Optimal Power Flow Derived Sparse Linear Solver Benchmarks

Jonathan Maack

National Renewable Energy Laboratory

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Technical Report
NREL/TP-2C00-86931
October 2023

Optimal Power Flow Derived Sparse Linear Solver Benchmarks

Jonathan Maack

National Renewable Energy Laboratory

Citation
Maack, Jonathan. 2023. *Optimal Power Flow Derived Sparse Linear Solver Benchmarks*. Golden, CO: National Renewable Energy Laboratory. NREL/TP-2C00-86931. https://www.nrel.gov/docs/fy24osti/86931.pdf.

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Technical Report
NREL/TP-2C00-86931
October 2023

National Renewable Energy Laboratory
15013 Denver West Parkway
Golden, CO 80401
303-275-3000 • www.nrel.gov

NOTICE

This work was authored by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. This research was supported by the Exascale Computing Project (17-SC-20-SC), a collaborative effort of the U.S. Department of Energy Office of Science and the National Nuclear Security Administration. A portion of this research was performed using computational resources sponsored by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy and located at the National Renewable Energy Laboratory. The views expressed herein do not necessarily represent the views of the DOE or the U.S. Government.

Executive Summary

Due to the changing nature of the power grid, it is increasingly important to be able to solve a high-fidelity optimal power-flow models on large power networks. This high-fidelity problem, called AC Optimal Power Flow (ACOPF), is a nonlinear, nonconvex optimization problem. One of the few reliable ways of solving such a problem is interior point methods. These methods result in sparse linear systems where the coefficient matrix is symmetric, indefinite and often ill-conditioned. As such, they are particularly challenging for sparse linear solvers and represent a considerable computational bottleneck in solving the ACOPF problem. In this paper, we introduce a repository of linear systems captured from ACOPF problems when solved by the open-source optimizer IPOPT. These matrices are meant to be used as a test suite for sparse linear solver development.

Acknowledgments

This work was authored by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. This research was supported by the Exascale Computing Project (17-SC-20-SC), a collaborative effort of the U.S. Department of Energy Office of Science and the National Nuclear Security Administration. A portion of this research was performed using computational resources sponsored by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy and located at the National Renewable Energy Laboratory. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

List of Acronyms

ACOPF	AC Optimal Power Flow
CPU	Central Processing Unit
DCOPF	DC Optimal Power Flow
GPU	Graphics Processing Unit
KKT	Karush-Kuhn-Tucker
TAMU	Texas A&M University

Table of Contents

Executive Summary . . . iv
Acknowledgments . . . v
List of Acronyms . . . vi
1 Introduction . . . 1
2 AC Optimal Power Flow . . . 2
3 Implementation Details . . . 4
4 Properties of the Linear Systems . . . 6
4.1 Sparsity . . . 6
4.2 Condition Numbers . . . 7
4.3 Eigenvalues and Singular Values . . . 8
5 Conclusions . . . 10
References . . . 11

List of Figures

Figure 1. Sparsity of matrices in stack for the given network. Sparsity includes explicit zeros and is constant across the optimization run. Note that as the system size increases, the matrices become more sparse. 6

Figure 2. Sparsity patterns of coefficient matrices. The sparsity pattern is constant across an optimization run (when explicit zeros are included). The nonzero diagonal in the lower-right is the result of a regularization added by IPOPT. . . . 7

Figure 3. Condition numbers of matrices from stack of IPOPT matrices. Matrix ID identifies the particular matrix in the stack. The optimization approaches a local minimum the further to the right in the plots. . . 8

Figure 4. Eigenvalues of first (left) and last (right) matrices in the RTS-GMLC stack. . . . 8

Figure 5. Eigenvalues of first (left) and last (right) matrices in the ACTIVSg200 stack. . . . 9

Figure 6. Singular values of first (left) and last (right) matrices in the RTS-GMLC stack. . . . 9

Figure 7. Singular values of first (left) and last (right) matrices in the RTS-GMLC stack. . . . 9

List of Tables

Table 1. Summary of Synthetic Grids . . . 4

Table 2. Characteristics of the Matrices. Nonzeros includes explicit zeros. Sparsity is measured as fraction of nonzero entries. The number of nonzeros and sparsity are constant across a case. . . . 6

1 Introduction

At present power grid planning and operations often rely on a linear programming formulation called DC Optimal Power Flow (DCOPF). As the power grid evolves through incorporating renewable generation, storage, and other new technologies, the reliance on the DC linear power flow approximation will likely no longer be possible as the required assumptions will be violated. This means that solving the full AC Optimal Power Flow (ACOPF) problem will be necessary.

ACOPF (as described in the next section) is a nonlinear, nonconvex programming problem. Currently, the most commonly used solver for nonlinear, nonconvex programs is IPOPT (Wächter and Biegler 2006). IPOPT uses an interior point method and a filtered-line search to find local extrema. The vast majority of time IPOPT spends solving a problem is in the solution of a linear system of equations arising from Newton's method applied to the (modified) Karush-Kuhn-Tucker conditions. In the case of the ACOPF problem, the matrices in the resulting linear systems are extremely sparse, symmetric, ill-conditioned and indefinite as will be discussed. These are examples of saddle point problems which are known to be difficult to solve (Benzi, Golub, and Liesen 2005). As discussed in Tasseff et al. 2019, the accepted method of solving these systems is to use sparse, direct solvers such as HSL's MA57 (Duff 2004). However, these sparse, direct solvers transfer poorly to GPUs due to the data movement required by pivoting. Furthermore, current GPU capable solvers show little or no speed up when running on accelerators compared to their CPU version for these matrices (Swirydowicz et al. 2021). As a result, solving these linear systems is a significant computational bottle neck and prevents the solution of large scale ACOPF problems needed in the power systems community.

In this report, we introduce a suite of linear systems captured from IPOPT optimization runs of ACOPF problems applied to the synthetic power networks developed by Texas A&M University (TAMU) (Birchfield et al. 2017; Xu et al. 2017; Li et al. 2018) and Grid Modernization Lab Consortium (Barrows et al. 2020) and discuss their properties. These linear systems are available to the public at https://github.com/NREL/opf_matrices for use as a benchmark for linear solvers.

The remainder of this report is organized as follows: In Section 2, we describe the general formulation of the ACOPF problem from which the linear systems originated. In Section 3, we describe the synthetic networks for which our ACOPF formulation was implemented. In Section 4, we discuss the sparsity and conditioning of the matrices in our benchmark suite. Section 5 gives our concluding remarks.

2 AC Optimal Power Flow

The ACOPF formulation that we used in generating the linear systems is given by

$$\min_{p_g,q_g,p_{sp},y_{ll},y_{ol},z_{ll},z_{ol}} c(p_g,q_g,p_{sp},y_{ll},y_{ol},z_{ll},z_{ol}) \tag{2.1}$$

subject to

$$P_g^{min} \le p_g \le P_g^{max} \qquad \forall g \in G \tag{2.2}$$
$$Q_g^{min} \le q_g \le Q_g^{max} \qquad \forall g \in G \tag{2.3}$$
$$P_w^{min} \le p_w \le \xi_w \le P_w^{max} \qquad \forall w \in W \tag{2.4}$$
$$p_{sp}^w = \xi_w - p_w \qquad \forall w \in W \tag{2.5}$$
$$(r_e)^2 + (s_e)^2 \le (F_e^{max})^2 \qquad \forall e \in E \tag{2.6}$$
$$V_b^{min} \le v_b \le V_b^{max} \qquad \forall b \in B \tag{2.7}$$
$$-\pi \le \theta_b \le \pi \qquad \forall b \in B \tag{2.8}$$
$$r_e = \alpha_e v_i^2 - \alpha_e v_i v_k \cos(\theta_i - \theta_k) - \beta_e v_i v_k \sin(\theta_i - \theta_k) \qquad \forall e = (i,k) \in E \tag{2.9}$$
$$s_e = -\beta_e v_i^2 + \beta_e v_i v_k \cos(\theta_i - \theta_k) - \alpha_e v_i v_k \sin(\theta_i - \theta_k) \qquad \forall e = (i,k) \in E \tag{2.10}$$
$$r_b = \mu_b v_b^2 \qquad \forall b \in B \tag{2.11}$$
$$s_b = -\nu_b v_b^2 \qquad \forall b \in B \tag{2.12}$$
$$\sum_{g \in G_b} p_g + \sum_{e \in E_b^{in}} r_e + y_{ll}^b = P_b^d + \sum_{e \in E_b^{out}} r_e + r_b + y_{ol}^b \qquad \forall b \in B \tag{2.13}$$
$$\sum_{g \in G_b} q_g + \sum_{e \in E_b^{in}} s_e + z_{ll}^b = Q_b^d + \sum_{e \in E_b^{out}} s_e + s_b + z_{ol}^b \qquad \forall b \in B \tag{2.14}$$
$$y_{ll}^b \ge 0 \qquad \forall b \in B \tag{2.15}$$
$$y_{ol}^b \ge 0 \qquad \forall b \in B \tag{2.16}$$
$$z_{ll}^b \ge 0 \qquad \forall b \in B \tag{2.17}$$
$$z_{ol}^b \ge 0 \qquad \forall b \in B \tag{2.18}$$

Here G is the set of thermal generators and W is the set of renewable generators. Note we consider any wind or solar generator to be part of W. All remaining generators are in G (that is, they are considered to be "thermal"). The set B is the set of buses, and E is the set of directed transmission lines. That is, $e = (i,k)$ is the line connecting bus i to bus k and $(i,k) \neq (k,i)$. Also, $G_b \subset G$ is the set of all generators located at bus b, $E_b^{in} \subset E$ are all lines directed toward bus b and $E_b^{out} \subset E$ are all lines directed away from bus b. The variables p_g and q_g represent the real and reactive set points for thermal generator g, respectively, while p_w is the real power set point for renewable generator w. The upper limit ξ_w is a random value constrained to lie between the renewable generator limits P_w^{min} and P_w^{max}. It is a chosen randomly in order to represent the variability of renewable generators. The value p_{sp}^w gives the amount of renewable generation that is not used or "spilled" (hence sp). The variables r_e and s_e are the real and reactive power flows on a transmission line $e = (i,k)$ from bus i to bus k measured at bus i. The voltage magnitude and angle at bus i are given by v_i and θ_i. The variables $y_{ll}, y_{ol}, z_{ll}, z_{ol}$ are slack variables and are present to ensure feasibility of the model. They roughly correspond to lost load (hence ll) or over load (hence ol) but are not strictly physical.

The objective function (2.1) represents the costs of operating the power grid. It includes terms such as generator fuel cost. For the grids used in this report, the objective function is a sum of terms for each generator and bus. Each thermal generator term is linear or quadratic in the real power set points p_g and has no dependence on the reactive power set points q_g. The renewable generator terms are linear in the renewable generation spilled p_{sp}^w. Every bus term is linear in the slack variables y_{ll}, y_{ol}, z_{ll}, and z_{ol}. It follows that the objective function takes the form

$$c(p_g,p_{sp},y_{ll},y_{ol},z_{ll},z_{ol}) = \sum_{g \in G} c_1^g p_g + c_2^g p_g^2 + \sum_{w \in W} c_{sp} p_{sp}^w + \sum_{b \in B} c_{ll}(y_{ll}^b + z_{ll}^b) + c_{ol}(y_{ol}^b + z_{ol}^b) \tag{2.19}$$

The constraints (mostly) represent the engineering limits of the grid elements or the power flow physics: Equations (2.2) and (2.3) give the real and reactive generator set point limits, respectively while (2.6) gives the transmission line flow limits, (2.7) gives the bus voltage magnitude limits and (2.8) gives the bus voltage angle limits. Equations (2.9) and (2.10) give the physical relation between real and reactive line flow and voltage magnitudes and angles. The quantities α_e and β_e are physical parameters of the transmission line or transformer. Note that for a transmission line, $\alpha_e = \alpha_f$ and $\beta_e = \beta_f$ for $e = (i,k)$ and $f = (k,i)$. This is generally not true for a transformer. Equations (2.11) and (2.12) describe real and reactive components of shunt connected elements where μ_b and ν_b are the physical parameters of the element. Constraints (2.13) and (2.14) are the real and reactive power balance equations. The parameter values P_b^d and Q_b^d represent real and reactive power demand at bus b.

For more details on ACOPF problems, other formulations as well as relaxations, see the comprehensive survey (Molzahn, Hiskens, et al. 2019). For a more fundamental exploration of power flow, see (Vaahedi 2014).

3 Implementation Details

In order to get a variety of sizes of linear systems, we implemented the ACOPF problem (2.1)-(2.18) for several different synthetic grids. Their names, geographic overlays, and sizes are summarized in Table 1. For more details on these synthetic grids and how they were created see (Barrows et al. 2020; Birchfield et al. 2017; Xu et al. 2017; Li et al. 2018). The grid data sets themselves can be downloaded at https://github.com/GridMod/RTS-GMLC or at https://electricgrids.engr.tamu.edu/.

The grid data sets include most of the required parameters to create the given ACOPF problem. All cases include generator cost information, non-renewable generator limits, line power flow limits, line parameters and real power demand. While each case has maximum capacities for the renewable generators, the upper limits in (2.4) are, generally speaking, time dependent and based on weather conditions at the generation site. This is captured by the value ξ_w. The RTS data set provides these time series data while the TAMU test systems do not. To fill in these values for the TAMU cases, we used random values distributed uniformly from zero to the maximum capacity of the renewable generator, that is we take $\xi_w \sim U(0, P_w^{max})$. The TAMU cases included reactive power demands and these were used as given. However, the RTS case does not. To replace this missing data, we set $Q_b^d = 0.1 P_b^d$, that is, we took the reactive power demand at a bus to be one-tenth the value of the real power demand.

We wanted ACOPF problems that not only converged to feasible solutions but also did so in a reasonable amount of time. With this goal in mind, we also altered the data in the larger cases. In particular, for the ACTIVSg2000 any line limit below 400 MW was raised to 400 MW. For ACTIVSg10k, ACTIVSg25k and ACTIVSg70k line limits below 1200 MW were raised to 1200 MW. Additionally, the voltage magnitude limits are normally set at $V_b^{min} = 0.9$ per unit (p.u.) and $V_b^{max} = 1.1$ p.u. We relaxed these to 0.8 p.u. and 1.2 p.u., respectively. (For an explanation of the per unit system see Vaahedi 2014.) All lower generator limits were reset to zero so that $P_g^{min} = P_w^{min} = 0$. While these alterations have numerical impacts on the linear systems, they do not materially affect the qualitative nature of the matrices.

Slack variable costs are not included in either data set. In all cases, we used \$10,000 per MWh for loss of load (both real and reactive) and \$100 per MWh for over load (again, both real and reactive). We also added a slight cost for unused available renewable generation. This was equal to \$1.20 per MWh. We take the cost function (2.19) to be the cost of operating the grid for 5 minutes (that is, the units are \$/5 minutes) so that $c_{sp} =$ \$0.1 per 5 minutes, $c_{ll} =$ \$833.33 per 5 minutes, and $c_{ol} =$ \$8.33 per 5 minutes. The generator cost functions are provided in the data sets and the units are converted accordingly.

The ACOPF problems were constructed using JuMP.jl (see Dunning, Huchette, and Lubin 2017). JuMP is a mathematical optimization language written in Julia that interfaces easily with IPOPT (Wächter and Biegler 2006), which we used to solve each ACOPF problem. IPOPT includes many parameters to control the algorithm. See the IPOPT manual available at https://coin-or.github.io/Ipopt/ for a description of these parameters as well as their default values. We set the control parameters as follows: The tolerance was set to 10^{-6} and the acceptable tolerance to 10^{-5}. We used HSL's MA57 (Duff 2004) as the internal linear solver provided as part of the COIN HSL library (*HSL. A collection of Fortran codes for large scale scientific computation*). All other control parameters were left at their default values.

In order to print the matrices, we intercepted them immediately before they were given to the linear solver. This was accomplised by inserting code in IPOPT's MA57 interface to print out the matrix in ASCII before the factorization

Table 1. Summary of Synthetic Grids

Grid Name	Geographic Area	# Buses	# Generators	# Branches
RTS-GLMC	Southwest US	73	102	120
ACTIVSg200	Central Illinois	200	49	245
ACTIVSg2000	Texas	2000	544	3206
ACTIVSg10k	Western US	10000	2485	12706
ACTIVSg25k	Northeastern US	25000	4834	32230
ACTIVSg70k	Eastern US	70000	10390	88207

call was made. Right hand side vectors were like-wise captured by printing in ASCII right before solve calls. Each floating point entry was printed in scientific notation with 16 decimal places.

For the RTS-GMLC, ACTIVSg200, and ACITVSg2000 test systems, every coefficient matrix and every right hand side vector were printed from every iteration of IPOPT. Due to size of the matrices and the number of iterations necessary to solve the ACOPF problem, only every 10 coefficient matrix starting with the first was saved for the ACTIVSg10k, ACTIVSg25k, and ACTIVSg70k test systems as well as the last matrix. All corresponding right hand side vectors for these matrices were saved.

4 Properties of the Linear Systems

In this section, we explore the properties of the coefficient matrices in this repository. We start by looking at the fraction of nonzeros along with the sparsity patterns. Then we examine the conditioning of the coefficients matrices. The section closes with an examination of the eigenvalues and singular values. This is restricted to the smaller systems for computational reasons.

4.1 Sparsity

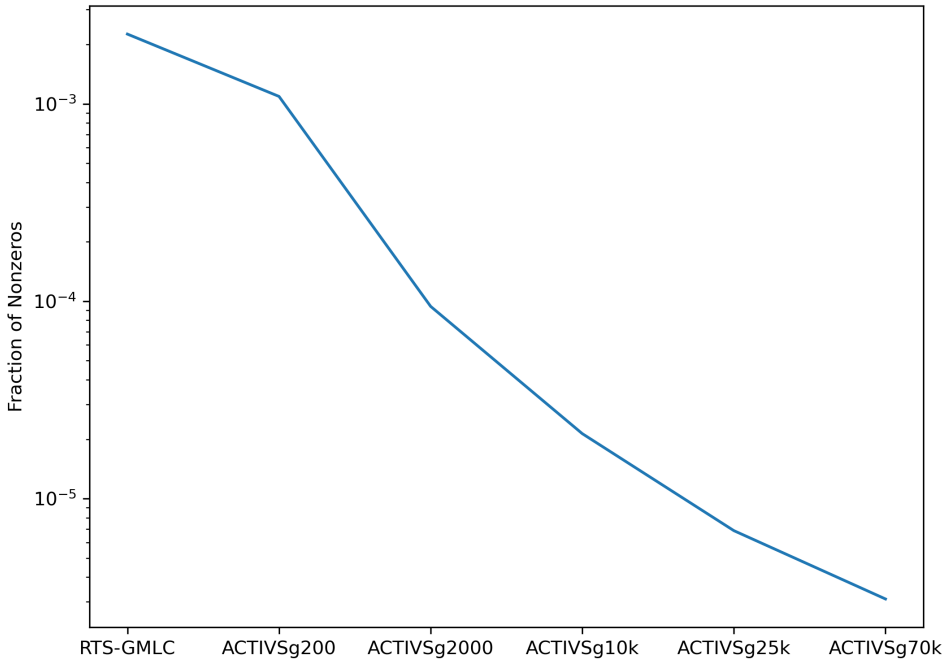


Figure 1. Sparsity of matrices in stack for the given network. Sparsity includes explicit zeros and is constant across the optimization run. Note that as the system size increases, the matrices become more sparse.

As mentioned in the introduction and seen in Table 2 and Figure 1, the coefficient matrices are sparse and grow more sparse as the underlying power network increases in size. This is hardly surprising as the number of nonzero terms in the Hessian of the objective function and the number of constraints in the ACOPF problem grows linearly in the number of buses, lines or generators on the system, whereas the number of entries in the coefficient matrix grows quadratically in these numbers. The sparsity level for a fixed test system is fixed across an optimization run.

Table 2. Characteristics of the Matrices. Nonzeros includes explicit zeros. Sparsity is measured as fraction of nonzero entries. The number of nonzeros and sparsity are constant across a case.

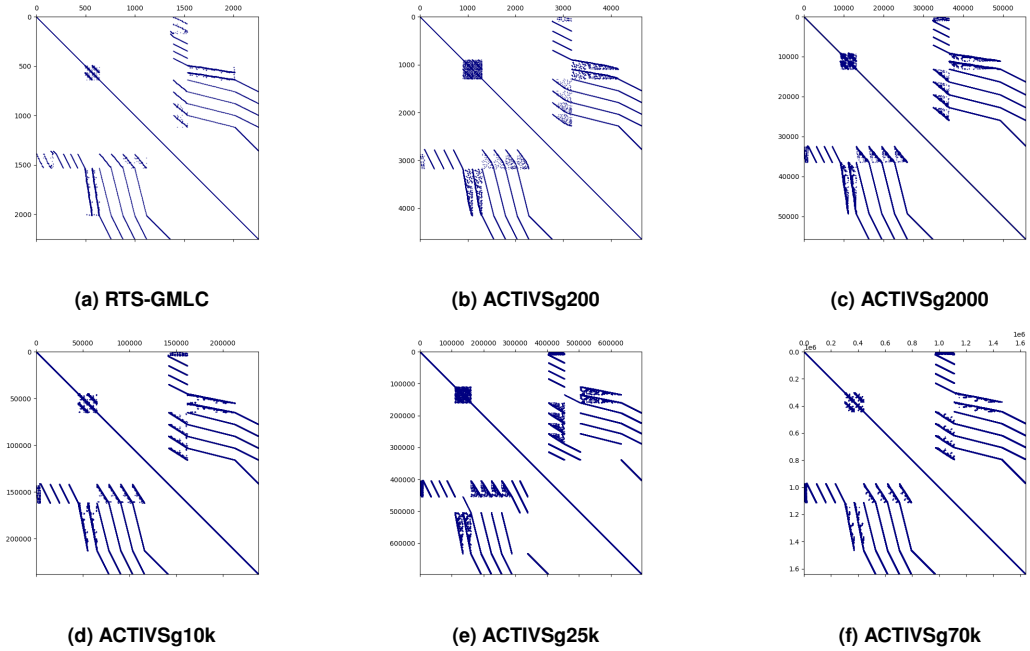
Grid Name	Dimension	Nonzeros	Sparsity
RTS-GLMC	2253	6852	2.256e-03
ACTIVSg200	4644	14078	1.090e-03
ACTIVSg2000	55667	173710	9.415e-05
ACTIVSg10k	238072	723720	2.134e-05
ACTIVSg25k	697161	2020647	6.880e-06
ACTIVSg70k	1640411	4991820	3.100e-06

The sparsity pattern of the coefficient matrices is given in Figure 2. Just like the sparsity level, this pattern is consistent across the optimization run. The nonzero diagonal in the lower-right corner is the result of a regularization applied by IPOPT to the coefficient matrix (Wächter and Biegler 2006).

The classic saddle-point matrix structure (see Benzi, Golub, and Liesen 2005) is clearly visible. This general form is

$$\begin{pmatrix} H & J^T \\ J & 0 \end{pmatrix} \tag{4.1}$$

where H is the Hessian of the Lagragian and J is the Jacobian of both the equality and inequality constraints (see Benzi, Golub, and Liesen 2005 or Wächter and Biegler 2006 for more details). While there appears to be more structure to the matrices, this is the result of adding variables and constraints of the same type in the same order for each grid. Furthermore, there is a part of the constraint Jacobian whose sparsity patterns directly reflects the structure of the underlying power grid. In our matrices, this is the first set of rows in the Jacobian. (For the RTS-GMLC case in Figure 2, these occur around row 1500. A similar block is visible in the other cases as well.) This block of the Jacobian arises from the power balance constraints given by (2.13) and (2.14). These equations involve all connected branches as well as generator and load at the bus. Therefore, this block directly reflects the underlying graph structure of the power network. As such, there is no general structure within the matrix that can be exploited beyond that given by (4.1).



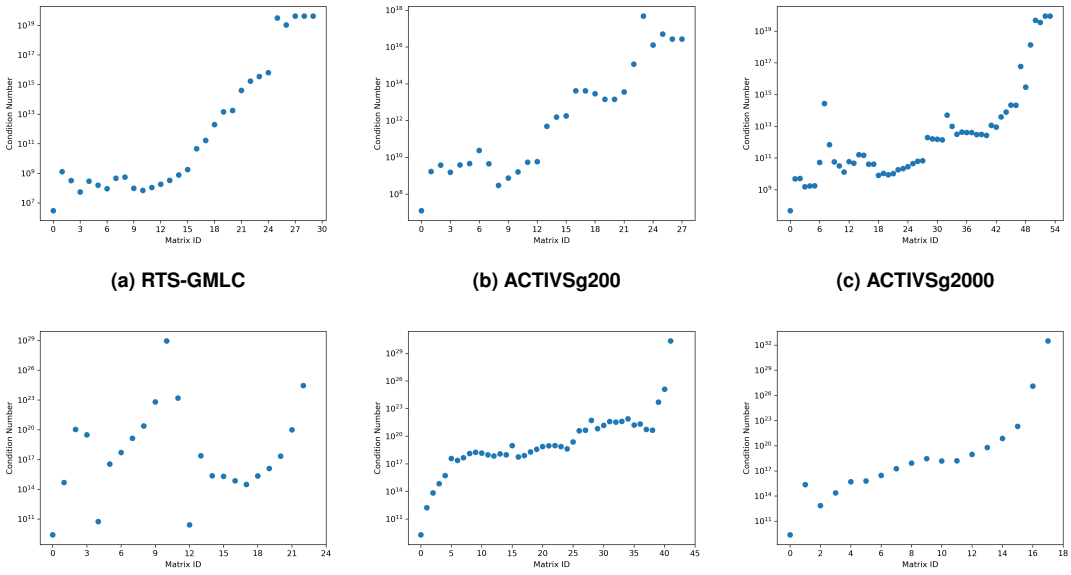
(a) RTS-GMLC **(b) ACTIVSg200** **(c) ACTIVSg2000**

(d) ACTIVSg10k **(e) ACTIVSg25k** **(f) ACTIVSg70k**

Figure 2. Sparsity patterns of coefficient matrices. The sparsity pattern is constant across an optimization run (when explicit zeros are included). The nonzero diagonal in the lower-right is the result of a regularization added by IPOPT.

4.2 Condition Numbers

We see several trends in the condition numbers of the coefficient matrices in Figure 3. First, note that these matrices are all ill-conditioned. The smallest condition number is a little less than 10^7. This occurs for the smallest system, the RTS-GMLC where the matrix is of dimension 2253. Second, the condition numbers grow as the network size grows. Again, this is not surprising since condition numbers tend to grow with the size of the matrix. Practically, this means that larger power networks make for more difficult ACOPF problems. Finally, the conditioning of the matrices worsens as the optimization proceeds with the largest condition numbers occurring near termination for all but one case. In general, we expect to lose a digit of accuracy in the solution for every order of magnitude in the condition number (Trefethen and Bau 2022). This means that we can expect linear solvers to struggle to provide precise solutions just as IPOPT requires greater precision in order to locate the local optimum.



(a) RTS-GMLC (b) ACTIVSg200 (c) ACTIVSg2000

(d) ACTIVSg10k (e) ACTIVSg25k (f) ACTIVSg70k

Figure 3. Condition numbers of matrices from stack of IPOPT matrices. Matrix ID identifies the particular matrix in the stack. The optimization approaches a local minimum the further to the right in the plots.

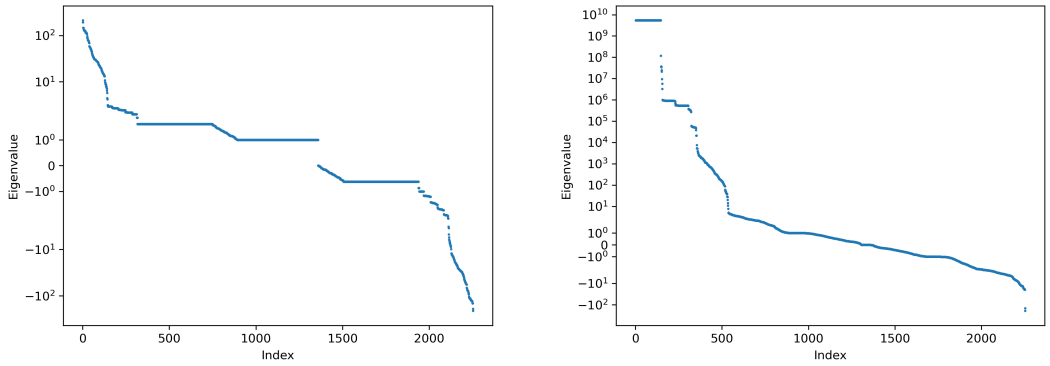


Figure 4. Eigenvalues of first (left) and last (right) matrices in the RTS-GMLC stack.

4.3 Eigenvalues and Singular Values

The eigenvalues of the first and last matrices of the RTS-GMLC and ACTIVSg200 cases are shown in Figures 4 and 5. We restrict ourselves to the two smallest test systems for computational considerations. We first note that the matrices are indefinite with both positive and negative eigenvalues. The growth in magnitude of the eigenvalues is also notable although perhaps expected given the growth in the condition number. The RTS-GMLC case jumps from 10^2 to 10^{10} while the ACTIVSg200 case jumps from 10^3 to 10^9. We also note that there are noticeable gaps in the spectrum. That is, there are places where the eigenvalues have a sudden decrease in magnitude.

To elaborate on the conditioning and explore the spectrum gaps further, we also plotted the singular values of the first and last matrices of the RTS-GMLC and ACTIVSg200 cases in Figures 6 and 7. In these plots, we see that there is also a significant decrease in the magnitude of the smallest singular value. The RTS-GMLC case drops from 10^{-4} to (roughly) 10^{-10} and the ACTIVSg200 case drops from 10^{-4} to (roughly) 10^{-8}. From this we see that the growth in the condition number results from both jumps in the largest singular values and drops in the smallest singular values. Furthermore, we see that the spectral gaps remain only at the tails of the spectrum. That is, either around the largest or smallest singular values (e.g., the right plot in Figure 6. Those gaps appearing in the middle have largely disappeared (e.g., compare the left plots of Figures 5 and 7).

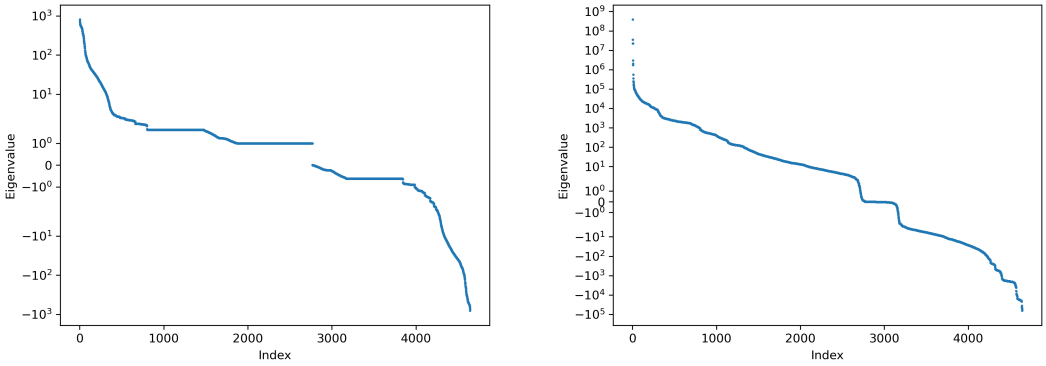


Figure 5. Eigenvalues of first (left) and last (right) matrices in the ACTIVSg200 stack.

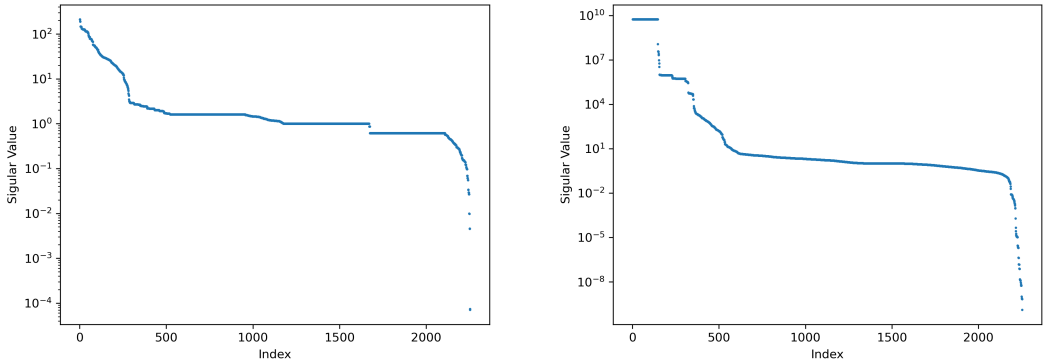


Figure 6. Singular values of first (left) and last (right) matrices in the RTS-GMLC stack.

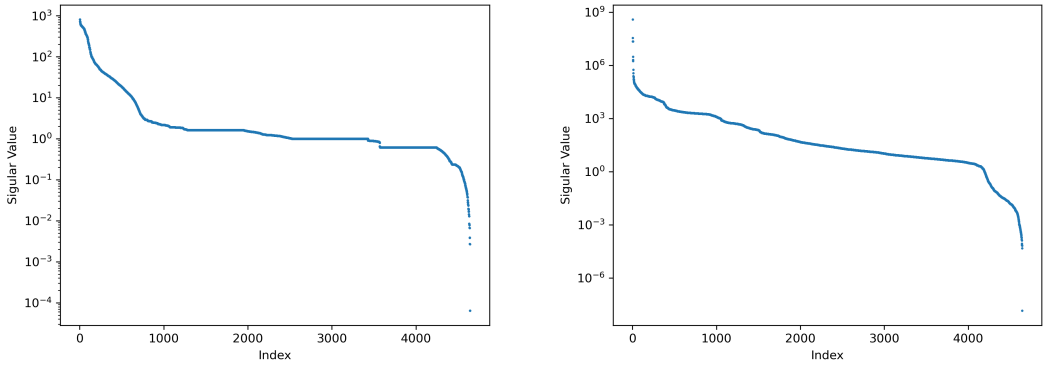


Figure 7. Singular values of first (left) and last (right) matrices in the RTS-GMLC stack.

5 Conclusions

Applying interior point optimization algorithms to ACOPF problems gives rise to sparse, symmetric, indefinite and particularly ill-conditioned linear systems. Solving these systems accurately and efficiently is challenging for present linear solvers and represents a significant computational bottleneck in the solution of ACOPF problems. Improving the linear solver reduces this bottleneck in two ways. First, the vast majority of time in an interior point method is spent solving saddle-point linear systems. Even slightly more efficient solutions (either by parallel computing or more efficient algorithms) can save significant time over the course of the whole optimization. Second, the accuracy of the solve impacts the number of iterations required for the solver to converge. Inaccurate solves require more iterations. This may be a particular problem around the optimal point where the condition number of the coefficient matrix is very large. Providing more accurate solutions will reduce the number of optimizer iterations and reduce run time.

In this paper, we have introduced a new set of test matrices to serve as a benchmark for linear solver development as well as discussing some of the properties of the matrices. We hope this benchmark suite will enable sparse linear solver development that eliminates or reduces the present computational bottleneck while giving accurate solutions.

References

Barrows, C., A. Bloom, A. Ehlen, J. Ikäheimo, J. Jorgenson, D. Krishnamurthy, J. Lau, et al. 2020. "The IEEE Reliability Test System: A Proposed 2019 Update." *IEEE Transactions on Power Systems* 35 (1): 119–127. https://doi.org/10.1109/TPWRS.2019.2925557.

Benzi, M., G. H. Golub, and J. Liesen. 2005. "Numerical solution of saddle point problems." *Acta numerica* 14:1–137.

Birchfield, A. B., T. Xu, K. M. Gegner, K. S. Shetye, and T. J. Overbye. 2017. "Grid Structural Characteristics as Validation Criteria for Synthetic Networks." *IEEE Transactions on Power Systems* 32 (4): 3258–3265. https://doi.org/10.1109/TPWRS.2016.2616385.

Coffrin, C., R. Bent, K. Sundar, Y. Ng, and M. Lubin. 2018. "PowerModels.jl: An Open-Source Framework for Exploring Power Flow Formulations." In *2018 Power Systems Computation Conference (PSCC),* 1–8. June. https://doi.org/10.23919/PSCC.2018.8442948.

Duff, I. S. 2004. "MA57—a code for the solution of sparse symmetric definite and indefinite systems." *ACM Transactions on Mathematical Software (TOMS)* 30 (2): 118–144.

Dunning, I., J. Huchette, and M. Lubin. 2017. "JuMP: A Modeling Language for Mathematical Optimization." *SIAM Review* 59 (2): 295–320. https://doi.org/10.1137/15M1020575.

HSL. A collection of Fortran codes for large scale scientific computation. Accessed on April 1, 2020 using https://www.hsl.rl.ac.uk/.

Li, H., A. L. Bornsheuer, T. Xu, A. B. Birchfield, and T. J. Overbye. 2018. "Load modeling in synthetic electric grids." In *2018 IEEE Texas Power and Energy Conference (TPEC),* 1–6. https://doi.org/10.1109/TPEC.2018.8312059.

Molzahn, D. K., I. A. Hiskens, et al. 2019. "A survey of relaxations and approximations of the power flow equations."

Swirydowicz, K., E. Darve, J. Maack, S. Regev, M. A. Saunders, S. J. Thomas, and S. Peles. 2021. "Linear solvers for power grid optimization problems: a review of GPU-accelerated linear solvers." Under Review, *Parallel Computing.*

Tasseff, B., C. Coffrin, A. Wächter, and C. Laird. 2019. "Exploring Benefits of Linear Solver Parallelism on Modern Nonlinear Optimization Applications." *arXiv preprint arXiv:1909.08104.*

Trefethen, L. N., and D. Bau. 2022. *Numerical linear algebra.* Vol. 181. Siam.

Vaahedi, E. 2014. *Practical power system operation.* John Wiley & Sons.

Wächter, A., and L. T. Biegler. 2006. "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming." *Mathematical programming* 106 (1): 25–57.

Xu, T., A. B. Birchfield, K. M. Gegner, K. S. Shetye, and T. J. Overbye. 2017. "Application of large-scale synthetic power system models for energy economic studies." In *Proceedings of the 50th Hawaii International Conference on System Sciences.*

Xu, Y., N. Myhrvold, D. Sivam, K. Mueller, D. J. Olsen, B. Xia, D. Livengood, V. Hunt, B. R. d'Orfeuil, D. Muldrew, et al. 2020. "US Test System with High Spatial and Temporal Resolution for Renewable Integration Studies." In *2020 IEEE Power & Energy Society General Meeting (PESGM),* 1–5. IEEE.