Global Optimal Power Flow over Large-Scale Power Transmission Networks

Ye Shi, Hoang Duong Tuan, Pierre Apkarian, Andrey Savkin

Global Optimal Power Flow over Large-Scale Power Transmission Networks*

Y. Shi[a], H. D. Tuan[a], P. Apkarian[b], A. V. Savkin[c]

[a]*School of Electrical and Data Engineering, University of Technology, Sydney, NSW 2007, Australia*
[b]*CERT-ONERA, Control System Department, 2, avenua Edouard Belin, 31055 Toulouse, France*
[c]*School of Electrical Engineering and Telecommunications, the University of New South Wales, NSW 2052, Australia*

Abstract

Optimal power flow (OPF) over power transmission networks poses challenging large-scale nonlinear optimization problems, which involve a large number of quadratic equality and indefinite quadratic inequality constraints. These computationally intractable constraints are often expressed by linear constraints plus matrix additional rank-one constraints on the outer products of the voltage vectors. The existing convex relaxation technique, which drops the difficult rank-one constraints for tractable computation, cannot yield even a feasible point. We address these computationally difficult problems by an iterative procedure, which generates a sequence of improved points that converge to a rank-one solution. Each iteration calls a semi-definite program. Intensive simulations for the OPF problems over networks with a few thousands of buses are provided to demonstrate the efficiency of our approach. The suboptimal values of the OPF problems found by our computational procedure turn out to be the global optimal value with computational tolerance less than 0.01%.

Keywords: Optimal power flow (OPF) problem; large-scale transmission networks; rank-one matrix constraint; nonsmooth optimization; semi-definite programming (SDP).

1. Introduction

Smart grids are operated by the advanced distribution management system (DMS), which is responsible for supervisory control and data acquisition in reactive dispatch, voltage regulation, contingency analysis, capability maximization and other smart operations. The optimal power flow (OPF) problem, which determines a steady state operating point that minimizes the cost of electric power generation or the transmission loss is the backbone of DMS (see e.g. [1, 2, 3, 4] and references therein). Mathematically, the OPF problem is highly nonlinear and nonconvex due to numerous quadratic equality and indefinite quadratic inequality constraints for bus interconnections, hardware operating capacity and the balance between power demand and supply. These nonlinear constraints are mathematically troublesome so the state-of-the-art nonlinear optimization solvers may converge to just stationary points (see [5] and references therein), which are not necessarily feasible. To handle these nonlinear constraints, it is common to reformulate them as linear constraints on the outer product $W = VV^H$ of the voltage vector $V = (V_1, V_2, ..., V_n)^T \in \mathbb{C}^n$. As a result, the OPF problem is recast by a semi-definite program (SDP)

plus the additional rank-one constraint on outer product matrix W [6, 7]. The rank-one constraint is then dropped for semi-definite relaxation (SDR). However, the optimal solution of SDR is of rank-more-than-one in general and cannot help retrieval of a feasible point or stationary point of the OPF problem [7, 8, 9]. In [10], we have extended the technique of [11] for solving the beamforming optimization problems in signal processing to optimize the outer product matrix W, which works very well and is practical for moderate-scale power distribution networks up to $n = 300$ buses. There is another approach (see e.g. [12] and references therein), which is based on hierarchies of moment-based relaxation for nonconvex quadratic problems to tackle large networks with simple nonconvex constraints.

Power transmission networks in modern smart grids are often devised with a few thousand buses [13, 14, 15]. Under a such large number n of buses it is impossible to use the single matrix $W \in \mathbb{C}^{n \times n}$, which involves $n(n+1)/2 \approx O(10^7)$ complex variables. On the other hand, the number of the flow lines for bus connection is relatively moderate so only a small portion of the crossed nonlinear terms $V_k V_m^*$ appears in the nonlinear constraints. The common approach is to use the outer products of overlapped groups of the voltage variables to cover them [16, 17, 9]. All rank-one constraints on these outer products are then dropped for SDR. Obviously, the optimal solution of this SDR usually is not of rank-one and thus does not have any physical meaning. There is no technique to retrieve a feasi-

---

*Corresponding author H.D. Tuan. Tel. +61.02.9405 9435. Fax +61.02.9405 9039

Email addresses: `ye.shi@student.uts.edu.au` (Y. Shi), `tuan.hoang@uts.edu.au` (H. D. Tuan), `Pierre.Apkarian@onera.fr` (P. Apkarian), `a.savkin@unsw.edu.au` (A. V. Savkin)

ble rank-one point from the rank-more-than-one solution of SDR.[1] Multiple matrix rank constrained optimization has received a great attention due to its potential application in robust control synthesis [19, 20] but to our best knowledge there is no effective computation so far. The contribution of this paper is two-fold:

- An effective decomposition for large-scale OPF problems, which involves essentially reduced numbers of the rank-one constraints on matrices of moderate size for expressing the network nonlinear constraints;

- A new iterative procedure for rank-one constrained optimization, which is practical for computational solutions of large-scale indefinite quadratic programming. Simulations for the large-scale OPF problems show that it is capable of finding the global optimal solution with the computational tolerance less than 0.01%.

The paper is structured as follows. Section 2 is devoted to the OPF problem formulation and its difficulties. Its computational solution is developed in Section 3. Section 4 provides simulation to show the efficiency of our method. The conclusions are drawn in Section 5.

Notation. j denotes the imaginary unit; $M \succeq 0$ means that M is a Hermitian symmetric positive semi-definite matrix; $\mathsf{rank}(M)$ is the rank of the matrix M; $\Re(\cdot)$ and $\Im(\cdot)$ denote the real and imaginary parts of a complex quantity; $a \leq b$ for two complex numbers a and b is componentwise understood, i.e. $\Re(a) \leq \Re(b)$ and $\Im(a) \leq \Im(b)$; $\langle .,.\rangle$ is the dot product of matrices, while $\mathsf{diag}\{A_i\}$ denotes the matrix with diagonal blocks A_i and zero off-diagonal blocks; the cardinality of a set $\mathcal{L}$ is denoted by $|\mathcal{L}|$.

2. Optimal power flow problem and challenges

Consider an AC electricity transmission network with a set of n buses $\mathcal{N} := \{1, 2, \cdots, n\}$. The buses are connected through a set of flow lines $\mathcal{L} \subseteq \mathcal{N} \times \mathcal{N}$, i.e. bus m is connected to bus k if and only if $(m, k) \in \mathcal{L}$. Accordingly, $\mathcal{N}(k) := \{m \in \mathcal{N} : (m, k) \in \mathcal{L}\}$. The power demanded at bus $k \in \mathcal{N}$ is $S_{L_k} = P_{L_k} + jQ_{L_k}$, where P_{L_k} and Q_{L_k} are the real and reactive power. A subset $\mathcal{G} \subseteq \mathcal{N}$ of buses is supposed to be connected to generators. Any bus $k \in \mathcal{N}\backslash\mathcal{G}$ is thus not connected to generators.

Other physical parameters are following [21, 22, 23]:

- $Y = [y_{km}]_{(k,m) \in \mathcal{N} \times \mathcal{N}} \in \mathbb{C}^{n \times n}$ is the admittance matrix [23]. Each y_{km} is the mutual admittance between bus k and bus m, so $y_{km} = y_{mk} \ \forall \ (k, m) \in \mathcal{L}$.

- V is the complex voltage vector, $V = [V_1, V_2, \cdots, V_n]^T \in \mathbb{C}^n$, where V_k is the complex voltage injected to bus $k \in \mathcal{N}$.

---
[1]There is an algorithm of finding a rank-one solution [18, Alg. 1], which however is applicable to simple nonconvex constraints and is not guaranteed to convergence

- I is the complex current vector, $I = YV = [I_1, I_2, \cdots I_n]^T \in \mathbb{C}^n$, where I_k is the complex current injected to bus $k \in \mathcal{N}$.

- I_{km} is the complex current in the power line $(k, m) \in \mathcal{L}$, $\sum_{m \in \mathcal{N}(k)} I_{km} = I_k = \sum_{m \in \mathcal{N}(k)} y_{km}V_m$.

- $S_{km} = P_{km} + jQ_{km}$ is the complex power transferred from bus k to bus m, where P_{km} and Q_{km} represent the real and reactive transferred power.

- $S_{G_k} = P_{G_k} + jQ_{G_k}$ is the complex power injected by bus $k \in \mathcal{G}$, where P_{G_k} and Q_{G_k} represent the real and reactive generated power.

For each bus k, it is obvious that

$$
\begin{aligned}
S_{G_k} - S_{L_k} &= (P_{G_k} - P_{L_k}) + j(Q_{G_k} - Q_{L_k}) \\
&= V_k I_k^* = V_k \sum_{m \in \mathcal{N}(k)} V_m^* y_{km}^*.
\end{aligned}
$$

Therefore, the real generated power P_{G_k} and reactive generated power Q_{G_k} at bus k are the following nonconvex quadratic functions of the bus voltage vector variable $V := (V_1, V_2, ..., V_n)^T \in \mathbb{C}^n$: $P_{G_k} = P_{L_k} + \Re(\sum_{m \in \mathcal{N}(k)} V_k V_m^* y_{km}^*)$

and $Q_{G_k} = Q_{L_k} + \Im(\sum_{m \in \mathcal{N}(k)} V_k V_m^* y_{km}^*)$.
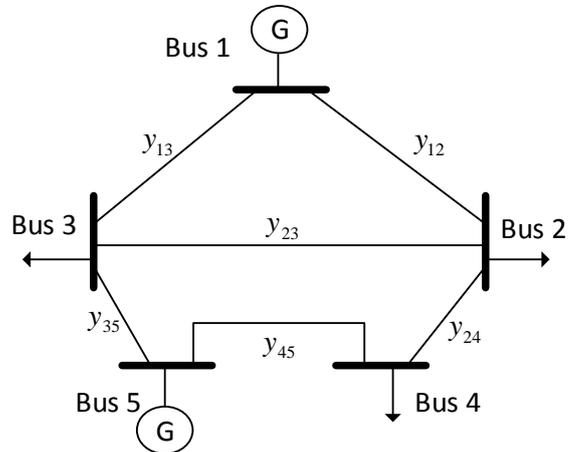
Figure 1: WB5-Five Bus Network [24]

For illustrative purpose, Figure 1 provides a diagram of WB5-Five Bus Network [24] with $\mathcal{N} = \{1, 2, 3, 4, 5\}$. It is seen from the bus connection in this Figure that $\mathcal{N}(1) = \{2, 3\}$ because buses 2 and 3 are connected to bus 1, $\mathcal{N}(2) = \{1, 3, 4\}$, $\mathcal{N}(3) = \{1, 5\}$, $\mathcal{N}(4) = \{2, 5\}$ and $\mathcal{N}(5) = \{3, 4\}$. Also $\mathcal{G} = \{1, 5\}$ because buses 1 and 5 are connected to generators.

The objective of OPF is to minimize either the total cost of the generated power defined by [22]

$$f(P_G) = \sum_{k \in \mathcal{G}} (c_{k2} P_{G_k}^2 + c_{k1} P_{G_k} + c_{k0}), \qquad (1)$$

with given $c_{k2} > 0$, c_{k1} and c_{k0} and real active generated power P_G, or the total transmission losses defined by

$$F_{loss}(P_G) = \sum_{k \in \mathcal{G}} (P_{G_k} - P_{L_k}) - \sum_{k \in \mathcal{N} \setminus \mathcal{G}} P_{L_k}. \qquad (2)$$

In this paper, we consider only the objective function (1). It is quite clear that our approach can be adjusted to solve the objective (2) in a straightforward manner.

Write the objective (1) as the following function of the bus voltages V:

$$f(V) = \sum_{k \in \mathcal{G}} [c_{k2}(P_{L_k} + \Re(\sum_{m \in \mathcal{N}(k)} V_k V_m^* y_{km}^*))^2$$
$$+ c_{k1}(P_{L_k} + \Re(\sum_{m \in \mathcal{N}(k)} V_k V_m^* y_{km}^*)) + c_{k0}]. \qquad (3)$$

Accordingly, the OPF problem is formulated as follows

$$\min_{V \in \mathbb{C}^n} f(V) \quad \text{s.t.} \quad (4a)$$

$$-P_{L_k} - jQ_{L_k} = \sum_{m \in \mathcal{N}(k)} V_k V_m^* y_{km}^*, k \in \mathcal{N} \setminus \mathcal{G}, \quad (4b)$$

$$P_{G_k}^{min} \le P_{L_k} + \Re(\sum_{m \in \mathcal{N}(k)} V_k V_m^* y_{km}^*) \le P_{G_k}^{max}, k \in \mathcal{G} \quad (4c)$$

$$Q_{G_k}^{min} \le Q_{L_k} + \Im(\sum_{m \in \mathcal{N}(k)} V_k V_m^* y_{km}^*) \le Q_{G_k}^{max}, k \in \mathcal{G} \quad (4d)$$

$$V_k^{min} \le |V_k| \le V_k^{max}, k \in \mathcal{N}, \quad (4e)$$

$$|S_{km}| = |V_k V_m^* y_{km}^*| \le S_{km}^{max}, \forall (k,m) \in \mathcal{L} \quad (4f)$$

$$|V_k - V_m| \le V_{km}^{max}, (k,m) \in \mathcal{L}, \quad (4g)$$

$$|\arg(V_k) - \arg(V_m)| \le \theta_{km}^{max}, (k,m) \in \mathcal{L}, \quad (4h)$$

where (4b) is the equation of the balance between the demand and supply power at bus $k \in \mathcal{N} \setminus \mathcal{G}$, (4c)-(4d) are the power generation bounds with the lower bound $P_{G_k}^{min}$, $Q_{G_k}^{min}$ and upper bound $P_{G_k}^{max}$, $Q_{G_k}^{max}$ of the real power reactive power generations, (4e) are the voltage amplitude bounds, (4f)-(4h) are capacity limitations with the line currents between the connected buses constrained by (4f), while (4g)-(4h) guarantee the voltage balance in terms of their magnitude and phases [23].

One can see that (4) is a highly nonlinear optimization subject to nonlinear inequality and equality constraints and thus poses a real challenge for its computation. A common approach is to use the slack variables

$$W_{km} = V_k V_m^*, k = 1, ..., n; m = 1, ..., n \qquad (5)$$

and then recast the problem (4) in $V \in \mathbb{C}^n$ to the following

problem in $W = [W_{km}]_{k,m=1,...,n} \in \mathbb{C}^{n \times n}$:

$$\min_{W \in \mathbb{C}^{n \times n}} F(W) \quad \text{s.t.} \quad W \succeq 0, \quad (6a)$$

$$-P_{L_k} - jQ_{L_k} = \sum_{m \in \mathcal{N}(k)} W_{km} y_{km}^*, \ k \in \mathcal{N} \setminus \mathcal{G}, \quad (6b)$$

$$P_{G_k}^{min} \le P_{L_k} + \Re(\sum_{m \in \mathcal{N}(k)} W_{km} y_{km}^*) \le P_{G_k}^{max}, \quad (6c)$$
$$k \in \mathcal{G},$$

$$Q_{G_k}^{min} \le Q_{L_k} + \Im(\sum_{m \in \mathcal{N}(k)} W_{km} y_{km}^*) \le Q_{G_k}^{max}, \quad (6d)$$
$$k \in \mathcal{G},$$

$$(V_k^{min})^2 \le W_{kk} \le (V_k^{max})^2, k \in \mathcal{N}, \quad (6e)$$

$$|W_{km} y_{km}^*| \le S_{km}^{max}, (k,m) \in \mathcal{L}, \quad (6f)$$

$$W_{kk} + W_{mm} - W_{km} - W_{mk} \le (V_{km}^{max})^2, \quad (6g)$$
$$(k,m) \in \mathcal{L},$$

$$\Im(W_{km}) \le \Re(W_{km}) \tan \theta_{km}^{max}, (k,m) \in \mathcal{L}, \quad (6h)$$

$$\text{rank}(W) = 1, \quad (6i)$$

where $F(W) = \sum_{k \in \mathcal{G}} [c_{k2}(P_{L_k} + \Re(\sum_{m \in \mathcal{N}(k)} W_{km} y_{km}^*))^2$
$+ c_{k1}(P_{L_k} + \Re(\sum_{m \in \mathcal{N}(k)} W_{km} y_{km}^*)) + c_{k0}]$, which is convex quadratic in W_{km}, while all constraints (6b)-(6h) are linear. All the problem nonconvexity is now concentrated in the single rank-one constraint (8b), which is then dropped for SDR. If the solution of this SDR is of rank-one, then it obviously leads to the global solution of the nonconvex optimization problem (6). Otherwise even a feasible point of (6) is hardly retrieved from the SDR solution. Moreover formulation (6) cannot be practically used for n more than a few thousands in large-scale networks as it involves $n(n+1)/2$ variables–a prohibitively large number. The next section will address and resolve all these issues.

3. Nonsmooth optimization based solution

The first issue is to decompose large-size matrix W in (6) into matrices of smaller size to make the problem tractable even with limited computational power. This is also prompted by the fact that there is only a small portion of the crossed terms $V_k V_m^*$ appearing in the nonlinear constraints (4b)-(4h) so the large-size matrix variable $W \in \mathbb{C}^{n \times n}$ contains many redundant terms $V_k V_m^*$. The main result of [16, 17, 9] is to decompose the set $\mathcal{N} := \{1, 2, ..., n\}$ of buses into $\mathcal{I}$ overlapped subsets $\mathcal{N}_i = \{i_1, ..., i_{N_i}\}$ of buses, called bags, such that $i_\ell \in \mathcal{N}(i_{\ell+1}), \ell = 1, ..., N_i - 1$ and $i_{N_i} \in \mathcal{N}(i_1)$, for each $i = 1, 2, ..., \mathcal{I}$, i.e. the buses in the same bag are serially connected. The set of bags can be reset to make bags of relatively same size. Define the Hermitian symmetric matrix variables

$$W^i = [W_{i_k i_m}]_{k,m=1,..,N_i} \in \mathbb{C}^{N_i \times N_i}, i = 1, 2, ..., \mathcal{I}. \quad (7)$$

By replacing $W_{km} = V_k V_m^*$ in (6) we have the following equivalent reformulation for (6)

$$\min_{W=\text{diag}\{W^i\}} F(W) \quad \text{s.t.} \quad (6b) - (6h), \quad (8a)$$

$$W^i \succeq 0, i = 1, ..., \mathcal{I}, \quad (8b)$$

$$\text{rank}(W^i) = 1, i = 1, ..., \mathcal{I}. \quad (8c)$$

Reference [16, 17] just dropped all rank-one constraints in (8c) for SDR without any justification. Reference [9] also dropped all rank-constraints in (8c) but then used a penalized SDR for locating low-rank semi-definite matrices W^i in (7). Based on these low-rank matrices, [9, Sec. IV] also proposed to find rank-one matrices, which however are not necessarily feasible to (8).

The variable number in (8) is $\sum_{i=1}^{\mathcal{I}} N_i(N_i+1)/2$. To keep this number reasonably moderate, it is desired that both $\mathcal{I}$ and N_i are sufficiently moderate. However, one can see that the above described decomposition [16, 17, 9] leads to a large number $\mathcal{I}$ of bags as well as few large size N_i that result in many rank-one constraints in (8c), which are much less probably satisfied by solving SDR.

Our first step toward to computation of (4) is to develop a new decomposition with many fewer bags involved. Recalling that $\mathcal{N}(k)$ is the set of the buses that are connected to bus k, the cardinality $|\mathcal{N}(k)|$ is small in large-scale networks. We resort $\mathcal{N} = \{1, 2..., n\}$ as $\mathcal{N} = \{N_1, ..., N_n\}$ such that the cardinality $|\mathcal{N}(N_k)|$ is in decreased order:

$$|\mathcal{N}(N_1)| \geq |\mathcal{N}(N_2)| \geq ... \geq |\mathcal{N}(N_n)|.$$

Accordingly, the first bag of buses is defined as $\mathcal{N}_1 = \mathcal{N}(N_1)$. The second bag is defined as

$$\mathcal{N}_2 = \{i \in \mathcal{N}(N_2) : \{i, N_2\} \not\subset \mathcal{N}_1\}.$$

Note that the crossed term $V_i V_{N_2}^*$ is already treated in the previous bag $\mathcal{N}_1$ whenever $\{i, N_2\} \subset \mathcal{N}_1$ so we exclude such bus i in defining bag $\mathcal{N}_2$.

Similarly, for $\ell \geq 3$ the ℓ-th bag is defined as

$$\mathcal{N}_\ell = \{i \in \mathcal{N}(N_\ell) : \{i, N_\ell\} \not\subset \mathcal{N}_{\ell'} \; \forall 1 \leq \ell' \leq \ell - 1\}$$

to exclude those buses i, whose crossed term $V_i V_{N_\ell}^*$ already is treated in a previous bag.

As each $|\mathcal{N}_i|$ is obviously small, such decomposition is very efficient, leading to a substantial reduction of involved bags in comparison with that used in [16, 17, 9].

Our next step is to tackle the numerous difficult rank-one constraints in (8c), not dropping them for SDR as in all the previous works.

Firstly we express $\mathcal{I}$ rank-one constraints in (8c) by the following single spectral constraint

$$\sum_{i=1}^{\mathcal{I}} (\text{Trace}(W^i) - \lambda_{\max}(W^i)) = 0, \quad (9)$$

where $\lambda_{\max}(W^i)$ stands for the maximal eigenvalue of $W^{(i)}$.

Indeed, (8b) implies $\text{Trace}(W^i) - \lambda_{\max}(W^i) \geq 0 \; \forall \, i$, so (9) means that $\text{Trace}(W^i) = \lambda_{\max}(W^i)$, i.e. W^i has only one nonzero eigenvalue so it is of rank-one. The nonnegative quantity $\sum_{i=1}^{\mathcal{I}}(\text{Trace}(W^i) - \lambda_{\max}(W^i))$ can therefore be used to measure the degree of satisfaction of the rank-one constraints (9). Without squaring, the penalization $\sum_{i=1}^{\mathcal{I}}(\text{Trace}(W^i) - \lambda_{\max}(W^i))$ is exact, meaning that (9) can be satisfied by a minimizer of the problem

$$\min_{W=\text{diag}\{W^i\}} F_\mu(W) := F(W) + \mu \sum_{i=1}^{\mathcal{I}} (\text{Trace}(W^i)$$
$$-\lambda_{\max}(W^i)) \quad \text{s.t.} \quad (6b) - (6h), (8b), \quad (10)$$

with a finite value of $\mu > 0$ (see e.g. [25, Chapter 16]). This is generally considered as a sufficiently nice property to make such exact penalization attractive.

For any $W^{i,(\kappa)}$ feasible for the convex constraints (6b)-(6h), (8b), function $\lambda_{\max}(W^i)$ is nonsmooth and is lower bounded by

$$\lambda_{\max}(W^i) = \max_{||w||=1} w^H W^i w \geq (w_{\max}^{i,(\kappa)})^H W^i w_{\max}^{i,(\kappa)}, \quad (11)$$

where $w_{\max}^{i,(\kappa)}$ is the normalized eigenvector corresponding to the eigenvalue $\lambda_{\max}(W^{i,(\kappa)})$, i.e.

$$\lambda_{\max}(W^{i,(\kappa)}) = (w_{\max}^{i,(\kappa)})^H W^{i,(\kappa)} w_{\max}^{i,(\kappa)}. \quad (12)$$

Accordingly, $\mu\lambda_{\max}(W^i) - \mu\lambda_{\max}(W^{i,(\kappa)}) \geq \mu((w_{\max}^{i,(\kappa)})^H W^i .w_{\max}^{i,(\kappa)} - (w_{\max}^{i,(\kappa)})^H W^{i,(\kappa)} w_{\max}^{i,(\kappa)}) = \mu\langle w_{\max}^{i,(\kappa)}(w_{\max}^{i,(\kappa)})^H, W^i - W^{i,(\kappa)}\rangle$, so $\mu w_{\max}^{i,(\kappa)}(w_{\max}^{i,(\kappa)})^H$ is a subgradient of the function $\mu\lambda_{\max}(W^i)$ at $W^{i,(\kappa)}$. Then $\mu\text{diag}\{w_{\max}^{i,(\kappa)}(w_{\max}^{i,(\kappa)})^H\}$ is a subgradient of the function $\mu\sum_{i=1}^{\mathcal{I}}\lambda_{\max}(W^i)$ at $\text{diag}\{W^{i,(\kappa)}\}$.

The following SDP provides an upper bound for the nonconvex optimization problem (10)

$$\min_{W=\text{diag}\{W^i\}} F^{(\kappa)}(W) := F(W) + \mu \sum_{i=1}^{\mathcal{I}} (\text{Trace}(W^i)$$
$$-(w_{\max}^{i,(\kappa)})^H W^i w_{\max}^{i,(\kappa)}) \quad \text{s.t.} \quad (6b) - (6h), (8b) \quad (13)$$

because $F^{(\kappa)}(\text{diag}\{W^i\}) \geq F_\mu(\text{diag}\{W^i\}) \quad \forall \, W^i \succeq 0$ according to (11). Suppose that $W^{(\kappa+1)} = \text{diag}\{W^{i,(\kappa+1)}\}$ is the optimal solution of SDP (13). Since $W^{(\kappa)} = \text{diag}\{W^{i,(\kappa)}\}$ is also feasible to (13) with $F_\mu(W^{(\kappa)}) = F^{(\kappa)}(W^{(\kappa)})$, it is true that

$$F_\mu(W^{(\kappa+1)}) \leq F^{(\kappa)}(W^{(\kappa+1)}) \leq F^{(\kappa)}(W^{(\kappa)}) = F_\mu(W^{(\kappa)}),$$

so $W^{(\kappa+1)}$ is a better feasible point of (10) than $W^{(\kappa)}$. Initialized by any feasible point $W^{(0)} = \text{diag}\{W^{i,(0)}\}$ of SDP constraint (8a)-(8b), the sequence $\{W^{(\kappa)}\} = \{\text{diag}\{W^{i,(\kappa)}\}\}$ with $W^{(\kappa+1)} = \text{diag}\{W^{i,(\kappa+1)}\}$ iteratively generated as the optimal solution of SDP (13) is a sequence of improved feasible points of the nonconvex optimization problem (10).

4

Since $W^{(\kappa)}$ are uniformly bounded, the sequence $\{W^{(\kappa)}\}$ has a limit point $\bar{W} = \mathsf{diag}\{\bar{W}^i\}$, which is the optimal solution of the optimization problem

$$\min_{W=\mathsf{diag}\{W^i\}} \quad F(W) + \mu \sum_{i=1}^{\mathcal{I}} (\mathsf{Trace}(W^i)$$
$$-(\bar{w}^i_{\max})^H W^i \bar{w}^i_{\max}) \quad \text{s.t.} \quad (6b)-(6h), (8b), \qquad (14)$$

where $\bar{w}^i_{\max}$ is the normalized eigenvector corresponding to the eigenvalue $\lambda_{\max}(\bar{W}^i)$ of $\bar{W}^i$. Particularly,

$$F(W) + \mu \sum_{i=1}^{\mathcal{I}} (\mathsf{Trace}(W^i) - (\bar{w}^i_{\max})^H W^i \bar{w}^i_{\max}) \quad \geq$$
$$F(W) + \mu \sum_{i=1}^{\mathcal{I}} (\mathsf{Trace}(\bar{W}^i) - (\bar{w}^i_{\max})^H \bar{W}^i \bar{w}^i_{\max}),$$

or equivalently, under the definition $g(W) = F(W) + \mu \sum_{i=1}^{\mathcal{I}} \mathsf{Trace}(W^i)$,

$$g(W) - g(\bar{W}) - \langle \mu \mathsf{diag}\{\bar{w}^i_{\max}(\bar{w}^i_{\max})^H\}, W - \bar{W}\rangle \geq 0$$

for all feasible points $W = \mathsf{diag}\{W^i\}$ in (6b)-(6h), (8b). As a result, $\bar{W}$ is the optimal solution of the convex optimization problem

$$\min_{W=\mathsf{diag}\{W^i\}} \quad g(W) - \langle \mu \mathsf{diag}\{\bar{w}^i_{\max}(\bar{w}^i_{\max})^H\}, W - \bar{W}\rangle$$
$$\text{s.t.} \quad (6b)-(6h), (8b),$$

so it must satisfy the optimality condition

$$\langle \nabla g(\bar{W}) - \mu \mathsf{diag}\{\bar{w}^i_{\max}(\bar{w}^i_{\max})^H\}, W - \bar{W}\rangle \geq 0$$

for all feasible points $\mathsf{diag}\{W^i\}$ in (6b)-(6h), (8b). The latter is also the first order necessary optimality condition for (10) because $\mu \mathsf{diag}\{\bar{w}^i_{\max}(\bar{w}^i_{\max})^H\}$ is a subgradient of the function $\mu \sum_{i=1}^{\mathcal{I}} \lambda_{\max}(W^i)$ at $\bar{W}$. As our simulations will show, $\bar{W}$ is indeed the global optimal solution of (10) and (8).

However, unlike [10] with only a single rank-one constrained matrix, although quantity

$$\sum_{i=1}^{\mathcal{I}} (\mathsf{Trace}(W^{i,(\kappa)}) - \lambda_{\max}(W^{i,(\kappa)})) \qquad (15)$$

in (10) is iteratively decreased, not all individual quantities

$$\mathsf{Trace}(W^{i,(\kappa)}) - \lambda_{\max}(W^{i,(\kappa)}) \qquad (16)$$

are iteratively decreased so the rank of each matrix $W^{i,(\kappa)}$ is no longer iteratively reduced to one as expected. Worse, $W^{i,(\kappa)}$ is rank-one but the rank of $W^{i,(\kappa+1)}$ in the next

iteration may turn to be more than one with

$$\mathsf{Trace}(W^{i,(\kappa+1)}) - \lambda_{\max}(W^{i,(\kappa+1)}) \quad >$$
$$\mathsf{Trace}(W^{i,(\kappa)}) - \lambda_{\max}(W^{i,(\kappa)}).$$

Consequently, it is very difficult to achieve rank-one for all $W^{i,(\kappa)}$ as desired. It is also impossible to add a "weight" to each term under the sum in the objective in (13) to handle the individual convergence of $\mathsf{Trace}(W^i) - \lambda_{\max}(W^i)$. We now develop a systematic way to resolve this issue as follows. For $\kappa = 0, 1, ...,$ and $W^{(\kappa)} = \mathsf{diag}\{W^{(i,(\kappa))}\}$ define

$$\mathcal{L}^{(\kappa)} = \{i \in \{1, ..., \mathcal{I}\} \; : \; \mathsf{rank}(W^{i,(\kappa)}) = 1\} \qquad (17)$$

and generate $W^{(\kappa+1)} = \mathsf{diag}\{W^{i,(\kappa+1)}\}$ as the optimal solution of the following SDP instead of SDP (13)

$$\min_{W=\mathsf{diag}\{W^i\}} \quad F(W) + \mu \sum_{i=1}^{\mathcal{I}} [\mathsf{Trace}(W^i)$$
$$-(w^{i,(\kappa)}_{\max})^H W^i w^{i,(\kappa)}_{\max}] \quad \text{s.t.} \quad (6b)-(6h), (8b), \quad (18a)$$
$$\mathsf{Trace}(W^i) - (w^{i,(\kappa)}_{\max})^H W^i w^{i,(\kappa)}_{\max} \leq \epsilon_{tol}, i \in \mathcal{L}^{(\kappa)}. (18b)$$

Note that $\mathsf{Trace}(W^i) \geq w^H W w$ for all $||w|| = 1$ and it is obvious that $\mathsf{rank}(W^i) = 1$ if and only if $\mathsf{Trace}(W^i) - w^H_{max} W^i w_{\max} = 0$ for some normalized $w_{\max}$. Therefore, the constraint (18b) for some tolerance ϵ_{tol} is introduced to warrant the rank-one of all $W^{i,(\kappa+1)}$, $i \in \mathcal{L}^{(\kappa)}$. As a result $\mathcal{L}^{(\kappa)} \subset \mathcal{L}^{(\kappa+1)}$ and $\mathcal{L}^{(\kappa)} \to \{1, ..., \mathcal{I}\}$ is expected to have all $W^{i,(\kappa)}$ of rank-one. Unlike (13), the iterations (18) leads to achieving rank-one of all W^i while the objective function F_μ is still decreased.

In summary, we propose the following *Large-Scale Nonsmooth Optimization Algorithm (Large-scale NOA)* for the multiple rank-one constrained optimization problem (10).

Initialization. Solve SDP

$$\min_{W=\mathsf{diag}\{W^i\}} \quad F(W) \quad \text{s.t.} \quad (6b)-(6h), (8b) \qquad (19)$$

to generate $W^{(0)} := \mathsf{diag}\{W^{i,(0)}\}$. If $\mathsf{rank}(W^{i,(0)}) \equiv 1$ stop: $W^{(0)}$ is the global solution of the nonconvex optimization problem (8). Otherwise set $\kappa = 0$ and define $\mathcal{L}^{(\kappa)}$ by (17).

κ-th iteration. For $\kappa = 0, 1, ..,$ solve (18) to generate $W^{(\kappa+1)} := \mathsf{diag}\{W^{i,(\kappa+1)}\}$. Reset $\kappa = \kappa + 1$ and define $\mathcal{L}^{(\kappa)}$ by (17). Stop whenever $\mathcal{L}^{(\kappa)} = \{1, ..., \mathcal{I}\}$. Otherwise go to the next iteration.

4. Simulation results

The computation facilities for our implementation are Processor Intel(R) Core i5-3470 CPU @3.20GHz, Matlab version R2013b and CVX with SDPT3. We set the tolerances $\epsilon = \epsilon_{tol} = 10^{-5}$ and the penalty parameter $\mu = 10^6$ which makes the penalty term $\mu \sum_{i=1}^{\mathcal{I}} (\mathsf{Trace}(W^{(i,(0))}) - \lambda_{\max}(W^{i,(0)}))$ at similar magnitude with the objective $F(W^{(0)})$. The data source for all examples is Matpower version 5.1

[23]. All examples were considered in [16, 17, 9] by SDR only. We recall that $\mathcal{T}$ is the number of matrix variables W^i in the OPF problem (8) and $\mathcal{L}^{(\kappa)}$ is defined by (17) is the set of rank-one matrices W^i found after κ-th iteration. The capability of our large-scale NOA in locating the global optimal solution of the OPF problems is demonstrated by showing that the global optimality tolerance (GOT) of its found solution defined as

$$\frac{\text{the found value- lower bound}}{\text{lower bound}}$$

is almost zero.

The numerical examples are presented as follows.

4.1. Polish-2383wp system

There are $n = 2383$ buses, 327 generators and 2896 transmission lines, leading to 2056 nonlinear constraints in (4b).
Initialization. A lower bound 1.8490×10^6 of (8) is found by solving SDP (19). $|\mathcal{L}^{(0)}| = 1210$ and there are 32 matrices $W^{i,(0)}$ of rank-more-than-one. Their largest size (smallest size, resp.) is 10×10 (2×2, resp.).
Stage 1. $|\mathcal{L}^{(10)}| = 1234$ is achieved. There are 8 matrices $W^{i,(10)}$ of rank-more-than-one. Their largest size (smallest size, resp.) is 10×10 and (3×3, resp.).
Stage 2. $|\mathcal{L}^{(19)}| = 1237$ is achieved. There are 5 matrices $W^{i,(20)}$ of rank-more-than-one. Their largest size (smallest size, resp.) is 9×9 (3×3, resp.).
Stage 3. $|\mathcal{L}^{(25)}| = \mathcal{I} = 1242$ is achieved. The found value of the objective is 1.8408×10^6 with GOT $4.3267e - 04$.

4.2. Polish-2736sp system

There are $n = 2736$ buses, 420 generators and 3504 transmission lines, which lead to 2316 nonlinear constraints in (4b).
Initialization. A lower bound 1.3041×10^6 of (8) is obtained by solving SDP (19). $|\mathcal{L}^{(0)}| = 1534$ and there are 4 matrices $W^{i,(0)}$ of rank-more-than-one. Their largest size (smallest size, resp.) is 6×6 (4×4, resp.).
Stage 1. $|\mathcal{L}^{(9)}| = \mathcal{I} = 1538$ is achieved. The found value of the objective is 1.3042×10^6 with GOT $7.6681e - 05$.

4.3. Polish-2737sop system

There are $n = 2737$ buses, 399 generators and 3506 transmission lines, which lead to 2338 nonlinear constraints in (4b).
Initialization. A lower bound 7.7571×10^5 of (8) is obtained by solving SDP (19). $|\mathcal{L}^{(0)}| = 1532$ and there are 6 matrices $W^{i,(0)}$ of rank-more-than-one. Their largest size (smallest size, resp.) is 6×6 (3×3, resp.).
Stage 1. $|\mathcal{L}^{(2)}| = \mathcal{I} = 1538$ is achieved. The found value of the objective is 7.7572×10^5 with GOT $1.2891e - 05$.

4.4. Polish-2746wop system

There are $n = 2746$ buses, 514 generators and 3514 transmission lines, which lead to 2232 nonlinear constraints in (4b).
Initialization. A lower bound 1.2039×10^6 of (8) is obtained by solving SDP (19). $|\mathcal{L}^{(0)}| = 1538$ and there are 8 matrices $W^{i,(0)}$ of rank-more-than-one. Their largest size (smallest size, resp.) is 6×6 (3×3, resp.).
Stage 1. $|\mathcal{L}^{(2)}| = \mathcal{I} = 1546$ is achieved. The found value of the objective is 1.2040×10^6 with GOT $8.3063e - 05$.

4.5. Polish-2746wp system

There are $n = 2746$ buses, 520 generators and 3514 transmission lines, which lead to 2226 nonlinear constraints in (4b).
Initialization. A lower bound 1.626590×10^6 of (8) is obtained by solving SDP (19). $|\mathcal{L}^{(0)}| = 1545$ and there are 2 matrices $W^{i,(0)}$ of rank-more-than-one. Their size is 4×4.
Stage 1. $|\mathcal{L}^{(1)}| = \mathcal{I} = 1547$ is achieved. The found value of the objective is 1.626591×10^6 with GOT $6.1478e - 07$.

4.6. Polish-3012wp system

There are $n = 3012$ buses, 502 generators and 3572 transmission lines, which lead to 2510 nonlinear constraints in (4b).
Initialization. A lower bound 2.5717×10^6 of (8) is obtained by solving SDP (19). $|\mathcal{L}^{(0)}| = 1682$ and there are 7 matrices $W^{i,(0)}$ of rank-more-than-one. Their largest size (smallest size, resp.) is 7×7 (2×2, resp.).
Stage 1. $|\mathcal{L}^{(4)}| = \mathcal{I} = 1689$ is achieved. The found value of the objective is 2.5727×10^6 with GOT $3.8885e - 04$.

4.7. Polish-3120sp system

There are $n = 3120$ buses, 505 generators and 3693 transmission lines, which lead to 2615 nonlinear constraints in (4b).
Initialization. A lower bound 2.1314×10^6 of (8) is obtained by solving SDP (19). $|\mathcal{L}^{(0)}| = 1749$ and there are 8 matrices $W^{i,(0)}$ of rank-more-than-one. Their largest size (smallest size, resp.) is 8×8 (2×2, resp.).
Stage 1. $|\mathcal{L}^{(9)}| = \mathcal{I} = 1757$ is achieved. The found value of the objective is 2.1391×10^6 with GOT 0.0036.

4.8. Numerical summary

One can observe that GOT of the solutions computed by the large-scale NOA is very small, proving its capability to provide the global solution of (8). Table 1 and Table 2 summarize the main points in our simulation. The second and third columns of Table 1 are the number $\mathcal{I}$ of bags in (7) by our decomposition and by that in [9], while the fourth and fifth columns give the maximum size N_i in (7). One can see that both $\mathcal{I}$ and the maximum N_i by our decomposition are substantially smaller than their counterparts by [9]. This leads to far smaller numbers of variables in (8), which are provided in the seventh and

Table 1: Comparison of bags number $\mathcal{I}$, largest bag size M_i and number of variables

System	$\mathcal{I}$	$\mathcal{I}$ by [9]	Max. N_i	Max. N_i by [9]	Var. # in (6)	Var. # in (8)	Var. # in (8) by [9]	Found value	Found by [9]
Polish-2383wp	1242	2383	10	23	2,840,536	23,199	89,893	1.8408×10^6	1.8742×10^6
Polish-2736sp	1538	2736	10	23	3,744,216	27,298	104,388	1.3042×10^6	1.3082×10^6
Polish-2737sop	1538	2737	10	23	3,746,953	27,034	103,720	7.7572×10^5	7.7766×10^5
Polish-2746wop	1546	2746	10	23	3,771,631	29,024	108,950	1.2040×10^6	1.2085×10^6
Polish-2746wp	1547	2746	10	24	3,771,631	28,257	107,148	1.6266×10^6	1.6324×10^6
Polish-3012wp	1689	3012	10	24	4,537,578	30,996	116,799	2.5727×10^6	2.6089×10^6
Polish-3120sp	1757	3120	10	24	4,868,760	32,637	121,869	2.1391×10^6	2.1608×10^6

Table 2: Performance comparison

System	Found value	Found by [9]	Found by [23]
Polish-2383wp	1.8408×10^6	1.8742×10^6	1.8685×10^6
Polish-2736sp	1.3042×10^6	1.3082×10^6	1.3078×10^6
Polish-2737sop	7.7572×10^5	7.7766×10^5	7.7763×10^5
Polish-2746wop	1.2040×10^6	1.2085×10^6	1.2083×10^5
Polish-2746wp	1.6266×10^6	1.6324×10^6	1.6317×10^6
Polish-3012wp	2.5727×10^6	2.6089×10^6	2.5917×10^6
Polish-3120sp	2.1391×10^6	2.1608×10^6	2.1427×10^6

eighth columns. The number $n(n + 1)/2$ of complex variables in (6) is also provided in the sixth column to contrast to the number of complex variables in (8) in the seventh column. Furthermore, the second column of Table 2 provides the best values of (4) found by our large-scale NOA, which are far smaller than ones in the third and fourth columns found by [9] and Matpower6.0 [23] (using an interior point method), respectively. In short, our computation approach to the OPF problem (8) outperforms other existing approaches in terms of computational efficiency and performance.

5. Conclusion

The OPF problems over power transmission network are large-scale optimization problems, which involve a large number of quadratic equality and indefinite quadratic inequality constraints and thus are difficult computationally. We have developed a large-scale nonsmooth optimization algorithm to compute their optimal solutions, which is efficient and practical for large-scale power transmission networks of a few thousands of buses. Applications of the developed large-scale NOA to the OPF problems over three-phase power transmission networks are currently under investigation.

References

[1] J. Carpentier, Contribution to the economic dispatch problem, Bulletin Society Francaise Electriciens 3 (8) (1962) 431–447.
[2] M. Huneault, F. D. Galiana, A survey of the optimal power flow literature, IEEE Trans. Power Systems 6 (2) (1991) 762–770.
[3] J. A. Momoh, M. E. El-Hawary, R. Adapa, A review of selected optimal power flow literature to 1993. part I: Nonlinear and quadratic programming approaches, Solar Energy 14 (1) (1999) 96–104.
[4] K. S. Pandya, S. K. Joshi, A survey of optimal power flow methods, J. of Theoretical and Applied Information Technology 4 (5) (2008) 450–458.
[5] W. Bukhsh, A. Grothey, K. McKinnon, P. Trodden, Local solutions of the optimal power flow problem, IEEE Trans. Power Systems 28 (4) (2013) 4780–4788.
[6] X. Bai, H. Wei, K. Fujisawa, Y. Wang, Semidefinite programming for optimal power flow problems, Elect. Power Energy Syst. 30 (6-7) (2008) 383–392.
[7] J. Lavaei, S. H. Low, Zero duality gap in optimal power flow problem, IEEE Trans. Power Systems 27 (1) (2012) 92–107.
[8] R. Madani, S. Sojoudi, J. Lavaei, Convex relaxation for optimal power flow problem: Mesh networks, IEEE Trans. Power Systems 30 (2015) 199–211.
[9] R. Madani, M. Ashraphijuo, J. Lavaei, Promises of conic relaxation for contigency-constrained optimal power flow problem, IEEE Trans. Power Systems 30.
[10] Y. Shi, H. D. Tuan, S. Su, H. H. M. Tam, Nonsmooth optimization for optimal power flow over transmission networks, in: Proc. 2015 IEEE Global Conf. on Signal and Information Processing (GlobalSIP), 2015.
[11] H. A. Phan, H. D. Tuan, H. H. Kha, D. T. Ngo, Nonsmooth optimization for efficient beamforming in cognitive radio multicast transmission, IEEE Trans. Signal Processing 60 (6) (2012) 2941–2951.
[12] C. Josz, D. K. Molzahn, Moment/sum-of-squares hierarchy for complex polynomial optimization, Available:http://arxiv.org/abs/1508.02068.
[13] Bamberger, et al, Vision and Strategy for Europes Electricity Networks of the Future, Office for Official Publications of the European Communities, 2006.
[14] P. Hallberg, et al, Smart Grids and Networks of the Future - EURELECTRIC Views, Ref: 2009-030-0440, 2009.
[15] M. Hashmi, S. Hnninen, K. Mki, Survey of smart grid concepts, architectures, and technological demonstrations worldwide, in: Proc. 2011 IEEE PES Conf. on Innovative Smart Grid Technologies (ISGT Latin America), 2011, pp. 1–7.
[16] D. K. Molzahn, J. T. Holzer, B. C. Lesieutre, C. L. DeMarco, Implementation of a large-scale optimal power flow solver based on semidefinite programming, IEEE Trans. Power Systems 28 (2013) 3987–3998.
[17] M. S. Andersen, A. Hansson, L. Vandenberghe, Reduced-complexity semidefinite relaxations of optimal power flow problems, IEEE Trans. Power Systems 29 (4) (2014) 1855–1863.
[18] D. K. Molzahn, C. Josz, I. A. Hiskens, P. Panciatici, A Laplacian-based approach for finding near globally optimal solutions to OPF problems, IEEE Trans. Power Systems 32 (1) (2017) 305–315.

[19] P. Apkarian, H. D. Tuan, A sequential SDP/Gauss-Newton algorithm for rank-constrained LMI problems, in: Proc. of the 38-th Conf. on Decision and Control (CDC), Phoenix, Arizona, 1999, pp. 2238–2245.

[20] P. Apkarian, D. Noll, O. Prot, A trust region spectral bundle method for nonconvex eigenvalue optimization, SIAM J. Optimization 19 (1) (2008) 281–306.

[21] S. Hadi, Power system analysis, PSA, 2010.

[22] H. W. Dommel, W. F. Tinney, Optimal power flow solutions, IEEE Trans. Power Apparatus Systems PAS-87 (10) (1968) 1866–1876.

[23] R. D. Zimmerman, C. E. Murillo-Sanchez, R. J. Thomas, MATPOWER: steady-state operations, planning, and analysis tools for power systems research and education, IEEE Trans. Power Systems 26 (1) (2011) 12–19.

[24] W. Bukhsh, A. Grothey, K. McKinnon, P. Trodden, Test case archive of optimal power flow (OPF) problems with local optima, http://www.maths.ed.ac.uk/optenergy/LocalOpt/introduction.html.

[25] J. F. Bonnans, J. C. Gilbert, C. Lemarechal, C. Sagastizabal, Numerical Optimization Theoretical and Practical Aspects (second edition), Springer, 2006.