Penalized Parabolic Relaxation for Optimal Power Flow Problem

Fariba Zohrizadeh, Mohsen Kheirandishfard, Edward Quarm Jnr., and Ramtin Madani

Abstract—This paper is concerned with optimal power flow (OPF), which is the problem of optimizing the transmission of electricity in power systems. Our main contributions are as follows: (i) we propose a novel parabolic relaxation, which transforms non-convex OPF problems into convex quadratically-constrained quadratic programs (QCQPs) and can serve as an alternative to the common practice semidefinite programming (SDP) and second-order cone programming (SOCP) relaxations, (ii) we propose a penalization technique which is compatible with the SDP, SOCP, and parabolic relaxations and guarantees the recovery of feasible solutions for OPF, under certain assumptions. The proposed penalized convex relaxation can be used sequentially to find feasible and near-globally optimal solutions for challenging instances of OPF. Extensive numerical experiments on small and large-scale benchmark systems corroborate the efficacy of the proposed approach. By solving a few rounds of penalized convex relaxation, fully feasible solutions are obtained for benchmark test cases from [1]–[3] with as many as 13,659 buses. In all cases, the solutions obtained are not more than 0.32% worse than the best-known solutions.

I. INTRODUCTION

The optimal power flow problem (OPF) is concerned with the optimization of voltages, power flows, and power injections across transmission and distribution networks. This problem can be formulated as the minimization of a cost function (e.g., generation cost) subject to nonlinear constraints on power and voltage variables. Due to the inherent complexity of physical laws that model the flow of electricity, some of these constraints are non-convex, which makes the OPF problem NP-hard in general [4], [5]. Substantial research efforts have been devoted to this fundamental problem since the 1960s [6]. Conventional methods for solving OPF include, linear approximations, local search algorithms, particle swarm optimization, fuzzy logic (see [7]–[9] and the references therein). However, the existing methods do not offer guaranteed recovery of globally optimal solutions or even feasible points [10].

One of the most promising approaches to OPF is semidefinite programming (SDP) relaxation, which is proven to be exact for a variety of benchmark instances [11]. In general, the solution of SDP relaxation offers a lower bound for the unknown globally optimal cost of OPF. In order to address the inexactness of SDP relaxation for challenging instances of OPF (e.g., [2], [12], [13]), further investigation and improvement are carried out in [14]–[21]. Since inexact convex relaxations may not lead to physically meaningful solutions for OPF, alternative strategies are proposed to infer OPF feasible points from inexact convex relaxations. For instance, branch-and-bound algorithms [22], [23] iteratively partition search spaces to find tighter relaxations. In [19], [24], [25], penalty terms are incorporated into the objective of convex relaxations in order ensure OPF feasibility. Moment relaxation algorithms [26]–[28] form hierarchies of SDP relaxations to obtain globally optimal solutions for OPF. Most recently, [29] proposes a sequential convex optimization method with the aim of recovering OPF feasible points.

In addition to the exactness issues, SDP relaxation suffers from high computational cost due to the presence of high-order conic constraints. This shortcoming limits the applicability of SDP relaxation especially for large-scale instances of the OPF problem. To overcome this issue and enhance the scalability of SDP relaxation, some studies propose computationally-cheaper relaxations including second-order cone programming (SOCP) [30], [31], quadratic programming (QP) [32], [33], linear programming (LP) [34], [35]. Some papers have leveraged the sparsity of power networks to decompose large-scale conic constraints into lower order ones [24], [36]–[39]. Additionally, several extensions of OPF have been recently studied under more general settings, to address considerations such as the security of operation [24], [40], [41], robustness [42], energy storage [33], distributed platforms [43], [44], and uncertainty of generation [45].

In this paper, we introduce a novel and computationally-efficient parabolic relaxation and investigate its relation with the common practice SDP and SOCP relaxations. The proposed parabolic relaxation relies on convex quadratic inequalities only, as opposed to conic constraints. A penalization method is introduced for finding feasible and near-globally optimal solutions, which is compatible with the SDP, SOCP, and parabolic relaxations. We offer theoretical guarantees for the recovery of feasible solutions for OPF using penalization.

A. Notations

Throughout this paper, matrices, vectors, and scalars are represented by bold uppercase, bold lowercase, and italic lowercase letters, respectively. The symbols $\mathbb{R}$, $\mathbb{C}$, and $\mathbb{H}_n$ denote the sets of real numbers, complex numbers, and $n \times n$ Hermitian matrices, respectively. The notation “i” is reserved for the imaginary unit. $\| \cdot \|$ represents the absolute value of a scalar or the cardinality of a set, depending on the context.
The symbols (·)* and (·)ᵀ represent the conjugate transpose and transpose operators, respectively. The notations `Iₙ` and `Oₘ×ₙ` refer to the n × n identity and m × n zero matrices, respectively. Given an n × 1 vector x, the notation [x] refers to the n × n diagonal matrix with the elements of x on the diagonal. The symbols λₘₐₓ(·) and λₘᵢₙ(·) denote the minimum and maximum eigenvalues, respectively. Given a matrix A, the notation A_ij refers to its (j,k) entry. A ≥ 0 means that A is symmetric/Hermitean positive semidefinite. Define A(∥) as the sub-matrix of A obtained by choosing the rows that belong to the index set ∥.

II. Problem Formulation

A power network can be modeled as a directed graph G = (V, E), with V and E as the set of buses and lines, respectively. For each bus k ∈ V, the demand forecast is denoted by d_k ∈ C, whose real and imaginary parts account for active and reactive demands, respectively. Define v_k ∈ C as the complex voltage at bus k. Let G be the set of generating units, each located at one of the buses. For each generating unit g ∈ G, the values p_g and q_g, respectively, denote the amount of active and reactive powers. The unit incidence matrix C ∈ {0, 1}^|G|×|V| is defined as a binary matrix whose (g, k) entry is equal to one, if and only if the generating unit g belongs to bus k. Additionally, define the pair of matrices C, C ∈ {0, 1}^|E|×|V| as the from and to incidence matrices, respectively. The (l, k) entry of Ĉ is equal to one, if and only if the line l ∈ E starts at bus k, while the (l, k) entry of Ĉ equals one, if and only if line l ends at bus k. Define Y ∈ C^|V|×|V| as the nodal admittance matrix of the network and Ŷ = ĈC ∈ C^|E|×|V| as the from and to branch admittance matrices. Define y_sh = g_sh + i b_sh ∈ C|V| as the vector of shunt admittances whose real and imaginary parts correspond to the shunt conductances and susceptances, respectively. The OPF problem can be formulated as,

\[
\begin{align*}
\text{minimize} & \quad h(p) \\
\text{subject to} & \quad \mathbf{d} + \text{diag}\{v v^* Y^*\} = C^T (p + i q) \\
& \quad \text{diag}\{\bar{C} v v^* \bar{Y}^*\} = \bar{s} \\
& \quad v^2 \leq \bar{v}^2 \\
& \quad p_{\text{min}} \leq p \leq p_{\text{max}} \\
& \quad q_{\text{min}} \leq q \leq q_{\text{max}} \\
& \quad |s|^2 \leq \bar{s}^2 \\
& \quad |\bar{s}|^2 \leq \bar{\bar{s}}^2
\end{align*}
\]

where h(p) = c_0^T 1 + c_1^T p + p^T c_2 p is the objective function, c_0, c_1, c_2 ∈ R^|V| are the vectors of fixed, linear, and quadratic cost coefficients, respectively. Constraint (lb) is the power balance equation, which accounts for conservation of energy at all buses of the network. Constraint (le) ensures that voltage magnitudes remain within pre-specified ranges, given by vectors v_{min}, v_{max} ∈ R^|V|. Power generation vectors are bounded by p_{min}, p_{max} ∈ R^|G| for active power and q_{min}, q_{max} ∈ R^|G| for reactive power. The flow of power entering the lines of the network from their starting and ending buses are denoted by s ∈ C^|G| and \bar{s} ∈ C^|G|, respectively, and upper bounded by the vector of thermal limits \bar{s}_{\text{max}} ∈ R^|E|.

III. Preliminaries and Sensitivity Analysis

Consider the standard π-model of line l = (f, t) ∈ E, with series admittance y_{str}, l = y_{str}, l + b_{str}, l and total shunt susceptance b_{prl}, l, in series with a phase shifting transformer whose tap ratio has magnitude τ_l and phase shift angle θ_{l} [3]. The model is shown in Figure 1

Define \( p_l = \bar{p}_{str}, l \), \( \bar{q}_l = \bar{q}_{str}, l + \frac{b_{prl}, l}{2} |v_l|^2 \), \( \bar{p}_l = \bar{p}_{str}, l \), \( \bar{q}_l = \bar{q}_{str}, l + \frac{b_{prl}, l}{2} |v_l|^2 \), \( \bar{p}_{str}, l + i \bar{q}_{str}, l \) and \( \bar{p}_{str}, l + i \bar{q}_{str}, l \) are complex powers passing through the series element from the two ends. Additionally,

\[
\begin{align*}
\bar{p}_{str}, l &= v_l^* \bar{Y}_l v_l, \\
\bar{q}_{str}, l &= v_l^* \bar{Y}_l q_l v_l,
\end{align*}
\]

where, \( \bar{Y}_l \), \( \bar{Y}_l q_l \), \( \bar{Y}_l p_l \), and \( \bar{Y}_l q_l \) are given as

\[
\begin{align*}
\bar{Y}_l p_l &\equiv
\begin{bmatrix}
\begin{bmatrix}
g_{str}, f & e^{i \theta_{l} y_{str}, f} & -2i \tau_l \\
-e^{-i \theta_{l} y_{str}, f} & y_{str}, f & -2i \tau_l e^{i \theta_{l} y_{str}, f} \\
0 & -2i \tau_l e^{-i \theta_{l} y_{str}, f} & y_{str}, t
\end{bmatrix}
\end{bmatrix}
\begin{bmatrix}
\begin{bmatrix}
b_{str}, f & e^{i \theta_{l} b_{str}, f} & -2i \tau_l \\
-e^{-i \theta_{l} b_{str}, f} & b_{str}, f & -2i \tau_l e^{i \theta_{l} b_{str}, f} \\
0 & -2i \tau_l e^{-i \theta_{l} b_{str}, f} & b_{str}, t
\end{bmatrix}
\end{bmatrix}
\begin{bmatrix}
g_{str}, f & e^{i \theta_{l} g_{str}, f} & -2i \tau_l \\
-e^{-i \theta_{l} g_{str}, f} & g_{str}, f & -2i \tau_l e^{i \theta_{l} g_{str}, f} \\
0 & -2i \tau_l e^{-i \theta_{l} g_{str}, f} & g_{str}, t
\end{bmatrix}
\end{bmatrix}
\end{align*}
\]

The next definition introduces the notion of sensitivity measure for power systems, which will be used later in the paper.

**Definition 1:** The sensitivity measure of the power system under study is defined as

\[
P \equiv 2|N| + 2|C| + \|y_{sh}\|_2 + \sum_{l \in E} \left( \frac{|b_{prl}, l|}{2} + \frac{|b_{prl}, l|}{2} \right)
\]

To derive the optimality conditions of the problem (13 - 11) we define the Jacobian of equality and inequality constraints.

**Definition 2:** For every arbitrary point \( x = (v, p + i q, \bar{s}, \bar{\bar{s}}) \in C^{|V|} \times C^{|G|} \times C^{|E|} \times C^{|E|} \), the Jacobian of equality constraints (1b) - (1d) is equal to \( J^L = \text{real}(J^L) \), where

\[
J^L = \begin{bmatrix}
2|g_{sh}|v & -2|g_{sh}|v & -C^T & 0 & C^T & 0 & C^T \\
-2|b_{sh}|v & 2|b_{sh}|v & 0 & -C^T & 0 & C^T & 0 & \bar{C}^T \\
2 U_f & -2 U_f & 0 & 0 & -I_{|G|} & 0 & 0 & 0 & 0 \\
2 U_t & -2 U_t & 0 & 0 & -I_{|E|} & 0 & 0 & 0 & 0 \\
2 U_1 & 2 U_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
2 U_1 & -2 U_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
2 U_1 & -2 i \bar{U}_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
2 i \bar{U}_1 & -2 i \bar{U}_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
2 i \bar{U}_1 & 2 i \bar{U}_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\]
and matrices $\tilde{U}_1$, $\tilde{U}_2$, $\hat{U}_1$, and $\hat{U}_2$ are defined as
\[
\tilde{U}_1 \triangleq \frac{1}{2}(v^*C^*\bar{Y} + [\bar{Y}v]C), \quad \tilde{U}_2 \triangleq \frac{1}{2}(v^*C^*\bar{Y} - [\bar{Y}v]C), \\
\hat{U}_1 \triangleq \frac{1}{2}(v^*C^*\bar{Y} + [\bar{Y}v]C), \quad \hat{U}_2 \triangleq \frac{1}{2}(v^*C^*\bar{Y} - [\bar{Y}v]C).
\]

Moreover, the Jacobian of inequality constraints (1e) – (1i) are, respectively, given as
\[
\begin{align*}
J_1^1 & \triangleq 2\text{real}[[v] \ -i[v] \ O_{|V| \times (2|G|+4|E|))}], \quad (5a) \\
J_2^1 & \triangleq [0_{|G| \times (2|V|)} \ I_{|G|} \ O_{|G| \times (2|G|+4|E|))}], \quad (5b) \\
J_3^1 & \triangleq [0_{|G| \times (2|V|+|G|)} \ I_{|G|} \ O_{|G| \times (4|E|))}], \quad (5c) \\
\tilde{J}_1^1 & \triangleq 2\text{real}[[0_{|E| \times (2|V|+2|G|)} \ [s] - i[s] \ 0_{|E| \times (2|E|))}], \quad (5d) \\
\tilde{J}_1^1 & \triangleq 2\text{real}[[0_{|E| \times (2|V|+2|G|)} \ [s] - i[s]]]. \quad (5e)
\end{align*}
\]

Given a feasible solution and its Jacobian, the well-known linear independence constraint qualification (LICQ) condition is used to characterize well-behaved feasible points.

**Definition 3 (LICQ):** Consider a feasible point $(v, p + iq, \hat{s}, \hat{s}) \in C^{|V|} \times C^{|G|} \times C^{|E|} \times C^{|E|}$ for the problem (1a) – (1l). The point $(v, p + iq, \hat{s}, \hat{s})$ is said to satisfy the LICQ condition if the gradient vectors of equality constraints (1b) – (1d) and those inequality constraints (1e) – (1l) that are active form a linearly independent set. In other words, the LICQ condition holds, if the matrix
\[
\begin{align*}
J_{B_1^1, B_2^1, B_3^1, B_4^1, B_5^1, B_6^1, B_7^1, B_8^1} & \triangleq [J^1_1 \ J^1_2 \ J^1_3 \ J^1_4] \\
J_{B_2^2, B_3^2, B_4^2, B_5^2, B_6^2, B_7^2, B_8^2} & \triangleq [J^2_1 \ J^2_2 \ J^2_3 \ J^2_4] \\
J_{B_3^3, B_4^3, B_5^3, B_6^3, B_7^3, B_8^3} & \triangleq [J^3_1 \ J^3_2 \ J^3_3 \ J^3_4] \\
J_{B_4^4, B_5^4, B_6^4, B_7^4, B_8^4} & \triangleq [J^4_1 \ J^4_2 \ J^4_3 \ J^4_4]
\end{align*}
\]
is full row rank, where
\[
\begin{align*}
B_1^1 & \triangleq \{k \in V \ | \ |v_k| = v_{\min, k}\}, \quad B_2^1 = \{k \in V \ | \ |v_k| = v_{\max, k}\}, \\
B_2^2 & \triangleq \{g \in G \ | \ p_g = p_{\min, g}\}, \quad B_2^2 = \{g \in G \ | \ p_g = p_{\max, g}\}, \\
B_3^1 & \triangleq \{q \in G \ | \ q_{\min, q}\}, \quad B_3^2 = \{q \in G \ | \ q_{\max, q}\}, \\
B_4^1 & \triangleq \{t \in E \ | \ |s_t| = f_{\max, t}\}, \quad B_4^2 = \{t \in E \ | \ |s_t| = f_{\max, t}\}.
\end{align*}
\]

Since the LICQ condition is only defined for feasible points, in this paper, we introduce a generalization of the LICQ condition that is applicable to infeasible points as well. To this end, we first need a measure for the distance between an arbitrary point $x_{0} = (v_{0}, p_{0} + iq_{0}, \hat{s}_{0}, \hat{s}_{0}) \in C^{|V|} \times C^{|G|} \times C^{|E|} \times C^{|E|}$ and the feasible set of the OPF problem (1a) – (1l).

**Definition 4 (Feasibility distance):** Define the OPF feasible set $F \subset C^{|V|} \times C^{|G|} \times C^{|E|} \times C^{|E|}$ as the set of all $x = (v, p + iq, \hat{s}, \hat{s})$ that satisfy (1b) – (1l). Moreover, for every arbitrary $x_{0} = (v_{0}, s_{0}, \hat{s}_{0}, \hat{s}_{0}) \in C^{|V|} \times C^{|G|} \times C^{|E|} \times C^{|E|}$, define the feasibility distance $\delta_{M}(x_{0})$ as
\[
\min_{x \in F} \{\|v-v_{0}\|_{M} + \|p+iq-s_{0}\|_{2} + \|\hat{s}-\hat{s}_{0}\|_{2} + \|\hat{s}-\hat{s}_{0}\|_{2}\}.
\]

where $M \in H_{|V|}$ is arbitrary.

**Definition 5 (Generalized LICQ):** The point $x = (v, p + iq, \hat{s}, \hat{s}) \in C^{|V|} \times C^{|G|} \times C^{|E|} \times C^{|E|}$ is said to satisfy the Generalized LICQ condition if the matrix $J_{B_1^1, B_2^1, B_3^1, B_4^1, B_2^2, B_3^2, B_4^2, B_5^1, B_6^1, B_7^1, B_8^1}(x)$ from the equation (6) is full row rank, where
\[
\begin{align*}
B_1^1 & = \{k \in V \ | \ |v_k|^{2} + v_{\min, k}^{2} + \delta_{M}(x)^{2} + 2\delta_{M}(x)|v_k| \geq 0\}, \\
B_1^2 & = \{k \in V \ | \ |v_k|^{2} + v_{\max, k}^{2} + \delta_{M}(x)^{2} + 2\delta_{M}(x)|v_k| \geq 0\}, \\
B_2^1 & = \{g \in G \ | \ -p_{g} + p_{\min, g} + \delta_{M}(x) \geq 0\}, \\
B_2^2 & = \{g \in G \ | \ -p_{g} + p_{\max, g} + \delta_{M}(x) \geq 0\}, \\
B_3^1 & = \{q \in G \ | \ q_{\min, q} + \delta_{M}(x) \geq 0\}, \\
B_3^2 & = \{q \in G \ | \ q_{\max, q} + \delta_{M}(x) \geq 0\}, \\
B_4^1 & = \{t \in E \ | \ |s_t|^{2} - f_{\max, t}^{2} + \delta_{M}(x)^{2} + 2\delta_{M}(x)|s_t| \geq 0\}, \\
B_4^2 & = \{t \in E \ | \ |s_t|^{2} - f_{\max, t}^{2} + \delta_{M}(x)^{2} + 2\delta_{M}(x)|s_t| \geq 0\}.
\end{align*}
\]

Additionally, for every $x$ that satisfies the Generalized LICQ condition, define $\sigma(x)$ as the minimum singular value of the matrix $J_{B_1^1, B_2^1, B_3^1, B_4^1, B_2^2, B_3^2, B_4^2, B_5^1, B_6^1, B_7^1, B_8^1}(x)$.

IV. Convexification of the OPF Problem

The proposed relaxation of the OPF problem involves three steps that are detailed in this section.

A. Lifting

The nonlinear constraints (1b) – (1c) and (1b) – (1l), as well as the objective function (1a) can be cast linearly by lifting the problem to a higher dimensional space. To this end, define the auxiliary variables $o, r \in R^{G}$ and $f, \bar{f} \in R^{E}$ accounting for $p^2$, $q^2$, $|s|^2$, and $|\hat{s}|^2$, respectively. Moreover, define the auxiliary matrix variable $W \in H_{|V|}$, accounting for $vv^*$. Observe that the constraints (1b) – (1c) can be cast linearly with respect to $W \in H_{|V|}$. To preserve the relation between the original and lifted formulations, the following additional constraint shall be imposed:
\[
W = vv^*.
\] (7)

The non-convexity of the lifted formulation is captured by the above constraint, which is addressed next.
B. Convex Relaxation

In order to make the OPF problem computationally tractable, it is common practice to relax the non-convex constraint (7) to
\[
W - v v^* \in \mathcal{C}, \tag{8}
\]
where \(\mathcal{C}\) is a proper convex cone. In what follows, we discuss two commonly-used conic relaxations, as well as a novel relaxation which transforms the OPF problem (13) -- (14) into a convex quadratically-constrained quadratic program.

1) SDP Relaxation: To derive a semidefinite programming (SDP) relaxation for the OPF problem (13) -- (14), we can use the cone of \(|V| \times |V|\) Hermitian positive semidefinite matrices:
\[
\mathcal{C}_1 \triangleq \{ H \in \mathbb{H}_{|V|} \mid H \succeq 0 \}.
\]
Unlike the original non-convex problem (13) -- (14), SDP relaxation is convex and is proven to result in a globally optimal solution for several benchmark cases of OPF [11]. Despite the advantages of SDP relaxation, imposing a high-dimensional conic constraint can be computationally challenging. For sparse QCQP problems, the complexity of solving SDP relaxation can be alleviated through a graph-theoretic analysis, namely tree decomposition [24], [36] -- [39], [46]. Using a simple greedy algorithm [24], \(V\) can be decomposed into several overlapping subsets \(A_1, A_2, \ldots, A_D \subseteq V\), and then the relaxation is formulated in terms of the reduced cone:

\[
\mathcal{C}_1^D \triangleq \{ H \in \mathbb{H}_{|V|} \mid H\{A_k, A_k\} \succeq 0, \forall k \in \{1, 2, \ldots, D\} \},
\]
where for each \(k\), \(H\{A_k, A_k\}\) represents the \(|A_k| \times |A_k|\) principal sub-matrix of \(H\) whose rows and columns are chosen from \(A_k\). The above decomposition leads to an equivalent but more tractable formulation of SDP relaxation. Nevertheless, solving large-scale instances of OPF on real-world systems can be still computationally challenging.

2) SOCP Relaxation: A computationally cheaper alternative to SDP relaxation is the second-order cone programming (SOCP) relaxation, which is formulated using the cone
\[
\mathcal{C}_2 \triangleq \{ H \in \mathbb{H}_{|V|} \mid H_{ii} \geq 0, H_{ii}H_{jj} \geq |H_{ij}|^2, \forall (i, j) \in \mathcal{E} \}.
\]
Incorporating \(\mathcal{C}_2\) into the constraint (8) leads to the SOCP relaxation of OPF. Note that although the SDP relaxation is generally tighter, the SOCP relaxation is far more scalable.

3) Parabolic Relaxation: In order to avoid conic constraints, in this paper, we propose a computationally efficient method, regarded as the parabolic relaxation, which transforms an arbitrary non-convex QCQP into a convex QCQP. The proposed method requires far less computational effort and can serve as an alternative to the common practice SDP and SOCP relaxations for solving large-scale OPF problems. To derive the parabolic relaxation, define:
\[
\mathcal{C}_3 \triangleq \{ H \in \mathbb{H}_{|V|} \mid H_{ii} \geq 0, H_{ii} + H_{jj} \geq 2 |\text{real}\{H_{ij}\}|, H_{ii} + H_{jj} \geq 2 |\text{imag}\{H_{ij}\}|, \forall (i, j) \in \mathcal{E} \}.
\]
If \(\mathcal{C}_3\) is used, the constraint (8) transforms to the following convex quadratic inequalities
\[
\begin{align}
|v_i - v_j|^2 &\leq W_{ii} + W_{jj} - (W_{ij} + W_{ji}) \quad \forall (i, j) \in \mathcal{E} \quad (9a) \\
|v_i + v_j|^2 &\leq W_{ii} + W_{jj} + (W_{ij} + W_{ji}) \quad \forall (i, j) \in \mathcal{E} \quad (9b) \\
|v_i - iv_j|^2 &\leq W_{ii} + W_{jj} + i(W_{ij} - W_{ji}) \quad \forall (i, j) \in \mathcal{E} \quad (9c) \\
|v_i + iv_j|^2 &\leq W_{ii} + W_{jj} - i(W_{ij} - W_{ji}) \quad \forall (i, j) \in \mathcal{E} \quad (9d) \\
|v_i|^2 &\leq W_{ii} \quad \forall i \in V \quad (9e)
\end{align}
\]
and there is no need to impose conic constraints.

Definition 6: For every \(k \in \{1, 2, 3\}\), define \(\mathcal{D}_k\) as the dual cone of \(\mathcal{C}_k\). Observe that the cone of Hermitian positive semidefinite matrices is self-dual, i.e., \(\mathcal{D}_1 = \mathcal{C}_1\). Moreover, \(\mathcal{D}_2\) and \(\mathcal{D}_3\) are, respectively, the sets of \(|V| \times |V|\) Hermitian scaled-diagonally-dominant (SDD) and diagonally-dominant matrices, defined as,
\[
\begin{align}
\mathcal{D}_2 &= \left\{ \sum_{(i, j) \in \mathcal{E}} [e_i, e_j] H_{ij} [e_i, e_j]^\dagger \mid H_{ij} \in \mathbb{H}_2, H_{ij} \succeq 0, \forall (i, j) \in \mathcal{E} \right\}, \\
\mathcal{D}_3 &= \left\{ H \in \mathbb{H}_{|V|} \mid |H_{ii}| \geq \sum_{j \in V \setminus \{i\}} |H_{ij}|, \forall i \in V \right\},
\end{align}
\]
where \([e_i]_{i \in |V|}\) represents the standard basis for \(\mathbb{R}^{|V|}\). Moreover, the interior of \(\mathcal{D}_k\) can be expressed as
\[
\text{int}\{\mathcal{D}_k\} = \left\{ M \in \mathbb{H}_{|V|} \mid \exists \varepsilon > 0; M - \varepsilon I_{|V|} \in \mathcal{D}_k \right\}, \quad (11)
\]
for every \(k \in \{1, 2, 3\}\).
In practice, the aforementioned convex relaxations are not necessarily exact, which means that solutions obtained by solving the relaxed problems may not be feasible for the OPF problem (13) -- (14). Next, we show that it is possible to resolve this issue and obtain near-optimal feasible points for OPF by incorporating a penalty term into the objective function of SDP, SOCP, and parabolic relaxations.

C. Penalization

To address the inexactness of convex relaxations, we revise objective functions by adding linear penalty terms of the form \(\kappa(W, o, r, \bar{f}, \bar{f}, v, p + iq, \bar{s}, \bar{s})\), using which the non-convex constraint (7) is implicitly imposed. Given an initial guess \(x_0 = (v_{i0}, p_{i0} + iq_{i0}, s_{i0}, \bar{s}_{i0})\) for the solution of the OPF problem (13) -- (14), the following definition introduces a family of penalty terms that guarantee the exactness of relaxation if \(x_0\) is sufficiently close to the set \(\mathcal{F}\).

Definition 7: Given an arbitrary initial point \(x_0 = (v_{i0}, p_{i0} + iq_{i0}, s_{i0}, \bar{s}_{i0}) \in \mathbb{C}^{|V|} \times \mathbb{C}^{|V|} \times \mathbb{C}^{|V|} \times \mathbb{C}^{|V|}\), the penalty function \(\kappa_{M, x_0}\) is defined as follows
\[
\kappa_{M, x_0}(W, o, r, \bar{f}, \bar{f}, v, p + iq, \bar{s}, \bar{s}) \triangleq \langle o^\dagger 1 - 2 p_{i0}^\dagger p + p_{i0}^\dagger p_{i0} + (r^\dagger 1 - 2 q_{i0}^\dagger q + q_{i0}^\dagger q_{i0}) + \\
(\bar{f}^\dagger 1 - \bar{s}_{i0}^\dagger \bar{s} - \bar{s}^\dagger \bar{s}_{i0} + \bar{s}_{i0}^\dagger \bar{s}_{i0}) + (\bar{f}^\dagger 1 - \bar{s}_{i0}^\dagger \bar{s} - \bar{s}^\dagger \bar{s}_{i0} + \bar{s}_{i0}^\dagger \bar{s}_{i0}) + \\
\text{tr}\{\text{WM}\} - v_0^\dagger Mv - v^\dagger Mv + v_0^\dagger Mv_0, \quad (12)
\]
where \(M \in \mathbb{H}_{|V|}\) is regarded as the penalty matrix.
Given $k \in \{1, 2, 3\}$, a penalty matrix $M$, and $\mu > 0$, the penalized relaxation problem equipped with the cone $C_k$ and the penalty term $\mu \times \kappa M_{\mathbf{x}_0}$ can be formulated as:

$$\begin{align*}
\text{minimize} & \quad h_1(\mathbf{o}, \mathbf{p}) + \mu \kappa M_{\mathbf{x}_0} \langle \mathbf{W}, \mathbf{o}, \mathbf{r}, \tilde{\mathbf{f}}, \mathbf{v}, \mathbf{p} + \mathbf{i}q, \mathbf{s}, \mathbf{\tilde{s}} \rangle \tag{13a} \\
\text{subject to} & \quad \mathbf{d} + \text{diag}(\mathbf{W}^* \mathbf{y}^*) = \mathbf{C}^T (\mathbf{p} + \mathbf{i}q) \tag{13b} \\
& \quad \text{diag}(\mathbf{C} \mathbf{W}^* \mathbf{y}^*) = \mathbf{s} \tag{13c} \\
& \quad \mathbf{v}_{\min}^2 \leq \text{diag}(\mathbf{W}) \leq \mathbf{v}_{\max}^2 \tag{13d} \\
& \quad \mathbf{p}_{\min} \leq \mathbf{p} \leq \mathbf{p}_{\max} \tag{13e} \\
& \quad |\mathbf{q}|^2 \leq \tilde{\mathbf{f}} \leq f_{\max}^2 \tag{13f} \\
& \quad |\tilde{\mathbf{s}}|^2 \leq \tilde{\mathbf{f}} \leq f_{\max}^2 \tag{13g} \\
& \quad \mathbf{p}^2 \leq \mathbf{o} \tag{13j} \\
& \quad \mathbf{W} - \mathbf{vv}^* \in C_k \tag{13k}
\end{align*}$$

where $h_1(\mathbf{o}, \mathbf{p}) \triangleq \mathbf{c}^T \mathbf{1} + \mathbf{c}_1^T \mathbf{p} + \mathbf{c}_2^T \mathbf{o}$ is the lifted objective function. The penalization is said to be tight if the problem (13a)–(13k) possesses a unique solution that satisfies the equation (7). The tightness of penalization guarantees the recovery of a feasible point for the OPF problem (1a)–(1i).

V. Theoretical Results

It is shown in [47] that the LICQ condition holds generally for OPF. According to the next theorem, if $\mathbf{x}_0$ is a feasible point for the OPF problem (1a)–(1i) that satisfies the LICQ condition, then the penalized convex relaxation problem (13a)–(13k) preserves the feasibility of $\mathbf{x}_0$ for appropriate choices of the penalty matrix $M$ and $\mu$.

**Theorem 1:** Let $\mathbf{x}_0 = (\mathbf{v}_0, \mathbf{p}_0 + \mathbf{i}q, \mathbf{s}_0, \mathbf{\tilde{s}}_0) \in \mathcal{F}$ be a feasible point for the OPF problem (1a)–(1i), which satisfies the LICQ condition. Assume that $M \in \text{int}\{D_k\}$, where $k \in \{1, 2, 3\}$. If $\mu$ is sufficiently large, then the penalized convex relaxation (13a)–(13k), equipped with the cone $C_k$ and the penalty term $\mu \times \kappa M_{\mathbf{x}_0}$, has a unique solution

$$(\mathbf{W}_{\text{opt}}, \mathbf{o}_{\text{opt}}, \mathbf{r}_{\text{opt}}, \tilde{\mathbf{f}}_{\text{opt}}, \mathbf{v}_{\text{opt}}, \mathbf{p}_{\text{opt}}, \mathbf{q}_{\text{opt}}, \mathbf{s}_{\text{opt}}, \mathbf{\tilde{s}}_{\text{opt}}),$$

such that $\mathbf{x}_{\text{opt}} \triangleq (\mathbf{v}_{\text{opt}}, \mathbf{p}_{\text{opt}} + \mathbf{i}q, \mathbf{s}_{\text{opt}}, \mathbf{\tilde{s}}_{\text{opt}})$ is feasible for the original OPF problem (1a)–(1i) and $h(\mathbf{p}_{\text{opt}}) \leq h(\mathbf{p}_0)$.

**Proof:** The theorem is proven in [48] for the more general case of optimization problems with bilinear matrix inequality (BMI) constraints. The proof for penalized SDP and SOCP relaxations of QCPs is given in [49].

Obtaining a feasible point for OPF may not be straightforward. The next theorem is concerned with the case where the initial point $\mathbf{x}_0$ is not feasible.

**Theorem 2:** Consider an arbitrary point $\mathbf{x}_0 = (\mathbf{v}_0, \mathbf{p}_0 + \mathbf{i}q, \mathbf{s}_0, \mathbf{\tilde{s}}_0) \in C^{|V|} \times C^{(|V|)} \times C^{(|E|)} \times C^{(|E|)}$, which satisfies the Generalized LICQ condition. Assume that $\mathbf{M} - I_{|V|} \in \text{int}\{D_k\}$ and

$$\mathbf{M} \leq \left(\frac{\sigma(\mathbf{x}_0)}{4\delta_M(\mathbf{x}_0)}\right) I_{|V|}, \tag{14}$$

where $k \in \{1, 2, 3\}$, and $P$, $\delta_M(\mathbf{x}_0)$ and $\sigma(\mathbf{x}_0)$ are given by Definitions 1, 4 and 5, respectively. If $\mu$ is sufficiently large, then the penalized relaxation problem (13a)–(13k), equipped with the cone $C_k$ and the penalty term $\mu \times \kappa M_{\mathbf{x}_0}$ has a unique solution

$$(\mathbf{W}_{\text{opt}}, \mathbf{o}_{\text{opt}}, \mathbf{r}_{\text{opt}}, \tilde{\mathbf{f}}_{\text{opt}}, \mathbf{v}_{\text{opt}}, \mathbf{p}_{\text{opt}}, \mathbf{q}_{\text{opt}}, \mathbf{s}_{\text{opt}}, \mathbf{\tilde{s}}_{\text{opt}}),$$

such that $\mathbf{x}_{\text{opt}} \triangleq (\mathbf{v}_{\text{opt}}, \mathbf{p}_{\text{opt}} + \mathbf{i}q, \mathbf{s}_{\text{opt}}, \mathbf{\tilde{s}}_{\text{opt}})$ is feasible for the original OPF problem (1a)–(1i).

**Proof:** The proof can be found in [48].

A. Choice of the Penalty Matrix

Motivated by the previous literature [19], [24], we propose to choose $M$ such that the term $\text{tr}\{\mathbf{W} M\}$ in the penalty function reduces the apparent power loss over the series element of every line in the network. According to (13a)–(13b), the apparent power loss over the series admittance $y_s$ can be expressed in terms of $\mathbf{v}_l$ and the admittance matrices $\mathbf{Y}_{p,l}, \mathbf{Y}_{q,l}, \mathbf{Y}_{q,l}^\dagger$, and $\mathbf{Y}_{p,l}$. Hence, in order to penalize the apparent power loss over all lines of the network, we choose the matrix $M$ as,

$$\mathbf{M} = \sum_{(ij) \in \mathcal{E}} [\mathbf{e}_i, \mathbf{e}_j] (\mathbf{M}_{ij} + \alpha \mathbf{I}_2)[\mathbf{e}_i, \mathbf{e}_j]^T, \tag{15}$$

where $e_1, \ldots, e_n$ denote the standard basis for $\mathbb{R}^n$, $\alpha$ is a positive constant and each $\mathbf{M}_{ij}$ is a $2 \times 2$ positive semidefinite matrix defined as,

$$\mathbf{M}_{ij} = \zeta_{ij} (\mathbf{Y}_{q,l} + \mathbf{\bar{Y}}_{q,l}) + \frac{\eta}{1 - \eta} (\mathbf{Y}_{p,l} + \mathbf{\bar{Y}}_{p,l}). \tag{16}$$

The parameter $\eta > 0$ sets the trade-off between active and reactive loss minimization and $\zeta_{ij} \in \{-1, +1\}$ is determined based on the inductive or capacitive behavior of the line $l \in \mathcal{E}$. More precisely, we set $\zeta_{ij} = 1$ if the series admittance $y_{as,l}$ is inductive (i.e., $b_{as,l} \leq 0$), and $\zeta_{ij} = -1$, otherwise. Observe that if $\alpha$ is sufficiently large, then $M$ belongs to the relative interior of the dual cones $D_1$, $D_2$, and $D_3$.

B. Sequential Convex Relaxation

The penalized convex relaxation (13a)–(13k) can be solved sequentially to find near-globally optimal solutions for OPF. The details of this sequential procedure are delineated by Algorithm 1. According to Theorem 1, once a feasible point for the OPF problem (1a)–(1i) is obtained, feasibility is preserved, and the objective value improves in each round.

VI. Experimental Results

In this section, we detail our experiments for verifying the efficacy of the proposed methods. We consider the IEEE and European test cases from MATPOWER [3], modified-IEEE test cases from [2], and test cases from the NESTA v0.7.0 archive [1]. All numerical experiments are performed in MATLAB using a 64-bit computer with an Intel 3.0 GHz, 12-core CPU, and 256 GB RAM. The CVX package version 3.0, SDPT3 version 4.0, and MOSEK version 8.0 are used for convex optimization.
Algorithm 1 Sequential Penalized Convex Relaxation.

Input: \( k \in \{1,2,3\}, \ M \in \text{int}\{D_k\}, \ \mu > 0, \) and \( x_0 = (v_0, p_0 + i q_0, \tilde{s}_o, \tilde{s}_i) \in \mathbb{C}^{|V|} \times \mathbb{C}^{|G|} \times \mathbb{C}^{|F|} \times \mathbb{C}^{|E|} \)

1: repeat
2: Obtain \( x_{\text{opt}} = (v_{\text{opt}}, p_{\text{opt}} + i q_{\text{opt}}, \tilde{s}_{\text{opt}}, \tilde{s}_{\text{opt}}) \) by solving the optimization (13a) – (13k), equipped with the cone \( C_k \) and the penalty term \( \mu \times \delta_M x_0 \).
3: \( x_0 \leftarrow x_{\text{opt}} \).
4: until stopping criteria is met.

Output: \( x_0 = (v_0, p_0 + i q_0, \tilde{s}_o, \tilde{s}_i) \)

### Table I: The lower bounds and run times of parabolic relaxation compared to the SDP and SOCP relaxations.

| Test Cases | SDP | SOCP | Parabolic |
|------------|-----|------|-----------|
|            | LB  | Time | LB  | Time | LB  | Time |
| 9          | 5296.67 | 1.14 | 5296.67 | 0.72 | 5216.03 | 0.66 |
| 14         | 8091.53 | 0.81 | 8095.12 | 0.71 | 7642.59 | 0.59 |
| 2869pegase | 5180.89 | 1.12 | 5173.88 | 0.69 | 506.31 | 0.04 |
| 39          | 41826.10 | 0.36 | 41845.05 | 0.24 | 41061.04 | 1.90 |
| 77          | 41737.79 | 1.38 | 41711.01 | 0.92 | 41006.74 | 0.90 |
| 118         | 129654.63 | 2.53 | 129341.96 | 1.68 | 125947.88 | 1.14 |
| 300         | 719711.69 | 6.56 | 718548.29 | 5.83 | 708314.84 | 2.64 |
| 898pegase   | 58190.04 | 2.91 | 58190.04 | 2.91 | 53905.86 | 1.15 |
| 1359pegase  | 74662.53 | 577.57 | 74521.39 | 439.98 | 73027.96 | 9.95 |
| 2869pegase  | 133988.93 | 267.37 | 133880.03 | 223.32 | 132381.10 | 24.91 |

Table [I] reports the optimal objective values for SDP, SOCP, and parabolic relaxations to the choice of penalty parameter \( \mu \). The results are shown in Figure [2] for the benchmark case, the best-known feasible cost is equal to 17551.89 [1]. The minimum values of \( \mu \) that offers tight penalization and its resulting percentage gap with the best-known cost value are, respectively, equal to 213.60 and 0.08% for SDP relaxation, 1288.88 and 0.29% SOCP relaxation, and 6628.91 and 1.02% for parabolic relaxation. For this experiment, the parameters \( \alpha \) and \( \eta \) in the equations (15) and (16) are set to 5 and 0, respectively. According to Figure [2], all of the proposed penalized convex relaxations result in near-globally optimal points for a wide range of \( \mu \) values. As shown by the figure at the bottom, a smaller choice of \( \mu \) leads to a lower objective values. The smallest value of \( \mu \), which produces a feasible solution for OPF is greater for parabolic relaxation compared to that of SDP and SOCP relaxations.

As our third experiment, we evaluate the performance of the proposed sequential scheme for solving OPF on several benchmark systems. The numerical results are reported in Tables [II] and [III]. For all of the test cases, we initialized Algorithm [I] with flat start, i.e., \( v_0 = 1, \ p_0 = p_{\text{min}}, \ q_0 = 0, \ \tilde{s}_o = \text{diag}\{\mathbf{C}_v V^* \ Y^*\}, \) and \( \tilde{s}_i = \text{diag}\{\mathbf{C}_v p V^* \ Y^*\}. \)

Note that the GFB% and GPB% values reported in Table [II] are calculated according to the best upper bounds provided by [1]–[3].
TABLE II: Result summary for several benchmark systems.

| Test Cases | SDP | SOCP | Parabolic | $c_f$ |
|------------|-----|------|-----------|------|
| 118        |     |      |           |      |
| 300        |     |      |           |      |
| 89pegase   |     |      |           |      |
| 135pegase  |     |      |           |      |
| 9mod       |     |      |           |      |
| 39mod1     |     |      |           |      |
| 39mod2     |     |      |           |      |
| 39mod3     |     |      |           |      |
| 39mod4     |     |      |           |      |
| 118mod     |     |      |           |      |
| 900mod     |     |      |           |      |
| nestagl    |     |      |           |      |
| nestal_gcr |     |      |           |      |
| nestal_30 |     |      |           |      |
| nestal_39 |     |      |           |      |
| nestal_gpr |     |      |           |      |
| nestal_gpee |     |      |           |      |
| nestal_30  |     |      |           |      |
| nestal_39  |     |      |           |      |
| nestal_gpr |     |      |           |      |
| nestal_gpee |     |      |           |      |
| test_cases |     |      |           |      |

For all of the test cases reported in Table II, Algorithm I equipped with any of the SDP, SOCP, and parabolic relaxations yields fully feasible points within the first few rounds. As shown in Table III, Algorithm I produces feasible points within 0.2% gap from the best reported solutions for benchmark systems case9mod and case39mod1, which is an improvement upon the existing penalization methods [26].

To verify the scalability of the proposed method, we conduct experiments on the largest available benchmark instances from [3] and [1]. The results are reported in Table III, Algorithm I equipped with SOCP relaxation finds fully feasible solutions that are not more than 0.4% away from the upper bounds obtained by solving OPF using MATPOWER.

VII. Conclusion

This paper is concerned with the AC optimal power flow (OPF) problem. We first consider two common practice semidefinite programming (SDP) and second order cone programming (SOCP) relaxations of OPF. Due to the computational complexity of conic optimization, we propose an efficient alternative, called parabolic relaxation, which transforms arbitrary non-convex quadratically constrained quadratic programs (QCQPs) to convex QCQPs. Additionally, we propose a novel penalization method which is guaranteed to provide feasible points for the original non-convex OPF, under certain assumptions. By applying the proposed penalized convex relaxations sequentially, we obtained fully feasible points with promising global optimality gaps for several challenging benchmark instances of OPF.

REFERENCES

[1] C. Coffrin, D. Gordon, and P. Scott, “NESTA, the Nicta Energy System Test Case Archive (2014),” arXiv preprint arXiv:1411.0839.
[2] W. A. Bukhsh, A. Grothey, K. I. McKinnon, and P. A. Trodden, “Local solutions of the optimal power flow problem,” IEEE Trans. Power Syst., vol. 28, no. 4, pp. 4780–4788, 2013.
[3] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, “MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education,” IEEE Trans. Power Syst., vol. 26, no. 1, pp. 12–19, 2011.
[4] K. Lehmann, A. Grastien, and P. Van Hentenryck, “AC-feasibility on tree networks is NP-hard,” IEEE Trans. Power Syst., vol. 31, no. 1, pp. 798–801, 2016.
[5] A. Verma, Power grid security analysis: An optimization approach. Columbia University, 2010.
[6] J.arpenter, “Contribution to the economic dispatch problem,” Bull. Soc. Françoise Elect., vol. 3, no. 8, pp. 431–447, 1962.
[7] K. Pandya and S. Joshi, “A survey of optimal power flow methods,” J. Theoret. Appl. Inf. Technol., vol. 4, no. 5, 2008.
[8] J. A. Momoh, R. Adapa, and M. El-Hawary, “A review of selected optimal power flow literature to 1993. I. nonlinear and quadratic programming approaches,” IEEE Trans. Power Syst., vol. 14, no. 1, pp. 96–104, 1999.
[9] J. A. Momoh, M. El-Hawary, and R. Adapa, “A review of selected optimal power flow literature to 1993. II. newton, linear programming and interior point methods,” IEEE Trans. Power Syst, vol. 14, no. 1, pp. 105–111, 1999.
