Correction: A Parallel Primal-Dual Interior-Point Method for DC Optimal Power Flow

Ariana Minot and Na Li
School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA
Contact: minot@fas.harvard.edu, nali@seas.harvard.edu

In our PSCC paper, the formulation we used is not equivalent to a DC-OPF problem but rather a network-flow problem. We thank our committee chair Professor Bakirtzis and Professor Chatzivasileiadis for bringing this to our attention. In this note, we update our formulation to a DC-OPF problem and demonstrate the original ideas from our PSCC still hold.

1 DC-OPF Formulation

In order for our method to work well in a distributed setting (*i.e.*, limited communication requirements), it is important that the system for solving the Newton step has a sparse structure. At the same time, the system matrix should be reasonably conditioned so that the iterative solver is well-behaved. For these two reasons, we utilize the following formulation of DC-OPF. Consider a power network with n buses in set $\mathcal{N}$ and m transmission lines in set $\mathcal{E}$. Let $\boldsymbol{\theta} \in \mathbb{R}^{n-1}$ be the voltage phase angle at each bus except bus 1 (the reference bus). Let b_l be the susceptance of line $l(i,j)$ from bus i to bus j, and $\boldsymbol{b}$ the vector of all line susceptances.

$$\min_{\boldsymbol{\theta}} \sum_i f_i(p_i) := f(\boldsymbol{B\theta}) \tag{1}$$

$$\text{s.t. } \underline{\boldsymbol{p}} \le \boldsymbol{B\theta} \le \bar{\boldsymbol{p}} \tag{2}$$

$$\underline{\boldsymbol{f}} \le \boldsymbol{DA\theta} \le \bar{\boldsymbol{f}}, \tag{3}$$

where the bus-susceptance $\bar{B}$ matrix is

$$[\bar{B}]_{ij} = \begin{cases} -b_l, & \text{if line } l \text{ connects buses } i \text{ and } j \\ 0, & \text{if buses } i \text{ and } j \text{ are not directly connected} \\ -\sum_{j\neq i}[\bar{B}]_{ij}, & \text{if } i = j. \end{cases} \tag{4}$$

The matrix $\boldsymbol{B}$ is obtained by removing the first column from $\bar{B}$. This results from setting the reference angle at bus 1 to 0. The line-bus incidence matrix $\bar{A}$ to be of dimension $m \times n$ with $[\bar{A}]_{li} = +1$ and $[\bar{A}]_{lj} = -1$ if the start bus of line l is i and the end bus of line l is j. The matrix $\boldsymbol{A}$ is $\bar{A}$ with the first column removed. The matrix $\boldsymbol{D}$ is an $(m \times m)-$matrix whose lth entry is b_l.

In our simulations, we use a quadratic cost function of the form

$$f(\boldsymbol{B\theta}) = (\boldsymbol{B\theta} - \boldsymbol{p}^*)^T \boldsymbol{W}(\boldsymbol{B\theta} - \boldsymbol{p}^*), \tag{5}$$

where $\boldsymbol{p}^*$ is a vector whose entries are the nominal power demand for demand buses and zero for generation buses. The weighting matrix $\boldsymbol{W}$ is diagonal and allows to parametrize the different generator costs and demand utility functions.

2 Proposed Method and Iterative Linear Solver Convergence

We verify that our proposed method for calculating the Newton step holds under the DC-OPF formulation in (1)-(3). The Newton step requires solving

$$
\begin{bmatrix}
2\boldsymbol{B}^T\boldsymbol{W}\boldsymbol{B} & \boldsymbol{B}^T & -\boldsymbol{B}^T & (\boldsymbol{D}\boldsymbol{A})^T & -(\boldsymbol{D}\boldsymbol{A})^T \\
-\boldsymbol{\lambda}_{\bar{\boldsymbol{p}}}\boldsymbol{B} & \mathrm{diag}(\boldsymbol{B}\boldsymbol{\theta}-\bar{\boldsymbol{p}}) & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\
\boldsymbol{\lambda}_{\underline{p}}\boldsymbol{B} & \boldsymbol{0} & \mathrm{diag}(\underline{\boldsymbol{p}}-\boldsymbol{B}\boldsymbol{\theta}) & \boldsymbol{0} & \boldsymbol{0} \\
-\boldsymbol{\lambda}_{\bar{f}}\boldsymbol{D}\boldsymbol{A} & \boldsymbol{0} & \boldsymbol{0} & \mathrm{diag}(\boldsymbol{D}\boldsymbol{A}\boldsymbol{\theta}-\bar{\boldsymbol{f}}) & \boldsymbol{0} \\
\boldsymbol{\lambda}_{\underline{f}}\boldsymbol{D}\boldsymbol{A} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \mathrm{diag}(\underline{\boldsymbol{f}}-\boldsymbol{D}\boldsymbol{A}\boldsymbol{\theta})
\end{bmatrix}
\begin{bmatrix}
\Delta\boldsymbol{\theta} \\ \Delta\boldsymbol{\lambda}_{\bar{p}} \\ \Delta\boldsymbol{\lambda}_{\underline{p}} \\ \Delta\boldsymbol{\lambda}_{\bar{f}} \\ \Delta\boldsymbol{\lambda}_{\underline{f}}
\end{bmatrix}
= -
\begin{bmatrix}
\boldsymbol{r}_{dual} \\ \boldsymbol{r}_{cent,\bar{p}} \\ \boldsymbol{r}_{cent,\underline{p}} \\ \boldsymbol{r}_{cent,\bar{f}} \\ \boldsymbol{r}_{cent,\underline{f}}
\end{bmatrix},
\tag{6}
$$

(7)

where $\boldsymbol{\lambda}_{\bar{p}}$ are the dual variables associated with the upper bounds on the power injections, $\boldsymbol{B}\boldsymbol{\theta} - \bar{\boldsymbol{p}} \leq \boldsymbol{0}$. Similar definition applies to the other dual variables. The right-hand side quantities are

$$
\boldsymbol{r}_{dual} = 2\boldsymbol{B}^T\boldsymbol{W}(\boldsymbol{B}\boldsymbol{\theta} - \boldsymbol{p}^*) + \boldsymbol{B}^T(\boldsymbol{\lambda}_{\bar{p}} - \boldsymbol{\lambda}_{\underline{p}}) + (\boldsymbol{D}\boldsymbol{A})^T(\boldsymbol{\lambda}_{\bar{f}} - \boldsymbol{\lambda}_{\underline{f}}) \tag{8}
$$

$$
\boldsymbol{r}_{cent} \overset{\mathrm{def}}{=}
\begin{bmatrix}
\boldsymbol{r}_{cent,\bar{p}} \\ \boldsymbol{r}_{cent,\underline{p}} \\ \boldsymbol{r}_{cent,\underline{p}} \\ \boldsymbol{r}_{cent,\underline{f}}
\end{bmatrix}
=
\begin{bmatrix}
\boldsymbol{\lambda}_{\bar{p}}(\boldsymbol{B}\boldsymbol{\theta} - \bar{\boldsymbol{p}}) - \frac{1}{t}\boldsymbol{1} \\
\boldsymbol{\lambda}_{\underline{p}}(\underline{\boldsymbol{p}} - \boldsymbol{B}\boldsymbol{\theta}) - \frac{1}{t}\boldsymbol{1} \\
\boldsymbol{\lambda}_{\underline{p}}(\underline{\boldsymbol{p}} - \boldsymbol{B}\boldsymbol{\theta}) - \frac{1}{t}\boldsymbol{1} \\
\boldsymbol{\lambda}_{\bar{p}}(\boldsymbol{D}\boldsymbol{A}\boldsymbol{\theta} - \bar{\boldsymbol{f}}) - \frac{1}{t}\boldsymbol{1} \\
\boldsymbol{\lambda}_{\underline{p}}(\underline{\boldsymbol{f}} - \boldsymbol{D}\boldsymbol{A}\boldsymbol{\theta}) - \frac{1}{t}\boldsymbol{1}
\end{bmatrix}
\tag{9}
$$

First, we eliminate the dual variables. Since the equality constraints are linear and separable in $\boldsymbol{\theta}$, this is computationally inexpensive and maintains sparsity of the system. Grouping all of the inequality constraints from (2)-(3) into $\boldsymbol{g}(\boldsymbol{\theta}) \leq \boldsymbol{0}$, let

$$
\boldsymbol{g}(\boldsymbol{\theta}) \overset{\mathrm{def}}{=}
\begin{bmatrix}
\boldsymbol{B}\boldsymbol{\theta} - \bar{\boldsymbol{p}} \\
\underline{\boldsymbol{p}} - \boldsymbol{B}\boldsymbol{\theta} \\
\boldsymbol{D}\boldsymbol{A}\boldsymbol{\theta} - \bar{\boldsymbol{f}} \\
\underline{\boldsymbol{f}} - \boldsymbol{D}\boldsymbol{A}\boldsymbol{\theta}
\end{bmatrix},
\; D\boldsymbol{g}(\boldsymbol{\theta}) =
\begin{bmatrix}
\boldsymbol{B} \\ -\boldsymbol{B} \\ \boldsymbol{D}\boldsymbol{A} \\ -\boldsymbol{D}\boldsymbol{A}
\end{bmatrix}
\tag{10}
$$

The reduced system is

$$
\boldsymbol{C}\Delta\boldsymbol{\theta} = \boldsymbol{w}, \tag{11}
$$

where the system matrix and right-hand side are

$$
\boldsymbol{C} = 2\boldsymbol{B}^T\boldsymbol{W}\boldsymbol{B} + D\boldsymbol{g}(\boldsymbol{\theta})^T \mathrm{diag}(-\boldsymbol{g}(\boldsymbol{\theta}))^{-1} \mathrm{diag}(\boldsymbol{\lambda}) D\boldsymbol{g}(\boldsymbol{\theta}) \tag{12}
$$

$$
\boldsymbol{w} = -\boldsymbol{r}_{dual} - D\boldsymbol{g}(\boldsymbol{\theta})^T \mathrm{diag}(\boldsymbol{g}(\boldsymbol{\theta}))^{-1} \boldsymbol{r}_{cent}. \tag{13}
$$

Since $\mathrm{diag}(\boldsymbol{g}(\boldsymbol{\theta}))$ is diagonal, $\boldsymbol{M}$ exhibits a sparse structure. The original matrix-splitting scheme can be applied to $\boldsymbol{C}$ in (12). Consider decomposing the matrix $\boldsymbol{C}$ into the difference of two matrices $\boldsymbol{M}$ and $\boldsymbol{N}$. Provided the spectral radius $\rho(\boldsymbol{M}^{-1}\boldsymbol{N})$ is less than one, let the fixed-point of the sequence

$$
\Delta\boldsymbol{\theta}^{t+1} = \boldsymbol{M}^{-1}\boldsymbol{N}\Delta\boldsymbol{\theta}^{t} + \boldsymbol{M}^{-1}\boldsymbol{w}, \tag{14}
$$

be $\Delta\boldsymbol{\theta}^*$, then $\boldsymbol{C}\Delta\boldsymbol{\theta}^* = \boldsymbol{w}$. The splitting (*i.e.* choice of $\boldsymbol{M}$ and $\boldsymbol{N}$) should be designed so that i) the matrix-splitting iterates converge and ii) the iterates in (14) are easy to calculate in a distributed way.

Without loss of generality, assume the bus ids and line ids are consecutively assigned across control areas. The matrix $\boldsymbol{C}$ can be decomposed into the sum of a block-diagonal matrix, $\boldsymbol{D}$, and a matrix containing the remaining off-diagonal entries $\boldsymbol{E}$. The consecutive ordering of the bus and line indices described above allows for all entries of $\boldsymbol{C}$ corresponding to buses within the same control area to be contained within a single diagonal block. Specifically, let

$$D_{ij} = \begin{cases} C_{ij} & \text{if buses } i \text{ and } j \\ & \text{belong to the same control area }, \\ 0 & \text{otherwise} \end{cases} \tag{15}$$

$$E_{ij} = \begin{cases} C_{ij} & \text{if buses } i \text{ and } j \\ & \text{belong to different control areas }, \\ 0 & \text{otherwise} \end{cases} \tag{16}$$

yielding $\boldsymbol{C} = \boldsymbol{D} + \boldsymbol{E}$. We use the following matrix-splitting design

$$\boldsymbol{M} = \boldsymbol{D} + \tau\bar{\boldsymbol{E}},\ \boldsymbol{N} = \tau\bar{\boldsymbol{E}} - \boldsymbol{E}, \tag{17}$$

where τ is a scalar parameter and $\bar{\boldsymbol{E}}$ is a diagonal matrix whose ith diagonal entry equals

$$\bar{E}_{ii} \stackrel{\text{def}}{=} \sum_{j \neq i} |A_{ij}|. \tag{18}$$

This is a block-Jacobi scheme modified to be diagonally dominant. We have the following proposition which ensures convergence of the matrix-splitting iterates in (14).

Proposition 1. *Using the splitting in (17), for $\tau \geq \frac{1}{2}$, the iterative updates in (14) converge.*

Proof. The sequence $\{\Delta\boldsymbol{v}^t\}$ in (14) converges to its limit $\Delta\boldsymbol{v}^*$ as $t \to \infty$ if and only if the spectral radius of the matrix $\boldsymbol{M}^{-1}\boldsymbol{N}$ is strictly less than 1 [1]. Furthermore, if the sequence converges, the limit $\Delta\boldsymbol{\theta}^*$ is the solution of the system, (i.e., $\boldsymbol{C}\Delta\boldsymbol{\theta}^* = \boldsymbol{w}$). In order to have the spectral radius $\rho(\boldsymbol{M}^{-1}\boldsymbol{N}) < 1$, it is sufficient to have $\boldsymbol{C} = \boldsymbol{M} - \boldsymbol{N} \succ \boldsymbol{0}$ and $\boldsymbol{M} + \boldsymbol{N} \succ \boldsymbol{0}$ [2].

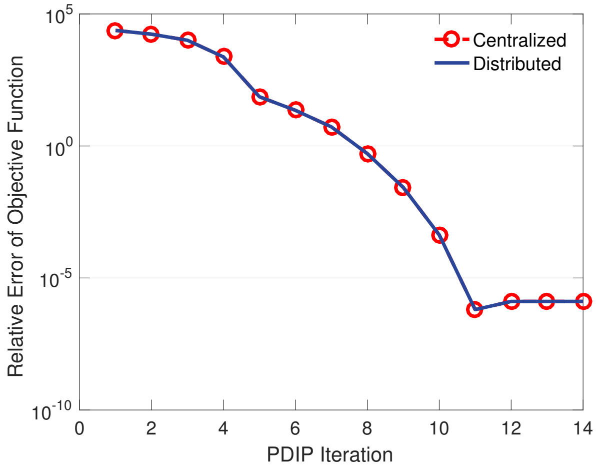
To show that $\boldsymbol{C} \succ 0$, first we note that the Hessian of the objective function $f_0(\boldsymbol{\theta}) \stackrel{\text{def}}{=} (\boldsymbol{B}\boldsymbol{\theta} - \boldsymbol{p}^*)^T \boldsymbol{W} (\boldsymbol{B}\boldsymbol{\theta} - \boldsymbol{p}^*)$ is diagonal and assumed to have strictly positive entries. Then, it is sufficient to show that 1) $D\boldsymbol{g}(\boldsymbol{\theta})$ has full column rank and 2) $\text{diag}(-\boldsymbol{g}(\boldsymbol{\theta}))^{-1}\text{diag}(\boldsymbol{\lambda})$ has strictly positive entries. Note from (10) that $D\boldsymbol{g}(\boldsymbol{\theta})$ has full column rank by construction since $\boldsymbol{B}$ is the bus susceptance matrix with the column corresponding to the reference angle eliminated. For each constraint i the entries

$$[\text{diag}(\boldsymbol{g}(\boldsymbol{\theta}))^{-1}\text{diag}(\boldsymbol{\lambda})]_{ii} = \frac{\lambda_i}{-g_i(\boldsymbol{\theta})} > 0 \tag{19}$$

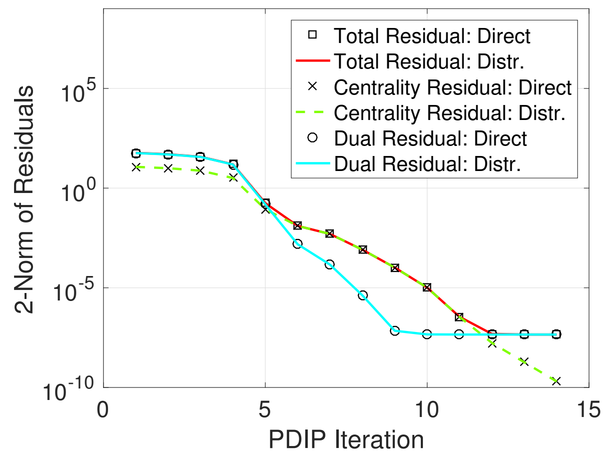
since $\lambda_i > 0$ and $g_i(\boldsymbol{\theta}) < 0$. Lastly, $\boldsymbol{M} + \boldsymbol{N}$ is positive definite due to its construction which makes it diagonally dominant. Details on this are provided in [3]. □

3 Updated Results

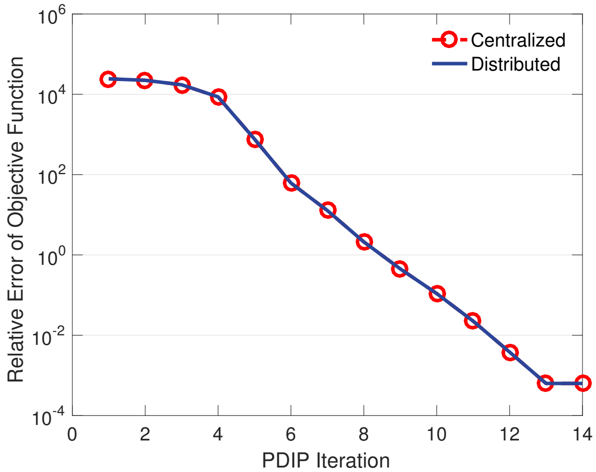
In Figure 3 (a)-(d), we show updated results using the new formulation. The IEEE 14-bus system is used in (a) and (b), and the IEEE 118-bus system is used in (c) and (d). The convergence of the relative error of the objective function for the centralized PDIP algorithm (*i.e.*, using direct inversion to calculate the Newton step) is compared to the distributed method (*i.e.*, using the proposed matrix-splitting method to calculate the Newton step). A tolerance of 10^{-10} is set to terminate the matrix-splitting iterates. For 118-bus system, between 10^5 and 10^6 matrix-splitting iterates are required per iteration. This is larger than in the original results using a simpler network-flow problem. While the number of inner-loop iterations required is high, each inner-loop iteration is a computationally inexpensive arithmetic calculation. Future work will be done to study the trade-off between the inner-loop tolerance and the outer-loop convergence. To study the quality of the solution, the convergence of the residuals using our distributed method is shown in (b) and (d) for the 14-bus and 118-bus systems, respectively.



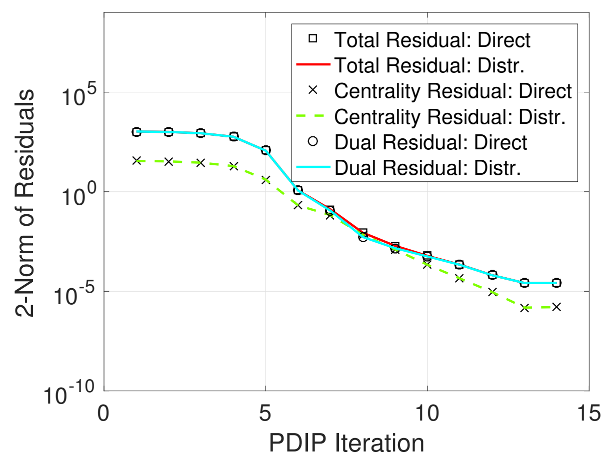
(a) Convergence of Obj. Function Relative Error: 14-Bus System



(b) Convergence of Residuals: 14-Bus Study



(c) Convergence of Obj. Function Relative Error: 118-Bus System



(d) Convergence of Residuals: 118-Bus Study

References

[1] R. S. Varga, *Matrix Iterative Analysis*, ser. Springer Series in Computational Mathematics. Berlin, Heidelberg, Paris: Springer Verlag, 2000.

[2] R. W. Cottle, J. Pang, and R. Stone, *Linear Complementarity Problem.* Academic Press, 1992.

[3] A. Minot and N. Li, "A Fully Distributed State Estimation Using Matrix Splitting Methods," *American Control Conference*, 2015.