Towards Optimal Kron-based Reduction Of Networks (Opti-KRON) for the Electric Power Grid

Samuel Chevalier[1] and Mads R. Almassalkhi[2]

Abstract— For fast timescales or long prediction horizons, the AC optimal power flow (OPF) problem becomes a computational challenge for large-scale, realistic AC networks. To overcome this challenge, this paper presents a novel network reduction methodology that leverages an efficient mixed-integer linear programming (MILP) formulation of a Kron-based reduction that is optimal in the sense that it balances the degree of the reduction with resulting modeling errors in the reduced network. The method takes as inputs the full AC network and a pre-computed library of AC load flow data and uses the graph Laplacian to constraint nodal reductions to only be feasible for neighbors of non-reduced nodes. This results in a highly effective MILP formulation which is embedded within an iterative scheme to successively improve the Kron-based network reduction until convergence. The resulting optimal network reduction is, thus, grounded in the physics of the full network. The accuracy of the network reduction methodology is then explored for a 100+ node medium-voltage distribution feeder example across a wide range of operating conditions. It is finally shown that a network reduction of 25-85% can be achieved within seconds and with worst-case voltage magnitude deviation errors within any super node cluster of less than 0.01pu. These results illustrate that the proposed optimization-based approach to Kron reduction of networks is viable for larger networks and suitable for use within various power system applications.

I. INTRODUCTION

Understanding how to best utilize resources distributed over a network has been and is an important question across many industries. For the power/energy industry, solving the centralized AC optimal power flow (OPF) problem is NP-hard and has been the focus of much research since the 1960s [1]–[3] and more recently, as optimization solvers matured [4], [5]. In some cases, the OPF problem is cast within the setting of (transmission) expansion planning and considers a large number of scenarios, decade-long prediction horizons, and many possible investment decisions [6]. In other cases, the focus of the OPF problem is near-term grid operations to determine active and/or reactive power set-points for PV inverters, batteries, and other controllable assets in the grid to minimize operating costs, line losses, voltage deviations from nominal, or to achieve a desired net-load profile that reflects whole-sale energy market conditions. These power set-points can be updated every minute or hour in a microgrid [7] or distribution grid applications [8], [9] corresponding to the timescales of grid or market conditions, e.g., renewable injections or frequency regulation signals. Thus, many applications of the AC OPF requires a mix of long prediction horizons and frequent re-computations, which for large networks are computationally challenging.

To overcome the computational challenges associated with solving practical (large-scale) AC OPF problems to (global) optimality, the power/energy community has often studied approximations of the AC physics, such as the so-called (linear) DC power flow [10], convex relaxations [11], [12], convex restrictions [13], and various distributed implementations [14]. However, if the AC network could be made a lot smaller, while representing the physics of the network sufficiently well, the computational challenge would decrease significantly [15]. Thus, in this paper, we focus on a novel method for (optimally) reducing the AC network, which could then be employed within an appropriate OPF setting.

Network reductions, not to be confused with model-order reduction from systems theory [16]–[18], have been studied extensively and employ a variety of methods, such as similarity or (electrical) distance measures for clustering, bus aggregations (e.g., REI), and equivalence techniques (e.g., Ward and Kron). In the case of reducing nodes belonging to an "external" area, which are nodes that are geographically or electrically distanced from the "internal" area, network reduction via Ward- or Kron-based methods can be readily applied and has been standard practice for decades [19]. However, recently techniques have focused inwards on the internal area or so-called "backbone-type" network reductions, where any nodes can be reduced in the network rather than just "external" nodes. These backbone-type equivalents rely on either an initial clustering approach (e.g., *k*-means clustering) to group nodes together into contiguous zones or a pre-defined set of zones. Once the nodes are assigned to specified zones (or subgrids), a network reduction can be readily applied to said zones (e.g., via Ward and Kron or heuristics) and possibly tuned based on some criteria. For example, network-preserving bus aggregation methods by [20] and [21] employ nonlinear and quadratic optimization, respectively, to tune (susceptance values in) the reduced admittance matrix so as to minimize tie line flow errors with respect to the full network. In [20], the method depends on pre-determined zones and a specific operating point to calculate the full network's power transfer distribution

factors (PTDFs). The algorithm [21] replaces the zonal input requirement with a list of pre-determined salient tie lines and also uses PTDFs, which inform a bus clustering algorithm that defines internal zones, which are then subject to bus aggregation. These methods can reduce 60,000-bus networks by up to 100X in the order of minutes (on a super computer) with small inter-zonal worst-case flow deviation errors - even under different operating conditions. Other approaches sidestep the dependence on operating points by employing DC load flow analysis in deriving independent PTDF values [22]. In this case, a 15,000-bus network is reduced by 85X after eight hours with relative line flow errors of less than 30%. Lastly, some methods are built around multiple clustering objectives and heuristics that preserve physical features and network structure, but are overly conservative (i.e., only reduce by 2-3X while line flow deviation errors are around 5-10%) [23].

Kron-based network reductions have been shown to be valuable across numerous applications in power system analysis [19], [24]. For example, comprehensive transmission planning schemes have been built around Kron-based equivalents that employ various optimization formulations whose solutions serve as seeds to identify a set of salient buses and lines to partition the network [19]. To speed up 3-phase distribution grid OPF, [9] presents a Kron-based network reduction, where a desired level of reduction informs a nodal clustering scheme that determines which nodes are reduced. This method was able to achieve 10-50X reductions in realistic distribution feeders with maximum voltage deviation errors (between reduced nodes and their corresponding non-reduced "super" node in the same cluster) of less than 0.015pu across a wide range of operating conditions.

Across these different approaches to network reductions in power systems, they all depend on pre-specified salient buses, tie-lines, and/or level of desired reduction as inputs. Clearly, these inputs affect the resulting network reduction and this is what motivates a simple, but interesting question: *is there an optimal Kron-reduction?* More precisely: are there a set of nodes and a level of reduction that is optimal (in some sense) when reducing a network? Thus, as a first step towards answering this question, the paper's key contribution is a novel network reduction methodology that leverages a mixed-integer linear programming (MILP) formulation to determine a Kron-based reduction that is optimal in the sense that it automatically balances the level of reduction (i.e., complexity) with resulting worst-case voltage deviation errors between the reduced and full networks. The method is based on a pre-computed library of AC load flow data (i.e., operating points) and guarantees that any feasible solution is a valid Kron reduction that preserves the network's structure. As far as the authors are aware, there is no other literature that casts a Kron-based network reduction entirely within an efficient MILP optimization formulation. To ensure tractability in the MILP formulation, we constrain nodes to only reduce to a "super node," if they are neighbors (as defined by the graph Laplacian). Then, we successively reduce the network via an iterative scheme to overcome the

nodal neighbor limitation. The entire methodology, denoted Opti-KRON, is validated via simulation-based analysis on a 115-node, radial and balanced IEEE test network, which represents a minor contribution as it provides insight on different optimal Kron-based network reductions.

The remaining paper is structured as follows. Section II presents the network model and summarizes Kron reductions. In Section II-A, the MILP formulation for Kron-based network reduction is presented. Simulation-based analysis is presented in Section III. Finally, the paper concludes in Section IV with a summary and a brief discussion on future directions and applications.

II. NETWORK MODEL AND KRON-REDUCTION

For the sake of notational simplicity, consider a single-phase power system network whose graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ has edge set $\mathcal{E}$, $|\mathcal{E}| = m$, vertex[1] set $\mathcal{V}$, $|\mathcal{V}| = n$, and signed nodal incidence matrix $E \in \mathbb{R}^{m \times n}$. The complex nodal admittance matrix (i.e., Y-bus matrix) $Y_b \in \mathbb{C}^{n \times n}$ associated with this system is constructed via

$$Y_b = E^T Y_l E + Y_s, \qquad (1)$$

where $Y_l \in \mathbb{C}^{m \times m}$ is the diagonal matrix of complex line admittances and $Y_s \in \mathbb{C}^{n \times n}$ is the diagonal matrix of complex nodal shunt admittances. In this paper, we generally assume $Y_s \neq 0$, implying Y_b is a nonsingular matrix. Leveraging this property, the so-called nodal impedance matrix $Z_b \in \mathbb{C}^{n \times n}$ can be directly computed as the inverse of (1): $Z_b = Y_b^{-1}$. The nodal admittance (and impedance) matrices directly relate complex nodal voltages and current injections via $I = Y_b V$ (and $Z_b I = V$).

A. The Kron-Reduction Procedure

Without loss of generality, we partition the network via

$$I = Y_b V \qquad (2)$$

$$\left[\frac{I_k}{I_r} \right] = \left[\begin{array}{c|c} Y_{b1} & Y_{b2} \\ \hline Y_{b3} & Y_{b4} \end{array} \right] \left[\frac{V_k}{V_r} \right], \qquad (3)$$

where subscripts "r" and "k" denote nodes which are ultimately reduced and kept, respectively. As in [24], Gaussian elimination of the nodal voltages V_r is achieved by

$$I_k = \left(Y_{b1} - Y_{b2} Y_{b4}^{-1} Y_{b3} \right) V_k + \left(Y_{b2} Y_{b4}^{-1} \right) I_r. \qquad (4)$$

The Kron reduction of (2), which is used to "eliminate" nodes with zero current injection (i.e., $I_r = 0$), is canonically given by the following Schur complement:

$$Y_K = Y_{b1} - Y_{b2} Y_{b4}^{-1} Y_{b3}. \qquad (5)$$

Alternatively, Y_K can be constructed using the network impedance matrix, whose associated partition is given by

$$\left[\begin{array}{c|c} Z_{b1} & Z_{b2} \\ \hline Z_{b3} & Z_{b4} \end{array} \right] \left[\frac{I_k}{0} \right] = \left[\frac{V_k}{V_r} \right]. \qquad (6)$$

---

[1] In power systems, a vertex in a power network is commonly denoted a node in distribution systems and a busbar (or bus) in transmission systems. Given the general discussion of power networks, we will use node and bus interchangeably.
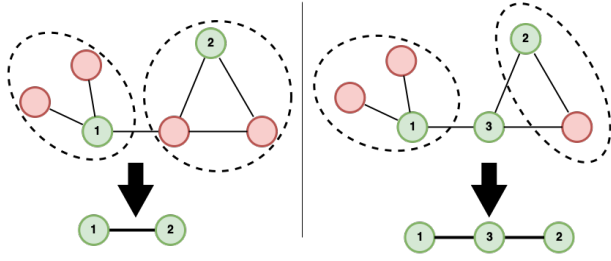
Fig. 1. Illustration of how a network can be partitioned in two different ways to yield two different Kron-reduced networks. The partition is based on reduced and kept (super) nodes. The green, numbered circles represent kept nodes or super nodes, while red circles are reduced nodes and eliminated in the reduced network. The dashed (---) ellipses illustrate which reduced nodes are assigned to which super nodes and define how injected currents are assigned to each super node.

Remark 1. *The Kron-reduced admittance Y_K is equal to the inverse of sub-impedance Z_{b1}.*

Proof. By construction, the Kron admittance relates $I_k = Y_K V_k$. From (6), $Z_{b1} I_k = V_k$. Therefore, $Z_{b1} = Y_K^{-1}$. ∎

In the following, we define the Kron impedance matrix as $Z_K \triangleq Z_{b1}$ from (6). Note that any removal of rows and columns from Z_K will result in a valid Kron impedance matrix, in the sense that it will relate nodal voltages and currents. Thus, if we can optimally select which nodes to reduce and where to assign them, we can effectively choose an optimal set of rows and columns to remove from Z_K. This would then allow us to define an optimal Kron impedance matrix through a set of binary decisions, which is illustrated in Fig. 1. This inspires the following mixed-integer formulation.

B. Mixed-Integer Approach for Constructing Kron Matrices

In the following, we define binary variable $s_i \in \{0, 1\}$, which selects the optimal Kron impedance matrix Z_k, where $s_i = 0$ or 1 indicates if the ith node is reduced or kept, respectively. We accordingly define binary vector $s \in \{0, 1\}^n$ and the associated diagonal selection matrix $S \triangleq \text{diag}\{s\}$. For any given binary values, the matrix product $S Z_b S$ thus yields a matrix which we refer to as a generalized Kron impedance, defined as $Z_{\hat{K}} \triangleq S Z_b S$. This generalized Kron impedance has the dimensions of the full nodal impedance matrix ($n \times n$), but a subset of its rows and columns are zeroed-out. As an example, the generalized Kron impedance of (6) is

$$Z_{\hat{K}} = \left[\begin{array}{c|c} I & 0 \\ \hline 0 & 0 \end{array}\right] \left[\begin{array}{c|c} Z_{b1} & Z_{b2} \\ \hline Z_{b3} & Z_{b4} \end{array}\right] \left[\begin{array}{c|c} I & 0 \\ \hline 0 & 0 \end{array}\right] = \left[\begin{array}{c|c} Z_{b1} & 0 \\ \hline 0 & 0 \end{array}\right], \quad (7)$$

where the diagonal binary values of S "kept" the top nodes and "reduced" the bottom ones.

We define an additional binary decision matrix, $A \in \{0, 1\}^{n \times n}$, which codifies where currents from reduced nodes are placed. Accordingly, $A_{i,j} = 1$ if the nodal current injection from bus j is placed at bus i, and $A_{i,j} = 0$ otherwise. Since current can only be assigned to a single bus, $\sum_i A_{i,j} = 1$ is always enforced. Furthermore, $\sum_j A_{i,j} \leq M_b S_{i,i}$ ensures that currents cannot be assigned to a reduced bus and $S_{i,i} = A_{i,i}$

guarantees that each non-reduced bus does not move its own current. Based on these rules, the matrix vector product $I_K = AI$ naturally and properly aggregates currents at non-reduced nodes (i.e., Kron currents), and the following product allows for the direct computation of Kron voltages:

$$V_K = S Z_b S A I. \quad (8)$$

We define the non-reduced nodes as "super nodes", and the Kron reduced voltages at these super nodes are given by (8).

In order to compute an *optimal* Kron reduction, we need to define an objective function, which balances the trade off between complexity (i.e., level of detail) and corresponding nodal voltage deviation error (i.e., performance of reduced network). As the number of reduced nodes is a measure of reduction in network complexity, we can capture this by minimizing the number non-zero binary values in the reduction vector s. In order to quantify voltage deviation error (which generally increases as network reduction increases), we take the infinity norm of the difference between the Kron voltage at the ith super node ($V_{K,i}$) and the voltages at all nodes within its cluster $C_{K,i}$, across all potential super nodes. The resulting objective function is then given by

$$\mathcal{L} = \underbrace{\left\| V_{K,i} - V_{j \in C_{K,i}} \right\|_\infty}_{\text{error}} - \underbrace{\alpha \frac{1}{n} \sum_{j=1}^{n} (1 - s_j)}_{\text{complexity}}, \quad (9)$$

where α balances these two terms. We note that the network currents I in (8) and the cluster voltages V_j in (9) are assumed to be given as input data libraries for the optimization problem. Ideally, these data vectors (or matrices) come from representative AC power flow solutions collected on the full network. Thus, an optimal (with respect to (9)), yet, naive MILP-based Kron reduction can be stated as

$$\min_{s,A} \ \left\| V_{K,i} - V_{j \in C_{K,i}} \right\|_\infty - \frac{\alpha}{n} \sum_{j=1}^{n} (1 - s_j) \quad (10a)$$

$$\text{s.t. } V_K = S Z_b S A I \quad (10b)$$

$$s_i, A_{i,j} \in \{0, 1\} \quad (10c)$$

$$S = \text{diag}(s) \quad (10d)$$

$$\sum_i A_{i,j} = 1 \quad (10e)$$

$$\sum_j A_{i,j} \leq M_b S_{i,i} \quad (10f)$$

$$S_{i,i} = A_{i,i}. \quad (10g)$$

While (10) will generally compute a valid Kron impedance matrix in (10b), the given formulation presents a variety of challenges. First, it does not formally constrain current injections associated with reduced nodes from being placed on super nodes which are electrically or geographically "far" from their physical location. Second, the product $S Z_b S A I$ in (10b) contains cubic binary terms. And third, this problem is generally intractable for large-scale systems, since matrix A is a binary matrix which engenders a large branch-and-bound search space for MILP solvers. In the following subsection, we address all three of these challenges to engender a tractable MILP-based Kron reduction.

C. Formulation Improvements

The issue of cubic binary terms in (10b) is sidestepped by first simplifying the product term SA.

Lemma 1. $SA = A$.

Proof. Constraint (10f) forces the j-th row of A to 0 if $S_{j,j} = 0$; thus, reduced nodes (i) must place their currents somewhere else ($A_{j,j} = 0$) and (ii) cannot receive currents from other reduced nodes ($A_{j,i} = 0$). If when a binary, whose value is 0, only multiplies other binaries whose value is also 0, then the original binary has no effect. Thus, $SA = A$. ∎

Using $SA = A$, we now have that $V_K = SZ_bAI$. However, we cannot apply the same trick again to simplify SZ_bA since Z_b is generally a dense matrix. This means that each element of A is multiplied by each diagonal element of S. Since the product of any two binaries can be reformulated in linear form (thus, "linearizing" the expression) by introducing a third auxiliary binary variable, directly linearizing SZ_bA will generally require n^3 binary auxiliary variables; e.g., binary matrices $B_1 = S_{1,1} \times A$, $B_2 = S_{2,2} \times A$, etc. Rather than directly linearizing this expression, however, we can leverage the physically motivated observation that any error accumulated by removing S from the Kron voltage equation (denoted with a tilde: $\tilde{V}_K = Z_bAI$) can be subsumed into auxiliary big-M slack factors. To do so, we reformulate the infinity norm in the objective function of (10) with a continuous slack variable δ:

$$\min_{s,A,\delta} \ \delta - \frac{\alpha}{n} \sum_{j=1}^{n} (1 - s_j) \tag{11a}$$

$$\text{s.t.} \ \tilde{V}_{K,i} - V_{j \in C_{K,i}} \le \delta + M_b(1 - A_{i,j}), \forall i \tag{11b}$$

$$V_{j \in C_{K,i}} - \tilde{V}_{K,i} \le \delta + M_b(1 - A_{i,j}), \forall i \tag{11c}$$

$$\tilde{V}_K = Z_bAI \tag{11d}$$

$$(10e) - (10g)$$

Lemma 2. *Despite the Kron voltage error in (11d) caused by the elimination of S, (11) and (10) have identical minimizers.*

Proof. Multiplying $S\tilde{V}_K$ yields V_K, so S effectively zeros-out non-super node voltages. However, it does not change the value of the super node voltage itself. $A_{i,j} = 1$ indicates that node j is inside the cluster associated with super node i. In this case, $M_b(1 - A_{i,j}) = 0$, and δ will be a supremum for the exact intra-cluster voltage deviations (since super node voltages are preserved in $\tilde{V}_K$). However, when $A_{i,j} = 0$, node j is not internal to the cluster associated with node i, which may or may not be a super node. Therefore, $M_b(1 - A_{i,j}) = M_b$ will safely upper bound any voltage deviation between $\tilde{V}_K$ and $V_{j \in C_{K,i}}$, thus leaving δ unaffected. Since δ accurately captures the infinity norm value from (10a), the programs must have identical minimizers. ∎

In order to avoid allowing the optimizer to add current from reduced nodes far from the super node itself, we employ the graph Laplacian to constrain current aggregation only at neighboring nodes. To accomplish this, we enforce the

binary values in matrix A (which chooses where currents are aggregated) to satisfy

$$A_{i,j} \le |E^T E|_{i,j}, \tag{12}$$

where $|E^T E|_{i,j}$ is the i,j-th entry of the absolute value of the graph Laplacian. Therefore, if two nodes are not direct neighbors, then their currents cannot be aggregated together. Not only does this prevent currents from being placed in non-physically meaningful places, it also greatly limits the size of the search space, thus greatly increasing the tractability of (10).

D. Successive Enhancement of Reduced Networks

While (11) & (12) jointly represent a highly tractable mathematical program, the degree of reduction it can achieve is limited by the graph Laplacian constraint in (12). Since nodes can only aggregate with their neighbors, the algorithm cannot typically achieve more than a 60% network reduction in a given solve. To overcome this hurdle, we propose an iterative implementation. That is, find an optimal network reduction, construct the reduced network, and then find another optimal network reduction of the pre-reduced network. This procedure is repeated until either (i) the desired level of reduction is achieved, or (ii), zero nodes are reduced. In order to control the maximal size β of network reduction (i.e., force optimizer to make small network reductions at each step), or control the maximal acceptable voltage deviations γ, we can embed associated constraints directly in the program.

$$\min_{A,\delta} \ \delta - \frac{\alpha}{n} \sum_{j=1}^{n} (1 - A_{j,j}) \tag{13a}$$

$$\text{s.t.} \ \sum_{j=1}^{n} (1 - A_{j,j}) \le n\beta \tag{13b}$$

$$\delta \le \gamma \tag{13c}$$

$$(10e) - (10g), (11b) - (11d), (12).$$

This iterative approach is illustrated in Fig. 2 and described algorithmically in Algorithm 1. We refer to the tractable mathematical program given by (13) as Opti-KRON. Next, we apply Opti-KRON and Algorithm 1 to optimally Kron reduce an IEEE test network, which represents a balanced, medium-voltage, radial distribution feeder.

III. EXPERIMENTAL RESULTS FOR 115-NODE RADIAL NETWORK

The 115-node radial, balanced IEEE distribution test feeder from [25] provides Y_b and is used herein to illustrate Algorithm 1. To balance complexity and error, $\alpha = 0.002$, while the maximum reduction in complexity for a single iteration of (11) is limited initially to 25% (i.e., $\beta = 0.25$). The worst-case voltage deviation error is effectively unconstrained by setting $\gamma = 1.0$pu. Finally, two distinct nodal net-injection profiles are applied to the network to beget the network's necessary data scenarios on complex branch currents, I, and nodal voltages, V. The corresponding voltage profiles, $|V_i| \ \forall i = 1, \ldots, n$ are shown in Fig. 3.
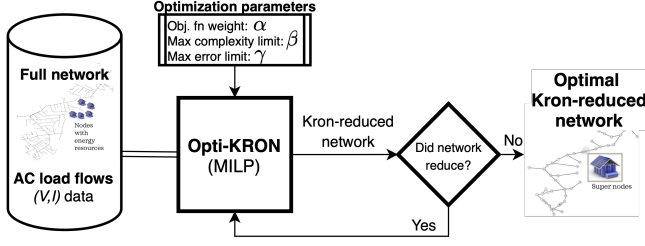
Fig. 2. The algorithm for successively enhancing the Kron-reduced network uses Opti-KRON, which is given by (13). The inputs are network and AC load flow data and the parameters that define the MILP formulation's objective function. The output is an optimal Kron-reduced network where kept nodes are denoted super nodes.

Algorithm 1: Opti-KRON Successive Enhancement

Data: $Y_b, V, I, \alpha, \beta, \gamma$
Result: Z_K (Optimal Kron-reduced network)
$p = 0, s^{(p)} = \mathbf{1}_n, \Delta s^{(p)} = n$;
$Z_b = Y_b^{-1}$;
while $\Delta s^{(p)} > 0$ **do**
$\quad s^{(p+1)} \leftarrow$ Solve Opti-KRON in (13);
$\quad \Delta s^{(p+1)} = \mathbf{1}_n^\top (s^{(p)} - s^{(p+1)})$;
$\quad p \leftarrow p+1$;
end
$S \leftarrow \text{diag}\{s^{(p)}\}$;
$Z_K \leftarrow S Z_b S A$;

With all input data now available, Algorithm 1 can be executed and converges in eight iterations and under five seconds total, which highlights tractability of Opti-KRON. The resulting optimal Kron-based network reduction has eliminated 85% of nodes, yet embodies a worst-case intra-cluster voltage deviation error across both load scenarios of less than 0.007pu. To investigate the accuracy of Opti-KRON, we subject the optimal Kron reduction *at each iteration* to operating conditions that sweep from low-load to high-load conditions (via a convex combination of the initial injection data). Then, we record the maximum intra-cluster (super node) voltage deviation errors, which are illustrated in Fig. 4. These results clearly show that despite subjecting the optimal Kron reduction to a wide range of operating conditions, the worst-case voltage deviation errors are still very small across all super nodes and loading conditions (i.e., all super node clusters deviate from their corresponding reduced nodes by less than 0.0065pu). The fact that errors do not increase away from known input data scenarios (V, I) (which are at either end in Fig. 4) may seem surprising. However, AC load flows are nonlinear, the optimal Kron-reduction minimizes the worst-case voltage errors, and the two load scenarios were low- and high-load conditions. This means that away from high-load conditions (which was in our initial set of data), the voltages at each node will become closer to 1.00pu and, thus, closer to each other, which reduces voltage deviation errors. Thus, including high net-load demand profiles to generate initial input data that has
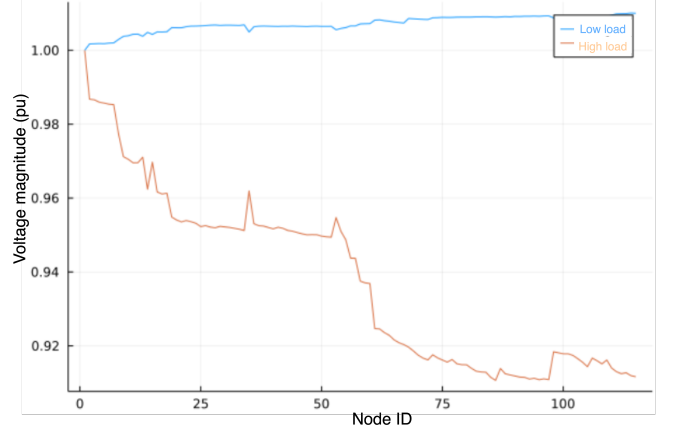

Fig. 3. The voltage profile resulting from two distinct net-load injections. The red line represents a heavily loaded scenario, while the higher voltage for the blue line represents a lightly loaded scenario with more solar PV injections.

large voltage deviations may help find an optimal network reduction that captures the full system behavior accurately. In addition, the structure-preserving nature of the optimal Kron reduction appears quite valuable to represent a wide range of operating conditions.

Lastly, to understand the effects of constraining the complexity at each iteration, we explored different upper bounds, $\beta = \{0.10, 0.25, 0.50, 0.75\}$. Then, we looked at the number of iterations required to achieve a converged optimal Kron-based network reduction, the level of the reduction, and the corresponding worst-case voltage errors. Results are summarized in Table I and show that smaller bounds can reduce overall errors, but at the cost of the reduction itself. The best overall point is $\beta = 0.25$, with high level of reduction and reasonably small voltage error (< 0.01pu).

TABLE I
DIFFERENT UPPER BOUNDS ON COMPLEXITY (β)

item	$\beta = 0.10$	0.25	0.50	0.75
Iterations (#)	17	8	7	7
Reduction (%)	75	85	83	83.5
Voltage error (pu)	0.0030	0.0065	0.0035	0.005

IV. CONCLUSION AND FUTURE WORK

This paper develops a novel and efficient mixed-integer linear optimization-based methodology for generating structure-preserving network reductions of electric power networks. The MILP formulation enables trading off complexity (in the number of reduced nodes) and errors (in terms of worst-case voltage deviations across all super node clusters) and uses the network's graph Laplacian to restrict nodal eliminations to only include neighbors of chosen super nodes. By leveraging the efficient MILP formulation, an iterative scheme is employed to successively enhance the network reduction while ensuring that each iterate is
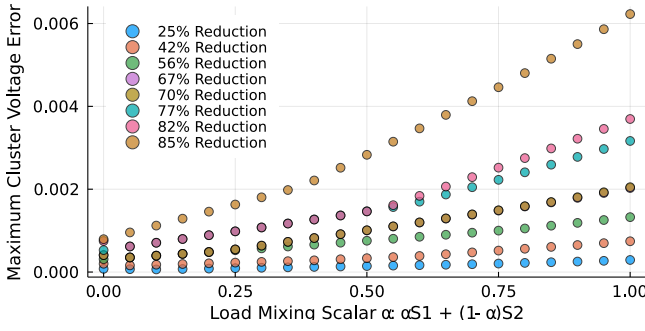
Fig. 4. Worst-case super node intra-cluster voltage errors for all iterate Kron-reduced versions of the 115-node test feeder is optimally reduced via eight iterations in Algorithm 1 and

a valid Kron reduction of the full network. Furthermore, simulation-based analysis is used to numerically explore the formulation and characterize and compare the optimal Kron reductions. The computational results illustrate that Opti-KRON can reduce full networks of more than 100 nodes by 25-90% at optimality and within seconds. These optimal network reductions engender worst-case intra-cluster voltage magnitude deviations of less than 0.01pu.

Future work will pursue a number of open questions resulting from discoveries herein. First, we will investigate the For example, while the iterative scheme is guaranteed to converge to a Kron-reduced network, we have not established global optimality guarantees at convergence. However, for radial networks, it may be possible to prove that the successive iterations will yield the globally optimal Kron reduced network [26]. Furthermore, we are interested in using the optimal Kron-reduced networks in OPF problems and want to incorporate the corresponding worst-case intra-cluster voltage deviations to yield robust OPF formulations (e.g., via tightened voltage bounds) whose solutions guarantee admissibility in the underlying full network [9]. Similarly, solving the OPF on a reduced network will require a disaggregation policy to lift the optimal solution on the reduced network (e.g., a dispatch of aggregated resources) to the full network's (individual) resources. This lifting may not be unique and can be carried out in numerous ways, which results in loss of optimality relative to solving the OPF on the full network [15]. Thus, developing disaggregation policies with optimality guarantees is of interest.

APPENDIX

While the Opti-KRON formulation is stated in complex variables in (13), it was decomposed into purely real rectangular coordinates to be solved. Decomposing the admittance and impedance matrices into their real and imaginary parts, we have

$$\langle Y_b \rangle = \begin{bmatrix} Y_G & -Y_B \\ Y_B & Y_G \end{bmatrix}, \langle Z_b \rangle = \begin{bmatrix} Z_G & -Z_B \\ Z_B & Z_G \end{bmatrix}. \quad (14)$$

To build the generalized Kron impedance, we extend the selection matrix into a block diagonal form:

$$\langle Z_K \rangle = \begin{bmatrix} S & 0 \\ 0 & S \end{bmatrix} \begin{bmatrix} Z_G & -Z_B \\ Z_B & Z_G \end{bmatrix} \begin{bmatrix} S & 0 \\ 0 & S \end{bmatrix}. \quad (15)$$

Likewise, for super node voltage selection and current aggregation, the following block diagonal expressions produce the desired quantities:

$$\begin{bmatrix} V_{s,r} \\ V_{s,i} \end{bmatrix} = \begin{bmatrix} S & 0 \\ 0 & S \end{bmatrix} \begin{bmatrix} V_r \\ V_i \end{bmatrix}, \begin{bmatrix} I_{s,r} \\ I_{s,i} \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix} \begin{bmatrix} I_r \\ I_i \end{bmatrix} \quad (16)$$

The remainder of the formulation follows likewise, but it is not included due to space limitations.

REFERENCES

[1] R. B. Squires, "Economic Dispatch of Generation Directly Rrom Power System Voltages and Admittances," en, *Transactions of the American Institute of Electrical Engineers. Part III: Power Apparatus and Systems*, vol. 79, no. 3, pp. 1235–1244, Apr. 1960.

[2] H. W. Dommel and W. F. Tinney, "Optimal Power Flow Solutions," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-87, no. 10, pp. 1866–1876, 1968.

[3] J. Carpentier, "Optimal power flows," *International Journal of Electrical Power & Energy Systems*, vol. 1, no. 1, pp. 3–15, 1979, Publisher: Elsevier.

[4] D. K. Molzahn, F. Dorfler, H. Sandberg, *et al.*, "A Survey of Distributed Optimization and Control Algorithms for Electric Power Systems," en, *IEEE Transactions on Smart Grid*, vol. 8, no. 6, pp. 2941–2962, Nov. 2017.

[5] D. K. Molzahn and I. A. Hiskens, "A Survey of Relaxations and Approximations of the Power Flow Equations," en, *Foundations and Trends® in Electric Energy Systems*, vol. 4, no. 1-2, pp. 1–221, 2019.

[6] Q. Ploussard, L. Olmos, and A. Ramos, "An Efficient Network Reduction Method for Transmission Expansion Planning Using Multicut Problem and Kron Reduction," en, *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 6120–6130, Nov. 2018.

[7] E. Dall'Anese, H. Zhu, and G. B. Giannakis, "Distributed Optimal Power Flow for Smart Microgrids," en, *IEEE Transactions on Smart Grid*, vol. 4, no. 3, pp. 1464–1475, Sep. 2013.

[8] S. R. Shukla, S. Paudyal, and M. R. Almassalkhi, "Efficient Distribution System Optimal Power Flow With Discrete Control of Load Tap Changers," en, *IEEE Transactions on Power Systems*, vol. 34, no. 4, pp. 2970–2979, Jul. 2019.

[9] M. Almassalkhi, S. Brahma, N. Nazir, *et al.*, "Hierarchical, Grid-Aware, and Economically Optimal Coordination of Distributed Energy Resources in Realistic Distribution Systems," en, *Energies*, vol. 13, no. 23, p. 6399, Dec. 2020.

[10] B. Stott, J. Jardim, and O. Alsac, "DC Power Flow Revisited," *IEEE Transactions on Power Systems*, vol. 24, no. 3, pp. 1290–1300, 2009.

[11] J. Lavaei and S. H. Low, "Zero Duality Gap in Optimal Power Flow Problem," en, *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 92–107, Feb. 2012.

[12] S. H. Low, "Convex Relaxation of Optimal Power Flow—Part I: Formulations and Equivalence," en, *IEEE Transactions on Control of Network Systems*, vol. 1, no. 1, pp. 15–27, Mar. 2014.

[13] D. Lee, K. Turitsyn, D. K. Molzahn, and L. A. Roald, "Robust ac optimal power flow with robust convex restriction," *IEEE Transactions on Power Systems*, vol. 36, no. 6, pp. 4953–4966, 2021.

[14] E. Dall'Anese and A. Simonetto, "Optimal Power Flow Pursuit," en, *IEEE Transactions on Smart Grid*, vol. 9, no. 2, pp. 942–952, Mar. 2018.

[15] D. F. Rogers, R. D. Plante, R. T. Wong, and J. R. Evans, "Aggregation and Disaggregation Techniques and Methodology in Optimization," en, *Operations Research*, vol. 39, no. 4, pp. 553–582, Aug. 1991.

[16] P. V. Kokotovic, R. E. O'Malley Jr, and P. Sannuti, "Singular perturbations and order reduction in control theory—an overview," *Automatica*, vol. 12, no. 2, pp. 123–132, 1976, Publisher: Elsevier.

[17] C. Sturk, L. Vanfretti, F. Milano, and H. Sandberg, "Structured model reduction of power systems," en, in *2012 American Control Conference (ACC)*, Montreal, QC: IEEE, Jun. 2012, pp. 2276–2282.

[18] S. Chevalier, L. Schenato, and L. Daniel, "Accelerated probabilistic power flow in electrical distribution networks via model order reduction and neumann series expansion," *IEEE Transactions on Power Systems*, pp. 1–1, 2021.

[19] Q. Ploussard, "Efficient reduction techniques for a large-scale Transmission Expansion Planning problem," en, Ph.D. dissertation, 2019.

[20] P. Fortenbacher, T. Demiray, and C. Schaffner, "Transmission Network Reduction Method using Nonlinear Optimization," en, *2018 Power Systems Computation Conference (PSCC)*, pp. 1–7, Jun. 2018, arXiv: 1711.01079.

[21] D. Shi and D. J. Tylavsky, "A Novel Bus-Aggregation-Based Structure-Preserving Power System Equivalent," en, *IEEE Transactions on Power Systems*, vol. 30, no. 4, pp. 1977–1986, Jul. 2015.

[22] H. Oh, "A New Network Reduction Methodology for Power System Planning Studies," en, *IEEE Transactions on Power Systems*, vol. 25, no. 2, pp. 677–684, May 2010.

[23] J. Sistermanns, M. Hotz, D. Hewes, R. Witzmann, and W. Utschick, "Feature- and Structure-Preserving Network Reduction for Large-Scale Transmission Grids," en, *2019 IEEE Milan PowerTech*, pp. 1–6, Jun. 2019, arXiv: 1903.11590.

[24] F. Dorfler and F. Bullo, "Kron Reduction of Graphs With Applications to Electrical Networks," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 60, no. 1, pp. 150–163, 2013.

[25] K. P. Schneider, B. A. Mather, B. C. Pal, *et al.*, "Analytic considerations and design basis for the ieee distribution test feeders," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 3181–3188, 2018.

[26] Y. Yuan, S. Low, O. Ardakanian, and C. Tomlin, "Inverse Power Flow Problem," en, *arXiv:1610.06631 [cs, eess, math]*, Oct. 2021, arXiv: 1610.06631.