\beta}$ measure (27) to construct coarse-grain model (45) but using compatible relaxation on $\mathbf{H}_f^{-\frac{1}{2}} \mathbf{L}_f \mathbf{H}_f^{-\frac{1}{2}}$ to determine the coarse DOFs, although the interpolation is formed with $\mathbf{L}_f$ only. The coarse-grain model contains only 53 buses, a reduction from 1,568 buses. Detailed tracking is obtained. DOFs = degrees of freedom

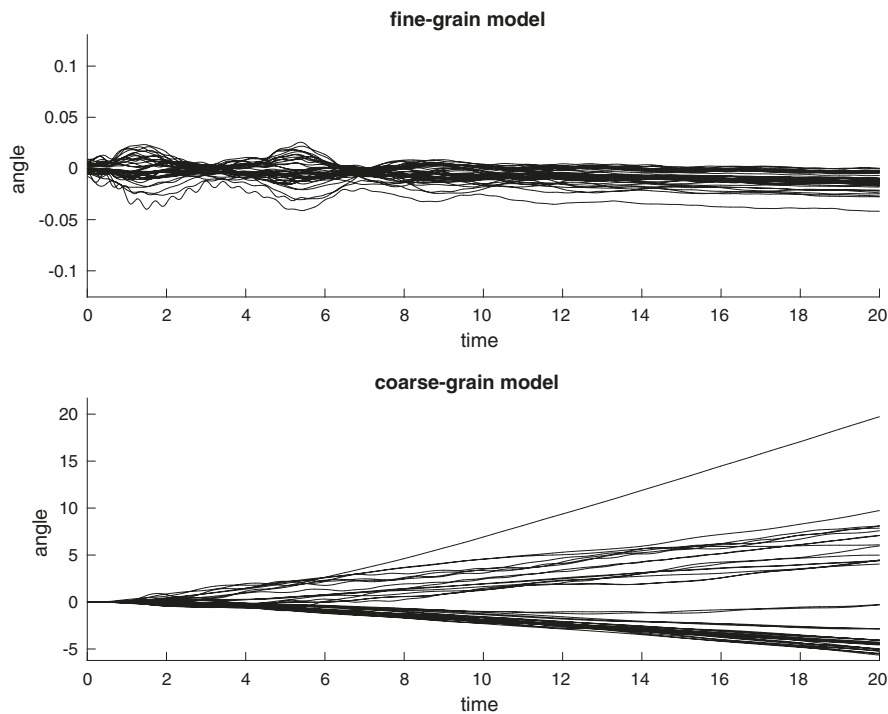


FIGURE 8 Tracking of coarse trajectories using five multigrid levels with caliber-4 interpolation and $c_{\alpha,\beta}$ measure (27) to construct coarse-grain model (49) with volume averaging to construct the coarse damping and inertial coefficients. The coarse-grain model contains only 41 buses, a reduction from 324 buses. Not only is poor coarse tracking obtained, but also significantly different fine-grain trajectories are obtained because of the invalid scaling $\tilde{\delta} = \mathbf{H}^{\frac{1}{2}} \delta$

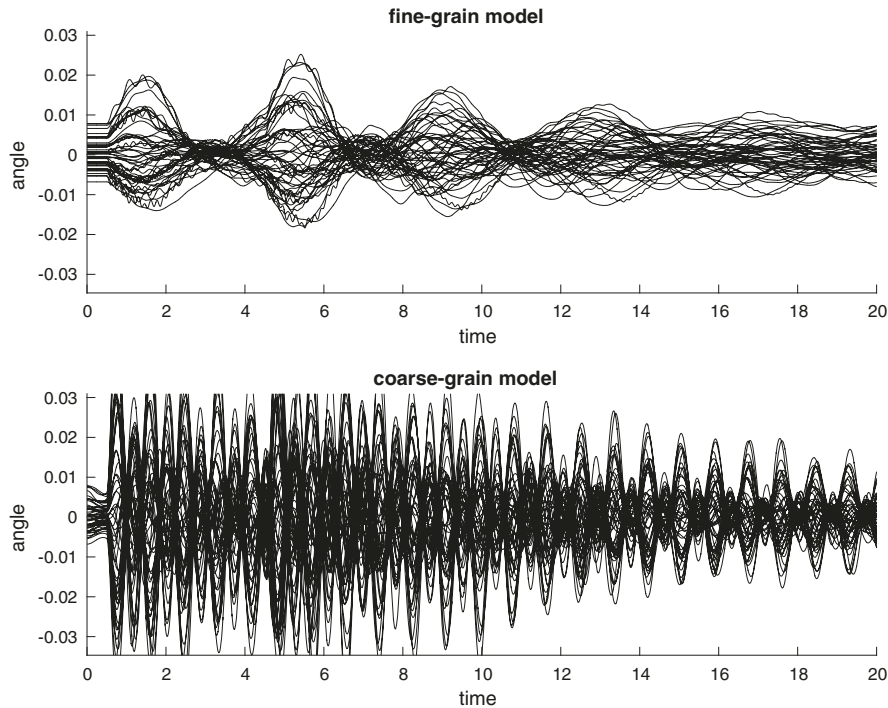


FIGURE 9 Tracking of coarse trajectories using seven multigrid levels with caliber-4 interpolation and $c_{\alpha,\beta}$ measure (27) to construct coarse-grain model (45) with volume averaging to construct the coarse damping and inertial coefficients. The coarse-grain model contains only 51 buses, a reduction from 784 buses. Poor tracking is obtained

example takes the first two-thirds of the buses to be generators. Caliber-5 interpolation is used, although caliber-4 gives similar results and six multigrid levels are used to reduce the number of buses from 648 to 56. Of these 56 coarse buses, the coarsening process selected 38 generators and 18 loads. Figure 10 shows the trajectories of the coarse bus phases. From the fine-grain plot, we can distinguish the load trajectories from the immediate damping of oscillations. From the coarse-grain plot, we see good tracking, clearly giving the qualitative feature of the dynamics. However, there are more oscillations in some of the coarse-grain trajectories, which is partially due to omitting the damping contributions of the load buses to the coarse generator buses. Figure 11 illustrates the trajectories when the load damping does contribute to the coarse generator buses. Although this leads to better tracking when the generators are consecutively in order, load damping contribution can lead to overdamping when the generators are randomly dispersed over the network.

The next set of experiments in fact takes the generators to be randomly dispersed over the network. This random dispersion leads to a changing ratio of the number of generator-to-load buses on the finest level. Figure 12 shows the tracking when the damping of the loads contribute to the damping of the coarse generators. We see that this contribution leads to overdamping even though there is roughly 4 times more coarse generator buses (45 to 12). Figure 13 shows the tracking when the damping of the loads do not contribute to the damping of the coarse generator buses. The oscillatory features are now more pronounced. Both approaches capture the global features of the dynamics.

The above experiments for the generator/load system used affinity measure $d_{\alpha,\beta}$. Figure 14 shows the results when $c_{\alpha,\beta}$ is used. Although the global dynamics are captured, there are more oscillations than when $d_{\alpha,\beta}$ is used. This extra oscillatory nature is a result of the extra coarse generator buses and fewer load buses selected by the $c_{\alpha,\beta}$ measure than the $d_{\alpha,\beta}$ measure (i.e., 55 coarse generators/18 coarse loads compared with 35 coarse generators/23 coarse loads). Since the differential equations for the load buses have large damping terms, less oscillations will occur in the dynamics of the load angles.

Statistics for Random Realizations. We now consider sets of experiments, each consisting of 500 realizations, with randomly selected event conditions and graph weights. Although this is a small number of realizations for the number of random variations, the results still demonstrate the overall performance of the coarse-grain technique. With respect to the event conditions, we apply two faults that begin at times randomly selected from $\{1.5, 2.5, 3.5, 4.5, 5.5\}$ seconds and

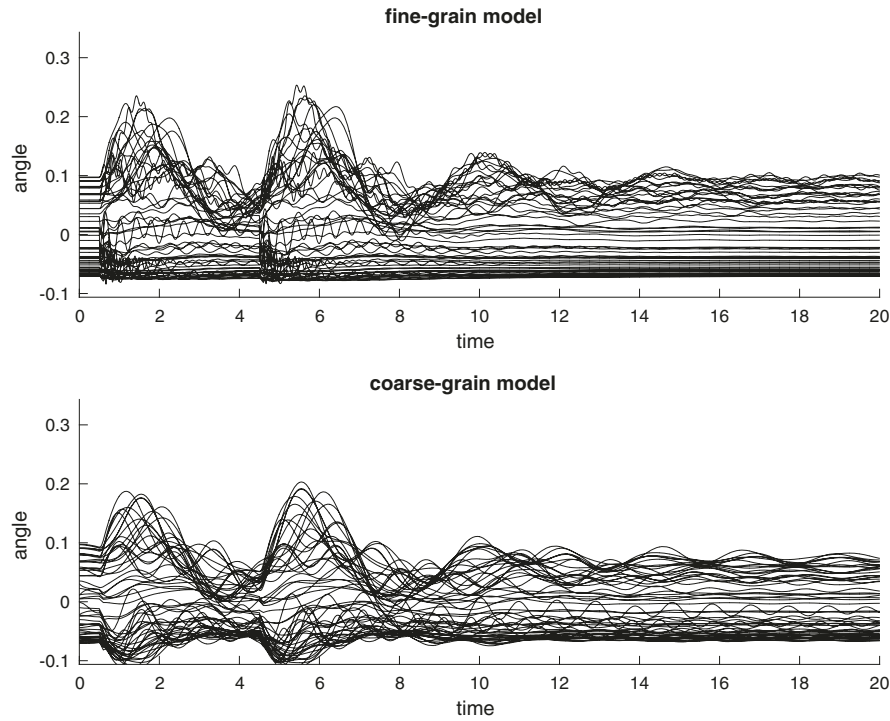


FIGURE 10 Tracking of coarse trajectories using six multigrid levels with caliber-5 interpolation and $d_{\alpha,\beta}$ measure (26) to construct the coarse-grain model for (50). The fine-grain model consists of 432 generators and 216 loads consecutively ordered; the coarse-grain model consists of 38 generators and 18 loads. Good tracking is obtained

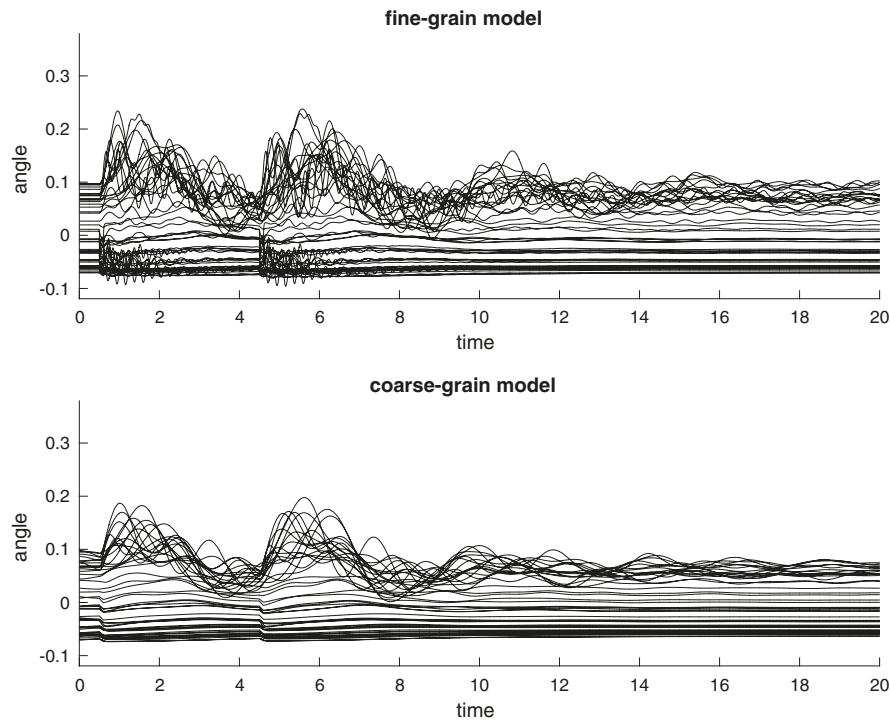


FIGURE 11 Tracking of coarse trajectories using six multigrid levels with caliber-5 interpolation and $d_{\alpha,\beta}$ measure (26) to construct the coarse-grain model for (50). The fine-grain model consists of 432 generators and 216 loads consecutively ordered; the coarse-grain model consists of 32 generators and 19 loads. Mild overdamping occurs because the damping of the load buses contributes to the damping of the coarse generator buses

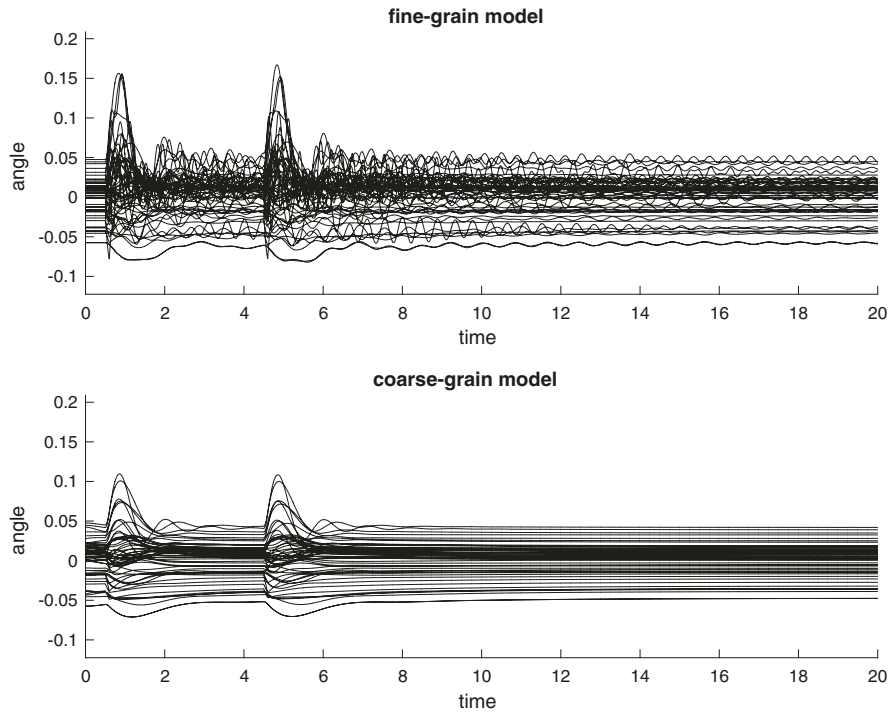


FIGURE 12 Tracking of coarse trajectories using six multigrid levels with caliber-5 interpolation and $d_{\alpha,\beta}$ measure (26) to construct the coarse-grain model for (50). The fine-grain model consists of 484 generators and 164 loads; the coarse-grain model consists of 45 generators and 12 loads. Strong overdamping occurs because the damping of the load buses contributes to the damping of the coarse generator buses

each lasting for 0.1 second. Furthermore, since the initial conditions are constructed by solving the randomly weighted Laplacians for the initial external force, the initial conditions are essentially random. Finally, for the generator and load cases, the locations of the generators are also randomly selected.

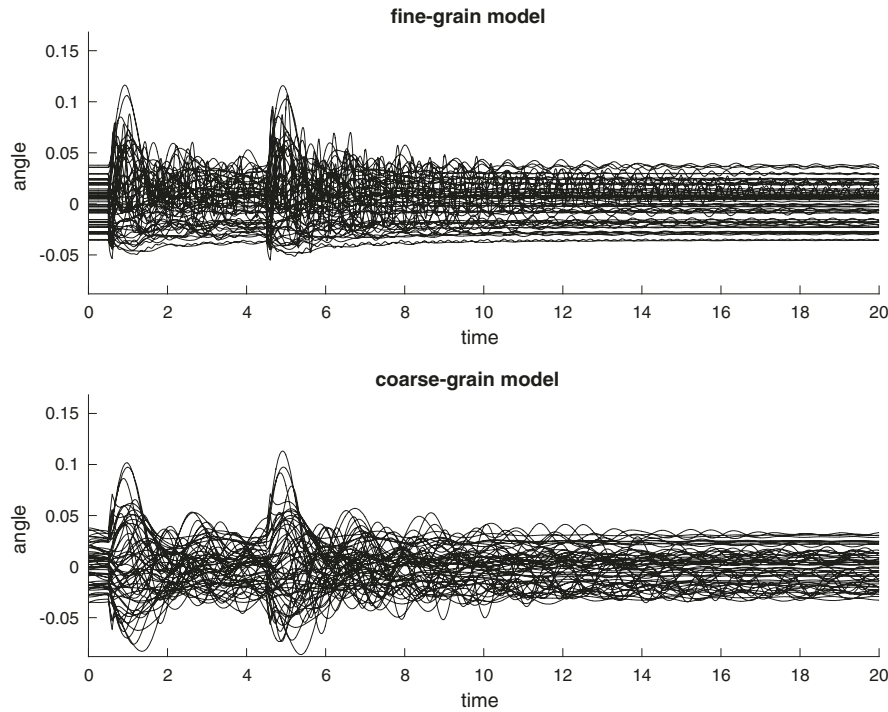


FIGURE 13 Tracking of coarse trajectories using six multigrid levels with caliber-5 interpolation and $d_{\alpha,\beta}$ measure (26) to construct the coarse-grain model for (50). The fine-grain model consists of 476 generators and 172 loads; the coarse-grain model consists of 35 generators and 23 loads. Fair tracking is obtained

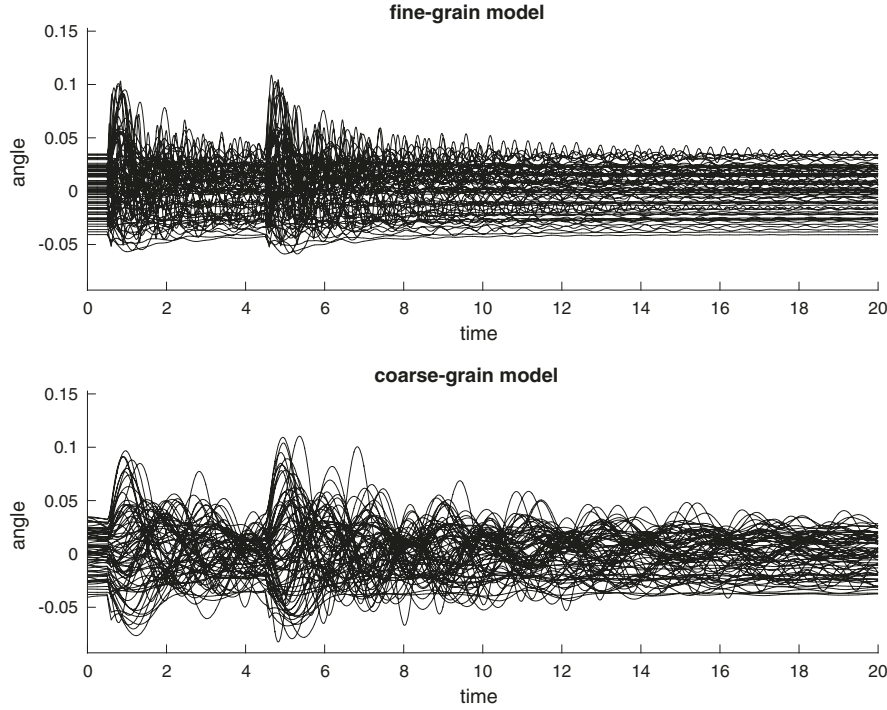


FIGURE 14 Tracking of coarse trajectories using six multigrid levels with caliber-5 interpolation and $c_{\alpha,\beta}$ measure (27) to construct the coarse-grain model for (50). The fine-grain model consists of 479 generators and 169 loads; the coarse-grain model consists of 55 generators and 18 loads. Fair tracking is obtained

The results of these experiments are displayed using the following relative error quantities computed by a simple quadrature rule:

$$\begin{aligned}
 \max &:= \max_{i \in C} \frac{\sqrt{\int_0^T [\delta_{k(i),f} - \delta_{i,c}]^2 dt}}{\sqrt{\int_0^T \delta_{k(i),f}^2 dt}} & \min &:= \min_{i \in C} \frac{\sqrt{\int_0^T [\delta_{k(i),f} - \delta_{i,c}]^2 dt}}{\sqrt{\int_0^T \delta_{k(i),f}^2 dt}} \\
 \text{mean} &:= \text{mean}_{i \in C} \frac{\sqrt{\int_0^T [\delta_{k(i),f} - \delta_{i,c}]^2 dt}}{\sqrt{\int_0^T \delta_{k(i),f}^2 dt}} & \text{std} &:= \text{std}_{i \in C} \frac{\sqrt{\int_0^T [\delta_{k(i),f} - \delta_{i,c}]^2 dt}}{\sqrt{\int_0^T \delta_{k(i),f}^2 dt}},
 \end{aligned}$$

where C is the set of coarse bus indices, $k(i)$ is the fine bus index corresponding to the coarse index i , *mean* is the mean of the relative errors over all the selected coarse buses, *std* is the standard deviation of the relative errors over the selected coarse buses, and T is the upper limit of the time integration of the dynamical system.

A relevant comment must be made about these and other quantitative measures. In power grid simulations, as well as many other dynamical system simulations, a reduced model is considered accurate if it captures the qualitative behavior of the dynamics, not the quantitative accuracy of the dynamics. Thus, these error measures may not do justice to a reduced model that accurately captures the qualitative behavior of the dynamics, as observed in the eye-ball norm. This may be different for some SVD/eigenvalue reduce model techniques, which simply truncate the eigenfunction expansion and integrate a transformed dynamical system in terms of the eigenfunction expansion. For linear systems, these techniques can give good quantitative accuracy, where such measures then would be appropriate. However, in this paper, we are interested in selecting the buses that lead to good reduced models since the buses are physical objects that the power grid engineers can retrofit.

The first set of experiments use coarse-grain model (44) with nonuniform inertia and damping coefficients drawn from uniform distributions over $[0.1, 50.1]$ and $[0.1, 10.1]$, respectively, and involving 784 generator buses. Caliber-4 interpolation and seven multilevels are chosen, which lead to coarse problems involving about 50 buses. Figure 15 illustrates the mean, $\text{mean} \pm \text{std}$, max, and min for the 500 realizations using affinity measure $c_{\alpha,\beta}$. We see that the *mean* is roughly about 0.4. However, the size of the standard deviation indicates that this large *mean* error may be due to a few trajectories (e.g., the trajectories corresponding to max. As mentioned above, even though these quantitative measures are large,

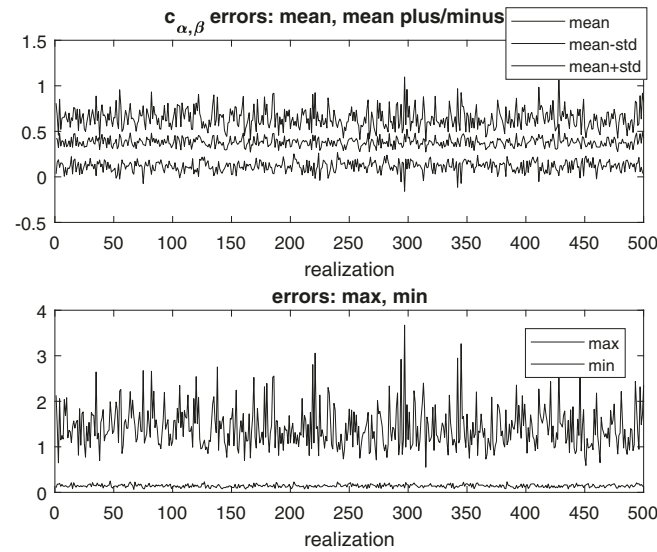


FIGURE 15 Tracking of coarse trajectories using seven multigrid levels with caliber-4 interpolation and $c_{\alpha,\beta}$ measure (27) to construct coarse-grain model (44)

good tracking of the qualitative behavior may still be occurring. Figure 16 illustrates the trajectories for one realization with mean = 0.37 and std = 0.25. From this plot, we do see that the qualitative behavior of the dynamics are tracked well even though the mean is relatively large.

Figure 17 shows the statistics using affinity measure $d_{\alpha,\beta}$. Now the mean is roughly 0.5 and, again, the standard deviation indicates that this may be due to the errors in a few trajectories. Figure 18 is a sample realization with mean = 0.53 and std = 0.27, although the qualitative behavior of the dynamics are tracked well.

The last set of experiments are conducted for model (50) with the generators and loads randomly dispersed over the network. The size of the fine-grain network is 800 buses, the damping coefficients of the finer loads do not contribute

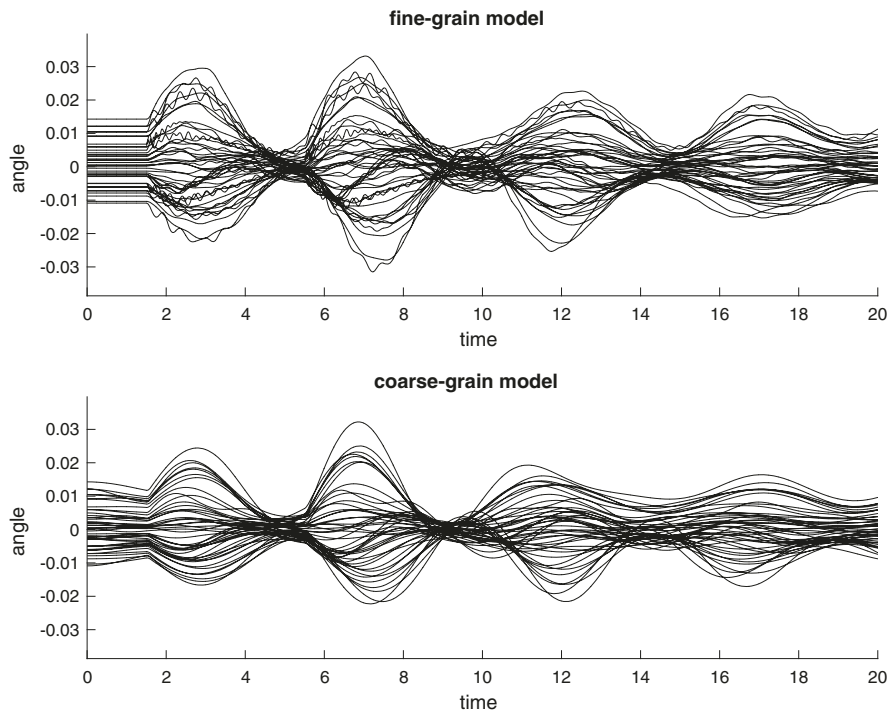


FIGURE 16 Tracking of coarse trajectories using seven multigrid levels with caliber-4 interpolation and $c_{\alpha,\beta}$ measure (27) to construct coarse-grain model (44). The statistics for this realization are max = 3.85, min = 0.15, mean = 0.37, and std = 0.25

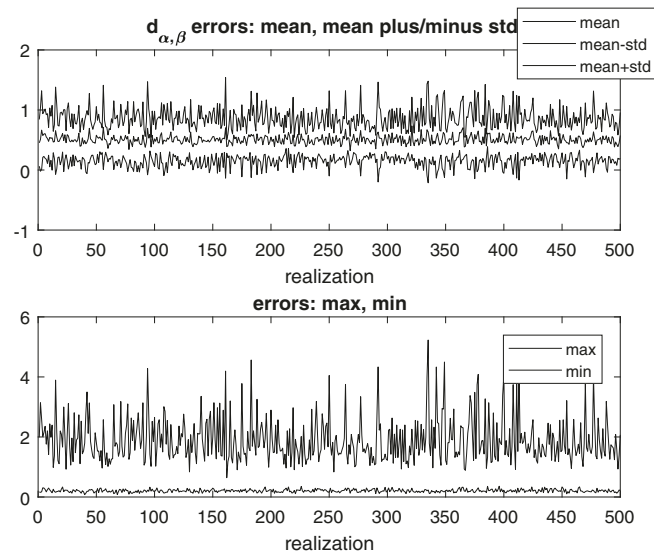


FIGURE 17 Tracking of coarse trajectories using seven multigrid levels with caliber-4 interpolation and $d_{\alpha,\beta}$ measure (26) to construct coarse-grain model (44)

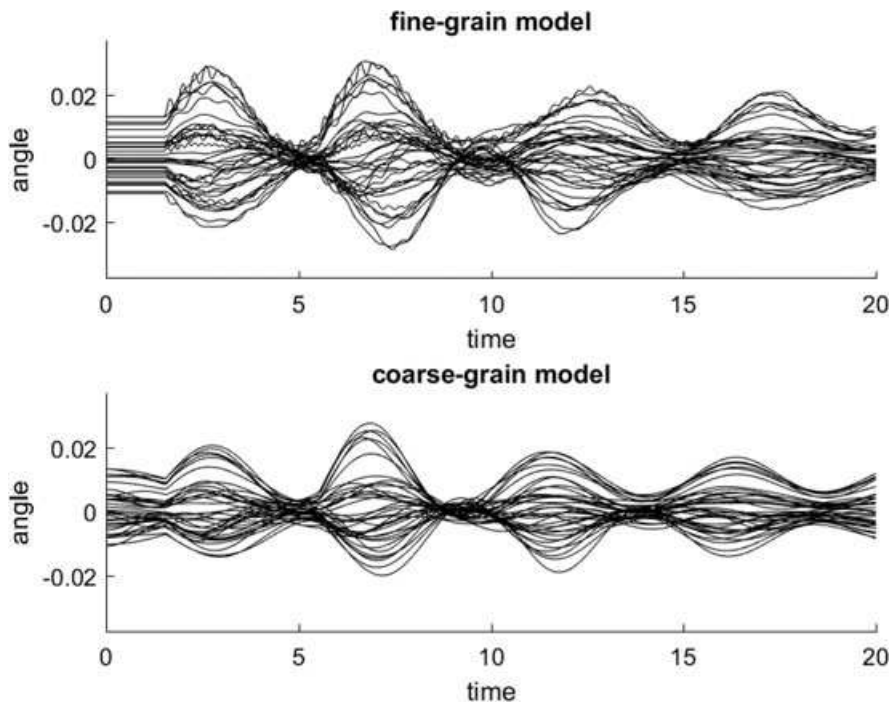


FIGURE 18 Tracking of coarse trajectories using seven multigrid levels with caliber-4 interpolation and $d_{\alpha,\beta}$ measure (26) to construct coarse-grain model (44). The statistics for this realization are max = 1.45, min = 0.18, mean1 = 0.53, and std = 0.27

to the damping coefficients of the coarse generators, caliber-5 interpolation with affinity measure $d_{\alpha,\beta}$ is used, and six multilevels are applied. Figure 19 shows the statistics: the mean is roughly 0.9 and the standard deviation indicates that the error is large for a few trajectories. Figure 20 is a sample realization with mean = 0.87 and std = 0.85, although the qualitative behavior of the dynamics is tracked but not as well as the generator-only cases. These results again motivate further investigation and development of methods to coarsen the damping coefficients of the fine load buses to contribute to the damping coefficients of the coarse generator buses.

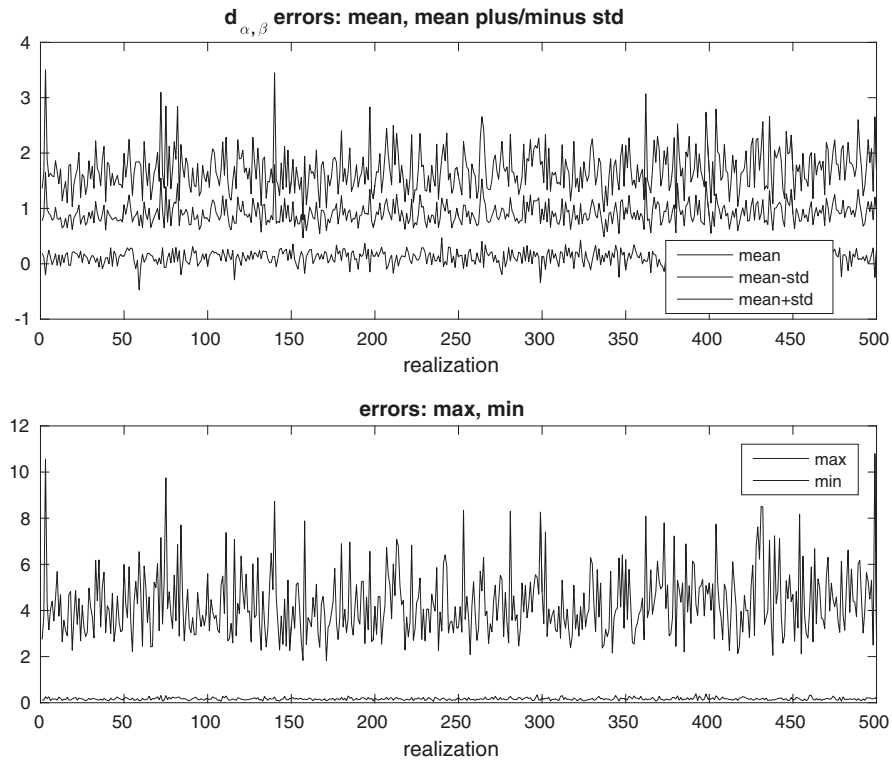


FIGURE 19 Tracking of coarse trajectories using six multigrid levels with caliber-5 interpolation and $d_{\alpha,\beta}$ measure (26) to construct the coarse-grain model for (50)

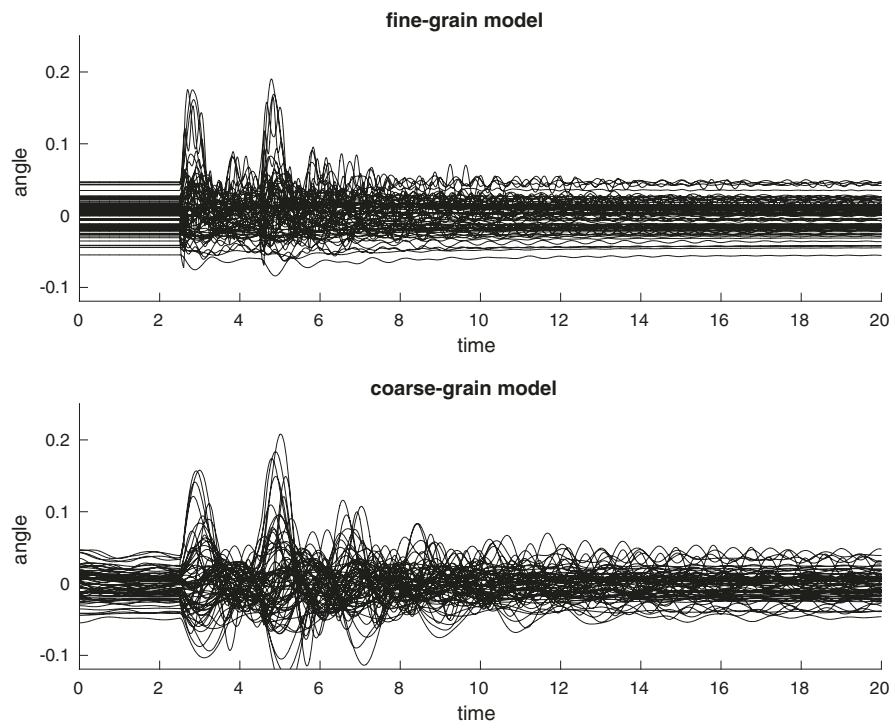


FIGURE 20 Tracking of coarse trajectories using six multigrid levels with caliber-5 interpolation and $d_{\alpha,\beta}$ measure (26) to construct the coarse-grain model for (50). Fine grain: 592 generators, 208 loads; coarse grain: 47 generators, 24 loads; statistics: max = 4.73, min = 0.16, mean = 0.87, and std = 0.85

7 | CONCLUSION

In this paper, we have presented a multigrid methodology for developing stable reduced models for power grid dynamical systems, which are characteristically different from equations that are traditionally handled by multigrid. One crux of this methodology is to exploit the graph Laplacian of these systems. Recently developed multigrid methods for graph Laplacians are then employed to determine the coarse DOFs and to form the coarse Laplacian for the reduced models, and the generated restriction operators are used to restrict the inertial and damping coefficients. However, these procedures alone do not lead to stable reduced models. In addition, some of the system structures of the fine-grain systems must be preserved in the reduced models. Numerical experiments were conducted to validate the effectiveness of this methodology.

Immediate future directions of the research include testing this methodology on real large-scale power grid systems and exploring techniques for restricting the load damping coefficients to coarse generator buses.

ACKNOWLEDGEMENT

The author would like to thank two anonymous referees for their comments and suggestions that help clarify the exposition of the technique and models.

ORCID

B. Lee  <http://orcid.org/0000-0003-1948-7384>

REFERENCES

1. Chow JH, Cullum J, Willoughby RA. A sparsity-based technique for identifying slow-coherent areas in large power systems. *IEEE Trans Power Appar Syst*. 1984;103(3):463–473.
2. Kundur P. *Power system stability and control*. New York, NY: McGraw-Hill; 1994.
3. Machowski J, Bialek JW, Bumby JR. *Power system dynamics: Stability and control*. 3rd ed. Chichester, UK: John Wiley & Sons; 2008.
4. Dimo P. *L'analyse nodale des réseaux d'énergie*. Paris, France: Eyrolles; 1971.
5. Zhukov LA. Simplified transformation of circuit diagrams of complex electric power systems. *Izvestia Akademii Nauk USSR Energ Transp*. 1964;2.
6. Barocio E, Pal BC, Thornhill NF, Messina AR. A dynamic mode decomposition framework for global power system oscillation analysis. *IEEE Trans Power Syst*. 2015;30(6):2902–2912.
7. Dörfler F, Bullo F. Synchronization and transient stability in power networks and non-uniform Kuramoto oscillators. *SIAM J Control Optim*. 2012;50:1616–1642.
8. Dörfler F, Chertkov M, Bullo F. Synchronization in complex oscillator networks and smart grids. *Proc Natl Acad Sci*. 2013;110(6):2005–2010.
9. Ron D, Saftro I, Brandt A. Relaxation-based coarsening and multiscale graph organization. *Multiscale Model Simul*. 2011;9:407–423.
10. Livne OE, Brandt A. Lean algebraic multigrid (LAMG): fast graph Laplacian linear solver. *SIAM J Sci Comput*. 2012;34:B499–B522.
11. John E, Saftro I. Single- and multi-level network sparsification by algebraic distance. *J Complex Netw*. 2016;5:352–388.
12. Dörfler F, Bullo F. On the critical coupling for Kuramoto oscillators. *SIAM J Appl Dyn Syst*. 2011;10:1070–1099.
13. Dörfler F, Bullo F. Exploring synchronization in complex oscillator networks. Paper presented at: 2012 IEEE 51st IEEE Conference on Decision and Control (CDC); 2012; Maui, HI. Piscataway, NJ: Institute of Electrical and Electronics Engineers; 2012.
14. Ruge JW, Stüben K. Algebraic multigrid. In: McCormick SF, editor. *Multigrid methods*. Philadelphia, PA: Society for Industrial and Applied Mathematics; 1987. *Frontiers in applied mathematics*.
15. Trottenberg U, Oosterlee CW, Schuller A. *Multigrid*. Cambridge, MA: Academic Press; 2001.
16. Livne OE. Coarsening by compatible relaxation. *Numer Linear Algebra Appl*. 2004;11:205–227.
17. Coifman RR, Lafon S. Diffusion maps. *Appl Comput Harmon Anal*. 2006;20:5–30.
18. Wang X, Pournaras E, Kooij RE, Van Mieghem P. Improving robustness of complex networks via the effective graph resistance. *Eur Phys J B*. 2014;87. <https://doi.org/10.1140/epjb/e2014-50276-0>
19. Falgout RD, Vassilevski PS. On generalizing the algebraic multigrid framework. *SIAM J Numer Anal*. 2004;42:1669–1693.

How to cite this article: Lee B. Multigrid for model reduction of power grid networks. *Numer Linear Algebra Appl*. 2018;e2201. <https://doi.org/10.1002/nla.2201>