Distributed Multi-Area Optimal Power Flow via Rotated Coordinate Descent Critical Region Exploration

Haitian Liu, Ye Guo, Hongbin Sun, Weisi Deng, Yifei Xu

Abstract—We consider the problem of distributed optimal power flow (OPF) for multi-area electric power systems. A novel coordination scheme using a two-layered structure is proposed, referred to as the rotated coordinate descent critical region exploration (RCDCRE). RCDCRE can reduce coordination, keep privacy and ensure convergence by stitching coordinate descent and parametric programming using coordinate system rotation. Each entity residing in the lower level can independently update its boundary information and optimally solve its local OPF in an asynchronous fashion. The coordinator in the upper level is only activated to provide a rotation matrix when reaching a coordinate-wise optimal point. The solution process does not require warm starts and can iterate from infeasible initial points using penalty-based formulations. The effectiveness of RCDCRE is verified based on IEEE 2-area 44-bus and 4-area 472-bus systems.

I. INTRODUCTION

Decarbonization is a common consensus to hedge against global warming, which pushes the need to increase the installed capacity and utilization level of renewable energy resources (RESs). As the spatial distribution of RESs is unbalanced, some areas cannot satisfy local demand unless a substantial amount of power is imported from neighboring areas. In contrast, some other areas lack enough flexibility to accommodate the RESs, leading to significant curtailment. Due to privacy and computation capability concerns, the modern vast interconnected power system is jointly operated by multiple utilities. Each entity is responsible for optimizing the dispatch of local RESs according to the interface power exchange. Hence, optimal tie-line scheduling is crucial for decarbonization. Moreover, it also contributes to the system’s overall efficiency and provides resilience to an area under extreme weather conditions.

As gathering all data at a central location may not be the best, a distributed computational paradigm is advocated to coordinate multi-area systems. The utilities collectively solve a distributed OPF to obtain the interchange schemes by optimizing internal assets and exchanging intermediate variables. Current distributed methods tailored for the joint power system operation are mainly derived from dual decomposition. See [1], [2] for a comprehensive review. Dual decomposition enjoys a near block-separable structure by adopting the (augmented) Lagrangian relaxation (LR) techniques. The distributed multi-area OPF can be solved by iteratively updating the local dispatch orders and coupling constraints’ dual multipliers. The dual update also makes economic sense, which indicates the price signals of net power exchange and the shadow prices of tie-line capacity.

However, since coupling constraints are relaxed, all intermediate solutions are not feasible until asymptotic convergence. Dual methods are generally first-order methods that optimize over a non-smooth dual function. They may suffer from slow convergence in large multi-area power systems. In practice, an infeasible dispatch may be returned as the final solution due to the limited computation time. Infeasible dispatch is less desirable for two reasons, 1) The economic performance is degraded if multiple automatic generation control units are activated to make up for the power mismatch. 2) Or even worse, the returned solutions cannot dispatch, causing security issues. Besides, the LR decomposition has strict assumptions on the problem structures to ensure feasible primal solutions can be obtained in the limit [3]. Divergent cases may occur when applying the alternating direction method of multipliers (ADMM) to a system more than 2-area [4].

Meanwhile, primal decomposition has also made progress in solving the multi-area OPF problem. Typical methods includes marginal equivalent decomposition (MED) [5], column and constraint generation (C&CG) [6], and critical region exploration (CRE) [7], [8]. MED identifies the marginal variables to exploit the finite structure of active constraint sets. C&CG improves the convergence of Benders decomposition by simultaneously submitting cutting planes and binding variables. CRE leverages multi-parameter programming (mP) theory, finding optima with the finite search of critical regions. Empirically, primal decomposition methods have better feasibility guarantees and convergence over the duals, as there is no constraint relaxed. However, some methods may exchange actual costs and network information (e.g., MED, C&CG), suffer from the strong coordination with synchronization requirements (e.g., C&C, CRE), or unable to dispatch when current boundary states are infeasible to the local problems (e.g., MED, CRE).

To date, solving the large-scale distributed multi-area OPF problem is still a challenging problem, both from the primal and dual decomposition. This paper proposed a rotated coordinate descent critical region exploration (RCDCRE) algorithm in terms of primal decomposition and multi-parameter programming theory. Two merits make RCDCRE tailored to multi-area OPF. 1) Compared to existing primal decompositions, RCDCRE preserves privacy by exchanging parametric mappings and reduces coordination by separat-
ing boundary jurisdiction with an asynchronous update. 2) Compared to existing dual decompositions, the proposed method can obtain feasible intermediate solutions quickly by parameterizing the penalized intra-regional OPF. Moreover, it also enjoys a faster convergence as more useful information is exchanged. Hence, RCDCRE is suitable for the solution process with an infeasible cold start or online implementation with fixed iteration requirements. Numerical results are encouraging and show convergence superiority over ADMM and Benders method.

II. PROBLEM SETUP

A. Multi-Area OPF formulation

Each area operates a transmission network in a multi-area system connected via tie-lines. The entire system is three-phase balanced and can be well approximated by a linear DC power flow (DCPF) model [9, Eq. (6.40)-(6.41)]. With a DCPF model, the linear map between real power injections and phase angles is captured by a suitable defined $B$-matrix simplified from the bus admittance matrix. Line flows are also linear to the phase angles. The relation is based on matrix $H$ simplified from the system branch admittance matrix $[10]$. An $N$-area OPF can be formulated as

$$\begin{align*}
\text{minimize} & \quad \sum_{i=1}^{N} \left( \frac{1}{2} g_i^T Q_i g_i + c_i^T g_i \right), \\
\text{subject to} & \quad B_i \delta_i + B_{i,i} \delta_i = g_i - d_i, \\
& \quad B_{i,j} \delta_i + B_{i,i} \delta_i + \sum_{j \in \text{nbd}(i)} B_{i,j} \delta_j = -d_j, \\
& \quad -f_i \leq H_i \delta_i + H_{i,i} \delta_i \leq f_i, \\
& \quad -f_i \leq H_i \delta_i + \sum_{j \in \text{nbd}(i)} H_{i,j} \delta_j \leq f_i, \\
& \quad \delta_i \leq \delta_i \leq \delta_i, \\
& \quad -\pi \leq \delta_i \leq \pi, \\
& \quad -\pi \leq \delta_i \leq \pi, \\
& \quad \delta_i^{\text{ref}} = 0, \quad \forall i = 1, \ldots, N, \\
\end{align*}$$

where, as shown in Fig. 1, decision variables include area $i$’s power generation $g_i$, internal and boundary phase angles $\delta_i, \delta_i$. Both loads $d_i$ and $d_j$ are constants.

![Fig. 1. An illustration for multi-area power system.](image)

The objective (1a) is to minimize the sum of area’s generation costs with coefficients $Q_i, c_i$. The DC model’s nodal power balance is divided into (1b)-(1c). Notation nbd($i$) collects all adjacent areas of area $i$. The subscripts $i, j$ and $j$ reflect bus partitions. Constraints (1d)-(1e) restrict internal and tie-line branch flow less than $f_i$ and $f_j$. Here, we slightly abuse the notations to let the row indices $i, i$ be the internal and tie-line of area $i$. And column indices $i, j$ of $H$ are still the bus partitions. The lower and upper bounds of generation capacities $g_i, g_i, \delta_i, \delta_i$ internal, and boundary phase angle limits are summarized in (1f)-(1h). Constraint (1i) artificially assigns a reference phase angle. To simplify notations, let $\theta_i = [\theta_i^T, \delta_i^T]^T, \delta_i = \delta_i$ and $\theta := [\theta_1^T, \ldots, \theta_N^T]^T$. Then we write (1b), (1c), (1d), (1f), (1g) as (2b), while (1e), (1h) become instances of (2c).

Problem (1) is converted into a compact form (2) as

$$\begin{align*}
\text{minimize} & \quad \sum_{i=1}^{N} \left( \frac{1}{2} x_i^T H_i x_i + f_i^T x_i \right), \\
\text{subject to} & \quad A_i x_i \leq b_i + C_i \theta, \quad \forall i = 1, \ldots, N, \\
& \quad \Theta_i := \{D_i \theta \leq r_i\}, \quad \forall i = 1, \ldots, N, \\
\end{align*}$$

B. Preliminaries on CRE

The compact form (2) can be interpreted as a primal decomposition formulation where $x_i, i = 1, \ldots, N$ are local variables, $\theta$ is the coupling variable. The architecture of the primal decomposition method is shown in Fig. 2, where the local and coordination problems are iteratively solved. Note that a central coordinator is needed, and it has jurisdiction over the coupling variable $\theta$.

![Fig. 2. The architecture of CRE. Red / green arrows indicate upward / downward communication links, and blue arrows are tie-lines among areas.](image)

CRE is also based on the structure above. By leveraging the multi-parametric programming theory, problem (2) can be seen as a multi-parametric linear or quadratic programming (mp-LP/QP). The local problem $i$ parametric in $\theta$ is

$$\begin{align*}
J^*_i(\theta) := \min_{x_i} & \quad \frac{1}{2} x_i^T H_i x_i + f_i^T x_i, \\
\text{subject to} & \quad A_i x_i \leq b_i + C_i \theta, \\
\end{align*}$$

The properties of the value function $J^*_i(\theta) = \sum_{i=1}^{N} J^*_i(\theta)$ are summarized in the following lemma.

**Lemma 1:** (cf. [11]): Consider the mp-LP/QP (3). The set of feasible parameters $\Theta^*$ is a polyhedral set, i.e.,

$$\Theta^* := \cap_{i=1}^{N} \Theta^*_i, \quad \Theta^*_i := \{\theta \mid \exists x_i : A_i x_i \leq b_i + C_i \theta\}.$$

The value function $J^*_i(\theta) : \Theta^* \to \mathbb{R}$ is convex piecewise linear / quadratic on a group of polyhedral critical regions (CR), i.e., $\Theta^* = \cap_{i=1}^{N} \cup_{r_i \in K_i} \text{CR}^*_r \cup K_i$, where $K_i$ is constant.

Under Lemma 1 and Fig. 2, CRE solves problem (2) by applying the following two steps recursively.

7187
1) **Local evaluation:** Each area solves (3) with \( \theta^k \) to get \( J^*_{i,k} \), a slice of value function \( J^*_{i} \) defined over \( CR^*_i \).

2) **Coordination update:** Coordinator gathers all \( J^*_{i,k}, CR^*_i \) to solve the coordination problem

\[
\theta^* := \{ \arg\min_{\theta} \sum_{i=1}^{N} J^*_i | \cap_{i=1}^{N} CR^*_{i,k}, \cap_{i=1}^{N} \Theta_i \}, \quad (4)
\]

and update \( \theta^* \) with a projected subgradient search

\[
\theta^{k+1} := \theta^* - \varepsilon \text{stepsize} \ n, \quad (5)
\]

\[
v := \{ \arg\min_v \|v\|^2 \ | v = \partial J^* \eta + N^\star \zeta, \ 1^\top \eta = 1 \}. \quad (6)
\]

Here, \( \varepsilon \text{stepsize} \) is constant, \( \partial J^* \), \( N^\star \) are the subdifferential and normal cone at \( \theta^* \) with \( \Theta = \cap_{i=1}^{N} \Theta_i \), \( I \) is an all one vector with proper dimensions.

**Lemma 2:** (cf. [7]): CRE obtains the optimal \( \theta^* \) after finite iterations when \( v \) in (6) satisfies \( \|v\| = 0 \).

### III. PROPOSED RCDCRE METHOD

#### A. Motivation and solution architecture

The coordinator in CRE gathers all local information and directly controls each area’s boundary. In practice, system operators prefer weaker coordination and update asynchronously. The proposed RCDCRE allows each area to reclaim jurisdiction of its boundary. As shown in Fig. 3, the coordinator’s jurisdiction is now intangible. Hence, we can deploy the coordination layer to any utility without privacy issues. Besides, recovering feasibility is a nontrivial task for all decomposition methods. The proposed RCDCRE method adopts penalty-based formulations to simultaneously take feasibility and optimality into account. Fast feasible recovery and convergence are achieved by implicitly exchanging the parametric mappings of the penalized intra-regional OPF.

**Fig. 3.** The architecture of RCDCRE when area \( i,i=1 \) is updating.

Two main procedures of RCDCRE are highlighted in Fig. 3, *i.e.*, block coordinate descent (BCD) and coordinate system rotation (CSR). The BCD is to realize asynchronization with a blockwise update to \( \theta \). The CSR is a coordination step to ensure the sequence of \( \theta \) generated by BCD converge to optimal. We use area 1’s update as an example to show the asynchronization. Area 1 itself conducts two steps similar to the CRE’s local evaluation and coordination update. The critical regions \( CR^\star_{j,s} \) and value functions \( J^\star_{j,s}(\theta) \) for the given parameter \( \theta^* \) from all areas are collected to area 1. Compared to CRE, two differences occur when it comes to updating \( \theta \). One is that the update in (4) is restricted to be blockwise, and the other is that a coordinatewise search with a small step size replaces the subgradient search in (5) to explore adjacent critical regions. In such a way, the asynchronous update within area 1 is realized, while all other areas’ jurisdictions will not be affected. By switching from one area to another under some pre-defined sequence (*e.g.*, cyclic), different blocks of \( \theta \) are updated, and a stationary point \( \theta^* \) will be attained. However, the convergence of \( \theta^* \) to optimal with merely BCD updates must rely on two additional assumptions to the coordination problem (4).

**Assumption 1:** The objective function with respect to all block coordinates is differentiable, or it is not differentiable but the non-differentiable part is \( N \) block-separable [12].

**Assumption 2:** The feasible set is the direct sum of \( N \) separable real Hilbert spaces of the block coordinates [13].

According to Lemma 1, for mp-LP/QP, \( J^* \) is a piecewise function, which is neither differentiable nor separable. The feasible set \( \Theta^\star \) is a polyhedral set that is also non-separable. Hence, the boundary phase angles \( \theta^* \) after BCD updates may get stuck at non-optimal solutions. This paper proposes a perspective that we can apply a matrix \( R \) to rotate the coordinate system of \( \theta \) with a certain angle such that the coordinatewise direction in the rotated system can avoid stucking. A direct way to find such an angle with high possibilities is to rotate a subgradient into a coordinatewise basis, serving as a rotation reference. Eq. (7) formalizes the above ideas as a CSR step performed by a central coordinator. Note that such CSR step (7) is only activated when \( \theta^* \) reaches stationary and has not fulfilled the optimality condition.

\[
\bar{v} = Rv, \quad \forall k : \bar{v} = e_k, \quad 1 \leq k \leq \dim(\theta). \quad (7)
\]

Here, we use a subgradient \( v \) obtained from the solution to problem (6). According to Lemma 2, \( \theta^* \) is non-optimal when \( \|v\| = 0 \). The notation \( \bar{v} \) indicates any coordinatewise direction, where \( e_k \) is the \( k^{th} \) standard basis for the parameters \( \theta \in \mathbb{R}^d \). The matrix \( R \) is what we want for conducting CSR, which can be obtained from compounding Givens matrices [14, Sec. 5.1.8]. The coordinator gathers the necessary information based on (6) to calculate \( v \) and generates \( R \) using Eq. (7). Then it send \( R \) back to each area. The \( \theta \)’s coordinate system can be rotated accordingly.

The recovery of infeasible parameters comprises two steps. We first adopt a big\( M \)-penalty based formulation to each area \( i \)’s intra-regional OPF problem (3).

\[
J^i_\star(\theta) := \min_{\bar{x}_i} \bar{x}_i^\top H_i \bar{x}_i + f_i^\top \bar{x}_i, \quad \text{subject to } \bar{A}_i \bar{x}_i \leq \bar{b}_i + \bar{C}_i \theta, \quad (8a)
\]

where the variables and coefficients are given by

\[
\bar{x}_i = \begin{bmatrix} x_i \\ s_i \end{bmatrix}, \quad \bar{H}_i = \begin{bmatrix} H_i & 0 \\ 0 & 0 \end{bmatrix}, \quad \bar{f}_i = \begin{bmatrix} f_i \\ M1 \end{bmatrix}, \\
\bar{A}_i = \begin{bmatrix} A_i & -I \\ 0 & -I \end{bmatrix}, \quad \bar{b}_i = \begin{bmatrix} b_i \\ 0 \end{bmatrix}, \quad \bar{C}_i = \begin{bmatrix} C_i \\ 0 \end{bmatrix}.
\]
The introduction of slack variable $s_i$ ensures an always feasible solution to problem (8) for any given parameters $\theta$. Note that to reduce the number of $s_i$, it is recommended to remove the redundant constraints to the jointly feasible region of $x_i, \theta$ space before the reformulation.

To adjust the tie-line schedules that physically connected to area $i$, area $i$ solve an interregional OPF problem (9) to update its boundary phase angles. Problem (9) can be viewed as problem (4) with $\ell_1$-penalization to the $\cap_{i=1}^N \Theta_i$ and fixed block coordinates to other area’s boundary state.

The introduction of slack variable $s_i$ ensures an always feasible solution to problem (8) for any given parameters $\theta$. For the non-optimal $\theta^*$, CSR has the same updating rule as $\partial J^*_i$, and is omitted for brevity.

\begin{align*}
\text{minimize} & \quad \sum_{i=1}^N \mathcal{J}^*_i, s_i + \sigma I^T \max \{ D_i, \theta - r_i, 0 \}, \\
\text{subject to} & \quad \cap_{i=1}^N \text{CR}^*_i, \quad \theta_{-i} = \theta^*_{-i},
\end{align*}

where $\sigma$ is a penalty coefficient. The boundary phase angles that are not controlled by area $i$ are denoted by $\theta_{-i}$. By Lemma 1, the value function segment $\mathcal{J}^*_i$ and critical region CR$_i^*$ can be explicitly expressed as

\begin{align*}
\mathcal{J}^*_i & := \frac{1}{2} \hat{H}_i \hat{\theta} + \hat{f}_i \hat{\theta} + \hat{e}_{i,s}, \quad (10a) \\
\text{CR}^*_i & := \{ \theta \mid \hat{D}_i \theta \leq \hat{r}_i \}. \quad (10b)
\end{align*}

Under non-degenerate assumptions, (10) can be directly obtained from KKT conditions. See [7], [8] for details. When the primal or dual solutions are non-unique, (10) can still be obtained with additional techniques, such as (a) projecting from a higher dimensional solution space [15], [16], (b) solving auxiliary problems of the optimal set [17], [18], or (c) searching the combinational tree of active constraints [19].

The $\ell_1$-penalty can return a feasible solution outside the initial feasible region. Hence, it allows RCDCRE to take an arbitrary initial $\theta$ as a cold start. By exchanging parametric mappings of the present boundary state, each area updates its boundary phase angles locally in a BCD fashion by solving penalty-based intra-regional OPF (8) and interregional OPF (9). The central coordinator conducts CSR if all areas finish BCD updates and the optimal condition has not been met. The proposed RCDCRE converges to optimal by iteratively applying the above steps.

Compared to existing techniques, LR decomposition needs feasible recovery steps for general problems [20]. Large penalty factor decreases ADMM’s performance, yet the intermediate solution is still asymptotic feasible. Benders decomposition with feasible cuts cannot track varied system conditions, and the convergence is poor if it adopts a big $M$-penalty formulation [21]. Our method considers an exact parametric relation between penalty formulations and feasibility. Sec. IV shows that it can recover feasible solutions in a few iterations from infeasible boundary phase angles.

Algo. 1 shows the details of RCDCRE. The algorithm sets an initial value $\theta^*$. By default, $\theta^*$ is not the coordinatewise optimal point to all the areas, i.e., set $A_w = \{1, \ldots, N\}$. Set $A_o$ is the complement of $A_w$. The coordinates in area $i$ are saved in $E_{i,w}$. For example, suppose area $i$ controls $p$th and $q$th of the boundary phase angle $\theta$, then we have

\begin{equation}
E_{i,w} = \{ e_p, -e_p, e_q, -e_p \}, \quad (11)
\end{equation}

where, $e_p$ is the $p$th basis, $-e_p$ is its negative direction. Set $E_{i,o}$ is the complement of $E_{i,w}$ for area $i$.

Step 3 obtain the current optimizing area and boundary state. Steps 4-13 update the local boundary phase angles. The coordinatewise search is conducted in step 5. The new explored parameter is denoted as $\theta^*$. In steps 6-8, all areas solve the big $M$-penalty intra-regional OPF problem (8) to dispatch internal assets given the explored parameter $\theta^*$. Area $i$ then gathers the parametric mappings of the present boundary state. Step 9 solves the $\ell_1$-penalty interregional OPF problem with other areas’ boundary phase angles $\theta_{-i}$ are fixed. The current optimal solution is denoted as $\theta_{\text{imp}}$.

The subdifferential $\partial J^*_i$, normal cone $\mathcal{N}^*_i$, and explored coordinate directions are updated in steps 10-13 when area $i$ optimizes the $\theta_i$.

1Note $\mathcal{N}^*_i$ has the same updating rule as $\partial J^*_i$, and is omitted for brevity.
is conducted in steps 21-23. Otherwise, Algo. 1 terminates with the optimal global solutions.

**Theorem 1:** For the mp-LP/QP in (2), RCDCRE with the two-layered scheme in Algo. 1 converges to the optimal solutions in finite iterations and rotations.

**B. Illustrative example**

We show how CSR avoids the sticking of BCD. Consider problem (12) with an initial point $(-1, -1)$.

\[
\begin{align*}
\text{minimize} & \quad \theta_1^2 + \theta_2^2, \\
\text{subject to} & \quad \theta_1 + \theta_2 \leq -1.
\end{align*}
\]  

(12a)

(12b)

In the left of Fig. 4, $\theta_2$ does the coordinatewise minimization first, the optimal solution is $(-1, 0)$. However, $\theta_1$ cannot find any further improvement from $(-1, 0)$ to the global optima $(-0.5, -0.5)$, as (12b) violates Assumption 2.

![Fig. 4. An example of coordinate system before and after rotation.](image)

Notice at $(-1, 0)$, a descent direction $[\sqrt{2}/2, -\sqrt{2}/2]^{\top}$ can be induced by the set plus of normal cone and gradient. Hence, we can use Eq. (7) to rotate it into $[1, 0]^{\top}$, which is the standard basis of $\theta_1$. The transformation is given by

\[
\begin{bmatrix}
1 \\
0
\end{bmatrix} = \begin{bmatrix}
\sqrt{2}/2 \\
\sqrt{2}/2
\end{bmatrix} \begin{bmatrix}
\sqrt{2}/2 \\
-\sqrt{2}/2
\end{bmatrix},
\]

(13)

where the matrix in (13) is $R$ in Eq. (7). By substituting $\theta = R^{-1}\tilde{\theta}$ into problem (12), the original problem is equivalently reformed in the new coordinate system as follows.

\[
\begin{align*}
\text{minimize} & \quad \tilde{\theta}_1^2 + \tilde{\theta}_2^2, \\
\text{subject to} & \quad \tilde{\theta}_2 \leq -\frac{\sqrt{2}}{2}.
\end{align*}
\]

(12a)

(12b)

As shown in the right of Fig. 4, current descent direction becomes $[1, 0]^{\top}$, which can be solved by BCD readily.

**IV. CASE STUDIES**

We used MATLAB 2021a with CPLEX version 12.9.0 to conduct all the simulations. Network data were obtained from the MATPOWER 7.1 [22]. For the illustrative purpose, feasible dispatch region $\Theta = \cap_{i=1}^2 \Theta_i$ in the two-area 44-bus system was divided into critical regions using MPT 3.0 [23]. A four-area 472-bus network was designed for comparative analysis. Unless specified, a zero cold start is adopted for the initial boundary phase angles $\theta$.

A. **Simulation on a two-area 44-bus network**

Consider the two-area power system in Fig. 5 that stitched IEEE 14 and 30 systems together with 2 tie-lines as shown. All tie-line capacities were set as 10 MW, and internal line capacities were 100 MW. Linear cost coefficients $c_i$ of the generators were perturbed to $c_i := c_i \circ (0.99 + 0.02\xi_i)$, for $i = 1, 2$, where entries of $\xi_i$ are independent $\mathcal{N}(0, 1)$ (standard normal) variables. Simulation results that neglect quadratic generation costs are shown in Fig. 7a. The proposed RCDCRE method can find the exact optimal solutions after switching the coordinates four times. The feasible space $\Theta$ contains three critical regions, and only two of them need to be explored for the entire RCDCRE procedure. If we do not neglect the quadratic costs, then $\Theta$ comprises four critical regions. As shown in Fig. 7b, the problem stuck at a non-optimal point after switching coordinates two times. After the CSR, RCDCRE can find the optima by another two switchings, shown in Fig. 7c.

![Fig. 5. Two-area 44-bus network.](image)

![Fig. 6. Simulations with linear costs and start point $(-0.1, -0.1)$.](image)

By changing the start point to $(-0.1, -0.1)$, the trajectory of feasibility recovery is illustrated in Fig. 6. Here, area 1 bus 2 is denoted by $\theta_1$, and $\theta_2$ represents area 2 bus 21. We adopt a linear generation cost for illustration. Note feasible parameter space $\Theta^* = \mathbb{R}^2$, as any parameter $\theta$ has at least a feasible solution to (8). Hence, by lemma 2, the value function $J^*$ is piecewise affine in $\mathbb{R}^2$, as shown in the left of Fig. 6. The converging process roughly comprises three stages for this case. In the first stage, the intra-regional OPF is infeasible to (3) but feasible to (8). Hence, RCDCRE returns a large objective defined over that critical region, mainly caused by big-M-penalty. After one BCD update along $J^*$, boundary phase angles $\theta$ become feasible to (3) for all areas rapidly. The $\ell_1$-penalty to (9) maximizes its effect in the second stage. We can observe that $\ell_1$-penalty allows several critical regions generated outside $\Theta$, shown in the shaded area in the right of Fig. 6. Within switching coordinate two times, $\theta$ converge back to $\Theta$. The final stage, when $\theta \in \Theta$ and feasible to (3) is identical to the trajectory in Fig. 7a. Similar convergence also applies to the quadratic generation cost case.
B. Simulation on a four-area 472-bus network

We next show the results on a large four-area 472-bus network with 50 MW tie-line and 500 MW internal line capacities. The topology is depicted in Fig. 8. A quadratic generation cost was adopted, where the linear coefficients were perturbed similar to the two-area one.

RCDCRE is compared with three other approaches, i.e., CRE taken from [7], Consensus ADMM taken from [2, Algo. 2], and Generalized Benders decomposition modified from [24, Sec. 5.1]. The objective convergence results are shown in Fig. 9. Compared to the centralized solution, the CRE method converges to the exact solution after 26 iterations. RCDCRE takes 114 iterations converging to the $10^{-3}$ relative optimality gap. Meanwhile, it needs 477 and 589 iterations for ADMM primal objective and Benders upper bound converging to the same accuracy level. RCDCRE allows each area to update its boundary states rather than the coordinated update like CRE and Benders. Moreover, it enjoys a better convergence over ADMM and Benders decomposition due to adequate information exchanged. Also note albeit ADMM dual enjoys a similar convergence rate as RCDCRE, the convergence superiority of RCDCRE over ADMM primal implies a better feasibility guarantee. Such an advantage is nontrivial for power system operations.

V. CONCLUSIONS

We have proposed RCDCRE as a novel privacy-preserving and cold start coordination approach to solve the multi-area OPF. RCDCRE adopts parametric penalty-based formulations to achieve rapid feasibility recovery and convergence.

The coordinator in RCDCRE does not control and update the boundary phase angles synchronously as in the Benders and CRE method. Instead, it conducts coordinate system rotation to ensure convergence and allows for an asynchronous update of each region. RCDCRE shows finite convergent property on a two-area 44-bus network. Results on a four-area 472-bus network show that RCDCRE enjoys a faster convergence rate than the ADMM and Benders decomposition methods.

The structure and simulations results indicate that RCDCRE is favorable to the operation of interconnected power systems, and it also applies to generic multi-agent optimization problems. Future work will entail degeneracy handling, co-ordination reduction, and applying the proposed method to general OPF problems in power system operations.

REFERENCES

[1] D. K. Molzahn, F. Dorfler, H. Sandberg, S. H. Low, S. Chakrabarti, R. Baldick, and J. Lavaei, “A Survey of Distributed Optimization and Control Algorithms for Electric Power Systems,” IEEE Trans. Smart Grid, vol. 8, no. 6, pp. 2941–2962, Nov. 2017.

[2] A. Kargarian, J. Mohammadi, J. Guo, S. Chakrabarti, M. Barati, G. Hug, S. Kar, and R. Baldick, “Toward Distributed/Decentralized DC Optimal Power Flow Implementation in Future Electric Power Systems,” IEEE Trans. Smart Grid, vol. 9, no. 4, pp. 2574–2594, Jul. 2018.

[3] I. Necoara and V. Nedeleva, “On linear convergence of a distributed dual gradient algorithm for linearly constrained separable convex problems,” Automatica, vol. 55, pp. 209–216, May 2015.

[4] C. Chen, B. He, Y. Ye, and X. Yuan, “The direct extension of ADMM for multi-block convex minimization problems is not necessarily convergent,” Math. Program., vol. 155, no. 1-2, pp. 57–79, Jan. 2016.
F. Zhao, E. Litvinov, and T. Zheng, “A Marginal Equivalent Decomposition Method and Its Application to Multi-Area Optimal Power Flow Problems,” IEEE Trans. Power Syst., vol. 29, no. 1, pp. 53–61, Jan. 2014.

B. Zeng and L. Zhao, “Solving two-stage robust optimization problems using a column-and-constraint generation method,” Oper. Res. Lett., vol. 41, no. 5, pp. 457–461, Sep. 2013.

Y. Guo, S. Bose, and L. Tong, “On Robust Tie-Line Scheduling in Multi-Area Power Systems,” IEEE Trans. Power Syst., vol. 33, no. 4, pp. 4144–4154, Jul. 2018.

Y. Guo, L. Tong, W. Wu, B. Zhang, and H. Sun, “Coordinated Multi-Area Economic Dispatch via Critical Region Projection,” IEEE Trans. Power Syst., vol. 32, no. 5, pp. 3736–3746, Sep. 2017.

A. J. Wood, B. F. Wollenberg, and G. B. Sheble, Power Generation, Operation, and Control, 3rd ed. New York: Wiley-Interscience, 2013.

R. D. Zimmerman and C. E. Murillo-Sánchez, “MATPOWER User’s Manual,” 2020.

F. Borrelli, A. Bemporad, and M. Morari, Convex Optimization, Cambridge: Cambridge University Press, 2017.

S. Salzo and S. Villa, “Parallel random block-coordinate forward backward algorithm: A unified convergence analysis,” Math. Program., Apr. 2021.

G. H. Golub and C. F. Van Loan, Matrix Computations, 4th ed. Baltimore, MD: Johns Hopkins University Press, 2013.

C. N. Jones, E. C. Kerrigan, and J. M. Maciejowski, “On Polyhedral Projection and Parametric Programming,” J. Optim. Theory Appl., vol. 138, no. 2, pp. 207–220, Aug. 2008.

A. Akhori and P. I. Barton, “An Improved Multi-parametric Programming Algorithm for Flux Balance Analysis of Metabolic Networks,” J. Optim. Theory Appl., vol. 178, no. 2, pp. 502–537, Aug. 2018.

C. N. Jones, E. C. Kerrigan, and J. M. Maciejowski, “Lexicographic perturbation for multiparametric linear programming with applications to control,” Automatica, vol. 43, no. 10, pp. 1808–1816, Oct. 2007.

J. Spjtvold, P. Tønnesen, and T. A. Johansen, “Continuous Selection and Unique Polyhedral Representation of Solutions to Convex Parametric Quadratic Programs,” J. Optim. Theory Appl., vol. 134, no. 2, pp. 177–189, Aug. 2007.

A. Gupta, S. Bhartiya, and P. S. V. Nataraj, “A novel approach to multiparametric quadratic programming,” Automatica, vol. 47, no. 9, pp. 2112–2117, Sep. 2011.

E. Gustavsson, M. Patriksson, and A.-B. Strömberg, “Primal convergence from dual subgradient methods for convex optimization,” Math. Program., vol. 150, no. 2, pp. 365–390, May 2015.

S. Candás, K. Zhang, and T. Hamacher, “A Comparative Study of Benders Decomposition and ADMM for Decentralized Optimal Power Flow,” in Proc. 2020 IEEE PES TSIG, Washington, DC, USA, Feb. 2020, pp. 1–5.

R. Zimmerman, C. Murillo-Sánchez, and R. Thomas, “Matpower: Steady-state operations, planning, and analysis tools for power systems research and education,” IEEE Trans. Power Syst., vol. 26, pp. 12–19, 2011, https://doi.org/10.1109/tpwrs.2010.2054713.

M. Herceg, M. Kvasnica, C. N. Jones, and M. Morari, “Multi-Parametric Toolbox 3.0,” in Proc. Eur. Control Conf., Zürich, Switzerland, July 17–19 2013, pp. 502–510.

J. R. Birge and F. Louveaux, Introduction to Stochastic Programming, ser. Springer Series in Operations Research and Financial Engineering, New York, NY: Springer New York, 2011.

T. Larsson, M. Patriksson, and A.-B. Strömberg, “Conditional subgradient optimization: Theory and applications,” Eur. J. Oper. Res., vol. 88, no. 2, pp. 382–403, Jan. 1996.

S. P. Boyd and L. Vandenberghe, Convex optimization. Cambridge University Press, 2004.

APPENDIX

A. Proof of Theorem 1

To prove theorem 1, we first impose an assumption about the stepsize in step 5 of algorithm 1.

Assumption 3: Given $\varepsilon_{\text{stepsize}}$, $\theta^\star$ can identify but not step across an adjacent critical region if $\theta^\star$ is on the boundary of current critical region $\cap_{i=1}^{N} C_{\theta_i}^\star$. Otherwise, $\varepsilon_{\text{stepsize}}$ ensures that $\theta^\star$ does not step out of $\cap_{i=1}^{N} C_{\theta_i}^\star$.

If $\|\theta^\text{mp} - \theta^\star\| \geq \varepsilon_{\text{stepsize}}$ such a stepsize can ensure $J(\theta^\text{mp}) < J(\theta^\star)$. Hence, the $J(\theta^\star)$ during the solution process is always non-increasing.

We proof Theorem 1 in three steps.

1) Without penalty formulation, RCDCRE converges when first-order non-differentiable optimality condition is fulfilled for feasible parameters $\theta \in \Theta^\star$.

2) The bigM-penalty formulation can return the same optimal solution when initial $\theta$ is infeasible to (3).

3) The $\ell_1$-penalty formulation returns the same optimal solutions when starting from $\theta \notin \Theta$.

1) Proof of 1): For block coordinate descent (BCD), it is well known that it converges to optimal in successive blockwise minimizations when assumptions 1-2 hold. However, when assumptions 1-2 fail and $\theta^\star$ gets stuck, the coordinate system rotation (CSR) conducts in RCDCRE is given by

$$v = R\theta^\star,$$  
$$\hat{\theta} = R^T \theta^\star,$$  

where $R$ is obtained from (7), the rotated parameters and coefficients are denoted in $\text{[\cdot]}$ values. Note $R^T = R^{-1}$ as $R$ is orthogonal. The subgradient direction $v$ is obtained from (6). The rotated subgradient direction is denoted by $\hat{v}$, and it is also a coordinate direction according to $R$. Hence, an exploration of $\theta^\star$ in step 5 of Algo. 1 is given by

$$\hat{\theta} = \theta^\star - \varepsilon_{\text{stepsize}} \hat{v}.$$  

Remark 1: The rotated subgradient direction $\hat{v}$ is also the subgradient in the rotated coordinate system.

Proof: According to (6), remark 1 implies the explored subdifferential and normal cones are also the subdifferential and normal cones in the rotated system. For each critical region with active and inactive set partitions $A, \bar{A}$, by KKT conditions we have

$$Hx + f + A^T_A \lambda_A = 0,$$  
$$A_A x = b_A + C_A \theta,$$  

where subscripts $i$ are dropped for simplicity. The $C_A \theta$ is replaced by $\bar{C}_A \theta$ because of (14). WLOG, we assume $H = 0$, the problem reduces to a mp-LP. If $A_A$ is full rank, then we have the primal mapping and value function as follows

$$x^\star(\theta) = A_A^{-1}(b_A + \bar{C}_A \theta),$$  
$$J^\star(\theta) = f^T A_A^{-1}(b_A + \bar{C}_A \theta).$$  

The gradient of current value function segment is

$$\nabla J^\star(\theta) = \bar{C}_A(A_A^{-1})^T f,$$  
$$= RC_A^T(A_A^{-1})^T f,$$  
$$= R\nabla J^\star(\theta),$$  

Similar argument also applies for mp-QP, degenerate conditions, as well as the rotated normal cones. By substituting $\partial J^\star(\theta), \mathcal{N}_\Theta^\star(\theta)$ into (6), we can obtain $\hat{v}$ is the optimal solution to (6), which completes the proof. ■
Under remark 1, the coordinatewise search is also a subgradient search, and we have
\[
\mathcal{J}(\theta^*) = \mathcal{J}(\hat{\theta}^* - \varepsilon_{\text{stepsize}} \hat{v}) 
\]

\[
= \sum_{i=1}^{N} \frac{1}{2} (\theta^* - \varepsilon_{\text{stepsize}} \hat{v})^T \hat{H}_{i,s} (\hat{\theta}^* - \varepsilon_{\text{stepsize}} \hat{v}) 
+ f_{i,s}^2 (\hat{\theta}^* - \varepsilon_{\text{stepsize}} \hat{v}) + c_{i,s} 
\]

\[
= \mathcal{J}(\hat{\theta}^*) + \frac{1}{2} (\varepsilon_{\text{stepsize}} \hat{v})^T \hat{H}_s (\varepsilon_{\text{stepsize}} \hat{v}) 
- \varepsilon_{\text{stepsize}} \hat{v}^T (H_s \hat{\theta}^* + f_s) 
\approx \mathcal{J}(\hat{\theta}^*) - \varepsilon_{\text{stepsize}} \hat{v}^T (H_s \hat{\theta}^* + f_s),
\]

where \( \hat{H}_s = \sum_{i=1}^{N} \hat{H}_{i,s} \) and \( f_s = \sum_{i=1}^{N} f_{i,s} \). When \( \varepsilon_{\text{stepsize}} \) is sufficiently small, the second term in (18c) can be cancelled. Note that (18d) comprises of two terms, \( \mathcal{J}(\hat{\theta}^*) \) indicates current optimal values. In the second term, \( \hat{H}, \hat{\theta}^* + f_s \) indicates the gradient of current critical region, \( \hat{v} \) is the subgradient induced by explored critical regions in (6).

Under the above notations, for every CSR conducted in RCDCRE, we have the following proposition

**Proposition 1:** (cf. [7, Appendix B]): In RCDCRE, suppose rotation matrix \( R \) generated from (7) has been applied to rotate the coordinate system. Then at least a new critical region of \( \hat{\theta}^* \) or a better parameter \( \theta^* \) can be found for block coordinate descent in the new coordinate system.

**Proof:** We first give a geometric interpretation of subgradient calculation in (6), shown in Fig. 10. In [25, Proposition 2.5], it is referred to as conditional steepest descent direction.

![Fig. 10. An illustration to (6).](image)

We let \( \hat{v} \) be solution to (6) with \( \partial f(x^*) \) and \( N_{\chi}^X \). Here, \( \partial f(x^*) \) is the convex hull of subgradients \( u_i, i = 1, \ldots, 4 \), and \( N_{\chi}^X \) is an empty set. As \( \hat{v} \) is the negative of the shortest conditional subgradient to the conditional subdifferential \( \partial f(x^*) + N_{\chi}^X \). Clearly, if the conditional subdifferential does not contain the origin, then \( \hat{v} \neq 0 \) and \( \hat{v}^T u > 0 \) for any \( u \in \partial f(x^*) + N_{\chi}^X \).

We distinguish two cases, if \( (\hat{H}, \hat{\theta}^* + f_s) \in \partial \mathcal{J}^* \), we have
\[
\hat{v}^T (\hat{H}, \hat{\theta}^* + f_s) > 0.
\]

Hence, by (18d), we have \( \mathcal{J}(\hat{\theta}^*) < \mathcal{J}(\hat{\theta}^*) \), which means \( \hat{v} \) is a descending direction. Now, suppose \( \hat{v}^T (\hat{H}, \hat{\theta}^* + f_s) \leq 0 \), implying \( \mathcal{J}(\hat{\theta}^*) > \mathcal{J}(\hat{\theta}^*) \). Although \( \hat{v} \) is not a descending direction, we can conclude that \( (\hat{H}, \hat{\theta}^* + f_s) \notin \partial \mathcal{J}^* \) and we have explored a new critical region at \( \hat{\theta}^* \). There are also cases that we can both find a better \( \theta^* \) and a new critical region, we omit the similar analysis for brevity.

**Corollary 1:** (cf. [7, Theorem 1]): RCDCRE terminates after finitely many steps, when \( \|v\| = 0 \).

**Proof:** As \( \partial \mathcal{J}^* \) is a subdifferential at \( \theta^* \), it is sufficient to prove optimality when \( v = 0 \), per first-order non-differentiable optimality condition \( 0 \in \partial \mathcal{J}^* + N_{\Theta}^2 \). We proof corollary 1 in two steps. Suppose \( \theta^* \) is non-optimal, then by proposition 1, a descending direction will be identified after finite CSRs as there are finite critical regions according to lemma 1. Hence, we can apply BCD as soon as a better \( \theta^* \) has been found. When \( \theta^* \) becomes the optimal solutions, also by proposition 1 and lemma 1, at least one new critical region will be identified for each CSR. Hence, RCDCRE converge in finite iterations.

2) Proof of 2): We first show that the optimal solutions are identical for (3) and (8) when \( \theta^* \) is feasible. Suppose the active set partition of (8) is given by \( \{A, \bar{A}, \{I, \bar{I}\} \} \), where \( A, I \) are the partition of (3), by KKT conditions we have

\[
H x + f + A^T \lambda_A = 0, \quad (19a)
\]

\[
M I + \lambda - \mu = 0, \quad (19b)
\]

\[
A_A x - s_A = b_A + C_A \theta, \quad (19c)
\]

\[
s_A = 0, \quad \mu_A = 0, \quad \lambda_A = 0, \quad (19d)
\]

\[
A_I x - s_I \leq b_I + C_I \theta, \quad (19e)
\]

\[
s_I \geq 0, \quad \mu_I \geq 0, \quad \lambda_I \geq 0, \quad (19f)
\]

Here, \( \lambda \) is the multiplies to the initial constraints (3b), \( \mu \) is the multiplies to the slack variables \( s \). As \( M \) is sufficiently large, \( s^* = 0 \) when \( \theta^* \) is feasible to (3), and \( \mu = M I \geq 0 \). Equation (19) reduces to

\[
H x + f + A^T \lambda_A = 0,
\]

\[
A_A x - s_A = b_A + C_A \theta, \quad \lambda_A \geq 0,
\]

\[
A_I x \leq b_I + C_I \theta, \quad \lambda_I = 0,
\]

which is exactly the KKT conditions of (3). Hence, the value function and critical regions of (3) and (8) are all the same for any feasible \( \theta^* \). When \( \theta^* \) is infeasible to (3), the problem can still be viewed as a feasible mp-LP/QP of (8). According to lemma 1, the value function is a continuous piecewise convex function, and there are finite critical region partitions. Clearly, the value function is continuous on the boundary of feasible and infeasible regions of (3). Hence, the effect of big\(M\)-penalty is merely to extend feasible space \( \Theta^* \). By corollary 1, we can conclude RCDCRE also converges to optimal solutions under big\(M\)-penalty formulation.

3) Proof of 3): The last concern of RCDCRE is the \( \ell_1 \)-penalty formulation to the convergence of \( \theta \) outside \( \Theta \), which is summarized in the following lemma.

**Lemma 3:** (cf. [26, 5.16]) Suppose \( \nu^* \) is the optimal Lagrange dual solutions of (9). If \( \sigma > \nu^* \), the any solution to \( \ell_1 \)-penalty problem \( (9) \) is also an optimal solution of the original problem.

So after several iterations when \( \Theta \cap \bigcup_{i=1}^{N} \mathcal{C}_{i,s} \neq \emptyset \), the algorithm can return the same optimal solutions. The above arguments combines to prove Theorem 1.