A Non-convex Alternating Direction Method of Multipliers Heuristic for Optimal Power Flow

Seungil You
California Institute of Technology
syou@caltech.edu

Qiuyu Peng
California Institute of Technology
qpeng@caltech.edu

Abstract—**The optimal power flow (OPF) problem is funda-mental to power system planing and operation. It is a non-convex optimization problem and the semidefinite programing (SDP) relaxation has been proposed recently. However, the SDP relaxation may give an infeasible solution to the original OPF problem. In this paper, we apply the alternating direction method of multiplier method to recover a feasible solution when the solution of the SDP relaxation is infeasible to the OPF problem. Specifically, the proposed procedure iterates between a convex optimization problem, and a non-convex optimization with the rank constraint. By exploiting the special structure of the rank constraint, we obtain a closed form solution of the non-convex optimization based on the singular value decomposition. As a result, we obtain a computationally tractable heuristic for the OPF problem. Although the convergence of the algorithm is not theoretically guaranteed, our simulations show that a feasible solution can be recovered using our method.**

NOTATION

i is the imaginary unit. W^* is the Hermitian of W, $\mathrm{Tr}\,(W)$ is a trace of W, and $\|W\|_F = \sqrt{\mathrm{Tr}\,(WW^*)}$ is the Frobenius norm of W. The generalized inequality, $W \succeq 0$, means W is a positive semidefinite matrix.

The projection operator $\Pi_S(W) = \underset{Z \in S}{\mathrm{argmin}} \|W - Z\|_F^2$, is the projection of W onto the set S.

I. INTRODUCTION

The optimal power flow (OPF) problem optimizes certain objective such as power loss and generation cost subject to power flow equations and operational constraints. It is a fundamental problem because it underlies many power system operations and planning problems such as economic dispatch, unit commitment, state estimation, stability and reliability assessment, volt/var control, demand response.

The OPF problem is hard to solve due to its nonconvex feasible set. Recently, convex relaxations of the OPF problem have been proposed. In [1], semidefinite programming (SDP) relaxation is proposed and the non-convex, rank 1 constraint is dropped such that the resulting problem becomes convex. If the solution of the SDP relaxation is feasible for the OPF problem, which means its rank is 1, then a global optimum of the OPF problem can be recovered. The SDP relaxation is called *exact* in this case. In [2], Lavaei, et. al numerically show that the SDP relaxation is exact for all the IEEE standard test networks and sufficient conditions have been derived in [3], [4] that guarantees the exactness of the SDP relaxation.

On the other hand, second order cone (SOCP) relaxation [5] is proposed to solve the nonconvex OPF problem for radial network that is more efficient and numerically stable than SDP relaxation. In [6], [7], sufficient conditions are derived that guarantee the exactness of the SOCP relaxation under mild conditions for radial network. These convex relaxation based approaches attain much of research interest, see [8], [9] and references therein.

However, SDP/SOCP relaxation is not always exact, especially when the underlying network is not radial. Especially, [10] points out the semidefinite relaxation may fail to recover the rank one optimal solution in the distribution network which implies we can't extract the feasible solution of the OPF.

To obtain a feasible optimal solution of the OPF, the polynomial optimization approach based on the moment relaxation has been proposed in [11] and [12], and the branch and bound method for OPF has been investigated in [13]. Although these global optimization approaches guarantee the global optimal solution which is also a feasible solution of OPF, but these algorithms may not be computationally tractable. For example, in the worst case, the moment relaxation requires the exponentially many decision variable which makes the optimization computationally intractable.

To keep the computational complexity tractable, while enjoying convexity of the problem, the convex heuristics for recovering the rank one solution has been proposed in [14], and [15]. In [14], the authors perturbed the objective function, and this technique is guaranteed to work in some cases, but in general case we still do not have a rank one feasible solution. In [15], the authors successively solve the approximate convex program to promote rank sparsity in the solution, but still there is no guarantee on the rank one recovery.

In this paper, we propose another heuristic to obtain a rank one solution of the OPF based on the alternating direction method of multipliers (ADMM) [16]. The ADMM method when applied to a convex program, generates a sequence of the convex programs combined with the dual update that approximately solves the targeting optimization. For a convex problem, it is guaranteed to converge to a global optimal solution and has been used to develop distributed algorithm for the OPF problem in [17], [18], [19]. In contrast, for a non-convex problem, each step requires the non-convex program which is hard to solve, and there is no convergence result, so it is not an interesting algorithm for a general non-convex

optimization. However, the rank one constraint in the OPF makes the ADMM particularly interesting.

When we separate the rank one constraint in the OPF problem by defining auxiliary variable and apply ADMM to solve this modified, but equivalent version of the OPF, then there exists a closed form solution to handle non-convexity in the OPF problem through the singular value decomposition. In other words, in each step, we need to solve a convex program without the rank constraint, and the polishing step tries to satisfy the rank one requirement through singular value decomposition. Therefore we still keep the convexity of the problem, but unlike other convex heuristics, we explicitly enforce the rank one requirement. If it converges, a rank one feasible solution can be obtained.

The main advantage is that we keep the convexity of the problem, and guaranteed to recover the feasible solution of the OPF, but the main drawback is no guarantee on the convergence. In this sense, our method can be seen as a local optimization heuristic for the OPF.

This paper is organized as follows. Section II introduces the bus injection model of the optimal power flow problem, and the alternating direction method of multipliers method. Section III proposes the ADMM based heuristic for the OPF, section IV contains a numerical simulation of our proposed algorithm and section V concludes the paper.

II. PROBLEM FORMULATION

A. Bus Injection Model

Consider a power network consist of a set of buses $\mathcal{N}$ and a set of lines $\mathcal{E}$ connecting those buses. For each bus $i \in \mathcal{N}$, let V_i and I_i represent the complex voltage and current respectively. Let $S_i := V_i I_i^* = P_i + \mathbf{i}Q_i$ denote the net complex power injection, which is defined the generation minus demand. Denote a variable without subscript as a column vector, e.g. $V = (V_1, V_2, \cdots, V_N)^T$ and $I = (I_1, I_2, \cdots, I_N)^T$. For each line $(i,j) \in \mathcal{E}$, let $y_{ij} = g_{ij} + \mathbf{i}b_{ij}$ represent its admittance. Let $\mathcal{N}_i := \{j \mid (i,j) \in \mathcal{E}\}$ represent the set of neighbors of i. The admittance matrix $Y \in \mathbb{C}^{|\mathcal{N}| \times |\mathcal{N}|}$ then is defined as

$$Y_{ij} = \begin{cases} \sum_{i \in \mathcal{N}_i} y_{ij} & \text{if } i = j \\ -y_{ij} & \text{if } (i,j) \in \mathcal{E} \\ 0 & \text{otherwise.} \end{cases}$$

The Kirchoff's law says that

$$I_i = \sum_{j \in \mathcal{N}_i} (V_i - V_j)y_{ij} \quad i \in \mathcal{N},$$

which is equivalent to $I = YV$. The bus injection model captures the relation between the power injection S_i and the voltage V_i at each bus i:

$$S_i = V_i I_i^* = \sum_{j \in \mathcal{N}_i} V_i(V_i - V_j)^* y_{ij}^* \quad i \in \mathcal{N}. \tag{1}$$

B. OPF and SDP relaxation

The OPF problem seeks to optimize certain objective, e.g. total line loss, or generation cost, subject to power flow equations (1) and various operational constraints. We consider an objective function of the following form:

$$\sum_{k \in \mathcal{N}} f_k(|V_i|, P_i).$$

For instance,

- To minimize total line loss, we can set $f_k(|V_i|, P_i) = P_i$ for each $i \in \mathcal{N}$.
- To minimize generation cost, we can set $f_k(|V_i|, P_i) = c_{i2}P_i^2 + c_{i1}P_i + c_{i0}$ for each $i \in \mathcal{N}$, where $c_{i2} > 0, c_{i1}, c_{i0}$ are given.[1]

or weighted sum of these functions.

From the operational requirement, the power injection at each bus $i \in \mathcal{N}$ is also restricted to some prescribed region:

$$P_i^{\min} \leq P_i \leq P_i^{\max}, \qquad Q_i^{\min} \leq Q_i \leq Q_i^{\max} \quad i \in \mathcal{N}. \tag{2}$$

For example, if the ith bus does not have generators, and the demand power is P_i^d, then the constraint becomes $-P_i^d \leq P_i \leq -P_i^d$, which implies the net power injection at the ith bus should match the demand. If the ith bus has a generator with operational constraint $\underline{P}_i^g \leq P_i^g \leq \overline{P}_i^g$, where P_i^g is the generation, then the constraint becomes $\underline{P}_i^g - P_i^d \leq P_i \leq \overline{P}_i^g - P_i^d$.

The voltage magnitude at each load bus $i \in \mathcal{N}$ needs to be maintained within a prescribed region, i.e.

$$V_i^{\min} \leq |V_i| \leq V_i^{\max}. \tag{3}$$

Typically the voltage magnitude is allowed to deviate by 5% from its nominal value, i.e. $V_i^{\min} = 0.95$ and $V_i^{\max} = 1.05$.

Finally, we can also include the line capacity constraint,

$$|I_{ij}| = |(V_i - V_j)y_{ij}| = |V_i - V_j||y_{ij}| \leq \overline{I}_{ij}$$

which is equivalent to

$$|V_i - V_j| \leq \Delta V_{ij}^{\max},$$

where $\Delta V_{ij}^{\max} = \overline{I}_{ij}/|y_{ij}|$.

The OPF problem can be written as

$$\begin{aligned} \min_{P,Q} \quad & \sum_k f_k(|V_i|, P_i) \\ \text{s.t.} \quad & S_i = P_i + \mathbf{i}Q_i = \sum_{j \in \mathcal{N}_i} V_i(V_i - V_j)^* y_{ij}^* \quad i \in \mathcal{N}, \\ & P_i^{\min} \leq P_i \leq P_i^{\max} \quad i \in \mathcal{N}, \\ & Q_i^{\min} \leq Q_i \leq Q_i^{\max} \quad i \in \mathcal{N}, \\ & V_i^{\min} \leq |V_i| \leq V_i^{\max} \quad i \in \mathcal{N}, \\ & |V_i - V_j| \leq \Delta V_{ij}^{\max} \quad (i,j) \in \mathcal{E}. \end{aligned} \tag{4}$$

---

[1]In principle, we formulate the quadratic cost as a function of generation, say $c_{i2}(P_i^g)^2 + c_{i1}(P_i^g) + c_{i0}$, where P_i^g is the generation at the ith bus. However, since the demand P_i^d is known, and $P_i = P_i^g - P_i^d$, we can manipulate the coefficient c_{i2}, c_{i1}, and c_{i0} to find an equivalent quadratic cost in terms of the net power injection P_i.

Note that this formulation is non-convex due to the first quadratic equality. To reformulate this problem as a rank constrained optimization, we need to express S_i in terms of the node voltage V_i. Denote $e_i = \begin{bmatrix} 0, \cdots, 1, \cdots, 0 \end{bmatrix}^T$ is the standard basis of $\mathbb{C}^n$, of which entries are zero except ith entry.

From the definition of the admittance matrix Y, we have

$$
\begin{aligned}
S_i &= e_i^* V (Y V e_i)^* = \mathrm{Tr}\left(VV^* Y^* e_i e_i^*\right) \\
|V_i|^2 &= (VV^*)_{ii}, \\
|V_i - V_j|^2 &= (VV^*)_{ii} + (VV^*)_{jj} - (VV^*)_{ij} - (VV^*)_{ji}.
\end{aligned}
$$

Therefore, by defining $\Phi_i := Y^* e_i e_i^*$, and $W = VV^*$, we arrive at

$$
\begin{aligned}
\min_W \quad & f_0(W) = \sum_k f_k(\mathrm{Tr}\,(W_{ii}), \mathrm{Re}\,\{\mathrm{Tr}\,(W\Phi_i)\}) \\
\text{s.t.} \quad & P_i^{\min} \le \mathrm{Re}\,\{\mathrm{Tr}\,(W\Phi_i)\} \le P_i^{\max} \quad i \in \mathcal{N}, \\
& Q_i^{\min} \le \mathrm{Im}\,\{\mathrm{Tr}\,(W\Phi_i)\} \le Q_i^{\max} \quad i \in \mathcal{N}, \\
& (V_i^{\min})^2 \le W_{ii} \le (V_i^{\max})^2 \quad i \in \mathcal{N}, \\
& W_{ii} + W_{jj} - W_{ij} - W_{ji} \le (\Delta V_{ij}^{\max})^2 \quad (i,j) \in \mathcal{E}, \\
& W \succeq 0, \ \mathbf{rank}\,W \le 1.
\end{aligned}
$$
(5)

Here the last requirement $W \succeq 0$, $\mathbf{rank}\,W \le 1$ is equivalent to $W = VV^*$.

Notice that this optimization is non-convex because of the rank constraint $\mathbf{rank}\,W \le 1$. The SDP relaxation [1] removes this rank one requirement to make the problem convex.

C. The alternating direction of the method of multiplier

ADMM blends the decomposability of dual decomposition with the superior convergence properties of the method of multipliers [16]. It solves optimization problem of the form:

$$
\begin{aligned}
\min_{x \in \mathcal{K}_x, z \in \mathcal{K}_z} \quad & f(x) + g(z) \\
\text{s.t.} \quad & Ax + Bz = c
\end{aligned}
$$

where $\mathcal{K}_x, \mathcal{K}_z$ are convex sets. Let λ denote the Lagrange multiplier for the constraint $Ax + Bz = c$. Then the augmented Lagrangian is defined as

$$
\begin{aligned}
L_\rho(x, z, \lambda) &:= f(x) + g(z) + \lambda^T (Ax + Bz - c) \\
&\quad + \frac{\rho}{2} \|Ax + Bz - c\|^2,
\end{aligned}
$$

where $\rho \ge 0$ is a constant. ADMM consists of the iterations:

$$
\begin{aligned}
x^{k+1} &\in \ \arg\min_{x \in \mathcal{K}_x} L_\rho(x, z^k, \lambda^k) && \text{(6a)} \\
z^{k+1} &\in \ \arg\min_{z \in \mathcal{K}_z} L_\rho(x^{k+1}, z, \lambda^k) && \text{(6b)} \\
\lambda^{k+1} &= \ \lambda^k + \rho(Ax^{k+1} + Bz^{k+1} - c) && \text{(6c)}
\end{aligned}
$$

Compared to dual decomposition, ADMM is guaranteed to converge to an optimal solution under less restrictive conditions. Let

$$
\begin{aligned}
r^k &:= \ \|Ax^k + Bz^k - c\| && \text{(7a)} \\
s^k &:= \ \rho\|A^T B(z^k - z^{k-1})\| && \text{(7b)}
\end{aligned}
$$

They can be viewed as the residuals for primal and dual feasibility. Assume:

- A1: f and g are closed proper and convex.
- A2: The unaugmented Lagrangian L_0 has a saddle point.

The correctness of ADMM is guaranteed by the following result in [16, Chapter 3].

Proposition 1 ([16])**.** *Suppose A1 and A2 hold. Let p^* be the optimal objective value. Then*

$$
\lim_{k \to \infty} r^k = 0, \quad \lim_{k \to \infty} s^k = 0
$$

implying

$$
\lim_{k \to \infty} f(x^k) + g(z^k) = p^*
$$

III. An ADMM heuristic for the OPF

In this section, we apply the ADMM method to derive a heuristic for the nonconvex OPF problem (5). Unlike ADMM for the convex problem, here we don't have a global convergence nor a convergence to a global optimum.

Let $\mathcal{C}$ be the feasible set of (5) without the rank constraint. To apply an ADMM heuristic, we transfer the rank constraint to another variable:

$$
\begin{aligned}
\underset{W, Z}{\text{minimize}} \quad & f_0(W) \\
\text{subject to} \quad & W \in \mathcal{C} \\
& \mathrm{Rank}(Z) \le 1 \\
& W = Z.
\end{aligned}
$$
(8)

Since $W = Z$, this formulation is exactly equivalent to (5).

ADMM tries to solve (5) using the Augmented Lagrangian:

$$
L(W, Z, U) := f_0(W) + \mathrm{Tr}\left(\Lambda^*(W - Z)\right) + \frac{\rho}{2}\|W - Z\|_F^2.
$$

Here ρ is a tuning parameter, and Λ is a dual variable corresponding to the equality constraint $W = Z$. To find a saddle point of L, ADMM iterates following steps:

1) W-minimization: $W^{k+1} = \underset{W \in \mathcal{C}}{\mathrm{argmin}} L(W, Z^k, \Lambda^k)$,

2) Z-minimization: $Z^{k+1} = \underset{\mathrm{Rank}(Z) \le 1}{\mathrm{argmin}} L(W^{k+1}, Z, \Lambda^k)$,

3) Λ-update: $\Lambda^{k+1} = \Lambda^k + \rho(W^{k+1} - Z^{k+1})$,

until some stopping criterion (which we will describe later) is satisfied.

Notice that the W-minimization step is a convex program, but the Z-minimization step is not a convex problem due to the rank constraint. This is the reason why we usually don't apply the ADMM method to a non-convex problem. It increases the problem complexity since we are required to solve a sequence of a non-convex problem, but we can't get any benefit from this procedure. However, the rank constraint helps us to come up with the tractable non-convex minimization, hence the ADMM provides a sequence of convex program which approximately solves the original non-convex OPF.

Let us formulate Z-minimization step explicitly. Since the objective function is given by,

$$
L(W^{k+1}, Z, \Lambda^k) = \frac{\rho}{2} \|Z - W^{k+1} - \frac{1}{\rho}\Lambda^k\|_F^2 + f_0(W^{k+1})
$$
$$
+ \mathrm{Tr}\left((\Lambda^k)^* W^{k+1}\right) - \frac{1}{2\rho}\|\Lambda^k\|_F^2,
$$

and only the first term depends on Z. Therefore Z-minimization is equivalent to the following minimization:

$$
\begin{aligned}
\underset{Z}{\text{minimize}} \quad & \|Z - (W^{k+1} + \frac{1}{\rho}\Lambda^k)\|_F^2 \\
\text{subject to} \quad & \mathrm{Rank}(Z) \leq 1.
\end{aligned}
\tag{9}
$$

This is precisely the projection of $W^{k+1} + \frac{1}{\rho}\Lambda^k$ onto a rank constrained set. The projection of a point onto a non-convex set is usually NP-hard, but in this case, the Eckart-Young theorem provides an analytical solution for this projection.

Theorem 2 (Eckart-Young theorem). *Let the singular value decomposition of $X = \sum_i \sigma_i u_i v_i^*$, where $\sigma_1 \geq \sigma_2 \geq \cdots$. Then, the projection, $\Pi_{\mathbf{rank}\, Z \leq 1}(X) = \sigma_1 u_1 v_1^*$.*

Proof. Since $\mathbf{rank}(Z) \leq 1$, there exists a vector $u, v \in \mathbb{C}^n$ such that $Z = uv^*$, where $\|u\|_2 = 1$. Since σ_1 is the largest singular value, we have, $\|Xu\|_2^2 \leq \sigma_1^2 \|u\|_2^2 = \sigma_1$. Therefore,

$$
\|X - Z\|_F^2 = \mathrm{Tr}\left((X - uv^*)(X - uv^*)^*\right)
$$
$$
= \mathrm{Tr}\left(XX^*\right) + (v - X^*u)^*(v - X^*u) - u^*X^*Xu
$$
$$
\geq \sum_{i=1} \sigma_i - \|Xu\|_2^2 \geq \sum_{i=1} \sigma_i - \sigma_1 = \sum_{i=2} \sigma_i,
$$

for all Z such that $\mathbf{rank}\, Z \leq 1$. Now it is routine to check that this lower bound is achieved by $Z^{\mathrm{opt}} = \sigma_1 u_1 v_1^*$. $\square$

Therefore, the hard, non-convex optimization can be done exactly. As we point out, this is not the usual case, and the special structure in the rank constraint helps us to solve the non-convex minimization step exactly.

Another important property of our algorithm is that if the initial iterates Z^0, Λ^0 are Hermitian matrices, then all W^k, Z^k, Λ^k are Hermitian matrices:

Proposition 3. *Suppose Z^0, Λ^0 are Hermitian matrices. Then for all $k \geq 0$, (W^k, Z^k, Λ^k) are Hermitian matrices.*

Proof. We can prove this by induction. Suppose Z^k, Λ^k are Hermitian matrices. Since the feasible set $\mathcal{C}$ requires $W^{k+1} \succeq 0$, W^{k+1} is Hermitian. Since $W^{k+1} + \Lambda^k/\rho$ is Hermitian, its singular value decomposition has the form $\sum_i \sigma_i u_i u_i^*$. Therefore the Z^{k+1} is also Hermitian. Since Λ^{k+1} is a sum of Hermitian matrices, it is also Hermitian. $\square$

Since we will choose $Z^0 = \Lambda^0 = 0$ as an initial iterate, we will assume that every matrices are Hermitian from now on. This implies the singular value decomposition becomes an eigenvalue decomposition.

A. Feasible point

The convergence of the ADMM heuristic for a non-convex problem is still an open question [16]. However, if it converges, then we are guaranteed to have a rank one feasible point of the OPF.

Proposition 4. *Let (W^*, Z^*, Λ^*) be the equilibrium point of the ADMM. Then W^* is a rank one feasible point of the OPF.*

Proof. By definition, the (W^*, Z^*, Λ^*) is an equilibrium point if and only if,

$$
\begin{aligned}
W^* &= \underset{W \in \mathcal{C}}{\arg\min}\ L(W, Z^*, \Lambda^*) \\
Z^* &= \underset{\mathbf{rank}(Z) \leq 1}{\arg\min}\ L(W^*, Z, \Lambda^*) \\
\Lambda^* &= \Lambda^* + \rho(W^* - Z^*).
\end{aligned}
$$

From the third equation, we can immediately conclude that $W^* = Z^*$. Since $\mathbf{rank}(Z^*) \leq 1$, so as $\mathbf{rank}(W^*) \leq 1$. Finally, since $W^* \in \mathcal{C}$, we can conclude that W^* is a rank one feasible point in the optimal power flow problem. $\square$

B. Stopping criterion and ρ

For the stopping criterion, we use the one from [16]. Let us define the primal, and dual residual:

$$
\begin{aligned}
R^k &= W^k - Z^k \\
S^k &= \rho(Z^{k+1} - Z^k).
\end{aligned}
$$

The primal residual R^k measures the difference between W^k and the rank one solution Z^k, and the dual residual S^k measures the progress on the Z^{k+1} and Z^k. When the algorithm converges, R^k and S^k should be zero. This gives rise to the following stopping criterion:

$$
\begin{aligned}
\|R^k\|_F &\leq \epsilon^{\mathrm{pri}}, \\
\|S^k\|_F &\leq \epsilon^{\mathrm{dual}}.
\end{aligned}
\tag{10}
$$

In other words, we terminate our algorithm when the above two stopping criterions are satisfied.

Lastly, the choice of ρ can be automated based on these residuals. For the details, refer [16].

C. Overall algorithm

By completing the square, we can summarize the W-minimization as follows:

$$
\begin{aligned}
\min_{W} \quad & f_0(W) + \frac{\rho}{2}\|W - (Z^k - \frac{1}{\rho}\Lambda^k)\|_F^2 \\
\text{s.t.} \quad & P_i^{\min} \leq \mathrm{Re}\left\{\mathrm{Tr}(W\Phi_i)\right\} \leq P_i^{\max} \quad i \in \mathcal{N}, \\
& Q_i^{\min} \leq \mathrm{Im}\left\{\mathrm{Tr}(W\Phi_i)\right\} \leq Q_i^{\max} \quad i \in \mathcal{N}, \\
& (V_i^{\min})^2 \leq W_{ii} \leq (V_i^{\max})^2 \quad i \in \mathcal{N}, \\
& W_{ii} + W_{jj} - W_{ij} - W_{ji} \leq (\Delta V_{ij}^{\max})^2 \quad (i,j) \in \mathcal{E}, \\
& W \succeq 0.
\end{aligned}
\tag{11}
$$

Although the objective function is not linear in W, we can make it linear by introducing slack variables. For f_0, refer [2], and for the Frobenius norm:

$$
\|W - T\|_F^2 \leq t \Leftrightarrow \begin{bmatrix} t & \mathrm{vec}(W - T) \\ (\mathrm{vec}(W - T))^* & I \end{bmatrix} \succeq 0,
$$

where vec($\cdot$) is the vectorization of the matrix. Therefore the optimization (11) is in fact a semidefinite program.

Notice that when $\rho = 0$, the optimization (11) is a SDP relaxation of the OPF. When $\rho > 0$, at each iteration, we reshape the objective function by adding "damping" term $\frac{\rho}{2}\|W - (Z^k - \frac{1}{\rho}\Lambda^k)\|_F^2$ to make W close enough to the $Z^k - \frac{1}{\rho}\Lambda^k$.

One can imagine that it would be more natural to enforce W^k close to Z^k rather than $Z^k - \frac{1}{\rho}\Lambda^k$, since $\mathbf{rank}\,Z^k \leq 1$. If we do this, then the entire procedure looks like an alternating projection which tries to find a point in the intersection of the convex set and the non-convex set, and we can imagine that it is highly likely to stuck at the point which is not in the intersection because one of the alternating set is not convex. The novelty in the ADMM method is the dual variable Λ^k which forces $W = Z$ in the limit. This helps us not to get trapped, but the price we pay is a possible oscillation.

Now we summarize our heuristic for the non-convex OPF.

An ADMM heuristic:
1) Set $\Lambda^0 = Z^0 = 0$, $\rho = \rho_0$, and $k = 0$.
2) Obtain W^{k+1} by solving SDP (11).
3) $Z^{k+1} = \sigma_1 v_1 v_1^*$, where σ_1, v_1 are the top singular value/vector of $W^{k+1} + \frac{1}{\rho}\Lambda^k$.
4) $\Lambda^{k+1} = \Lambda^k + \rho(W^{k+1} - Z^{k+1})$.
5) Terminate if the stopping criterion (10) is satisfied.
6) Let $k = k + 1$, and go to the step 2.

IV. CASE STUDIES

For numerical calculation ,we use SDPT3 [20], combined with the problem parser, YALMIP [21]

A. Two bus network

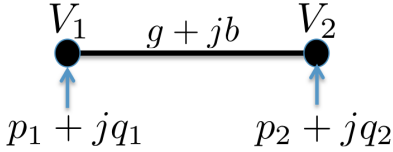$$V_1 \quad g + jb \quad V_2$$

$$p_1 + jq_1 \qquad p_2 + jq_2$$

Fig. 1: Two Bus network

Consider the two bus network in Fig. 1, where $(p_1, q_1), (p_2, q_2)$ are the complex power injection, V_1, V_2 are the complex voltage at bus 1 and 2 and Y_{12} is the line impedance. For simplicity, assume $|V_1| = |V_2| = 1$ and then the feasible power injections are constrained on a ellipse given by

$$\frac{(p_1 + p_2 - 2g)^2}{4g^2} + \frac{(p_1 - p_2)^2}{4b^2} = 1,$$

which is equivalent to the rank 1 constraint on W. In Fig. 2a, the feasible ellipse is illustrated using blue doted lines. It is shown in [3] that the feasible region becomes the two disjoint regions under some reactive power constraints on q_1, q_2, as shown by the black lines on the ellipse. Then the feasible set after relaxing the rank 1 constraints on W becomes the red
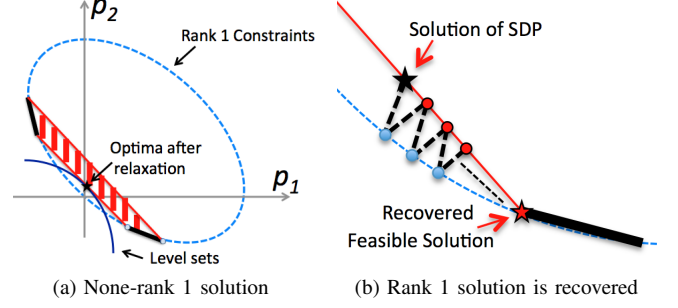

Fig. 2: Illustration of two bus network.

Bus	1	2	3	4	5	6	7	8	9	10
$P^{\min}$	-16	-14	-18	0	0	0	-20	-20	0	0
$P^{\max}$	-16	-14	-18	∞	∞	∞	-20	-20	20	∞
$Q^{\min}$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$
$Q^{\max}$	∞	∞	∞	∞	∞	∞	∞	∞	∞	∞

TABLE I: Power constraint for the 10 bus ring network.

region. In this case, if the level set of the objective hits the boundary of the feasible set as in Fig. 2a, we cannot obtain a rank 1 solution and the SDP relaxation fails to work.

In Fig. 2b, we show how the feasible solution is recovered by our heuristic algorithm. We begin with the solution with $\rho = 0$, denoted by W. Then we project W onto the ellipse, whose solution is Z. We find that the distance between W and Z decrease as we proceed our algorithm and eventually they coincide on the feasible power injection region on the ellipse, denoted by the red star in Fig. 2b. Hence, a feasible power injection (rank 1 solution) is received in this simple network.

B. 10 bus system

We consider the 10 bus system from [14]. Here the mesh network is a ring with 10 nodes and 10 links. All units are in 100 MVA. We assume 4,5,6,9, and 10 bus have generators, and the corresponding cost function is a linear function of P_i, $f_k(P_i) = c_i P_i$, where

$$c_4 = c_5 = c_{10} = 1, \qquad c_6 = c_9 = 2.$$

The line admittance values are $y_{1,2} = y_{3,4} = y_{4,5} = y_{10,1} = -\mathbf{i}$, $y_{2,3} = y_{6,7} = y_{9,10} = -2\mathbf{i}$, and $y_{5,6} = -3\mathbf{i}$, $y_{7,8} = -0.5\mathbf{i}$, and $y_{8,9} = -0.7\mathbf{i}$. The power demand at the load bus is given by

$$P_1^d = 16,\ P_2^d = 14,\ P_3^d = 18,\ P_7^d = P_8^d = 20.$$

The voltage constraint is $V_i^{\min} = 0.95$, and $V_i^{\max} = 1.05$. We put the generator operating constraint at the 9th bus, $P_9 \leq 20$. See the Table I for the summary of constraints on the power injection.

Finally we put the line constraint on all lines: $|V_i - V_j| \leq \Delta V^{\max}$ for all $(i, j) \in \mathcal{E}$. By changing $\Delta V^{\max}$, we control the size of the feasible set. In our simulation, we parameterize $\Delta V^{\max} = 2(1 - \cos(\theta^{\max}))$, where $\theta^{\max}$ is the maximum allowable phase difference between buses.

$\theta^{\max}$	$\mathbf{rank}\,W^{\mathrm{sdp}}$	$f_0(W^{\mathrm{sdp}})$	$\mathbf{rank}\,W^{\mathrm{admm}}$	$f_0(W^{\mathrm{admm}})$
10°	10	110.6	N/A	N/A
20°	10	88.0	1	88.0
30°	10	88.0	1	88.0
40°	10	88.0	1	88.0
50°	10	88.0	1	88.0

TABLE II: Simulation results, W^{sdp} is the solution from the SDP relaxation, and W^{admm} is the solution from our method. When $\theta^{\max} = 10°$, our method fails to converge.

$\theta^{\max}$	P_4	P_5	P_6	P_9	P_{10}
10°	18.1	19.2	15.7	6.9	28.0
20°	24.1	31.9	0.0	0.0	32.0
30°	26.8	30.7	0.0	0.0	30.5
40°	27.7	30.1	0.0	0.0	30.3
50°	28.1	29.7	0.0	0.0	30.2

TABLE III: Generation profile from the SDP relaxation. This profile is not implementable because the optimal solution is not feasible (rank 10).

$\theta^{\max}$	P_4	P_5	P_6	P_9	P_{10}
10°	N/A	N/A	N/A	N/A	N/A
20°	25.1	33.1	0.0	0.0	29.8
30°	28.4	31.9	0.0	0.0	27.7
40°	31.8	31.1	0.0	0.0	25.1
50°	35.4	30.7	0.0	0.0	21.9

TABLE IV: Generation profile from our method. This profile is implementable because the rank of the solution is 1

Table II shows the rank of the solution from the SDP relaxation, and the solution from our heuristic method. Indeed, when our heuristic converges, it recovers the rank one solution although the SDP relaxation always generate a full rank solution.[2] Moreover, since our heuristic attains the lower bound of the objective value, we can claim that it recovers the global optimal solution of which rank is one.

This example shows that our heuristic can be used to find a *hidden* rank one solution of the OPF where the semidefinite relaxation fails to find an existing rank one optimal solution.

V. CONCLUSION

In this paper, we propose a non-convex ADMM heuristic for the OPF. By introducing a redundant variable whose rank is one, we can split the minimization into two steps, where the first step is a convex optimization, and the second step is a rank constrained minimization. Then, we show that the second step, a non-convex optimization, can be carried out analytically. The convergence is still an open question, but our numerical studies show that the effectiveness of our method, that is the SDP relaxation of the OPF returns a full rank solution whereas our method returns a rank one solution when it converges. Moreover, we observe the convergence of our heuristic under the existence of hidden rank one solution in the SDP relaxation of the OPF. Inspired by this, the convergence proof under this

assumption, the existence of rank one global optimal solution, is currently investigated to show that our heuristic can be used to recover the hidden rank one solution in the SDP relaxation.

REFERENCES

[1] X. Bai, H. Wei, K. Fujisawa, and Y. Wang, "Semidefinite programming for optimal power flow problems," *International Journal of Electrical Power & Energy Systems*, vol. 30, no. 6, pp. 383–392, 2008.

[2] J. Lavaei and S. H. Low, "Zero duality gap in optimal power flow problem," *Power Systems, IEEE Transactions on*, vol. 27, no. 1, pp. 92–107, 2012.

[3] B. Zhang and D. Tse, "Geometry of feasible injection region of power networks," in *Communication, Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on*. IEEE, 2011, pp. 1508–1515.

[4] S. Bose, D. F. Gayme, K. M. Chandy, and S. H. Low, "Quadratically constrained quadratic programs on acyclic graphs with application to power flow," *arXiv preprint arXiv:1203.5599*, 2012.

[5] M. Farivar, R. Neal, C. Clarke, and S. Low, "Optimal inverter var control in distribution systems with high pv penetration," in *Power and Energy Society General Meeting, 2012 IEEE*. IEEE, 2012, pp. 1–7.

[6] M. Farivar and S. H. Low, "Branch flow model: relaxations and convexification (parts I, II)," *IEEE Trans. on Power Systems*, vol. 28, no. 3, pp. 2554–2572, August 2013.

[7] L. Gan, N. Li, U. Topcu, and S. H. Low, "Optimal power flow in distribution networks," in *Proc. 52nd IEEE Conference on Decision and Control*, December 2013, in arXiv:12084076.

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

[9] ——, "Convex Relaxation of Optimal Power FlowPart II: Exactness," *Control of Network Systems, IEEE Transactions on*, 2014, to be appeared.

[10] B. C. Lesieutre, D. K. Molzahn, A. R. Borden, and C. L. DeMarco, "Examining the limits of the application of semidefinite programming to power flow problems," in *Communication, Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on*. IEEE, 2011, pp. 1492–1499.

[11] C. Josz, J. Maeght, P. Panciatici, and J. C. Gilbert, "Application of the Moment-SOS Approach to Global Optimization of the OPF Problem," *arXiv preprint arXiv:1311.6370*, 2013.

[12] D. K. Molzahn and I. A. Hiskens, "Moment-based relaxation of the optimal power flow problem," *arXiv preprint arXiv:1312.1992*, 2013.

[13] A. Gopalakrishnan, A. U. Raghunathan, D. Nikovski, and L. T. Biegler, "Global optimization of optimal power flow using a branch & bound algorithm," in *Communication, Control, and Computing (Allerton), 2012 50th Annual Allerton Conference on*. IEEE, 2012, pp. 609–616.

[14] R. Madani, S. Sojoudi, and J. Lavaei, "Convex relaxation for optimal power flow problem: Mesh networks," in *Submitted to the 2013 Asilomar Conference*, 2013.

[15] R. Louca, P. Seiler, and E. Bitar, "A rank minimization algorithm to enhance semidefinite relaxations of optimal power flow," in *51st Annu. Allerton Conf. Commun., Control, and Comput*, 2013, pp. 2–4.

[16] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[17] M. Kraning, E. Chu, J. Lavaei, and S. Boyd, "Dynamic network energy management via proximal message passing," *Optimization*, vol. 1, no. 2, pp. 1–54, 2013.

[18] E. Dall'Anese, H. Zhu, and G. B. Giannakis, "Distributed optimal power flow for smart microgrids," *arXiv preprint arXiv:1211.5856*, 2012.

[19] Q. Peng and S. H. Low, "Distributed algorithm for optimal power flow on a radial network," *arXiv preprint arXiv:1404.0700*, 2014.

[20] K.-C. Toh, M. J. Todd, and R. H. Tütüncü, "Sdpt3a matlab software package for semidefinite programming, version 1.3," *Optimization methods and software*, vol. 11, no. 1-4, pp. 545–581, 1999.

[21] J. Lofberg, "Yalmip: A toolbox for modeling and optimization in matlab," in *Computer Aided Control Systems Design, 2004 IEEE International Symposium on*. IEEE, 2004, pp. 284–289.

[2] Minimum eigenvalue of the solution from the SDP relaxation is greater than 0.01 in all cases.