Check for updates

Received: 22 January 2020 | Revised: 5 October 2020 | Accepted: 13 October 2020

DOI: 10.1002/nla.2347

RESEARCH ARTICLE

WILEY

Algebraic multigrid for the nonlinear powerflow equations

Barry Lee | **Enrique Pereira Batista**

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

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

Funding information
National Science Foundation, Grant/Award Number: DMS-1734727

Abstract

In a recent article, one of the authors developed a multigrid technique for coarse-graining dynamic powergrid models. A key component in this technique is a relaxation-based coarsening of the graph Laplacian given by the powergrid network and its weighted graph, which is represented by the admittance matrix. In this article, we use this coarsening strategy to develop a multigrid method for solving a static system of nonlinear equations that arises through Ohm's law, the so-called powerflow equations. These static equations are tightly knitted to the dynamic model in that the full powergrid model is an algebraic-differential system with the powerflow equations describing the algebraic constraints. We assume that the dynamic model corresponds to a stable operating powergrid, and thus, the powerflow equations are associated with a physically stable system. This stability permits the coarsening of the powerflow equations to be based on an approximate graph Laplacian, which is embedded in the powerflow system. By algebraically constructing a hierarchy of approximate weighted graph Laplacians, a hierarchy of nonlinear powerflow equations immediately becomes apparent. This latter hierarchy can then be used in a full approximation scheme (FAS) framework that leads to a nonlinear solver with generally a larger basin of attraction than Newton's method. Given the algebraic multigrid (AMG) coarsening of the approximate Laplacians, the solver is an AMG-FAS scheme. Alternatively, using the coarse-grid nodes and interpolation operators generated for the hierarchy of approximate graph Laplacians, a multiplicative-correction scheme can be derived. The derivation of both schemes will be presented and analyzed, and numerical examples to demonstrate the performance of these schemes will be given.

KEYWORDS

FAS, graph Laplacians, multigrid, multiplicative correction, powerflow equations, relaxation-based coarsening

1 | INTRODUCTION

A power grid is composed of generators (sources) and loads (sinks), referred to as buses, all connected through a network of transmission lines. Mathematically, and abstractly, it is represented as a system of differential-algebraic equations (DAEs) defined on the graph with the buses being the nodes and the transmission lines being the edges:

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

 © 2020 John Wiley & Sons, Ltd.

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

where $\mathbf{x}$ contains the differential variables describing the dynamics of the generators and $\mathbf{y}$ contains the algebraic variables governed by algebraic constraints. To provide a concrete example of this system, assume that the powergrid network is composed of n buses of which m are generators, $(n-m-1)$ are loads, and 1 is a so-called slack bus, which is used as a point of reference ("origin") for the system. Also, let N_g and N_l represent the set of nodes corresponding to the generator and load buses, respectively. Then a simple model describing a dynamic powergrid is

$$\mathbf{H}_i\ddot{\delta}_i + \mathbf{D}_i\dot{\delta}_i + \mathbf{P}_{e,i} - \mathbf{P}_{m,i} = 0 \quad i \in N_g, \tag{3}$$

$$\mathbf{P}_i = \sum_{j=1, j\neq i}^{n} \hat{\mathbf{V}}_i\hat{\mathbf{V}}_j\mathbf{B}_{ij}\sin(\delta_i - \delta_j) \quad i \in N_g \cup N_l, \tag{4}$$

$$\mathbf{Q}_i = -\hat{\mathbf{V}}_i^2\mathbf{B}_{ii} - \sum_{j=1, j\neq i}^{n} \hat{\mathbf{V}}_i\hat{\mathbf{V}}_j\mathbf{B}_{ij}\cos(\delta_i - \delta_j) \quad i \in N_l, \tag{5}$$

$$\hat{\mathbf{V}}_i = 1 \quad i \in N_g. \tag{6}$$

Here, $\hat{\mathbf{V}}_i$ and δ_i are respectively the voltage magnitude and phase angle for bus i; $\mathbf{H}_i$, $\mathbf{D}_i$, $\mathbf{P}_{e,i}$, and $\mathbf{P}_{m,i}$ are respectively the inertia, damping, generated electric power output, and mechanical power of generator i; P_i and Q_i are the real and imaginary components of the injected power at bus i; and $\mathbf{B}_{ij}$ is the ijth element of the admittance matrix. More details of this general model will be given later. By introducing the speed/frequency variable $\omega_i = \dot{\delta}_i$, Equation (3) can be converted into a first-order differential system, and after a minor rearrangement of the terms, DAE form (1)-(2) is obtained:

$$\begin{pmatrix} \dot{\delta}_i \\ \dot{\omega}_i \end{pmatrix} = \begin{pmatrix} \omega_i \\ -\mathbf{H}_i^{-1}[\mathbf{D}_i\omega_i - \mathbf{P}_{e,i} + \mathbf{P}_{m,i}] \end{pmatrix} \quad i \in N_g,$$

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{cases} \mathbf{P}_i - \displaystyle\sum_{j=1, j\neq i}^{n} \hat{\mathbf{V}}_i\hat{\mathbf{V}}_j\mathbf{B}_{ij}\sin(\delta_i - \delta_j) & i \in N_g \cup N_l \\ \mathbf{Q}_i + \hat{\mathbf{V}}_i^2\mathbf{B}_{ii} + \displaystyle\sum_{j=1, j\neq i}^{n} \hat{\mathbf{V}}_i\hat{\mathbf{V}}_j\mathbf{B}_{ij}\cos(\delta_i - \delta_j) & i \in N_l \\ \hat{\mathbf{V}}_i - 1 & i \in N_g. \end{cases}$$

In this form, the differential variables are (δ_i, ω_i), and as will be explained later, the algebraic variables are $(\delta_i, \mathbf{P}_i)$ for $i \in N_g$ and $(\delta_i, \hat{\mathbf{V}}_i)$ for $i \in N_l$ (note that for the generator buses, δ_i must satisfy the differential and algebraic equations).

The algebraic equations in the powergrid model are called the powerflow equations, for example, Equations (4)–(6) in the above example. To numerically solve the system of DAEs, a common approach is to process the powerflow and differential subsystems separately by lagging one of them, for example, repeatedly solving the powerflow equations first and then time-marching the differential equations for a while, at which an updated powerflow subsystem is solved. Solving the powerflow equations is thus a relevant component in powergrid engineering. In fact, in addition to dynamic powergrid modeling, the numerical solution of these algebraic equations is required in other powergrid analysis, for example, voltage stability calculations, where the loads are stressed in order to determine how much the powergrid system can safely handle without breaching the system (see References 1,2 where the voltage stability is examined for the North American Western Interconnection [WECC]).

Numerical methods for solving the powerflow equations have been investigated since the mid-1950s.[3-6] The first methods were the basic nonlinear Gauss–Seidel and successive overrelaxation iterations. Advancement over the basic schemes quickly followed a decade later when Newton's method was applied to these equations.[7] It can be accurately stated that Newton's method is the method of choice in the powergrid community, and because of this, most of the development in powerflow solvers has been limited to strategies for improving the computational efficiency and robustness of Newton's method/strategies and its variants. A relevant variant that was developed early on is the fast-decoupled powerflow method.[8] This is a quasi-Newton method where, under some powergrid assumptions, the Jacobian system for the full Newton's method is replaced with a system that decouples the corrections to the algebraic variables δ_i and $\hat{\mathbf{V}}_i$. A relevant strategy is the continuation/homotopy approach, where a difficult powerflow system is solved

by consecutively applying Newton's method to a sequence of systems, starting with an easy system and progressively increasing the complexity of the system until the target system is reached (see References 2,9 where a continuation method was used in a voltage stability analysis). A more recent and natural strategy is to utilize high-performance computers (HPC) for faster wall-clock computation (see Reference 10 for a description of an HPC framework for powergrid applications, which includes solving the powerflow equations with Newton's method). HPC has also been used to solve the linear Jacobian systems in Newton's iteration. In Reference 11, a domain decomposition preconditioner is used in a GMRES iteration to solve these linear systems. This preconditioner lends itself well to parallel efficiency on HPC.

It is worthwhile noting some novel research in the solution of the powerflow equations. One approach is based on constructive algebra, that is, Groebner basis techniques. As will be seen in Section 2, the powerflow equations can be expressed as a quadratic polynomial matrix system (see Equation 9). Because of this, constructive algebra techniques for polynomial matrices have been used to analytically solve and analyze powerflow systems.[12,13] However, given the immense computational cost in solving large polynomial matrices, only small systems can be handled with this approach. Thus, this approach is more of a tool for qualitative analysis of the powerflow equations. Another approach, which is more tractable but still limited, formulates the powerflow equations as an optimization problem involving a monotone variational inequality.[14] Semidefinite programming algorithms then are used to solve the optimization problem. Again, the method has been applied only to small test systems (IEEE 14-bus and 39-bus scenarios).

There has been limited research in multigrid methods for the powerflow equations and powergrid problems in general. An initial attempt was to use AMG as a preconditioner for solving the Jacobian systems in Newton's method.[15] However, even then it was recognized that AMG cannot be directly applied as a preconditioner to the Jacobian system, but has to be applied to a modified linear system that more closely satisfies the requirements of AMG. It should be noted that this approach does not apply multigrid directly to the powerflow equations. Indeed, multigrid for solving these equations faces several barriers. First, these equations are purely algebraic in nature without having any elliptic partial differential equation (PDE) structures, from which multigrid strategies can be extracted. Specifically, elliptic PDE features can reveal multiscale structures that can be used to guide the construction of the multigrid components, that is, intergrid operators and relaxation schemes. For the powerflow equations, the mathematical structures that can expose the multiscale features are not obvious, and hence, the construction of a multilevel hierarchy is not obvious. Of course, a hierarchy can be constructed from the historical evolution of the power network. But given the many transformative engineering devices that have been incorporated into the network, such a hierarchy may not be sufficient for multigrid processing since the mathematical features necessary for fast multigrid performance may not exist in this hierarchy. One also can use the physical structure of the grid to partition it for a domain decomposition approach, albeit not exploiting multiscale features.

Progress in multigrid for the powergrid has been achieved only recently. In Reference 16, a multigrid-based coarse-graining of the dynamic powergrid was developed. This method can be applied to the differential equations of the DAEs. Moreover, by taking a frequency-dependent model for the loads,[17,18] the powerflow equations can be approximated with a system of differential equations, for which this coarse-graining approach can be applied to. Note that these differential equations are not elliptic, and thus, standard AMG techniques cannot be employed. Instead, this coarse-graining method applies a relaxation-based coarse grid selection scheme to a graph Laplacian that is embedded in the powerflow equations. The network characteristics of this graph Laplacian is given through the admittance matrix, which physically describes the conductances and susceptances of the transmission lines. This technique will lead to both generator and load degrees of freedom on the coarser levels, which physically makes more sense than eliminating all the loads on the coarse levels as is done in some engineering-based, coarse-graining approaches used in the powergrid community (see Reference 18 for other approaches), that is, a physical powergrid, especially a modernized one, should consists of both generators and loads. With the selection of coarse buses, interpolation operators are constructed, and in turn, a coarse-grain model is formed. Hence, the hierarchical structure of the powergrid is inherent in the system's associated graph Laplacian.

Progress in multigrid for powergrid applications was also achieved in Reference 19. This article presents the first *authentic* multigrid method for the powerflow equations. The key feature of this method is the utilization of a multiplicative coarse-grid correction rather than the usual additive coarse-grid correction (see Reference 20,21 for more details on multiplicative-correction multigrid). This approach exploits the quadratic nature of the powerflow equations (see Equation 9) to derive equations for the multiplicative correction. As for the coarsening strategy, an aggregation scheme that aggregates nodes based on the entries of the admittance matrix is used. This multigrid method also relies on a number of powergrid specific approximations and assumptions, making the scheme rather complex and nonstandard, that is,

does not have "black-box" features in the sense that the multigrid solver can be generated from only the fine-grid system without application specifics that are exploited in the design of the multigrid components.

In this article, the method in Reference 16 will be extended to handle the algebraic powerflow equations with as little powergrid specifics as possible. In particular, since the admittance matrix can be viewed as an approximate graph Laplacian, the relaxation-based coarsening will be applied directly to the admittance matrix to produce a hierarchy of admittance matrices. This hierarchy is all that is needed to construct a hierarchy of powerflow equations. With the coarse-grid powerflow equations, a FAS approach,[22-28] can be readily developed to solve the powerflow system. However, notice that unlike most FAS research in the literature, the resulting FAS is applied to a nonlinear system that does not arise from a PDE. This leads to a distinguishing feature of the FAS method of this article: the coarse-grid problems are not derived from a PDE or its discretization.

As a comparison with the multiplicative correction scheme of Reference 19, this FAS method is less complex, essentially requiring only a technique to coarsen the admittance matrix. But this method does require an implicit assumption on the powergrid system: its admittance matrix must be an approximate graph Laplacian. This assumption allows us to apply the relaxation-based coarsening to generate a hierarchy of coarse degrees of freedom (CDOFs) and interpolation operators. This hierarchy, in fact, can be used to construct a multiplicative-correction scheme similar to the method of Reference 19. Like other multiplicative-correction methods, this scheme requires an updating of the coarse-grid operators after each multigrid cycle since these operators explicitly depend on the solution approximation (see Subsection 4.2). This updating is indispensable; omitting it can lead to divergence in the iteration. The added costs in these updates are the major disadvantage of the multiplicative-correction schemes. Indeed, this updating puts a computational premium on multigrid μ-cycles, $\mu > 1$.

Common features in both the FAS and multiplicative-correction schemes of this article are the coarse aggregrates consist of only the same bus type (i.e., loads or generators) and the coarse-grid corrections are intravariably interpolated (i.e., load coarse-grid corrections interpolate only to fine load buses, and generator coarse-grid corrections interpolate only to fine generator buses). Given the physical disparity between the loads and generators, these restrictions are appropriate. Furthermore, these restrictions eliminate the need to derive physically relevant equations for CDOFs associated with aggregates involving both types of buses if such aggregating were permitted.

Another common feature in both schemes is that the generator buses are selected as coarse nodes on all levels. This is done because voltage magnitude constraints are prescribed on these buses (Equation 6). By keeping the generator buses on all levels, these constraints are enforced on the coarsest level, which trivially implies they are enforced on the finest level. On the other hand, if some of the generator buses were not selected on the coarser levels, then a coarse-grid correction to the actual generator degrees of freedom (DOFs) can lead to constraint violations on the finer levels. Keeping these buses on all levels naturally leads to a degradation in the efficiency of the multigrid solver because the system cannot be coarsened down to a few buses. However, given that approximately only 10% of the buses are generators, the number DOFs on the coarsest grid can be kept to a reasonably small size. In fact, coarsening to very few DOFs may not lead to accurate and stable coarse-grid systems, particularly since these systems are nonlinear. Furthermore, given that the generators are often separated by long pathlength distances, keeping these buses on all levels eliminates the need to aggregrate generators that are connected through long sequences of transmission lines with loads in between the end-node generators of the path.

The article proceeds as follows. In Section 2, the powerflow equations will be described for both the polar and cartesian formulations. The latter formulation will allow clean multigrid processing when the admittance matrix is complex-valued rather than purely imaginary-valued. Given the important role the admittance matrix plays in the schemes to be developed, a description of it will be given in Subsection 2.1. In Section 3, multigrid for powergrids will be reviewed with emphasis placed on a relaxation-based coarsening of weighted graph Laplacians. In Section 4, the FAS and multiplicative-correction algorithms will be developed. Subsection 4.1 will examine the FAS scheme, particularly looking at the interpolation operator applied in the multigrid iteration (i.e., the solve phase) in relation to the interpolation operator generated for coarsening the admittance matrix. This comparison will reveal how strongly the complementary smoothing/coarse-grid principle is satisfied in the FAS method. Subsection 4.2 will examine this principle for the multiplicative-correction scheme which uses the intergrid operators generated for coarsening the admittance matrices in the solve phase. Finally, Subsection 4.3 will discuss the basin of attraction for Newton's method and the AMG-FAS scheme. In Section 5, numerical experiments for both methods will be presented to illustrate their effectiveness. In addition, experiments will be presented to compare the robustness of the AMG-FAS and Newton's methods to demonstrate the larger basin of attraction for the AMG-FAS method.

2 | THE POWERFLOW EQUATIONS

We assume the same powergrid setting as in Section 1: a total of n buses with m generators, $(n-m-1)$ loads, and 1 slack bus. Also, let E denote the set of edges of this interconnected grid and $\mathcal{G}(n,E)$ the graph. The powerflow equations are based on the nodal network equations (Ohm's law)

$$\mathbf{I} = \mathbf{YV}, \tag{7}$$

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. Since the properties of the admittance matrix is relevant to the coarsening algorithm, we will examine the details of it later. The power injected at node i is given by

$$\mathbf{V}_i \mathbf{I}_i^* := \mathbf{P}_i + \hat{\imath}\mathbf{Q}_i, \tag{8}$$

where the asterisk denotes the complex conjugate and the symbol $\hat{\imath}$ denotes the imaginary number $\sqrt{-1}$. The real ($\mathbf{P}_i$) and imaginary ($\mathbf{Q}_i$) components of the power are called the active and reactive powers, respectively. For the full system, the injected power is given by

$$\mathrm{diag}(\mathbf{V})(\mathbf{YV})^* = \mathbf{P} + \hat{\imath}\mathbf{Q}, \tag{9}$$

from which we observe a quadratic expression in $\mathbf{V}$. Separating the admittance matrix into its real and imaginary components, that is,

$$\mathbf{Y}_{ij} = \mathbf{G}_{ij} + \hat{\imath}\mathbf{B}_{ij},$$

and expressing the voltages in polar form, that is, $\mathbf{V}_i = \hat{\mathbf{V}}_i e^{\hat{\imath}\delta_i}$, we have

$$\mathbf{P}_i = \hat{\mathbf{V}}_i^2 \mathbf{G}_{ii} + \sum_{j=1, j\neq i}^{n} \hat{\mathbf{V}}_i \hat{\mathbf{V}}_j [\mathbf{B}_{ij} \sin(\delta_{ij}) + \mathbf{G}_{ij} \cos(\delta_{ij})], \tag{10}$$

$$\mathbf{Q}_i = -\hat{\mathbf{V}}_i^2 \mathbf{B}_{ii} + \sum_{j=1, j\neq i}^{n} \hat{\mathbf{V}}_i \hat{\mathbf{V}}_j [\mathbf{G}_{ij} \sin(\delta_{ij}) - \mathbf{B}_{ij} \cos(\delta_{ij})], \tag{11}$$

where $\hat{\mathbf{V}}_j$ is now the voltage magnitude at the jth 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 nearly symmetric matrices $\mathbf{G}$ and $\mathbf{B}$, and $\delta_{ij} := (\delta_i - \delta_j)$ with δ_i denoting the phase angle at node i. 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. With this categorization of the buses, the specified voltage magnitudes at the PV buses are explicitly included in the system as constraints. Assuming that the slack bus is located at node n, all the specified PV voltage magnitudes take a normalized per unit quantity, and the PV and PQ buses are respectively numbered from 1 to r and $(r+1)$ to $(n-1)$, the powerflow equations in polar form are

$$\mathbf{P}_i = \hat{\mathbf{V}}_i^2 \mathbf{G}_{ii} + \sum_{j=1, j\neq i}^{n} \hat{\mathbf{V}}_i \hat{\mathbf{V}}_j [\mathbf{B}_{ij} \sin(\delta_{ij}) + \mathbf{G}_{ij} \cos(\delta_{ij})] \quad i = 1, \ldots, n-1, \tag{12}$$

$$\mathbf{Q}_i = -\hat{\mathbf{V}}_i^2 \mathbf{B}_{ii} + \sum_{j=1, j\neq i}^{n} \hat{\mathbf{V}}_i \hat{\mathbf{V}}_j [\mathbf{G}_{ij} \sin(\delta_{ij}) - \mathbf{B}_{ij} \cos(\delta_{ij})] \quad i = r+1, \ldots, n-1, \tag{13}$$

$$\hat{\mathbf{V}}_i = 1 \quad i = 1, \ldots, r. \tag{14}$$

The unspecified quantities are determined by solving these powerflow equations. In many applications, and in most mathematics literature on powergrids, for example, References 17,29-31, the conductances are assumed negligible due to physical properties of the transmission lines. In these scenarios, the terms involving $\mathbf{G}_{ij}$ in (12)–(13) drop out to produce a mathematically analyzable system.[17,30,31] With regard to the example powerflow equations in Section 1,

negligible conductances were assumed and the PV and PQ buses were assumed to be generator buses and load buses, respectively.

For ease of computation in the developed multigrid algorithms, a cartesian form of the powerflow equations also will be needed. Decomposing the complex voltage $\mathbf{V}_i$ into its real and imaginary components $\mathbf{V}_i^{\mathrm{r}}$ and $\mathbf{V}_i^{\mathrm{im}}$, this system is

$$\mathbf{P}_i = \mathbf{V}_i^{\mathrm{r}} \sum_{j=1}^{n} (\mathbf{G}_{ij}\mathbf{V}_j^{r} - \mathbf{B}_{ij}\mathbf{V}_j^{\mathrm{im}}) + \mathbf{V}_i^{\mathrm{im}} \sum_{j=1}^{n} (\mathbf{B}_{ij}\mathbf{V}_j^{\mathrm{r}} + \mathbf{G}_{ij}\mathbf{V}_j^{\mathrm{im}}) \quad i = 1, \ldots, n-1, \tag{15}$$

$$\mathbf{Q}_i = -\mathbf{V}_i^{\mathrm{r}} \sum_{j=1}^{n} (\mathbf{B}_{ij}\mathbf{V}_j^{\mathrm{r}} + \mathbf{G}_{ij}\mathbf{V}_j^{\mathrm{im}}) + \mathbf{V}_i^{\mathrm{im}} \sum_{j=1}^{n} (\mathbf{G}_{ij}\mathbf{V}_j^{\mathrm{r}} - \mathbf{B}_{ij}\mathbf{V}_j^{\mathrm{im}}) \quad i = r+1, \ldots, n-1, \tag{16}$$

$$1 = (\mathbf{V}_i^{\mathrm{r}})^2 + (\mathbf{V}_i^{\mathrm{im}})^2 \quad i = 1, \ldots, r. \tag{17}$$

More engineering and physics details on the powerflow equations can be found in References 1,18,32.

2.1 | The admittance matrix

Because of the significant role the admittance matrix plays in the nonlinear multigrid solvers, this subsection describes the derivation of this matrix. Consider an edge $(i,j) \in E$ representing a transmission line between buses i and j. This line will have an impedance

$$z_{ij} = r_{ij} + \hat{\imath} x_{ij},$$

where r_{ij} and x_{ij} are respectively the resistance and reactance of the line. Although these quantities are generally nonnegative for actual transmission lines,[33] there are special cases where they can be negative but of small magnitude.[34] To see how this impedance contributes to the admittance matrix, by Ohm's laws, the current at node i is

$$\begin{aligned} \mathbf{I}_i &= \sum_{(i,j)\in E, j\neq i} \frac{\mathbf{V}_i - \mathbf{V}_j}{z_{ij}} \\ &= \sum_{(i,j)\in E, j\neq i} w_{ij}(\mathbf{V}_i - \mathbf{V}_j) = \sum_{(i,j)\in E, j\neq i} (\mathbf{G}_{ij} + \hat{\imath}\mathbf{B}_{ij})(\mathbf{V}_i - \mathbf{V}_j), \end{aligned} \tag{18}$$

where w_{ij} is the reciprocal of z_{ij} giving $w_{ij} := (\mathbf{G}_{ij} + \hat{\imath}\mathbf{B}_{ij})$, the conductance and susceptance for edge (i,j). We see that if the resistance and reactance are positive, then $\mathbf{G}_{ij}$ and $\mathbf{B}_{ij}$ will have opposite signs, and that even if the resistance and reactance are small in magnitude (in particular, the negative resistance and reactance case), $\mathbf{G}_{ij}$ and $\mathbf{B}_{ij}$ can have large magnitudes. Expressing (18) in terms of the admittance matrix, denoted by $\mathbf{Y}^1$, we have

$$\mathbf{I}_i = \sum_{j=1}^{n} \mathbf{Y}_{ij}^1 \mathbf{V}_j,$$

where we see that $\mathbf{Y}^1$ is a complex graph Laplacian, that is,

$$\mathbf{Y}_{ij}^1 = \begin{cases} \sum\limits_{(i,j)\in E, j\neq i} w_{ij} & i = j \\ -w_{ij} & i \neq j \\ 0 & (i,j) \notin E. \end{cases}$$

$\mathbf{Y}^1$ is the simplest form of the admittance matrix. A realistic power system may include charging capacitances, transformers, phase shifters, and shunts, all of which are needed for the transmission of electricity through the power network. These devices are placed along a fraction of the edges or nodes in the power system (see Table 1, which provides the number of transformers and phase shifters in some real-world power systems). To see how these devices change the admittance matrix of the system, we consider continuously adding one of the different types of devices along edge (i,j) or node i. The

TABLE 1 Realistic power grid systems: Average number of $V(1,1)$ AMG-FAS and multiplicative-correction cycles to reduce the initial residual by six orders of magnitude

Case	#Buses	#Gens	#Transf/#pshifters	FAS its	MultCorr its
IEEE 57	57	7	15/0	11	10
IEEE 68	68	15	16/0	9	11
IEEE 118	118	54	9/0	9	12
IEEE 145	145	49	52/0	10	11
IEEE 300	300	68	62/0	32	48
Illinois	200	48	0/0	16	17
S. Carolina	500	89	0/0	17	19
Wisconsin	1664	77	0/0	15	22
Texas	2007	281	0/0	12	16
Half Europe	1354	74	234/6	20	20
Polish-1999	2383	326	170/6	12	16
Polish-2004	2746	381	171/1	16	20
Mod. Polish	3374	441	383/2	21	24
Mod. Europe	9241	1445	1339/66	24	28

Note: Caliber 3 interpolation was used, and the average was taken over 10 simulations.

modifications to the corresponding rows and columns of the admittance matrix are indicative of the changes to other rows and columns that correspond to edges and nodes that possess these devices.

So, first consider a charging capacitance along edge (i,j). The modified admittance matrix is

$$\begin{bmatrix} \mathbf{Y}^1_{11} & \cdots & \mathbf{Y}^1_{1i} & \cdots & \mathbf{Y}^1_{1j} & \cdots & \mathbf{Y}^1_{1n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{Y}^1_{i1} & \cdots & \left(\mathbf{Y}^1_{ii} + \hat{\imath}\frac{b_{ij}}{2}\right) & \cdots & \mathbf{Y}^1_{ij} & \cdots & \mathbf{Y}^1_{in} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \mathbf{Y}^1_{j1} & \cdots & \mathbf{Y}^1_{ji} & \cdots & \left(\mathbf{Y}^1_{jj} + \hat{\imath}\frac{b_{ij}}{2}\right) & \cdots & \mathbf{Y}^1_{jn} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{Y}^1_{n1} & \cdots & \mathbf{Y}^1_{ni} & \cdots & \mathbf{Y}^1_{nj} & \cdots & \mathbf{Y}^1_{nn} \end{bmatrix}, \tag{19}$$

where b_{ij} is a nonnegative value. Let $\mathbf{Y}^2$ denote the modified matrix that results from all the charge capacitances in the system. Next, assuming that there is also a transformer between buses i and j having a phase shift ϕ and tap ratio τ (a ratio that changes the magnitude of the voltage, currents, impedance between nodes i and j), the admittance matrix becomes

$$\begin{bmatrix} \mathbf{Y}^1_{11} & \cdots & \mathbf{Y}^1_{1i} & \cdots & \mathbf{Y}^1_{1j} & \cdots & \mathbf{Y}^1_{1n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{Y}^1_{i1} & \cdots & \left[\sum_{(i,k)\in E, k\neq j} w_{ik} + \left(w_{ij} + \hat{\imath}\frac{b_{ij}}{2}\right)\frac{1}{\tau^2}\right] & \cdots & \mathbf{Y}^1_{ij}\frac{1}{\tau e^{-\hat{\imath}\phi}} & \cdots & \mathbf{Y}^1_{in} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \mathbf{Y}^1_{j1} & \cdots & \mathbf{Y}^1_{ji}\frac{1}{\tau e^{\hat{\imath}\phi}} & \cdots & \left(\mathbf{Y}^1_{jj} + \hat{\imath}\frac{b_{ij}}{2}\right) & \cdots & \mathbf{Y}^1_{jn} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{Y}^1_{n1} & \cdots & \mathbf{Y}^1_{ni} & \cdots & \mathbf{Y}^1_{nj} & \cdots & \mathbf{Y}^1_{nn} \end{bmatrix}. \tag{20}$$

Let $\mathbf{Y}^3$ denote the modified matrix resulting from all the charge capacitances, transformers, and phase shifters along specific edges of the system, possibly with some of these edges having only a selection of these devices (e.g., if edge (i,j) does not involve a charging capacitance, then $b_{ij}=0$ in (20)). Finally, assuming a shunt at node i also, the admittance matrix is

$$\begin{bmatrix} \mathbf{Y}^1_{11} & \cdots & \mathbf{Y}^1_{1i} & \cdots & \mathbf{Y}^1_{1j} & \cdots & \mathbf{Y}^1_{1n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{Y}^1_{i1} & \cdots & \left[\sum_{(i,k)\in E, k\neq j} w_{ik} + \left(w_{ij} + \hat{\imath}\frac{b_{ij}}{2}\right)\frac{1}{\tau^2} + \alpha_i + \hat{\imath}\beta_i\right] & \cdots & \mathbf{Y}^1_{ij}\frac{1}{\tau e^{-\hat{\imath}\phi}} & \cdots & \mathbf{Y}^1_{in} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \mathbf{Y}^1_{j1} & \cdots & \mathbf{Y}^1_{ji}\frac{1}{\tau e^{\hat{\imath}\phi}} & \cdots & \left(\mathbf{Y}^1_{jj} + \hat{\imath}\frac{b_{ij}}{2}\right) & \cdots & \mathbf{Y}^1_{jn} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{Y}^1_{n1} & \cdots & \mathbf{Y}^1_{ni} & \cdots & \mathbf{Y}^1_{nj} & \cdots & \mathbf{Y}^1_{nn} \end{bmatrix}, \tag{21}$$

where α_i is nonnegative and β_i can be positive or negative. Let $\mathbf{Y}^4$ denote the modified matrix due to all of the charge capacitances, transformers, phase shifters, and shunts in the system, possibly with some of the edges involving only some of these devices.

In the remainder of the article, $\mathbf{Y}$ will denote the admittance matrix for either $\mathbf{Y}^1$, $\mathbf{Y}^2$, $\mathbf{Y}^3$, or $\mathbf{Y}^4$. An important feature is that $\mathbf{Y}$ is almost symmetric when there are a few phase shifters. Moreover, notice that when there are a few transmission lines with negative resistances and reactances, the rows and columns of $\mathbf{Y}$ that correspond to these lines will have positive and negative entries. Unfortunately, this can eliminate possible diagonal dominance in $\mathbf{Y}$ if these resistances and reactances were instead positive. In Subsection 5.1, we will see how this can have detrimental effects on the relaxation for such admittance matrices.

3 | COARSENING OF THE POWERGRID

With the admittance matrix described and the powerflow equations given by (12)–(14) or (15)–(17), we now consider developing a nonlinear multigrid solver for these equations. However, unlike most systems that multigrid has been successfully applied to, these equations are purely algebraic. In particular, these equations do not correspond to an elliptic PDE, which can provide the intrinsic hierarchical structures in the problem and thus be used to guide the construction of the coarse-grid problems. Furthermore, since these equations are nonlinear, an AMG approach is not directly applicable (cf., see Reference 27 where an AMG approach is developed for a nonlinear PDE exploiting the PDE and its discretization structures, that is, the exact de Rham sequence). To expose the hierarchical structures, the full DAE system describing the powergrid will be considered.

To this end, assume that the conductances are negligible ($\mathbf{G}_{ij}=0$). The differential equations in the DAEs are defined on the generator buses and are given by (3), which we repeat for convenience

$$\mathbf{H}_i\ddot{\delta}_i + \mathbf{D}_i\dot{\delta}_i + \mathbf{P}_{e,i} - \mathbf{P}_{m,i} = 0 \quad i \in N_g. \tag{22}$$

Now, since the conductances are negligible, the electric power is given by

$$\mathbf{P}_{e,i} = \sum_{j=1}^{n} \hat{\mathbf{V}}_i\hat{\mathbf{V}}_j\mathbf{B}_{ij}\sin(\delta_{ij}). \tag{23}$$

As for the algebraic equations in the DAEs, they are the powerflow equations. Consider (12) at the load buses, where the active power is drawn because the loads are sinks in the system. These algebraic equations can be converted to differential equations by taking a frequency-dependent model for the drawn power:[17,18] for small frequency variations about $\mathbf{P}^o_i$, a so-called stable operating point,

$$\mathbf{P}_i = -\mathbf{P}^o_i - \mathbf{D}_i\dot{\delta}_i, \quad \mathbf{D}_i > 0, \quad i \in N_l. \tag{24}$$

Substituting (23) into (22) and (24) into (12) leads to

$$\mathbf{H}_i\ddot{\delta}_i + \mathbf{D}_i\dot{\delta}_i + \sum_{j=1}^{n} a_{ij}\sin(\delta_{ij}) = \mathbf{P}_{m,i} \quad i \in N_g, \tag{25}$$

$$\mathbf{D}_i\dot{\delta}_i + \sum_{j=1}^{n} a_{ij}\sin(\delta_{ij}) = -\mathbf{P}_i^o \quad i \in N_l, \tag{26}$$

where $a_{ij} = \hat{\mathbf{V}}_i\hat{\mathbf{V}}_j\mathbf{B}_{ij}$.

In Reference 16, a coarse-graining technique based on multigrid was developed for dynamic powergrid models of the form (25)-(26). In these models, the $\hat{\mathbf{V}}_i$ and $\hat{\mathbf{V}}_j$ in a_{ij} are assumed to be constant so that the unknowns are only the phase angles. In practice, this assumption is reasonable since the DAEs are numerically processed by first solving the powerflow equations for the voltage magnitudes, angles, and possibly $\mathbf{P}_i^o$, with the resulting angles taken as the initial conditions for the differential equations with the voltage magnitudes held constant. These differential equations then are integrated for a time interval, after which the powerflow equations are solved for an updated set of voltage magnitudes and angles, and the integration process repeated.

The hierarchical structure in (25)-(26) is hidden in a graph Laplacian that reveals the stability of a dynamic system. To intuitively/physically see this, one can view the dynamic powergrid as an oscillating mass–spring system with all the masses connected to a base and to each other. When the strength of the connections to each other are sufficiently strong, a "stable" scenario is obtained in that the frequencies of the oscillations will synchronize and the phases will be cohesive;[17,29-31] when the connections are weak, an "unstable" scenario is obtained where each mass oscillates close to its own natural frequency and the phases generally will be disordered. Mathematically, the different scenarios are determined by the equilibrium points of dynamic system (25)-(26) (see Reference 35 for a background on the stability of dynamic systems). These points are described through the graph Laplacian defined by the coefficients a_{ij}.

Let $\mathbf{L}_f$ (subscript f denotes the fine grid) be the graph Laplacian with elements

$$\mathbf{L}_{f,ij} = \begin{cases} -a_{ij} & i \neq j \\ \sum\limits_{j\neq i} a_{ij} & i = j. \end{cases} \tag{27}$$

This graph Laplacian was used in Reference 16 to guide the coarse-graining of a stable dynamic powergrid, that is, to construct a stable coarse-grain model. In particular, this coarsening is based on a synchronization requirement on the dynamic model. Specifically, the solution $\boldsymbol{\delta}$ of (25)-(26) is said to be phase cohesive if there exists a $\gamma \in [0, \pi)$ such that

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

The solution is also said to be frequency synchronized if all the frequencies $\dot{\delta}_i$ converge to a common frequency as $t \to \infty$. Following Reference 31, a powergrid system is said to be synchronized if it is phase cohesive and frequency synchronized with $\gamma < \frac{\pi}{2}$. Like in the spring-mass analogy, synchrony in the dynamic powergrid system implies stability. For example, physically the frequencies are the same over the whole network (e.g., in North America, the common household current is 60 Hz) and the phases are cohesive when the powergrid is functioning normally. However, if over part of the network, the frequencies are asynchronous and the phases are incohesive, a wide-area blackout can be triggered.

It is clear from (25) and (26) that the bus angles couple only through the network term $\sum_{j=1}^{n} a_{ij}\sin(\delta_{ij})$. Consider the system

$$\sum_{j=1}^{n} a_{ij}\sin\delta_{ij} = \tilde{\mathbf{P}}_i \quad i \in N_g \cup N_l, \tag{28}$$

where

$$\tilde{\mathbf{P}}_i = \begin{cases} \mathbf{P}_{m,i} & i \in N_g \\ -\mathbf{P}_i^o & i \in N_l. \end{cases}$$

The solutions of this system are the equilibrium points of (25)-(26). With $\mathbf{L}_f^\dagger$ denoting the pseudoinverse of $\mathbf{L}_f$, if

$$\|\mathbf{L}_f^\dagger \tilde{\mathbf{P}}\|_{E,\infty} \leq \sin\gamma, \tag{29}$$

where $\tilde{\mathbf{P}} = (\tilde{\mathbf{P}}_i)_{i=1}^n$, $\|\mathbf{x}\|_{E,\infty} = \max_{(i,j)\in E}|x_i - x_j|$, and $\gamma < \frac{\pi}{2}$, then for almost all powergrid networks, there exists a solution δ_{ij} of (28) that satisfies

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

see Reference 31, and also References 17,29,30. This solution implies that (25)-(26) is stable.

Thus, to construct a coarse-grain model of a stable powergrid, ideally the coarse buses should be chosen to ensure an accurate coarse-grid approximation of $\mathbf{L}_f$ with the resulting coarse-grid Laplacian satisfying an analogous synchrony condition (29). Implicit in this requirement is that the coarse-grid Laplacian is a good approximation $\mathbf{L}_f$. Thus, in Reference 16, the coarse buses are selected to ensure good approximations to the eigenvectors of $\mathbf{L}_f$ corresponding to the smaller eigenvalues, since $\mathbf{L}_f^\dagger \tilde{\mathbf{P}}$ is often dominated by these eigenvectors. Checking whether an analogous synchrony condition holds on the coarse-grid can be performed as an optional second step. To realize the coarse bus selection, a relaxation-based AMG procedure was used.[36-38]

Turning to powerflow equations (12)–(14) or (15)–(17), assuming that the powerflow equations are associated with a stable powergrid (otherwise it would be questionable why such a powergrid should even be solved), synchrony condition (29) will guide the coarsening procedure of the powerflow equations. Specifically, the coarsening will be guided by the implicit condition that an accurate coarse-grid Laplacian must be obtained. The actual coarsening procedure will be applied to the admittance matrix, which may only be an approximate graph Laplacian with nonzero row-sum. Applying a relaxation-based procedure on the admittance matrix will produce a set of coarse buses, for which a coarse admittance matrix can be formed. As will be seen in Section 4, this coarse matrix immediately leads to a coarse powerflow problem for use in a nonlinear FAS framework.

Notice that the coarsening procedure is applied directly to the admittance matrix rather than to $[\hat{\mathbf{V}}_i \mathbf{Y}_{ij}^* \hat{\mathbf{V}}_j^*]_{i,j=1}^n$ or its associated graph Laplacian. This is done because the $\hat{\mathbf{V}}_i$'s at the load buses are unknown in the powerflow equations. If the coarsening were based on $[\hat{\mathbf{V}}_i \mathbf{Y}_{ij}^* \hat{\mathbf{V}}_j^*]_{i,j=1}^n$, then a new coarsening must be conducted after each FAS iteration since $\hat{\mathbf{V}}$ would have been updated. Coarsening on the admittance matrix equivalently corresponds to coarsening $[\hat{\mathbf{V}}_i \mathbf{Y}_{ij}^* \hat{\mathbf{V}}_j^*]_{i,j=1}^n$ only at the initial approximation $\hat{\mathbf{V}}^0 = \mathbf{1}$. This means that the coarsening is performed only once in the FAS algorithm.

3.1 | Relaxation-based coarsening

The goal of the coarsening scheme is to determine the buses that will enable good coarse-grid approximation to the algebraically smooth errors that cannot be efficiently resolved on the fine grid. Since there is a large number of these errors, this goal is achieved by determining the buses that will allow good coarse-grid approximation to the algebraically smooth eigenvectors of the graph Laplacian or susceptance matrix (i.e., $\mathbf{Y} = \hat{\imath}\mathbf{B}$), which accurately span these errors. These buses are determined using a relaxation-based procedure and a measure that quantifies the appropriateness of each bus to allow an accurate approximation of the algebraically smooth eigenvectors. Such relaxation procedure is used because the algebraically smooth modes are often poorly attenuated by relaxation. The performance of the relaxation thus can itself expose good candidates for the coarse buses (see References 36-41 and references listed therein. In the following, to select the buses, we use the algorithm in Reference 38 but modified by using the correlation-related measure of Reference 37 since this measure has physical relevance in the power system. We describe the algorithm for a graph Laplacian, although it also can be applied to the admittance matrix.

Hence, consider applying the relaxation to the homogeneous problem

$$\mathbf{L}_f \mathbf{x} = \mathbf{0} \tag{30}$$

with a random initial guess. After several sweeps, the resulting approximation will be dominated by these algebraically smooth eigenvectors. By repeating this process for a collection of random initial guesses that are each orthogonal to the

previously relaxed vectors, a good approximation to these eigenvectors can be obtained. From this collection of relaxed vectors, the coarse buses are selected based on a correlation-related measure.[37] Let $\{\mathbf{x}^l\}_{l=1}^K$ be the set of relaxed vectors, let α, β denote two arbitrary components of the vector $\mathbf{x}^l$, and define the nodal inner product

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

This measure is

$$c_{\alpha\beta} = \frac{|(\mathbf{x}_\alpha, \mathbf{x}_\beta)|^2}{(\mathbf{x}_\alpha, \mathbf{x}_\alpha)(\mathbf{x}_\beta, \mathbf{x}_\beta)}, \tag{31}$$

which has physical relevance: correlations between buses can reveal coherency patterns. For a given threshold θ, for example, $0.5 \leq \theta < 1$, nodes α and β are said to have *close* affinity if $c_{\alpha\beta}$ is greater than this threshold. Nodes that are close have a better chance of being "aggregated" together with the same coarse node, which is referred to as a *seed*. This affinity measure can be used to determine the seeds and the grouping of nonseed nodes with the seeds in a manner similar to how the strength-of-connection measure is used to achieve these goals in the classical Ruge–Stuben AMG method.[24,42] Specifically, in a manner similar to how the strength-of-connection is used to "fractionally aggregate" the fine nodes (i.e., fractionally in the sense that each FDOF at a fine node can interpolate from CDOFs at coarse nodes of several aggregates), the affinity measure can be used to fractionally aggregate the nonseed nodes with the seeds. This will lead to an overlapping partitioning of the nodes in that a nonseed node can be part of several aggregates. For the powerflow equations, because of the physical differences between the PV and PQ buses, we will restrict the buses to aggregate only with the same bus type.

As developed in Reference 38, to determine which nodes become the seeds, the affinity measure is used to construct the projected volume of the resulting aggregate if a node were 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 $(i,j) \in E$) and let v_i be its volume which is initialized to 1. The projected volume v_i for node i is defined as

$$v_i = v_i + \sum_{(i,j)\in E} v_j \frac{c_{ij}}{\sum_{(j,k)\in E} c_{jk}}. \tag{32}$$

Notice that v_i is the volume of node i plus a weighted sum of the volumes of the nodes neighboring i. The weight on neighboring node j's volume is the ratio between the affinity measure between nodes i and j and the cumulative sum of node j's affinity measures with all of its neighbors. Hence, if node i were selected as a seed, then v_i estimates how large the associated aggregate will be if only its neighboring nodes can contribute a portion of their volumes to the aggregate's volume.

As for the actual selection of the seeds, all nodes that have projected volumes greater than a constant threshold multiple κ of the average projected volume of 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

$$\frac{\sum_{j\in(C\cap N(i))} c_{ij}}{\sum_{j\in N(i)} c_{ij}} \leq Q \quad or \quad \frac{\sum_{j\in(C\cap N(i))} |\mathbf{L}_{f,ij}|}{\sum_{j\in N(i)} |\mathbf{L}_{f,ij}|} \leq Q, \tag{33}$$

then move i to C. Interpreting, the first inequality holds if the c_{ij}'s to its neighboring nonseed nodes, that is, $N(i) \setminus (C \cap N(i))$, are large compared with the affinity measures to its neighboring C nodes, that is, $(C \cap N(i))$. The second inequality has a similar interpretation but now 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 nonseed buses form the F nodes. For $i \in F$, i can be associated with several aggregates/seeds. The maximal number of seeds that each nonseed can interpolate from is called the caliber of the interpolation. Given caliber l, node i can interpolate only from the l seeds with the largest affinity $c_{ij}, j \in C$, or the largest graph connection $|\mathbf{L}_{f,ij}|, j \in C$.[38] Finally, denoting the seeds that F-node i can interpolate from by $N^{C_l}(i)$, the caliber l interpolation operator

$\mathcal{I}$ is computed as

$$\mathcal{I}_{ij} = \begin{cases} \mathbf{L}_{f,ij} / \sum_{k \in N^{C_l}(i)} \mathbf{L}_{f,ik} & i \in F, j \in N^{C_l}(i) \\ 1 & j = i, i \in C \\ 0 & \textit{elsewhere}. \end{cases} \tag{34}$$

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

$$\mathbf{L}_c = \mathcal{I}^t \mathbf{L}_f \mathcal{I}, \tag{35}$$

which is often referred to as the RAP or Galerkin coarsening of $\mathbf{L}_f$. Note that the coarse constant vector $\mathbf{1}_c$ is interpolated to $\mathbf{1}_f$. For graph Laplacian $\mathbf{L}_f$, this implies that $\mathbf{L}_c$ is row-sum zero. However, for the admittance matrices that are only approximately graph Laplacian and nearly row-sum zero (i.e., matrices $\mathbf{Y}^2$, $\mathbf{Y}^3$, and $\mathbf{Y}^4$), their coarse forms are also approximately row-sum zero.

4 | AMG FOR THE POWERFLOW EQUATIONS

With the relaxation-based coarsening of the admittance matrix described, we now present two AMG schemes for the powerflow equations. The first is a FAS approach, which lends itself to some simple analysis (e.g., to show that the interpolation operator constructed for coarsening the admittance matrix can be used to interpolate the coarse-grid corrections). The other is a multiplicative coarse-grid correction scheme that is very similar to the method of Reference 19.

In the following, the algorithms and justification for the choice of interpolation in the FAS cycle will be given.

4.1 | AMG-FAS multigrid

The above relaxation-based coarsening scheme was applied to purely imaginary graph Laplacians in Reference 16. Assuming that the conductances are negligible, this scheme can be applied to admittance matrix $\mathbf{Y}^1$, and possibly to the other admittance forms under this negligible conductance assumption. Although the coarse admittance matrix is not the coarse-grid problem per se, the coarse-grid problem comes out naturally once this matrix is available. Indeed, with $\mathbf{Y}_c = \hat{\imath}\mathbf{B}_c$ denoting the coarse admittance matrix, by simply replacing $\mathbf{Y}$ with $\mathbf{Y}_c$ in the powerflow equations for the CDOFs, we obtain the coarse-grid equations

$$\mathbf{P}_{c,I} = \sum_{J=1, J \neq I}^{N} \hat{\mathbf{V}}_{c,I} \hat{\mathbf{V}}_{c,J} \mathbf{B}_{c,IJ} \sin(\delta_{c,IJ}) \quad I = 1, \ldots, N-1, \tag{36}$$

$$\mathbf{Q}_{c,I} = -\hat{V}_{c,I}^2 \mathbf{B}_{c,II} + \sum_{J=1, J \neq I}^{N} \hat{\mathbf{V}}_{c,I} \hat{\mathbf{V}}_{c,J} \mathbf{B}_{c,IJ} \cos(\delta_{c,IJ}) \quad I = r+1, \ldots, N-1, \tag{37}$$

$$\hat{\mathbf{V}}_{c,I} = 1 \quad I = 1, \ldots, r, \tag{38}$$

where all the PV buses are included in the set of coarse buses. The coarse system for more levels, for the cartesian formulation, and for more general admittance matrix forms can be constructed similarly. For example, assuming that coarsening scheme constructs accurate coarse admittance matrices, the coarse system for the cartesian formulation is

$$\begin{aligned} \mathbf{P}_{c,I} = \mathbf{V}_{c,I}^{\mathrm{r}} \sum_{J=1}^{N} (\mathbf{G}_{c,IJ} \mathbf{V}_{c,J}^{\mathrm{r}} - \mathbf{B}_{c,IJ} \mathbf{V}_{c,J}^{\mathrm{im}}) \\ + \mathbf{V}_{c,I}^{\mathrm{im}} \sum_{J=1}^{N} (\mathbf{B}_{c,IJ} \mathbf{V}_{c,J}^{\mathrm{r}} + \mathbf{G}_{,IJ} \mathbf{V}_{c,J}^{\mathrm{im}}) \quad I = 1, \ldots, N-1, \end{aligned} \tag{39}$$

$$\mathbf{Q}_{c,I} = -\mathbf{V}^{\mathrm{r}}_{c,I}\sum_{J=1}^{N}(\mathbf{B}_{c,IJ}\mathbf{V}^{\mathrm{r}}_{c,J} + \mathbf{G}_{c,IJ}\mathbf{V}^{\mathrm{im}}_{c,J}) \tag{40}$$

$$+ \mathbf{V}^{\mathrm{im}}_{c,I}\sum_{J=1}^{N}(\mathbf{G}_{c,IJ}\mathbf{V}^{\mathrm{r}}_{c,J} - \mathbf{B}_{c,IJ}\mathbf{V}^{\mathrm{im}}_{c,J}) \quad I = r+1, \ldots, N-1,$$

$$1 = (\mathbf{V}^{\mathrm{r}}_{c,I})^2 + (\mathbf{V}^{\mathrm{im}}_{c,I})^2 \quad I = 1, \ldots, r. \tag{41}$$

These coarse-grid problems correspond to a Galerkin-style coarsening of the fine-grid problem. To see this, denote the interpolation operator generated for the graph Laplacian by $\mathcal{I}_c^f$ and the injection operator by $\hat{\mathcal{I}}_f^c$. Then the right-hand sides of (39) and (40) correspond to a coarsening of the power $\mathbf{V}_i(\mathbf{YV})_i^*$ with

$$(\hat{\mathcal{I}}_f^c\mathbf{V}_f)_I[(\mathcal{I}_f^c\mathbf{Y}_f\mathcal{I}_c^f)\hat{\mathcal{I}}_f^c\mathbf{V}_f]_I^*.$$

Note that since $\mathbf{Y}_f$ is complex, $\mathcal{I}_c^f$ will generally be complex. If $\mathbf{Y}_f$ is complex symmetric, $\mathcal{I}_f^c$ is taken to be $[\mathcal{I}_c^f]^t$; otherwise, $\mathcal{I}_f^c$ is taken to be the complex conjugate of $[\mathcal{I}_c^f]^t$ (see Reference 43 for more details).

With the coarse nonlinear functionals defined, an FAS scheme employing a nonlinear smoother can be applied. Letting $\mathcal{N}_l(\mathbf{V}_l) = \mathbf{g}_l$ denote the level l powerflow equations, the algorithm is

$\mathbf{V}_l$ = AMG-FAS($\mathcal{N}_l, \mathbf{V}_l, \mathbf{g}_l$):

(a) presmooth: nonlinear relaxation $\mathcal{S}_l, \mathbf{V}_l \leftarrow \mathcal{S}_l(\mathbf{V}_l, \mathbf{g}_l)$
(b) restricted solution: $\hat{\mathbf{V}}_{l+1} = \hat{\mathcal{I}}_l^{l+1}\mathbf{V}_l$
(c) coarse-grid rhs: $\mathbf{g}_{l+1} = \mathcal{I}_l^{l+1}(\mathbf{g}_l - \mathcal{N}_l(\mathbf{V}_l)) + \mathcal{N}_{l+1}(\hat{\mathbf{V}}_{l+1})$
(d) coarse-grid recursion: $\mathbf{V}_{l+1}$ = AMG-FAS($\mathcal{N}_{l+1}, \hat{\mathbf{V}}_{l+1}, \mathbf{g}_{l+1}$)
(e) coarse-grid correction: $\mathbf{V}_l \leftarrow \mathbf{V}_l + \mathcal{I}_{l+1}^l(\mathbf{V}_{l+1} - \hat{\mathbf{V}}_{l+1})$
(f) postsmooth: nonlinear relaxation $\mathbf{V}_l \leftarrow \mathcal{S}_l(\mathbf{V}_l, \mathbf{g}_l)$.

(See References 22-24,26 for further details on the FAS algorithm.)

4.1.1 | Interpolation in the FAS solve phase

In the above algorithm, the chosen interpolation operator in this solve phase is $\mathcal{I}_{l+1}^l$, the interpolation generated in the construction of the coarse admittance matrices and given by (34). There are several reasons for this choice. An intuitive, physical reason is provided by nodal network system (7). From this system, we see that an appropriate coarse current is a weighted average of the fine voltage, with the weights determined by the entries of the admittance matrix. Since the powerflow equations are defined through the nodal network system, the weighted averaging effected by (34) on the admittance matrix can be appropriate for interpolating the voltage in the FAS solve phase.

More precise justifications of this choice of interpolation in the solve phase can be derived. We consider justifications based on the complementary smoothing/coarse-grid correction principle and on the approximation of the coarse-grid problem in FAS to its counterpart fine-grid problem for the interpolated coarse-grid unknowns. This second justification reflects a consistency in the FAS algorithm and the coarse-grid problems derived from the admittance matrix coarsening.

Complementary smoothing/coarse-grid correction perspective: The complementary smoothing/coarse-grid correction principle of multigrid states that errors which cannot be handled by relaxation must be handled through the coarse-grid correction. These errors must be in the range of interpolation. Now, since $\mathcal{I}_{l+1}^l$ was generated using a relaxed-based method on the admittance matrix, the algebraically smooth components, or near-nullspace, of $\mathbf{Y}_l$ accurately span the range of $\mathcal{I}_{l+1}^l$. Thus, if $\mathcal{I}_{l+1}^l$ is an appropriate choice for the solve phase, then by the complementary principle, the near-nullspace of $\mathbf{Y}_l$ must be poorly attenuated by the nonlinear relaxation. Because the problem is nonlinear, one cannot show this by observing how relaxation acts on a near-nullspace component by itself. Rather, we must observe how relaxation acts on a full approximation that has an error in the near-nullspace of $\mathbf{Y}_l$. To show the poor attenuation, powergrid properties will be needed.

To this end, we note that locally at node i, the nonlinear relaxation iterates on

$$0 = \mathbf{P}_{l,i} - \hat{\imath}\mathbf{Q}_{l,i} - \mathbf{V}_{l,i}^{*}\sum_{j}\mathbf{Y}_{l,ij}\mathbf{V}_{l,j}, \tag{42}$$

where we have used the complex-conjugated form of (8). For example, a nonlinear Gauss–Seidel commonly used in the powergrid community is

$$\mathbf{V}_{l,i}^{k} = \frac{1}{\mathbf{Y}_{l,ii}}\left[\frac{\mathbf{P}_{l,i} - \hat{\imath}\mathbf{Q}_{l,i}}{(\mathbf{V}_{l,i}^{(k-1)})^{*}} - \sum_{j\neq i}\mathbf{Y}_{l,ij}\mathbf{V}_{l,j}^{(k-1)}\right] \tag{43}$$

for iterate k.[6] Suppose that the initial iterate is

$$\mathbf{V}_{l}^{0} = \mathbf{V}_{l}^{\text{exact}} + \mathbf{e}_{l},$$

where $\mathbf{V}_{l}^{\text{exact}}$ is the exact solution and $\mathbf{e}_{l}$ is the error. For stable systems, the voltages are close to 1 ($\|\mathbf{V}_{l}^{\text{exact}}\|_{\infty} \approx 1$), and hence, we take a normalized initial approximation. This is possible for the finest level. Moreover, assuming that the coarser-level admittance matrices sufficiently approximate the finest-level admittance matrix, since the FAS formulation solves for a coarse-grid approximation to the solution of the fine-grid problem rather than an error correction, this is also possible for the coarser levels. Now taking $\mathbf{V}_{l}^{0}$ in this Gauss–Seidel iteration, we have

$$\begin{aligned}
\mathbf{V}_{l,i}^{1} &= \frac{1}{\mathbf{Y}_{l,ii}}\left[\frac{\mathbf{P}_{l,i} - \hat{\imath}\mathbf{Q}_{l,i}}{(\mathbf{V}_{l,i}^{0})^{*}} - \sum_{j\neq i}\mathbf{Y}_{l,ij}\mathbf{V}_{l,j}^{0}\right] \\
&= \frac{1}{\mathbf{Y}_{l,ii}}\left[\frac{\mathbf{P}_{l,i} - \hat{\imath}\mathbf{Q}_{l,i}}{(\mathbf{V}_{l,i}^{\text{exact}} + \mathbf{e}_{l,i})^{*}} - \sum_{j\neq i}\mathbf{Y}_{l,ij}(\mathbf{V}_{l,j}^{\text{exact}} + \mathbf{e}_{l,j})\right] \\
&= \frac{1}{\mathbf{Y}_{l,ii}}\left[\frac{\mathbf{P}_{l,i} - \hat{\imath}\mathbf{Q}_{l,i}}{(\mathbf{V}_{l,i}^{\text{exact}} + \mathbf{e}_{l,i})^{*}} - \sum_{j}\mathbf{Y}_{l,ij}\mathbf{V}_{l,j}^{\text{exact}} - \sum_{j}\mathbf{Y}_{l,ij}\mathbf{e}_{l,j}\right] + \mathbf{V}_{l,i}^{\text{exact}} + \mathbf{e}_{l,i} \\
&= \frac{1}{\mathbf{Y}_{l,ii}}\left[\frac{\mathbf{P}_{l,i} - \hat{\imath}\mathbf{Q}_{l,i}}{(\mathbf{V}_{l,i}^{\text{exact}} + \mathbf{e}_{l,i})^{*}} - \frac{\mathbf{P}_{l,i} - \hat{\imath}\mathbf{Q}_{l,i}}{(\mathbf{V}_{l,i}^{\text{exact}})^{*}} - \sum_{j}\mathbf{Y}_{l,ij}\mathbf{e}_{l,j}\right] + \mathbf{V}_{l,i}^{0} \\
&= \frac{\mathbf{P}_{l,i} - \hat{\imath}\mathbf{Q}_{l,i}}{\mathbf{Y}_{l,ii}}\frac{-\mathbf{e}_{l,i}^{*}}{(\mathbf{V}_{l,i}^{\text{exact}} + \mathbf{e}_{l,i})^{*}(\mathbf{V}_{l,i}^{\text{exact}})^{*}} - \frac{\sum_{j}\mathbf{Y}_{l,ij}\mathbf{e}_{l,j}}{\mathbf{Y}_{l,ii}} + \mathbf{V}_{l,i}^{0}.
\end{aligned}$$

In particular, if $\mathbf{e}_{l}$ is a near-nullspace component of $\mathbf{Y}_{l}$ then

$$\mathbf{V}_{l,i}^{1} \approx \frac{\mathbf{P}_{l,i} - \hat{\imath}\mathbf{Q}_{l,i}}{\mathbf{Y}_{l,ii}}\frac{-\mathbf{e}_{l,i}^{*}}{(\mathbf{V}_{l,i}^{0}\mathbf{V}_{l,i}^{\text{exact}})^{*}} + \mathbf{V}_{l,i}^{0},$$

and using the norm conditions on $\mathbf{V}_{l}^{\text{exact}}$ and $\mathbf{V}_{l}^{0}$, the correction to $\mathbf{V}_{l,i}^{0}$ satisfies

$$\left|\frac{\mathbf{P}_{l,i} - \hat{\imath}\mathbf{Q}_{l,i}}{\mathbf{Y}_{l,ii}}\frac{-\mathbf{e}_{l,i}^{*}}{(\mathbf{V}_{l,i}^{0}\mathbf{V}_{l,i}^{\text{exact}})^{*}}\right| \lesssim 2\left|\frac{\mathbf{P}_{l,i} - \hat{\imath}\mathbf{Q}_{l,i}}{\mathbf{Y}_{l,ii}}\right|.$$

For real powergrid systems, on the finest level, the ratio $\left|\frac{\mathbf{P}_{l,i} - \hat{\imath}\mathbf{Q}_{l,i}}{\mathbf{Y}_{l,ii}}\right| \ll 1$ for most nodes.

Figure 1 illustrates this ratio for some real powergrid systems. In fact, for all levels, this can be intuitively seen from

$$(i)\ \mathbf{P}_{l,i} - \hat{\imath}\mathbf{Q}_{l,i} = (\mathbf{V}_{l,i}^{\text{exact}})^{*}\sum_{j}\mathbf{Y}_{l,ij}\mathbf{V}_{l,j}^{\text{exact}},$$

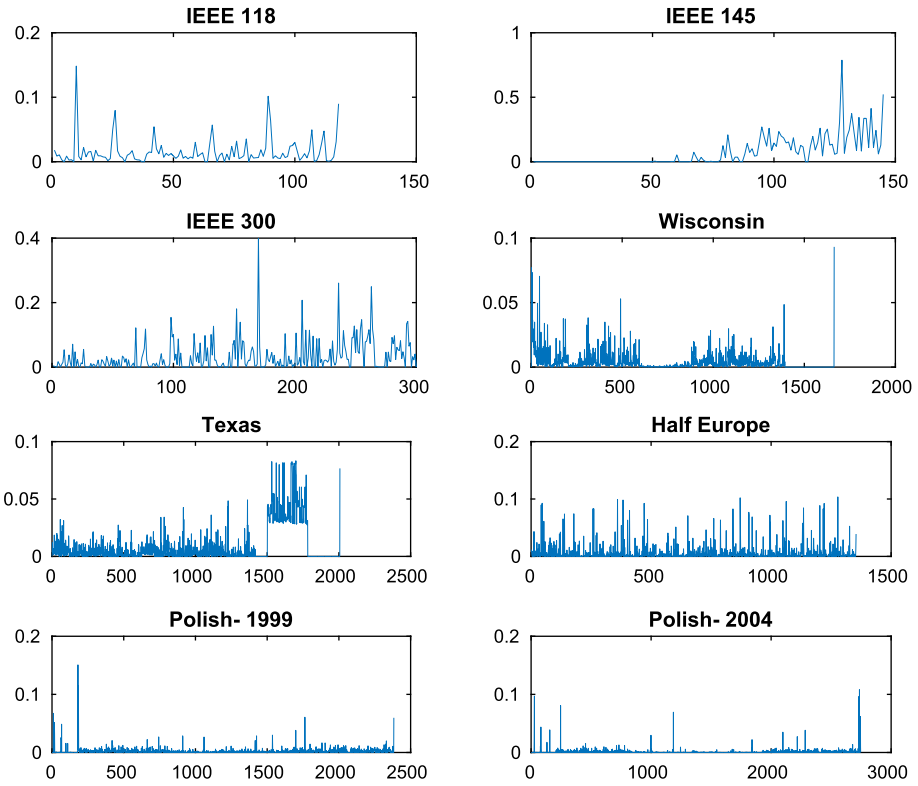

FIGURE 1 Ratio $\left|\frac{\mathbf{P}_{l,i}-\hat{\imath}\mathbf{Q}_{l,i}}{\mathbf{Y}_{l,ii}}\right|$ for some real powergrid systems. The horizontal axis corresponds to the bus indices, and the vertical axis corresponds to the ratio

(*ii*) $\mathbf{V}_{l,i}^{\text{exact}} \approx 1$, and (*iii*) $\mathbf{Y}_l$ being an approximate graph Laplacian (nearly row-sum zero). Hence, with $\left|\frac{\mathbf{P}_{l,i}-\hat{\imath}\mathbf{Q}_{l,i}}{\mathbf{Y}_{l,ii}}\right| \ll 1$, the nonlinear Gauss–Seidel is ineffective on the near-nullspace components of $\mathbf{Y}_l$ (i.e., $\mathbf{V}_{l,i}^1 \approx \mathbf{V}_{l,i}^0$ so that $\mathbf{e}_{l,i}$ is not effectively attenuated). To eliminate these components, these errors must be approximately resolved on a coarse-grid and prolongated with an interpolation operator whose range is accurately spanned by the near-nullspace of $\mathbf{Y}_l$. Such an operator is $\mathcal{I}_{l+1}^l$. Thus, the interpolation operators used to form the coarse admittance matrices are appropriate in the FAS solve phase.

Consistency in the FAS algorithm and the admittance matrix coarsening: By comparing the actual coarse-grid system used in FAS to its counterpart fine-grid system for the interpolated coarse-grid unknowns, another justification for using $\mathcal{I}_{l+1}^l$ in the solve phase can be derived. These two systems should sufficiently approximate each other. As mentioned earlier, such an approximation reflects a consistency in the FAS algorithm and the coarse-grid problems derived from the admittance matrix coarsening.

To show this, we assume that the coarse-grid unknown generates a coarse-grid correction that leads to a solution to the fine-level powerflow equations. We consider the cartesian formulation since the nonlinearity is only quadratic, and hence, the nonlinear system for the coarse-grid correction can be explicitly derived. Let the corrected approximation have the form

$$\mathbf{V}_l + \tilde{\mathbf{I}}_{l+1}^l(\mathbf{V}_{l+1} - \hat{\mathcal{I}}_l^{l+1}\mathbf{V}_l) = \mathbf{V}_l + \tilde{\mathbf{I}}_{l+1}^l\mathbf{e}_{l+1}, \tag{44}$$

where $\tilde{\mathbf{I}}_{l+1}^l$ denotes a suitable interpolation operator and $\mathbf{e}_{l+1} := (\mathbf{V}_{l+1} - \hat{\mathcal{I}}_l^{l+1}\mathbf{V}_l)$. The fine-level problem is

$$\mathbf{0} = \mathbf{P}_l - \hat{\imath}\mathbf{Q}_l - \text{diag}([\mathbf{V}_l + \tilde{\mathbf{I}}_{l+1}^l\mathbf{e}_{l+1}]^*)\mathbf{Y}_l(\mathbf{V}_l + \tilde{\mathbf{I}}_{l+1}^l\mathbf{e}_{l+1}).$$

Expanding this out, we obtain

$$\begin{aligned}\text{diag}([\tilde{\mathbf{I}}_{l+1}^l\mathbf{V}_{l+1}]^*)\mathbf{Y}_l(\tilde{\mathbf{I}}_{l+1}^l\mathbf{V}_{l+1}) = {}& (\mathbf{P}_l - \hat{\imath}\mathbf{Q}_l - \text{diag}(\mathbf{V}_l^*)\mathbf{Y}_l\mathbf{V}_l)\\ &+ \text{diag}([\tilde{\mathbf{I}}_{l+1}^l\hat{\mathcal{I}}_l^{l+1}\mathbf{V}_l]^*)\mathbf{Y}_l\mathbf{V}_l + \{\text{diag}([\tilde{\mathbf{I}}_{l+1}^l\hat{\mathcal{I}}_l^{l+1}\mathbf{V}_l]^*) - \text{diag}(\mathbf{V}_l^*)\}\mathbf{Y}_l\tilde{\mathbf{I}}_{l+1}^l\mathbf{e}_{l+1}\\ &+ \text{diag}([\tilde{\mathbf{I}}_{l+1}^l\mathbf{V}_{l+1}]^*)\mathbf{Y}_l(\tilde{\mathbf{I}}_{l+1}^l\hat{\mathcal{I}}_l^{l+1} - \mathbf{I})\mathbf{V}_l\end{aligned}$$

$$
\begin{aligned}
&= \mathbf{r}_l + \mathrm{diag}([\tilde{\mathbf{I}}_{l+1}^{l}\hat{\mathcal{I}}_{l}^{l+1}\mathbf{V}_l]^*)\mathbf{Y}_l\mathbf{V}_l \\
&\quad + \{\mathrm{diag}([\tilde{\mathbf{I}}_{l+1}^{l}\hat{\mathcal{I}}_{l}^{l+1}\mathbf{V}_l]^*) - \mathrm{diag}(\mathbf{V}_l^*)\}\mathbf{Y}_l\tilde{\mathbf{I}}_{l+1}^{l}\mathbf{e}_{l+1} \\
&\quad + \mathrm{diag}([\tilde{\mathbf{I}}_{l+1}^{l}\mathbf{V}_{l+1}]^*)\mathbf{Y}_l(\tilde{\mathbf{I}}_{l+1}^{l}\hat{\mathcal{I}}_{l}^{l+1} - \mathbf{I})\mathbf{V}_l,
\end{aligned} \tag{45}
$$

where $\mathbf{r}_l := (\mathbf{P}_l - \hat{\imath}\mathbf{Q}_l - \mathrm{diag}(\mathbf{V}_l^*)\mathbf{Y}_l\mathbf{V}_l)$ is the residual. On the other hand, the coarse-grid problem in FAS is

$$
\mathrm{diag}(\mathbf{V}_{l+1}^*)\mathbf{Y}_{l+1}\mathbf{V}_{l+1} = \mathbf{r}_{l+1} + \mathrm{diag}([\hat{\mathcal{I}}_{l}^{l+1}\mathbf{V}_l]^*)\mathbf{Y}_{l+1}\hat{\mathcal{I}}_{l}^{l+1}\mathbf{V}_l. \tag{46}
$$

Comparing (46) with (45), to have good agreement between these two expressions, one requirement is that

$$
\mathrm{diag}(\mathbf{V}_{l+1}^*)\mathbf{Y}_{l+1}\mathbf{V}_{l+1} = \mathrm{diag}(\mathbf{V}_{l+1}^*)\mathcal{I}_{l}^{l+1}\mathbf{Y}_l\mathcal{I}_{l+1}^{l}\mathbf{V}_{l+1} \tag{47}
$$

approximates

$$
\mathrm{diag}([\tilde{\mathbf{I}}_{l+1}^{l}\mathbf{V}_{l+1}]^*)\mathbf{Y}_l\tilde{\mathbf{I}}_{l+1}^{l}\mathbf{V}_{l+1}. \tag{48}
$$

That is, viewing (47) and (48) as components of the nonlinear operator defined over the coarse space and the fine space restricted to the range of $\tilde{\mathbf{I}}_{l+1}^{l}$ respectively, the goal is to select $\tilde{\mathbf{I}}_{l+1}^{l}$ so that the coarse nonlinear operator acting on $\mathbf{V}_{l+1}$ accurately approximates the fine nonlinear operator acting on $\tilde{\mathbf{I}}_{l+1}^{l}\mathbf{V}_{l+1}$. Consider the case when $\mathcal{I}_{l+1}^{l}$ is the piecewise constant interpolation. In Reference 19, it was shown that this interpolation satisfies

$$
\mathcal{I}_{l}^{l+1}\mathrm{diag}([\mathcal{I}_{l+1}^{l}\mathbf{w}_{l+1}]^*) = \mathrm{diag}(\mathbf{w}_{l+1}^*)\mathcal{I}_{l}^{l+1} \tag{49}
$$

for all coarse vectors $\mathbf{w}_{l+1}$. Applying this to (47), we have

$$
\mathrm{diag}(\mathbf{V}_{l+1}^*)\mathcal{I}_{l}^{l+1}\mathbf{Y}_l\mathcal{I}_{l+1}^{l}\mathbf{V}_{l+1} = \mathcal{I}_{l}^{l+1}\mathrm{diag}([\mathcal{I}_{l+1}^{l}\mathbf{V}_{l+1}]^*)\mathbf{Y}_l\mathcal{I}_{l+1}^{l}\mathbf{V}_{l+1},
$$

from which we see that this is a good approximation to (48) if $\tilde{\mathbf{I}}_{l+1}^{l}$ is chosen to be $\mathcal{I}_{l+1}^{l}$.

For higher caliber $\mathcal{I}_{l+1}^{l}$'s, a natural choice for $\tilde{\mathbf{I}}_{l+1}^{l}$ is $\mathcal{I}_{l+1}^{l}$. However, the approximation of (47) to (48) does not come out as cleanly. Nevertheless, comparing these two expressions, we see that $\mathcal{I}_{l+1}^{l}$ is an appropriate choice when

$$
\mathrm{diag}(\mathbf{V}_{l+1}^*)\mathcal{I}_{l}^{l+1} \quad \text{approximates} \quad \mathrm{diag}([\mathcal{I}_{l+1}^{l}\mathbf{V}_{l+1}]^*) \tag{50}
$$

with sufficient accuracy.

Thus, the left-hand sides of (45) and (46) can sufficiently agree if $\tilde{\mathbf{I}}_{l+1}^{l}$ is taken to be $\mathcal{I}_{l+1}^{l}$. We now show that the right-hand sides of these expressions also sufficiently agree if $\tilde{\mathbf{I}}_{l+1}^{l}$ is $\mathcal{I}_{l+1}^{l}$. Comparing the right-hand sides of (46) and (45), good agreement can further be achieved if

$$
\{\mathrm{diag}([\tilde{\mathbf{I}}_{l+1}^{l}\hat{\mathcal{I}}_{l}^{l+1}\mathbf{V}_l]^*) - \mathrm{diag}(\mathbf{V}_l^*)\}\mathbf{Y}_l\tilde{\mathbf{I}}_{l+1}^{l}\mathbf{e}_{l+1} \approx 0 \tag{51}
$$

and

$$
\mathrm{diag}([\hat{\mathcal{I}}_{l}^{l+1}\mathbf{V}_l]^*)\mathbf{Y}_{l+1}\hat{\mathcal{I}}_{l}^{l+1}\mathbf{V}_l \tag{52}
$$

approximates

$$
\mathrm{diag}([\tilde{\mathbf{I}}_{l+1}^{l}\mathbf{V}_{l+1}]^*)\mathbf{Y}_l(\tilde{\mathbf{I}}_{l+1}^{l}\hat{\mathcal{I}}_{l}^{l+1} - \mathbf{I})\mathbf{V}_l + \mathrm{diag}([\tilde{\mathbf{I}}_{l+1}^{l}\hat{\mathcal{I}}_{l}^{l+1}\mathbf{V}_l]^*)\mathbf{Y}_l\mathbf{V}_l. \tag{53}
$$

Condition (51) can be satisfied if $\tilde{\mathbf{I}}_{l+1}^{l}$ is selected to be $\mathcal{I}_{l+1}^{l}$ since $\tilde{\mathbf{I}}_{l+1}^{l}\mathbf{e}_{l+1}$ then will be in the near-nullspace of $\mathbf{Y}_l$. To show the second condition, assume that

$$
\mathbf{V}_{l+1} \approx \hat{\mathcal{I}}_{l}^{l+1}\mathbf{V}_l, \tag{54}
$$

which is a reasonable assumption in the FAS setting (recall that the coarse-grid problem solves for an approximation to the solution rather than for an error correction). With this, we have

$$\begin{aligned}&-\text{diag}([\tilde{\mathbf{I}}_{l+1}^{l}\mathbf{V}_{l+1}]^{*})\mathbf{Y}_l\mathbf{V}_l + \text{diag}([\tilde{\mathbf{I}}_{l+1}^{l}\hat{\mathcal{I}}_{l}^{l+1}\mathbf{V}_l]^{*})\mathbf{Y}_l\mathbf{V}_l\\ &\quad\approx -\text{diag}([\tilde{\mathbf{I}}_{l+1}^{l}\mathbf{V}_{l+1}]^{*})\mathbf{Y}_l\mathbf{V}_l + \text{diag}([\tilde{\mathbf{I}}_{l+1}^{l}\mathbf{V}_{l+1}]^{*})\mathbf{Y}_l\mathbf{V}_l = 0\end{aligned}$$

for any $\tilde{\mathbf{I}}_{l+1}^{l}$. Thus, we need (52) to approximate only

$$\text{diag}([\tilde{\mathbf{I}}_{l+1}^{l}\mathbf{V}_{l+1}]^{*})\mathbf{Y}_l\tilde{\mathbf{I}}_{l+1}^{l}\hat{\mathcal{I}}_{l}^{l+1}\mathbf{V}_l, \tag{55}$$

in the sense these two nonlinear operators acting on the appropriate elements approximately agree. Again, taking $\mathcal{I}_{l+1}^{l}$ to be the piecewise constant interpolation, (52) and (49) lead to

$$\begin{aligned}\text{diag}([\hat{\mathcal{I}}_{l}^{l+1}\mathbf{V}_l]^{*})\mathbf{Y}_{l+1}\hat{\mathcal{I}}_{l}^{l+1}\mathbf{V}_l &= \mathcal{I}_{l}^{l+1}\text{diag}([\mathcal{I}_{l+1}^{l}\hat{\mathcal{I}}_{l}^{l+1}\mathbf{V}_l]^{*})\mathbf{Y}_l\mathcal{I}_{l+1}^{l}\hat{\mathcal{I}}_{l}^{l+1}\mathbf{V}_l\\ &\approx \mathcal{I}_{l}^{l+1}\text{diag}([\mathcal{I}_{l+1}^{l}\mathbf{V}_{l+1}]^{*})\mathbf{Y}_l\mathcal{I}_{l+1}^{l}\hat{\mathcal{I}}_{l}^{l+1}\mathbf{V}_l.\end{aligned} \tag{56}$$

Comparing this to (55), a good choice for $\tilde{\mathbf{I}}_{l+1}^{l}$ is again $\mathcal{I}_{l+1}^{l}$. For higher caliber $\mathcal{I}_{l+1}^{l}$'s, a natural choice for $\tilde{\mathbf{I}}_{l+1}^{l}$ is $\mathcal{I}_{l+1}^{l}$. To be more precise, using assumption (54), we have

$$\text{diag}([\hat{\mathcal{I}}_{l}^{l+1}\mathbf{V}_l]^{*}) \approx \text{diag}(\mathbf{V}_{l+1}^{*}),$$

and so (52) becomes

$$\text{diag}([\hat{\mathcal{I}}_{l}^{l+1}\mathbf{V}_l]^{*})\mathbf{Y}_{l+1}\hat{\mathcal{I}}_{l}^{l+1}\mathbf{V}_l \approx \text{diag}(\mathbf{V}_{l+1}^{*})\mathcal{I}_{l}^{l+1}\mathbf{Y}_l\mathcal{I}_{l+1}^{l}\hat{\mathcal{I}}_{l}^{l+1}\mathbf{V}_l.$$

Comparing this to (55), $\mathcal{I}_{l+1}^{l}$ is then an appropriate choice if (50) is satisfied to sufficient accuracy.

4.2 | Multiplicative correction

In References 20,21, multigrid methods with multiplicative coarse-grid correction are presented for computing stationary probability vectors for Markov chains that involve column-stochastic transition-probability matrices. This multiplicative-correction strategy guarantees the physical property that the components of the iterates are nonnegative, as they should be since they represent probabilities. In Reference 19, a multiplicative correction scheme is developed for the powerflow equations. This type of correction guarantees that the coarse-grid systems have the same structure as the fine-grid system, and hence, the multiplicative correction again preserves some physics: the coarse-grid systems represent physical powerflow systems. Although the AMG-FAS algorithm generates coarse-grid systems that have the same form as the fine-grid system, we now consider a multiplicative-correction scheme. Following Reference 19, a multiplicative correction scheme will be developed using the relaxation-based coarsening of the admittance matrix. Using caliber one interpolation (i.e., piecewise constant) in this coarsening, the scheme is very similar to the one in Reference 19. The main differences are in the coarsening of the admittance matrix and in the smoother. For the smoother, we use only Gauss–Seidel (43) on the PQ buses rather than the involved smoother of Reference 19. These differences remove some of the powergrid specifics assumed in that article.

In a two-level setting, this multiplicative correction approach updates the fine-level approximation according to

$$\mathbf{V}_f \leftarrow \text{diag}(\mathbf{V}_f)(\tilde{\mathbf{I}}_{c}^{f}\mathbf{e}_c), \tag{57}$$

where $\tilde{\mathbf{I}}_{c}^{f}$ is an appropriate interpolation operator. The defining system for $\mathbf{e}_c$ follows from powerflow equation (9):

$$\begin{aligned}\mathbf{P}_f + i\mathbf{Q}_f &= \text{diag}[\text{diag}(\mathbf{V}_f)(\tilde{\mathbf{I}}_{c}^{f}\mathbf{e}_c)][\mathbf{Y}_f\text{diag}(\mathbf{V}_f)(\tilde{\mathbf{I}}_{c}^{f}\mathbf{e}_c)]^{*}\\ &= \text{diag}(\tilde{\mathbf{I}}_{c}^{f}\mathbf{e}_c)\text{diag}(\mathbf{V}_f)[\mathbf{Y}_f\text{diag}(\mathbf{V}_f)\tilde{\mathbf{I}}_{c}^{f}]^{*}\mathbf{e}_c^{*}\end{aligned} \tag{58}$$

and restricting

$$\tilde{\mathbf{I}}_f^c(\mathbf{P}_f + \hat{\imath}\mathbf{Q}_f) = \{\tilde{\mathbf{I}}_f^c \text{diag}(\tilde{\mathbf{I}}_c^f \mathbf{e}_c)\text{diag}(\mathbf{V}_f)[\mathbf{Y}_f \text{diag}(\mathbf{V}_f)\tilde{\mathbf{I}}_c^f]^*\}\mathbf{e}_c^*. \tag{59}$$

For a piecewise constant $\tilde{\mathbf{I}}_c^f$, identity (49) gives

$$\begin{aligned}\tilde{\mathbf{I}}_f^c(\mathbf{P}_f + \hat{\imath}\mathbf{Q}_f) &= \text{diag}(\mathbf{e}_c)\{\tilde{\mathbf{I}}_f^c \text{diag}(\mathbf{V}_f)[\mathbf{Y}_f \text{diag}(\mathbf{V}_f)\tilde{\mathbf{I}}_c^f]^*\}\mathbf{e}_c^* \\ &:= \text{diag}(\mathbf{e}_c)[\mathcal{Y}_c \mathbf{e}_c]^*,\end{aligned} \tag{60}$$

where

$$\mathcal{Y}_c^* := \{\tilde{\mathbf{I}}_f^c \text{diag}(\mathbf{V}_f)[\mathbf{Y}_f \text{diag}(\mathbf{V}_f)\tilde{\mathbf{I}}_c^f]^*\}, \tag{61}$$

that is, $\mathcal{Y}_c$ is the Galerkin coarsening of the operator $\mathbf{Y}_f$ using the interpolation operator $\text{diag}(\mathbf{V}_f)\tilde{\mathbf{I}}_c^f$. Equivalently, $\mathcal{Y}_c$ is the Galerkin coarsening of a row and column scaled $\mathbf{Y}_f$ using $\tilde{\mathbf{I}}_c^f$. Note that, even if $\mathbf{Y}_f$ is approximately row-sum zero, the row and column scaling means that $\mathcal{Y}_c$ is generally not row-sum zero unless $\mathbf{V}_f \equiv \mathbf{1}$ (cf., References 20,21 where it was extensively shown that, like the fine-grid operator, the coarse-grid operators are column-sum zero and row-sum zero at the exact solution). Thus, we distinguish $\mathcal{Y}_c$ from an admittance matrix by referring to it as a coarse-grid matrix. As mentioned earlier and highlighted in Reference 19, the nice feature of (60) is that it has the same form as the fine-grid system. Because of this, we will use this form even when $\tilde{\mathbf{I}}_c^f$ is of higher-order caliber.

The above two-level method can be recursively applied to more levels. Taking $\mathcal{Y}_0 = \mathbf{Y}_0$, a hierarchy of coarse-grid matrices $\{\mathcal{Y}_l\}$ is formed. However, since $\text{diag}(\mathbf{V}_{l-1})\tilde{\mathbf{I}}_l^{l-1}$ is used in constructing $\mathcal{Y}_l$, the explicit dependence on the approximation $\mathbf{V}_{l-1}$ means that a new hierarchy must be formed after each multigrid cycle. From the numerics, this reconstruction is very important for the convergence of the multiplicative-correction scheme. Updating the hierarchy at every other cycle, or only until the approximate solution has achieved moderate accuracy, is not sufficient for the scheme to converge. This required updating is the major disadvantage of this scheme when compared with the AMG-FAS method, with the disadvantage more pronounced for more general μ-cycles, for example, W-cycles. Moreover, there is the question of how to determine the coarse buses and how to construct $\tilde{\mathbf{I}}_{l+1}^l$. The smoothing/coarse-grid correction principle suggests that these should be based on $\mathcal{Y}_l$, which implies that the CDOFs and $\tilde{\mathbf{I}}_{l+1}^l$ must be adjusted after each multigrid cycle. However, numerical experiments indicate that generating these upfront with the relaxation-based technique applied to the hierarchy of admittance matrices is adequate, that is, $\tilde{\mathbf{I}}_{l+1}^l = \mathcal{I}_{l+1}^l$. An explanation for this has yet to be determined, and may require assumptions not needed in the AMG-FAS algorithm. For example, justifying that the slowly damped multiplicative corrections of Gauss–Seidel are in the near-nullspace of $\mathbf{Y}_l$ requires additional assumptions. For consider the finest level $l = 0$ with initial approximation

$$\mathbf{V}_l^0 = \mathbf{V}_l^{\text{exact}} \mathbf{e}_l^0,$$

where $\mathbf{e}_l^0$ is the initial multiplicative error correction. Then taking the same form for $\mathbf{V}_l^1$, Gauss–Seidel (43) produces

$$\begin{aligned}\mathbf{V}_{l,i}^{\text{exact}} \mathbf{e}_{l,i}^1 &= \frac{1}{\mathbf{Y}_{l,ii}} \left[\frac{\mathbf{P}_{l,i} - \hat{\imath}\mathbf{Q}_{l,i}}{(\mathbf{V}_{l,i}^{\text{exact}} \mathbf{e}_{l,i}^0)^*} - \sum_{j \neq i} \mathbf{Y}_{l,ij} \mathbf{V}_{l,j}^{\text{exact}} \mathbf{e}_{l,j}^0 \right] \\ &= \frac{1}{\mathbf{Y}_{l,ii}} \left[\frac{\mathbf{P}_{l,i} - \hat{\imath}\mathbf{Q}_{l,i}}{(\mathbf{V}_{l,i}^{\text{exact}} \mathbf{e}_{l,i}^0)^*} - \sum_{j} \mathbf{Y}_{l,ij} \mathbf{V}_{l,j}^{\text{exact}} \mathbf{e}_{l,j}^0 \right] + \mathbf{V}_{l,i}^{\text{exact}} \mathbf{e}_{l,i}^0\end{aligned}$$

or

$$\mathbf{e}_{l,i}^1 = \frac{1}{\mathbf{Y}_{l,ii}} \left[\frac{\mathbf{P}_{l,i} - \hat{\imath}\mathbf{Q}_{l,i}}{|\mathbf{V}_{l,i}^{\text{exact}}|^2 (\mathbf{e}_{l,i}^0)^*} - \sum_{j} \mathbf{Y}_{l,ij} \frac{\mathbf{V}_{l,j}^{\text{exact}}}{\mathbf{V}_{l,i}^{\text{exact}}} \mathbf{e}_{l,j}^0 \right] + \mathbf{e}_{l,i}^0.$$

Using the boundness of $|\mathbf{V}_{l,i}^{\text{exact}}|$ and $|\mathbf{e}_{l,i}^0|$, and $\left| \frac{\mathbf{P}_{l,i} - \hat{\imath}\mathbf{Q}_{l,i}}{\mathbf{Y}_{l,ii}} \right| \ll 1$, the first bracketed term is negligible. Thus, we see that $\mathbf{e}_l^0$ is slowly attenuated if it is in the near-nullspace of

$$[\text{diag}(\mathbf{V}_l^{\text{exact}})]^{-1} \mathbf{Y}_l [\text{diag}(\mathbf{V}_l^{\text{exact}})] \tag{62}$$

rather than $\mathbf{Y}_l$. Nevertheless, with the additional assumption that $\mathbf{V}_l^{\text{exact}}$ does not highly vary in the neighborhood of each node $\left(\text{i.e., } \frac{\mathbf{V}_{l,j}^{\text{exact}}}{\mathbf{V}_{l,i}^{\text{exact}}} \approx 1\right)$, $\mathbf{e}_l^0$ will be approximately in the near-nullspace of $\mathbf{Y}_l$.

For coarser levels, the operators involve $\mathcal{Y}_l$ instead of $\mathbf{Y}_l$. Now the multiplicative correction iterates as

$$\mathbf{e}_{l,i}^1 = \frac{1}{\mathcal{Y}_{l,ii}}\left[\frac{\mathbf{P}_{l,i} - \hat{\imath}\mathbf{Q}_{l,i}}{(\mathbf{e}_{l,i}^0)^*} - \sum_j \mathcal{Y}_{l,ij}\mathbf{e}_{l,j}^0\right] + \mathbf{e}_{l,i}^0.$$

Assuming that $\left|\frac{\mathbf{P}_{l,i} - \hat{\imath}\mathbf{Q}_{l,i}}{\mathcal{Y}_{l,ii}}\right| \ll 1$, if $\mathbf{e}_l^0$ is slowly attenuated by the smoother, then it must be in the near-nullspace of $\mathcal{Y}_l$. This again indicates that taking $\tilde{\mathbf{I}}_{l+1}^l$ to be an interpolation based on the near-nullspace of $\mathbf{Y}_l$ is not ideal. Nevertheless, numerical experiments demonstrate that basing $\tilde{\mathbf{I}}_{l+1}^l$ on this near-nullspace can produce an effective solver.

4.3 | Robustness: Basin of attraction

FAS applied to PDEs is known to generally have a larger basin of attraction than Newton's method (see Reference 26 for a comparison of the convergence bounds for a scalar nonlinear equation). However, assuming that the iterates remain within their basins, Newton's method converges quadratically, whereas FAS converges only linearly. To achieve this quadratic rate, Newton's method requires tighter conditions that often results in a small basin. In particular, for the general fine-level system $\mathcal{N}_f(\mathbf{V}_f) = \mathbf{g}_f \equiv \mathbf{0}$, Newton's method requires $\mathcal{N}_f$ and all the second-order partial derivatives of its components be continuous in some neighborhood of the root, and requires the Jacobian of $\mathcal{N}_f$ evaluated at the root to be nonsingular.[44] With these conditions, the kth iterate satisfies the bound

$$\|\mathbf{V}_f^k - \mathbf{V}_f^{\text{exact}}\|_\infty \leq \frac{1}{M}(M\|\mathbf{V}_f^0 - \mathbf{V}_f^{\text{exact}}\|_\infty)^{2^k}, \tag{63}$$

where $\mathbf{V}_f^{\text{exact}}$ is the exact root and M is an upper bound on the Hessian of $\mathcal{N}_f$ and the inverse of the Jacobian of $\mathcal{N}_f$ in the neighborhood of the root.[44] From (63), one sees that convergence is guaranteed only if $\mathbf{V}_f^0$ is sufficiently close to $\mathbf{V}_f^{\text{exact}}$.

Establishing the convergence of FAS is more involved. We consider only a two-level method, follow the presentation in Reference 23, and only summarize some of the steps. The first step is to introduce divided-difference operators to represent the nonlinear procedures in the FAS algorithm. With $\mathcal{U}_f$ denoting the fine-level space where the nonlinear problem is posed, the fine-level divided-difference operators DS_f^μ and $D\mathcal{N}_f$ for the μ-sweep smoother S_f^μ and nonlinear operator $\mathcal{N}_f$ are defined by

$$DS_f^\mu(\mathbf{V}_f, \mathbf{V}_f'; \mathbf{g}_f)[\mathbf{V}_f - \mathbf{V}_f'] = S_f^\mu(\mathbf{V}_f; \mathbf{g}_f) - S_f^\mu(\mathbf{V}_f'; \mathbf{g}_f) \quad \forall \mathbf{V}_f, \mathbf{V}_f' \in \mathcal{U}_f, \tag{64}$$

$$D\mathcal{N}_f(\mathbf{V}_f, \mathbf{V}_f')[\mathbf{V}_f - \mathbf{V}_f'] = \mathcal{N}_f(\mathbf{V}_f) - \mathcal{N}_f(\mathbf{V}_f') \quad \forall \mathbf{V}_f, \mathbf{V}_f' \in \mathcal{U}_f. \tag{65}$$

These operators are not unique. An example is

$$DS_f^\mu(\mathbf{V}_f, \mathbf{V}_f'; \mathbf{g}_f) = \int_0^1 \frac{\partial}{\partial \mathbf{V}}[S_f^\mu(\mathbf{V}_f + t(\mathbf{V}_f' - \mathbf{V}_f); \mathbf{g}_f)]dt,$$

where the partial derivative is taken with respect to first argument of S_f^μ. Identity (64) follows from

$$\begin{aligned}
&(\mathbf{V}_f - \mathbf{V}_f')\int_0^1 \frac{\partial}{\partial \mathbf{V}}[S_f^\mu(\mathbf{V}_f + t(\mathbf{V}_f' - \mathbf{V}_f); \mathbf{g}_f)]dt \\
&= \int_0^1 \frac{\partial}{\partial \mathbf{V}}[S_f^\mu(\mathbf{V}_f + t(\mathbf{V}_f' - \mathbf{V}_f); \mathbf{g}_f)](\mathbf{V}_f - \mathbf{V}_f')dt \\
&= -\int_0^1 \frac{\partial}{\partial \mathbf{V}}[S_f^\mu(\mathbf{V}_f + t(\mathbf{V}_f' - \mathbf{V}_f); \mathbf{g}_f)]\frac{d\mathbf{V}}{dt}dt
\end{aligned}$$

$$= -\int_{\mathbf{V}_f}^{\mathbf{V}'_f} \frac{\partial}{\partial \mathbf{V}}[\mathcal{S}_f^{\mu}(\mathbf{V};\mathbf{g}_f)]d\mathbf{V}$$
$$= \mathcal{S}_f^{\mu}(\mathbf{V}_f;\mathbf{g}_f) - \mathcal{S}_f^{\mu}(\mathbf{V}'_f;\mathbf{g}_f).$$

With $\mathbf{V}_f^j$ denoting the jth iterate of the FAS algorithm, let $\mathbf{V}_f^{\text{sm}}$ denote the result of applying $\mathcal{S}_f^{\mu}$ to $\mathbf{V}_f^j$ (i.e., $\mathbf{V}_f^{\text{sm}} := \mathcal{S}_f^{\mu}(\mathbf{V}_f^j;\mathbf{g}_f)$), $\mathbf{V}_c^{\text{exact}}$ denote the exact solution of the coarse-grid problem, and $\mathbf{V}_f^{\text{corr}}$ denote the coarse-grid corrected approximation $\mathbf{V}_f^{\text{corr}} := \mathbf{V}_f^{\text{sm}} + \mathcal{I}_c^f[\mathbf{V}_c^{\text{exact}} - \hat{\mathcal{I}}_f^c\mathbf{V}_f^{\text{sm}}]$. Using the divided-difference operators, the two-grid FAS operator with μ pre- and postsmoothings can be shown to be

$$DS_f^{\mu}(\mathbf{V}_f^{\text{corr}}, \mathbf{V}_f^{\text{exact}};\mathbf{g}_f)\Big\{\mathbf{I} - \mathcal{I}_c^f D\mathcal{N}_c^{-1}[\mathbf{V}_c^{\text{exact}}, \mathcal{N}_c(\hat{\mathcal{I}}_f^c\mathbf{V}_f^{\text{sm}})]$$
$$\mathcal{I}_f^c D\mathcal{N}_f(\mathbf{V}_f^{\text{exact}}, \mathbf{V}_f^{\text{sm}})\Big\} DS_f^{\mu}(\mathbf{V}_f^j, \mathbf{V}_f^{\text{exact}};\mathbf{g}_f) \tag{66}$$

(see Reference 23 for more details on the derivation for a closely related nonlinear multigrid method to FAS). A bound on this iteration operator can be obtained using linear multigrid theory (i.e., smoothing and approximation properties for DS_f^{μ} and $\mathcal{I}_c^f D\mathcal{N}_c^{-1}\mathcal{I}_f^c$) to show convergence with an appropriate number of smoothing sweeps.

Notice that the above procedures do not require an *explicit* condition on the closeness of the initial approximation to the solution. Rather the procedures depend on how particular operators evaluated at specific values satisfy the smoothing and approximation properties. However, implicit in the argument is that every step of the FAS algorithm is well defined, and this is related to the closeness of the initial approximation. Specifically, consider the ϵ_1-neighborhood of the exact solution

$$\mathcal{B}_f^{\epsilon_1}(\mathbf{V}_f^{\text{exact}}) = \{\mathbf{V}_f \in \mathcal{U}_f : \|\mathbf{V}_f - \mathbf{V}_f^{\text{exact}}\| \leq \epsilon_1\}$$

and its image $\text{Im}_{\mathcal{N}_f}(\mathcal{B}_f^{\epsilon_1}(\mathbf{V}_f^{\text{exact}}))$ under $\mathcal{N}_f$. Radius ϵ_1 must be sufficiently small so that the map

$$\mathcal{N}_f : \mathcal{B}_f^{\epsilon_1}(\mathbf{V}_f^{\text{exact}}) \to \text{Im}_{\mathcal{N}_f}(\mathcal{B}_f^{\epsilon_1}(\mathbf{V}_f^{\text{exact}}))$$

is a homeomorphism. The coarse-grid corrected approximation $\mathbf{V}_f^{\text{corr}}$ must remain in this neighborhood to ensure convergence. This in turn, leads to conditions on the intermediate steps in the FAS algorithm for constructing $\mathbf{V}_f^{\text{corr}}$. For example, on the coarse-grid, since a $\mathcal{B}_c^{\epsilon_2}(\mathbf{V}_c^{\text{exact}})$ constraint is required to guarantee that $\mathcal{N}_c : \mathcal{B}_c^{\epsilon_2}(\mathbf{V}_c^{\text{exact}}) \to \text{Im}_{\mathcal{N}_c}(\mathcal{B}_c^{\epsilon_2}(\mathbf{V}_c^{\text{exact}}))$ is a homeomorphism (in order for existence of a unique coarse solution), the right-hand side

$$\mathbf{g}_c = \mathcal{I}_f^c[\mathbf{g}_f - \mathcal{N}_f(\mathbf{V}_f^{\text{sm}})] + \mathcal{N}_c(\hat{\mathcal{I}}_f^c\mathbf{V}_f^{\text{sm}})$$

must be in $\text{Im}_{\mathcal{N}_c}(\mathcal{B}_c^{\epsilon_2}(\mathbf{V}_c^{\text{exact}}))$. This leads to conditions relating ϵ_2 to an upper bound on the norm of $\mathbf{g}_c$.

Overall, to obtain a concrete expression for the basin of attraction for FAS is nontrivial. Nevertheless, the radius of this basin is generally larger than that of Newton's method because only linear convergence is pursued.

5 | NUMERICAL RESULTS

In this section, we examine the performance of the two multigrid solvers on some realistic powergrid systems, and pseudosystems constructed by taking their admittance matrices from finite element discretizations of two elliptic partial differential equations. Because the realistic systems are relatively small-scale, the experiments on the pseudosystems illustrate the scalability of the solvers. For the realistic- and pseudosystems, 20 test vectors with 10 Gauss–Seidel smoothing sweeps are used in the relaxation-based coarsening procedure on the admittance matrices. The actual nonlinear solver involves $V(1,1)$ cycles with nonlinear Gauss–Seidel smoothing on the finer levels, and two and five sweeps of Newton's method on the coarsest level for the realistic- and pseudosystems, respectively, with the Jacobian linear systems solved with a direct method. The additional coarse Newton sweeps for the pseudosystems were performed to provide robustness

to the methods. The stopping criterion for the multigrid iteration is a six-order reduction in the initial residual, that is, $\left\| \frac{\mathbf{r}^k}{\mathbf{r}^0} \right\| < 10^{-6}$. The prototype software was implemented in Matlab.

5.1 | Realistic powergrid systems

The realistic test-suite systems are IEEE57, IEEE68, IEEE118, IEEE145, IEEE300; the synthetic Illinois, South Carolina, Texas and Wisconsin systems; and several European systems. These systems were obtained from several sources. Systems IEEE57, IEEE118, IEEE145, and IEEE300 are from University of Washington's Power System Test Case Archive.[45] System IEEE68 is from an Office of Science Technical Report.[46] Systems Illinois, South Carolina, Texas and Wisconsin are from Texas A&M University's Electric Grid Test Case Repository[47] and described in Reference 48. These are synthetic systems that are statistically similar to the actual transmission systems in the corresponding states. Systems Half-Europe and Modified Europe are from MatPower,[49] and were originally described in References 34,50. These systems accurately represent the European high-voltage transmission network. The first system consists of 1354 buses and operates at 380 and 220 kV, and the second consists of 9241 buses and operates at 750, 400, 380, 330, 220, 154, 150, 120, and 110 kV. A peculiar feature of the 9241-bus system is the presence of a few transmission lines with negative resistance/reactance. These lines dramatically affect the conditioning of the admittance matrix. As mentioned in Subsection 2.1, and observed in this system, diagonal dominance of the admittance matrix is lost because of these lines. Indeed, the rows corresponding to these lines have very small diagonal elements, leading to divergence in the relaxation process for generating the hierarchy of admittance matrices. Thus, we modified these lines to have positive resistance/reactance leading to a modified European system. Finally, we have three high-voltage Polish system scenarios also from MatPower. The scenarios describe different winter peak conditions: scenario 1 involves 2383 buses for the 1999–2000 winter, scenario 2 involves 2383 buses for the 1999–2000 winter, and scenario 3 involves 3374 buses for the 2007–2008 winter (one isolated bus has been removed to avoid singularities). As in the 9241-bus European system, there are a few transmission lines with negative resistance/reactance in scenario 3. Again, we modified these lines to have positive resistance/reactance to produce a modified Polish system.

Table 1 tabulates the average number of multigrid iterations over 10 simulations for the AMG-FAS and multiplicative-correction schemes with caliber 3 interpolation. Five levels were taken with the coarsening rate automatically determined by the relaxation-based coarsening algorithm. Because of the nonlinearity of the problem, the coarsening must be performed conservatively to achieve robustness in the solver (threshold parameters κ and Q of the coarsening algorithm were respectively set to 1.5 and 0.5, which are reasonable choices: node i is a seed if its projected volume is 1.5 times greater than the average projected volume, and fine node i becomes a seed if its cumulative affinity measures or coefficient connections to the coarse nodes is less than a half of its cumulative measures/connections to all of its neighboring nodes). From Table 1 the AMG-FAS scheme performed slightly better than the multiplicative-correction scheme, and at less cost per cycle since the coarse-grid operators must be updated after each cycle in the multiplicative-correction scheme.

Although parameter tuning is not the focus of this article, we consider a selection of κ-Q parameters to see how they affect the convergence of the AMG-FAS method. Table 2 displays the average number of $V(1,1)$ cycles over 10 simulations for each parameter pair, using caliber 3 interpolation. We see that slower convergence generally occurs when κ is large and Q is small. By choosing larger κ's, fewer nodes are selected as seeds in the first phase of the coarse-grid selection procedure. Smaller Q's then lead to fewer fine nodes being converted to seeds in the second phase. Hence, larger κ's and smaller Q's result in more aggressive coarsening, which often gives slower convergence. On the other hand, smaller κ's and larger Q's result in more conservative coarsening and faster convergence, but at the expense of having denser coarse-grid operators.

5.2 | Pseudopowergrid systems

The admittance matrices for the pseudosystems were generated via a finite element discretization of 2 two-dimensional elliptic boundary-value problems with homogeneous Dirichlet boundary conditions:

$$\text{Laplace}: \begin{cases} \nabla \cdot \nabla u = g_1 & \mathbf{x} \in \Omega \\ u = 0 & \mathbf{x} \in \partial\Omega \end{cases} \qquad \text{Diffusion}: \begin{cases} \nabla \cdot \mathbf{A}(\mathbf{x}) \nabla u = g_2 & \mathbf{x} \in \Omega \\ u = 0 & \mathbf{x} \in \partial\Omega, \end{cases}$$

TABLE 2 Average number of $V(1,1)$ AMG-FAS cycles using different coarsening parameters

Case	κ	Its for $Q = 0.25$	κ	Its for $Q = 0.5$	κ	Its for $Q = 0.75$
IEEE 57	1.0	13.1	1.0	9.0	1.0	10.5
	1.5	13.5	1.5	14.5	1.5	8.8
	2.0	12.7	2.0	8.3	2.0	9.6
IEEE 68	1.0	7.7	1.0	7.3	1.0	6.0
	1.5	14.0	1.5	10.4	1.5	6.2
	2.0	16.0	2.0	8.3	2.0	8.0
IEEE 118	1.0	8.9	1.0	8.3	1.0	5.9
	1.5	10.0	1.5	7.9	1.5	5.0
	2.0	10.1	2.0	7.8	2.0	6.2
IEEE 145	1.0	10.5	1.0	10.1	1.0	9.4
	1.5	11.5	1.5	11.3	1.5	7.4
	2.0	12.6	2.0	9.4	2.0	9.6
IEEE 300	1.0	36.1	1.0	31.8	1.0	20.8
	1.5	41.9	1.5	31.2	1.5	30.9
	2.0	41.7	2.0	32.6	2.0	31.4
Illinois	1.0	23.0	1.0	16.2	1.0	14.4
	1.5	24.4	1.5	16.6	1.5	14.3
	2.0	24.1	2.0	15.3	2.0	14.6
S. Carolina	1.0	17.6	1.0	15.9	1.0	16.9
	1.5	25.9	1.5	18.2	1.5	18.1
	2.0	29.0	2.0	19.3	2.0	16.8
Wisconsin	1.0	20.2	1.0	14.3	1.0	12.8
	1.5	23.5	1.5	14.5	1.5	13.8
	2.0	26.6	2.0	18.0	2.0	17.3
Texas	1.0	15.6	1.0	11.9	1.0	9.9
	1.5	20.0	1.5	12.7	1.5	11.0
	2.0	22.3	2.0	16.0	2.0	11.1
Half Eur.	1.0	26.1	1.0	18.8	1.0	19.6
	1.5	27.8	1.5	21.8	1.5	16.9
	2.0	30.6	2.0	21.2	2.0	17.8
Polish-1999	1.0	21.7	1.0	13.3	1.0	11.2
	1.5	23.7	1.5	12.1	1.5	11.0
	2.0	20.9	2.0	18.1	2.0	11.2
Polish-2004	1.0	24.8	1.0	15.1	1.0	12.7
	1.5	25.0	1.5	15.8	1.5	14.0
	2.0	23.6	2.0	16.9	2.0	13.9
Mod. Polish	1.0	29.8	1.0	24.3	1.0	19.1
	1.5	35.8	1.5	24.6	1.5	19.8
	2.0	36.1	2.0	22.9	2.0	21.8

Note: Caliber 3 interpolation was used, and the average was taken over 10 simulations for each parameter selection.

TABLE 3 Pseudopowergrid systems: Average number of $V(1,1)$ AMG-FAS cycles to reduce the initial residual by six orders of magnitude

				Oper complx		FAS its	
Case	**#Buses**	**#PV buses**	**#Levs**	**Cal. 1**	**Cal. 3**	**Cal. 1**	**Cal. 3**
Laplace	225	17	5	2.72	4.68	19	9
	961	65	6	2.95	5.03	24	10
	3969	257	6	2.86	5.34	27	10
	16,129	1025	7	2.98	5.38	31	9
Diffusion	225	17	5	2.65	4.59	16	10
	961	65	6	2.78	4.98	22	12
	3969	257	6	2.86	5.20	27	10
	16,129	1025	7	2.88	5.31	33	16

Note: Average taken over 10 runs.

where

$$\mathbf{A}(\mathbf{x}) = \begin{bmatrix} 1 + [\sin(\pi x_1)\sin(\pi x_2)]^2 & 0 \\ 0 & 1 + [\cos(3\pi x_1)\cos(\pi x_2)]^2 \end{bmatrix}$$

and Ω is the unit square. Ω was uniformly partitioned with gridpoints $(x_{1i}, x_{2j}), x_{1i} = ih_{x1}, x_{2j} = jh_{x2}$, where h_{x1}, h_{x2} are the meshsizes in the x_1 and x_2 coordinates. Applying a bilinear finite element discretization, with $\mathbf{D}$ denoting the stiffness matrix of the discretization, the admittance matrix was formed as

$$\mathbf{Y} = (0.01 - 0.1\hat{\imath})\mathbf{D}.$$

The magnitudes of the real and imaginary components in this scaling provide a reasonable ratio between the conductances and susceptances. Because of the uniform tessellation, $\mathbf{Y}$ is a structured matrix with some variation in the edge weighting (cf., Reference 19, which used an admittance matrix with a similar structured pattern but with uniform edge weighting). The PV buses were selected at gridpoints (x_{1i}, x_{2j}) with i and j striding by 4 to give roughly $\frac{n}{16}$ PV buses, and the slack bus was selected at node n. Finally, with the exact voltage magnitudes and angles chosen as

$$\hat{\mathbf{V}}^{\text{exact}} = [1 + 0.05x_{1i}(1 - x_{1i})\sin(\pi x_{2j})]_{ij} \quad \delta^{\text{exact}} = \frac{\pi}{4}[x_{1i}(1 - x_{1i})\sin(\pi x_{2j})]_{ij},$$

the power was generated using Equation (8). The initial approximation is set to $\mathbf{V}^0 = \mathbf{1}$.

Table 3 shows the average number of multigrid cycles needed to reduce the initial residual by six orders of magnitude for the AMG-FAS method using calibers 1 and 3 interpolation. This comparison of the calibers is conducted to illustrate that using higher caliber interpolation can be effective even though the analysis in Subsection 4.1.1 does not come out as clearly for higher caliber as it does for caliber 1 interpolation. From Table 3, we see that a higher-order caliber improves the convergence rate and scalability. Computationally though, this comes at a cost since higher caliber interpolation leads to denser coarse-grid operators. Using the operator complexity measure $\frac{\sum_l nnz(\mathbf{Y}_l)}{nnz(\mathbf{Y}_1)}$, where $nnz(\mathbf{Y}_l)$ is the number of nonzeros in $\mathbf{Y}_l$, we can gauge the costs of a FAS cycle. Columns 5 and 6 give the operator complexity using calibers 1 and 3 interpolation. We see that the complexity using caliber 3 interpolation is about 1.8 times that in using caliber 1. Comparing the number of FAS cycles, we see that using caliber 3 is slightly more efficient than using caliber 1 since the number of iterations using caliber 1 is greater than 1.8 times the number of iterations using caliber 3.

Table 4 tabulates the results for the multiplicative-correction scheme. The additional cost in reconstructing the coarse-grid operators after each cycle was clearly noticeable in these experiments, and hence, the experiments for this scheme were not scaled up to 16,129 buses. Since the same coarsening procedure is used to generate the hierarchy

Case	#Buses	#PV buses	#Levs	MultCorr its	
				Cal. 1	Cal. 3
Laplace	225	17	4 \ 5	20 \ 21	10 \ 13
	961	65	5 \ 6	25 \ 29	14 \ 12
	3969	257	5 \ 6	24 \ 30	10 \ 12
Diffusion	225	17	4 \ 5	15 \ 19	10 \ 15[a]
	961	65	5 \ 6	22 \ 29	10 \ 21[a]
	3969	257	5 \ 6	24 \ 31	11 \ 25[a]

TABLE 4 Pseudopowergrid systems: Average number of $V(1,1)$ multiplicative-correction cycles to reduce the initial residual by six orders of magnitude

Note: Average taken over 10 runs.

[a]Average taken over only runs that completed and converged (some runs diverged).

Case	FAS: caliber 1		FAS: caliber 3		Newton	
	#Conv/div	Its	#Conv/div	Its	#Conv/div	Its
IEEE 118	100/0	7.8	100/0	6.0	96/4	6.0
IEEE 145	100/0	13.2	100/0	12.4	0/100	–
IEEE 300	91/9	23.9	76/24	21.7	0/100	–
Texas	94/6	13.4	96/4	8.4	0/100	–
Half Europe	87/13	18.6	85/15	12.9	0/100	–
Polish-1999	89/11	10.9	88/12	9.7	0/100	–
Polish-2004	38/62	19.4	28/72	16.4	0/100	–
Mod. Polish	7/93	20.6	2/98	18.0	0/100	–

TABLE 5 Robustness experiments: The initial voltage at each bus except the slack bus is $1 + 0.1\sigma_1 + 0.1\hat{\imath}\sigma_2$, where σ_1, σ_2 are drawn from a normal distribution with mean 0 and SD 1

Note: The slack bus is set to 1. One hundred realizations were conducted for each scenario. "Its" is the average number of iterations over the converged runs. The results reflect a larger basin of attraction for the FAS method than for Newton's method.

of admittance matrices for the FAS and multiplicative-correction schemes, the operator complexities given in Table 3 provide a gauge on the cost of a cycle using the different caliber interpolations. With this, we again see that caliber 3 interpolation achieves slightly better performance than caliber 1 interpolation. However, with more variation in the edge weights (i.e., the admittance matrix generated from the diffusion equation), some runs diverged with caliber 3 interpolation. This issue was ameliorated by reducing the number of coarse levels. It also can be ameliorated by switching to a lower caliber on the coarser levels. This issue may also reflect problems arising from inadequately satisfying the complementary smoothing/coarse-grid principle when using higher caliber $\mathcal{I}_{l+1}^{l}$ in the solve phase of the multiplicative scheme (see Subsection 4.2). Finally, we see that the AMG-FAS method performs better than the multiplicative-correction method.

5.3 | Robustness: Basin of attraction

To compare the robustness of the AMG-FAS and Newton's methods, a selection of the realistic problems were considered. Except for the slack bus, at each bus, the initial approximation was taken to be a random perturbation of 1, that is, $\mathbf{V}_i = 1 + c\sigma_1 + c\hat{\imath}\sigma_2$ where σ_1, σ_2 were drawn from a normal probability distribution with mean 0 and SD 1. One hundred realizations were performed, and again convergence was taken to be a six-order reduction in the norm of the initial residual. Tables 5 and 6 give the results when $c = 0.1$ and 0.025, respectively. In these tables, "Its" is the average number of iterations to reach convergence for only the converged runs. We see that the AMG-FAS scheme is more robust than Newton's scheme, reflecting a larger basin of attraction for the former method. However, for scenarios where Newton's method converges, a much faster rate of convergence is achieved for Newton's method.

TABLE 6 Robustness experiments: Same setup as in Table 5 but now with the perturbations being $0.025\sigma_1 + 0.025\hat{\imath}\sigma_2$

	FAS: caliber 1		FAS: caliber 3		Newton	
Case	**#Conv/div**	**Its**	**#Conv/div**	**Its**	**#Conv/div**	**Its**
IEEE 118	100/0	8.8	100/0	10.3	100/0	4.0
IEEE 145	100/0	10.4	100/0	8.9	95/5	6.2
IEEE 300	100/0	31.4	100/0	24.6	0/100	–
Texas	100/0	14.8	100/0	8.0	100/0	6.0
Half Europe	100/0	17.2	100/0	11.8	27/73	9.8
Polish-1999	100/0	14.0	100/0	11.0	0/100	–
Polish-2004	100/0	14.1	100/0	12.3	0/100	–
Mod. Polish	100/0	21.2	100/0	21.4	0/100	–

Note: The results reflect a larger basin of attraction for the FAS method than for Newton's method.

6 | CONCLUSION

In this article, we explored two nonlinear multigrid methods for solving the powerflow equations. The key procedure in these solvers is constructing the hierarchy of powerflow equations using a hierarchy of admittance matrices, which is obtained by a relaxation-based coarsening algorithm. The interpolation operators generated for the hierarchy of admittance matrices can be used in the multigrid solve phase. Justification of this can be illustrated for the AMG-FAS method. Numerical experiments demonstrate that the methods can be effective, although not as robust as multigrid methods for linear problems. Moreover, these experiments show that the AMG-FAS scheme performs better than the multiplicative-correction scheme, and at a lower computational cost since the latter scheme requires a reconstruction of the coarse-grid operators after each multigrid cycle. Finally, numerical experiments comparing the AMG-FAS scheme to Newton's method demonstrate a larger a basin of attraction for the AMG-FAS scheme, albeit its convergence rate is only linear. Nevertheless, applying a few cycles of this scheme can produce a sufficiently accurate initial approximation for quadratic convergence to be achieved in Newton's method.

ACKNOWLEDGMENTS

The authors would like to thank two anonymous referees for their thoughtful comments and suggestions that improved this paper. This work was supported in part by NSF grant DMS-1734727.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ORCID

Barry Lee https://orcid.org/0000-0003-1948-7384

REFERENCES

1. Van Cutsem T, Vournas C. Voltage stability of electric power systems. New York, NY: Springer, 2008.
2. Makarov YV, Vyakaranam B, Wu D, Lee B, Hou A, Elbert S, Huang Z. On the configuration of the US Western Interconnection voltage stability boundary. Proceedings of the IEEE Power Engineering Society Transmission and Distribution Conference, New York, NY: Curran Associates, Inc; 2014. p. 1–5.
3. Ward JB, Hale HW. Digital computer solution of powerflow problems. AIEE Trans Power App Syst. 1956;75:398–404.
4. Brown RJ, Tinney WF. Digital solutions for large power networks. AIEE Trans Power App Syst. 1957;76:347–355.
5. Glimn AF, Stagg GW. Automatic calculation of load flows. AIEE Trans Power App Syst. 1957;76:817–828.
6. Stott B. Review of load-flow calculation methods. Proc IEEE. 1974;62:916–929.
7. Tinney WF, Hart CE. Power flow solution by Newton's method. IEEE Trans Power App Syst. 1967;PAS-86:1449–1456.
8. Stott B. Decoupled Newton load flow. IEEE Trans Power App Syst. 1972;PAS-91:1955–1959.
9. Murray W, De Rubira TT, Wigington A. Improving the robustness of Newton-based power flow methods to cope with poor initial points. Proceedings of the 2013 North American Power Symposium (NAPS), Manhattan, KS; 2013. p. 1–6.
10. Palmer B, Perkins W, Chen Y, et al. GridPACK: A framework for developing power grid simulations on high-performance computing platforms. Int J High Perform Comput Apps. 2016;30:223–240.

11. Abhyankar S, Smith B, Constantinescu E. Evaluation of overlapping restricted additive Schwarz preconditioning for parallel solution of very large power flow problems. Proceedings of the International Conference for High Performance Computing, Network: Storage and Analysis (SC 2013), New York, NY: ACM; 2013.
12. Lindberg J, Zachariah A, Boston N, Lesieutre B. The geometry of real solutions to the power flow equations. Poster session presented at 2018 56th Annual Allerton Conference on Communication, Control, and Computing (Allerton), Monticello, IL; 2018. p. 596–603.
13. Lesieutre B, Lindberg J, Zachariah A, Boston N. On the distribution of real-valued solutions to the power flow equations. Poster session presented at 2019 57th Annual Allerton Conference on Communication, Control, and Computing (Allerton), Monticello, IL; 2019. p. 165–170.
14. Krishnamurthy D, Low S, Chertkov M. Solving the power flow equations: A monotone operator approach; 2015. arXiv preprint arXiv:1506.08472. doi:https://doi.org/10.2172/1210207, https://www.osti.gov/servlets/purl/1210207,
15. Idema R, Papaefthymious G, Lahaye D, Vuik C, Van Der Sluis L. Towards faster solution of large power flow problems. IEEE Trans Power Syst. 2013;28:4918–4925.
16. Lee B. Multigrid for model reduction of power grid networks. Numer Lin Alg Appl. 2018;25(6):e2201. https://doi.org/10.1002/nla.2201.
17. Dorfler F, Bullo F. Exploring synchronization in complex oscillator networks. Poster session presented at 2012 IEEE 51st Annual Conference on Decision and Control (CDC), New York, NY: IEEE; 2012. p. 7157–7170.
18. Machowski J, Bialek JW, Bumby JR. Power system dynamics stability and control. 2nd ed. West Sussex, UK: John Wiley, 2008.
19. Ponce C, Bindel DS, Vassilevski PS. A nonlinear algebraic multigrid framework for the power flow equations. SIAM J Sci Comput. 2018;40:B812–B833.
20. De Sterck H, Manteuffel TA, McCormick SF, et al. Smoothed aggregation multigrid for Markov chains. SIAM J Sci Comput. 2010;32:40–61.
21. De Sterck H, Manteuffel TA, McCormick SF, Miller K, Ruge J, Sanders G. Algebraic multigrid for Markov chains. SIAM J Sci Comput. 2010;32:544–562.
22. Brandt A. Multigrid Techniques: 1984 Guide, with Applications to Fluid Dynamics, (GMD Studien Nr. 85, GMD-AIW, Postfach 1240, D-5205, St. Augustin 1, W. Germany, 1984); 1984.
23. Hackbusch W. Multi-Grid methods and applications. Berlin, Germany: Springer-Verlag, 1985.
24. Trottenberg U, Oosterlee CW, Schuller A. Multigrid. London, UK: Academic Press, 2001.
25. Reusken A. Convergence of the multigrid full approximation scheme for a class of elliptic mildly nonlinear boundary value problems. Numer Math. 1988;52:251–277.
26. Yavneh I, Dardyk G. A multilevel nonlinear method. SIAM J Sci Comput. 2006;28:24–46.
27. la Cour Christensen M, Vassilevski PS, Villa U. Nonlinear multigrid solvers exploiting AMGe coarse spaces with approximation properties. J Comput Appl Math. 2018;340:691–708.
28. Brabazona KJ, Hubbard ME, Jimack PK. Nonlinear multigrid methods for second order differential operators with nonlinear diffusion coefficient. Comput Math Appl. 2014;16:1619–1634.
29. Dorfler F, Bullo F. On the critical coupling for Kuramoto oscillators. SIAM J Appl Dyn Syst. 2011;10:1070–1099.
30. Dorfler F, Bullo F. Synchronization and transient stability in power networks and non-uniform Kuramoto oscillators. SIAM J Control Optim. 2012;50:1616–1642.
31. Dorfler F, Chertkov M, Bullo F. Synchronization in complex oscillator networks and smart grids. Proc Nat Acad Sci. 2013;110:2005–2010.
32. Kundur P. Power system stability and control. New York, NY: McGraw-Hill, 1994.
33. Andersson G. Modeling and analysis of electric power systems. Lecture 227-0526-00. Zurich: ITET ETH Zurich, 2008.
34. Josz C, Fliscounakis S, Maeght J, Panciatici P. AC power flow data in MATPOWER and QCQP format: iTesla, RTE snapshots, and PEGASE; 2016. arXiv 1603.01533.
35. Guckenheimer J, Holmes P. Nonlinear oscillations, dynamical systems, and bifurcations of vector fields. New York, NY: Springer-Verlag, 1983.
36. Livne OE. Coarsening by compatible relaxation. Numer Lin Alg Apps. 2004;11:205–227.
37. Livne OE, Brandt A. Lean algebraic multigrid (LAMG): Fast graph Laplacian linear solver. SIAM J Sci Comput. 2012;34:B499–B522.
38. Ron D, Safro I, Brandt A. Relaxation-based coarsening and multiscale graph organization. Multiscale Model Simul. 2011;9:407–423.
39. Brezina M, Falgout RD, MacLachlan S, Manteuffel T, McCormick S, Ruge J. Adaptive algebraic multigrid. SIAM J Sci Comput. 2006;27:1261–1286.
40. Manteuffel T, McCormick S, Park M, Ruge J. Operator-based interpolation for bootstrap algebraic multigrid. Numer Lin Alg Appl. 2010;17:519–537.
41. Brandt A, Brannick J, Kahl K, Livshitz I. Bootstrap AMG. SIAM J Sci Comput. 2011;32:612–632.
42. Ruge JW, Stuben K. Algebraic multigrid (AMG). In: McCormick SF, editor. Multigrid Methods, Frontiers in Applied Mathematics. Philadelphia, PA: SIAM, 1987.
43. MacLachlan SP, Oosterlee CW. Algebraic multigrid solvers for complex-values matrices. SIAM J Sci Comput. 2008;30:1548–1571.
44. Suli E, Mayers D. An introduction to numerical analysis. Cambridge, UK: Cambridge University Press, 2006.
45. University of Washington Department of Electrical Engineering. https://labs.ece.uw.edu/pstca/.
46. Department of Energy Singular perturbations, coherency and aggregation of dynamic systems Office of science technical report. U.S. Department of Energy, Office of Science and Technical Information, 1983. https://www.osti.gov/.
47. Department of electrical engineering, electric grid test case repository; Texas A&M University, College Station, TX: https://electricgrids.engr.tamu.edu/electric-grid-test-cases/.

48. Birchfield AB, Xu T, Gegner KM, Shetye KS, Overbye TJ. Grid structural characteristics as validation criteria for synthetic networks. IEEE Trans Power Syst. 2017;32:3258–3265.
49. MatPower Software. https://matpower.org/.
50. Fliscounakis S, Panciatici P, Capitanescu F, Wehenkel L. Contingency ranking with respect to overloads in very large power systems taking into account uncertainty, preventive, and corrective actions. IEEE Trans Power Sys. 2013;28:4909–4917.

How to cite this article: Lee B, Pereira Batista E. Algebraic multigrid for the nonlinear powerflow equations. *Numer Linear Algebra Appl*. 2021;28:e2347. https://doi.org/10.1002/nla.2347