Convex Relaxation for Mixed-Integer Optimal Power Flow Problems

Chin-Yao Chang Sonia Martínez Jorge Cortés

Abstract— **Recent years have witnessed the success of employing convex relaxations of the AC optimal power flow (OPF) problem to find global or near-global optimal solutions. The majority of the effort has focused on solving problem formulations where variables live in continuous spaces. Our focus here is in the extension of these results to the co-optimization of network topology and the OPF problem. We employ binary variables to model topology reconfiguration in the standard semidefinite programming (SDP) formulation of the OPF problem. This makes the problem non-convex, not only because the variables are binary, but also because of the presence of bilinear products between the binary and other continuous variables. Our proposed convex relaxation to this problem incorporates the bilinear terms in a novel way that improves over the commonly used McCormick approximation. We also address the exponential complexity associated with the discrete variables by partitioning the network graph in a way that minimizes the impact on the optimal value of the relaxation. As a result, the problem is broken down into several parallel mixed-integer problems, reducing the overall computational complexity. Simulations in the IEEE 118-bus test case demonstrate that our approach converges to solutions which are very close to the lower bound of the mixed-integer OPF problem.**

I. INTRODUCTION

Mixed integer programming (MIP) appears in the context of optimal power flow (OPF) problems, where integer variables are introduced for switching control. Operators commonly establish switching control laws off-line to handle contingency situations and maintain the system stability. Such prescribed long-term decisions separate the OPF problem, that usually focus on economic dispatch or loss minimization, from topology design or other forms of discrete control. While the co-optimization of network topology and power dispatch may give rise to significant potential benefits, the complexity of such highly non-convex problems makes progress along this direction difficult.

Literature review: Transmission switching commonly serves as a corrective mechanism in response to system contingency, see [1], [2] and references therein for existing methods. In [3], [4], linearized OPF, also known as DCOPF, is used for the fast co-optimization of network topology and OPF. Despite its relative low complexity, DCOPF may lead, especially in congested systems, to poor solutions that can even lead to voltage collapse [5]–[7]. [8] proposes quadratic convex (QC) relaxations for the MIP-OPF problem, which provides more accurate results than DCOPF, while still retaining a fast computation time. Recent studies have found that semidefinite programming (SDP) [9]–[11] provides global or near global optimal solutions for many classes of OPF problems. The SDP-based convex relaxation of ACOPF is tighter than the ones derived from DCOPF and QC. However, the handling of discrete decision variables in the SDP-relaxed OPF problem is challenging and not well understood, while definitely worthy of further exploration given its potential impact. The challenges in the MIP-OPF problem not only stem from the use of discrete variables, but also from the presence of bilinear terms, which arise in two different ways. One is the product of the voltage of the terminal buses when computing the line power flow between the buses, which is naturally convexified in SDP. The other terms appear in the product of a discrete (binary) variable with a continuous decision variable, which has the physical interpretation of whether the control is active (line is connected) or not. The paper [12] uses lift-and-branch-and-bound procedure to deal with the SDP formulation of MIP-OPF, but the scalability of the approach is very limited. The work [13] also considers using SDP to solve the MIP-OPF, where the bilinear terms associated with line connections is partially addressed by assuming certain nominal network topology. For the bilinear terms of other discrete decision variables, [13] tackles them using the standard McCormick relaxation [16].

Statement of contributions: We formulate the co-optimization of topology design and OPF by defining a SDP convexified OPF problem with binary variables. Each binary variable corresponds to the decision of whether its associated line is connected or not, which naturally multiplies the voltage of the terminal nodes. Our contributions are twofold. First, we propose a new relaxation method to approximate the bilinear terms involving binary and continuous variables. Our method introduces a positive semidefinite (PSD) matrix for each switchable line that encodes the bilinear products and has the physical interpretation of virtual voltages at the terminal buses. Inspired by its physical meaning, we impose various linear constraints that lead to a convexification of the MIP-OPF problem, as all binary variables appear only in the bilinear products of the original problem. We show that the resulting optimal solution satisfies certain physically-meaningful inequalities regarding power losses and illustrate how the proposed relaxation can provide better approximations than the McCormick relaxation. Our second contribution is the introduction of a graph partitioning method to refine the solution obtained from the convex relaxation. Our method is based on the observation that the optimal dual variables correspond to the sensitivity of the optimal

C.-Y. Chang, Sonia Martínez, and Jorge Cortés are with the Department of Mechanical and Aerospace Engineering, University of California, San Diego, CA, USA. Email: {chc433,soniamd,cortes}@ucsd.edu

value with respect to perturbing the associated constraints. This leads us to define an undirected graph whose edge weights are the sum of the dual variables associated with the terminal nodes, which we partition using a minimum weight edge-cut set to minimize the impact on the value of the optimal solution of the relaxation. Our last step consists of solving the mixed-integer-SDP problem associated with each subnetwork in parallel, then reconstructing the solution of the original problem by combining the solutions of the sub-networks. Since each sub-network involves less integer variables than the original problem, it is amenable to standard integer programming techniques. Simulations on IEEE 118 bus test case demonstrate that the proposed method converges to a near optimum value with a tight difference from the lower bound. All proofs are omitted due to the space reasons and will appear elsewhere.

II. PRELIMINARIES

This section introduces basic notation and concepts from graph theory and optimization.

A. Notation

We denote by $\mathbb{N}$, $\mathbb{R}$, $\mathbb{R}_+$, and $\mathbb{C}$ the sets of positive integers, reals, positive real, and complex numbers, respectively. We denote by $|\mathcal{N}|$ the cardinality of the set $\mathcal{N}$. For a complex number $a \in \mathbb{C}$, we let $|a|$ and $\angle a$ be the complex modulus and angle of a. The real and imaginary parts of a are represented as $\Re(a)$ and $\Im(a)$. The 2-norm of a complex vector $v \in \mathbb{C}^n$ is written as $\|v\|$. Let $\mathbb{S}_+^n \subset \mathbb{C}^{n\times n}$ and $\mathcal{H}^n \subset \mathbb{S}_+^n$ be the set of positive semidefinite and n-dimensional Hermitian matrices, respectively. For $A \in \mathbb{C}^{n\times n}$, let A^* and $\mathbf{Tr}\{A\}$ denote its conjugate transpose and trace, respectively.

B. Graph Theory

We review basic notions of graph theory following [14]. A graph is a pair $\mathcal{G} = (\mathcal{N}, \mathcal{E})$, where $\mathcal{N} \subseteq \mathbb{N}$ is its set of vertices or nodes and $\mathcal{E} \subseteq \mathcal{N} \times \mathcal{N}$ is its set of edges. A *self-loop* is an edge that connects a vertex to itself. Two nodes $i, k \in \mathcal{N}$ are *connected* if $\{i, k\} \in \mathcal{E}$. The graph is *undirected* if $\{i, k\} = \{k, i\} \in \mathcal{E}$. A *path* in a graph is a sequence of vertices such that any two consecutive nodes correspond to an edge of the graph. An *orientation* of an undirected graph is an assignment of exactly one direction to each of its edges. A *simple graph* is a graph with neither self-loops nor multiple edges connecting any pair of two vertices. A *vertex-induced subgraph* of $\mathcal{G} = (\mathcal{N}, \mathcal{E})$, written as $\mathcal{G}_s[\mathcal{N}_s]$, is a subgraph of $\mathcal{G}$ with the set of nodes $\mathcal{N}_s \subseteq \mathcal{N}$ and set of edges $\mathcal{E}_s = \mathcal{E} \cap (\mathcal{N}_s \times \mathcal{N}_s)$. An *edge cut set* is a set of edges of the graph which, if removed, disconnects it. A *weighted graph* is a graph in which each branch $\{i, k\}$ is given a numerical weight, $w_{ik} \in \mathbb{R}_+$. Given the weights of all the edges, $w \in \mathbb{R}_+^{|\mathcal{E}|}$, the *weighted adjacency matrix* of a simple graph, A, has $A(i, k) = A(i, k) = w_{ik}$ and $A(i, k) = 0$ otherwise. A n*-optimal partition* of $\mathcal{G}(\mathcal{N}, \mathcal{E}, A)$ divides $\mathcal{N}$ into n number of disjoint sets such that $\cup_{i=1}^n \mathcal{V}_i = \mathcal{N}$ and at the same time, has $\sum_{\{i,k\}\in\mathcal{E}_c} w_{ik}$ minimized, where $\mathcal{E}_c$ is the edge cut set.

C. Strong Duality of Convex Optimization

Here, we review some fundamental concepts in convex optimization following [15]. Consider a convex optimization problem of the form

$$\min_x f_0(x), \quad \text{s.t. } Ax = b,\ f_i(x) \le 0,\ i = 1, \dots, m, \quad (1)$$

where $f_0, \dots, f_m : \mathbb{R}^n \to \mathbb{R}$ are convex functions, $A \in \mathbb{R}^{n\times r}$, $b \in \mathbb{R}^r$, and $Ax = b$ defines affine equality constraints. The dual problem of optimization (1) is given as

$$\max_{\lambda\ge 0,\mu} \Big(\min_x f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \mu^\top (Ax - b) \Big), \quad (2)$$

where $\lambda \in \mathbb{R}^m$ and $\mu \in \mathbb{R}^r$ are known as Lagrange multipliers. Let $p^\star$ and $d^\star$ be the optimal value of the primal and dual problems, respectively. Strong duality holds if $p^\star = d^\star$. Under strong duality, the Karush-Kuhn-Tucker (KKT) conditions are a necessary and sufficient characterization of the optimality of the primal-dual solution $(x^\star, \lambda^\star, \mu^\star)$,

$$\begin{cases} 0 \in \bigtriangledown f_0(x^\star) + \sum_{i=1}^m \lambda_i^\star \bigtriangledown f_i(x^\star) + (\mu^\star)^\top A x^\star, \\ \lambda_i^\star f_i(x^\star) = 0, \quad \forall i = 1, \dots, m, \\ (\mu^\star)^\top (Ax^\star - b) = 0, \\ Ax^\star = b, \quad f_i(x^\star) \le 0, \quad \forall i = 1, \dots, m, \\ \lambda_i^\star \ge 0, \quad \forall i = 1, \dots, m. \end{cases}$$

These conditions correspond to stationarity, complementary slackness, and primal and dual feasibility, respectively. The (refined) Slater's condition holds if there exists $x \in \mathbb{R}^n$ with

$$Ax = b \text{ and } f_i(x) < 0, \quad \forall i = 1, \dots, m.$$

Slater's condition implies that strong duality holds. The analogous results hold for a more general SDP formulation.

D. McCormick Relaxation

The McCormick envelopes [16] provide linear relaxations for optimization problems that involve bilinear terms. Consider a bilinear term on the variables $x, y \in \mathbb{R}$, xy, for which there exist upper and lower bounds of the form,

$$\underline{x} \le x \le \overline{x}, \quad \underline{y} \le y \le \overline{y}.$$

Then the McCormick relaxation consists of substituting the term xy by its surrogate $w \in \mathbb{R}$ in the optimization problem and adding the following McCormick envelopes on w,

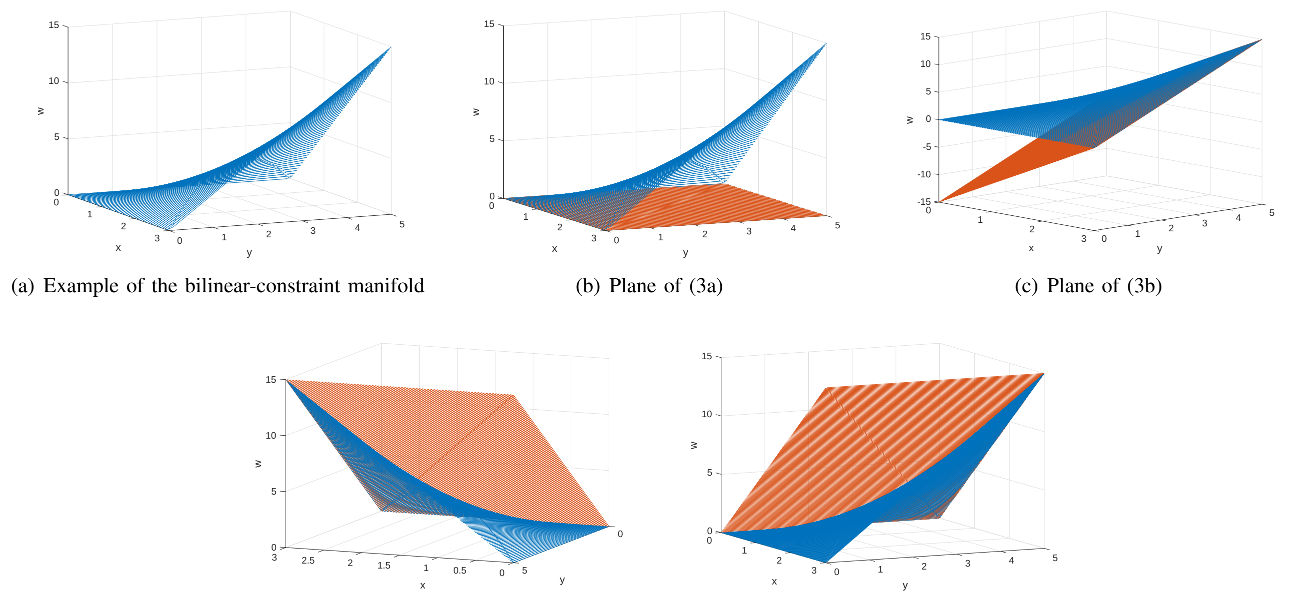
$$w \ge \underline{x}y + x\underline{y} - \underline{x}\underline{y}, \quad (3a)$$

$$w \ge \overline{x}y + x\overline{y} - \overline{x}\overline{y}, \quad (3b)$$

$$w \le \overline{x}y + x\underline{y} - \overline{x}\underline{y}, \quad (3c)$$

$$w \le x\overline{y} + \underline{x}y - \underline{x}\overline{y}, \quad (3d)$$

As illustrated in Figure 1, constraint (3) is tight in the sense that each plane in (3) is tangent to the bilinear-constraint manifold at two boundary lines. Figure 1 also illustrates that the convex polyhedral of (x, y, w) encloses the actual bilinear-constraint manifold.



(a) Example of the bilinear-constraint manifold

(b) Plane of (3a)

(c) Plane of (3b)

(d) Plane of (3c)

(e) Plane of (3d)

Fig. 1. Illustration of all the planes of McCormick envelopes.

III. Problem Statement

This section introduces the problem of interest. We begin with the formulation of the OPF problem over an electrical network and its SDP convex relaxation. The formulation is built according to [17], [18]. In the latter formulation, we introduce binary variables for the co-optimization of topology design, control of tap changers, and capacitor banks.

Consider an electrical network graph with generation buses $\mathcal{N}_G$, load buses $\mathcal{N}_L$, and electrical interconnections described by an undirected edge set $\mathcal{E}$. Let $\mathcal{N} = \mathcal{N}_G \cup \mathcal{N}_L$ and denote its cardinality by N. We denote the phasor voltage at bus i by $V_i = E_i e^{j\theta_i}$, where $E_i \in \mathbb{R}$ and $\theta_i \in [-\pi, \pi)$ are the voltage magnitude and phase angle, respectively. When convenient, we let $V = \{V_i \mid i \in \mathcal{N}\}$ denote the collection of voltages at all buses. The active and reactive power injections at bus i are given by the power flow equations

$$\begin{aligned} P_i &= \mathbf{Tr}\{Y_i V V^*\} + P_{D_i}, \\ Q_i &= \mathbf{Tr}\{\overline{Y}_i V V^*\} + Q_{D_i}, \end{aligned} \tag{4}$$

where $P_{D_i}, Q_{D_i} \in \mathbb{R}$ are the active and reactive power demands[1] at bus i, and $Y_i, \overline{Y}_i \in \mathcal{H}^N$ are derived from the admittance matrix $\mathbf{Y} \in \mathbb{C}^{N\times N}$ as follows

$$Y_i = \frac{(e_i e_i^\top \mathbf{Y})^* + e_i e_i^\top \mathbf{Y}}{2}, \tag{5a}$$

$$\overline{Y}_i = \frac{(e_i e_i^\top \mathbf{Y})^* - e_i e_i^\top \mathbf{Y}}{2j}. \tag{5b}$$

[1]Some buses may have generation and load simultaneously. For buses with only generators, P_{D_i}, Q_{D_i} are both zero.

Here $\{e_i\}_{i=1,\ldots,N}$ denotes the canonical basis of $\mathbb{R}^N$. The OPF problem also involves the following box constraints

$$\begin{aligned} &\underline{V}_i^2 \le |V_i|^2 \le \overline{V}_i^2, \ \forall i \in \mathcal{N}, \\ &\underline{P}_i \le P_i \le \overline{P}_i, \ \ \underline{Q}_i \le Q_i \le \overline{Q}_i, \ \forall i \in \mathcal{N}, \\ &|V_i - V_k|^2 \le \overline{V}_{ik}, \ \forall \{i,k\} \in \mathcal{E}, \end{aligned} \tag{6}$$

where $\overline{V}_{ik}$ is the upper bound of the voltage difference between buses i, k, and $\underline{V}_i$ and $\overline{V}_i$ are the lower and upper bounds of the voltage magnitude at bus i, respectively. The quantities $\underline{P}_i, \underline{Q}_i, \overline{P}_i, \overline{Q}_i$, are defined similarly. The objective function for the OPF problem is typically given as a quadratic function of the active power injection,

$$\sum_{k \in \mathcal{N}_G} c_{i2} P_i^2 + c_{i1} P_i, \tag{7}$$

where $c_{i2} \ge 0$, and $c_{i1} \in \mathbb{R}$. The OPF problem is usually formulated as minimization over (7) subject to (4) and (6). Such optimization is non-convex in general due to the quadratic terms on V. To address this, one can equivalently define $W = VV^* \in \mathcal{H}^N$ (or $W \in \mathcal{H}^N$ and $\mathrm{rank}(W) = 1$) as the decision variable (note all the terms in (4), (6) and (7) are quadratic in V). Dropping the rank constraint on W makes the OPF problem convex, giving rise to the following SDP convex relaxation,

(P1)
$$\min_{W \succeq 0} \sum_{i \in \mathcal{N}_G} c_{i2} P_i^2 + c_{i1} P_i,$$

subject to

$$P_i = \mathbf{Tr}\{Y_i W\} + P_{D_i}, \ \forall i \in \mathcal{N}, \tag{8a}$$

$$Q_i = \mathbf{Tr}\{\overline{Y}_i W\} + Q_{D_i}, \ \forall i \in \mathcal{N} \tag{8b}$$

$$\underline{P}_i \le P_i \le \overline{P}_i, \ \forall i \in \mathcal{N}, \tag{8c}$$

$$\underline{Q}_i \le Q_i \le \overline{Q}_i, \ \forall i \in \mathcal{N}, \tag{8d}$$

$$\underline{V}_i^2 \le \mathbf{Tr}\{M_i W\} \le \overline{V}_i^2,\ \forall i \in \mathcal{N}, \tag{8e}$$
$$\mathbf{Tr}\{M_{ik} W\} \le \overline{V}_{ik},\ \forall \{i,k\} \in \mathcal{E}. \tag{8f}$$

where $M_i, M_{ik} \in \mathcal{H}^N$ are defined so that $\mathbf{Tr}\{M_i W\} = |V_i|^2$ and $\mathbf{Tr}\{M_{ik} W\} = |V_i - V_k|^2$.

We are interested in solving the OPF problem with binary variables modeling potential topology reconfigurations. Assume that we can add a set of edges, $\mathcal{E}_s \subset (\mathcal{N} \times \mathcal{N}) \setminus \mathcal{E}$, to the network $\mathcal{G} = (\mathcal{N}, \mathcal{E})$. With an option for topological changes, the power system can be made more robust while lowering the generation cost or power losses. For each line $\{i,k\} \in \mathcal{E}_s$, we define a binary variable $\alpha_{ik} \in \{0,1\}$. The line is connected if $\alpha_{ik} = 1$ and disconnected otherwise. Taking this into account, the active and reactive power injections of each node become

$$P_i = \mathbf{Tr}\{Y_i W\} + P_{D_i} + \sum_{k \in \mathcal{N}_{i,s}} \alpha_{ik} P_{ik}, \tag{9}$$
$$Q_i = \mathbf{Tr}\{\overline{Y}_i W\} + Q_{D_i} + \sum_{k \in \mathcal{N}_{i,s}} \alpha_{ik} Q_{ik},$$

where $\mathcal{N}_{i,s} := \{k \mid \{i,k\} \in \mathcal{E}_s\}$, $P_{ik} = \mathbf{Tr}\{Y_{ik} W\}$, $Q_{ik} = \mathbf{Tr}\{\overline{Y}_{ik} W\}$, $Y_{ik}, \overline{Y}_{ik} \in \mathbb{C}^{N \times N}$, P_{ik} and Q_{ik} are the active and reactive power flowing from edge i to k, respectively. We define **(P1)**-α to be the OPF problem with the topology determined by a specific $\alpha \in [0,1]^{|\mathcal{E}_s|}$. We also define **(P2)** to be the optimization **(P1)** with constraints (8a) and (8b) replaced by (9). The problem **(P2)** is non-convex for two reasons: the binary variables α_{ik} and the bilinear products of α_{ik} and W. The first problem is usually addressed using existing integer programming solvers [13], [19]. McCormick relaxation procedure described in Section II-D is the standard way to deal with the second problem. We will instead provide alternative routes to address each of these problems for the optimization **(P2)**.

IV. Approximation of the Bilinear Terms

Here, we discuss an alternative procedure to the McCormick relaxation. We start by noting that every binary variable α only appears in the bilinear products in (9) with another continuous variable W. If we convexify the binary variables by allowing them to take values in the interval $[0,1]$, then we can interpret each bilinear term of $\{i,k\} \in \mathcal{E}_s$ as the line power flow from i to k, with the magnitude less than or equal to what W indicates. If the direction of power flow of every line $\{i,k\} \in \mathcal{E}_s$ was known, then the bilinear term would no longer be an issue. For example, if we knew that $P_{ik} = \mathbf{Tr}\{Y_{ik} W\} \in \mathbb{R}_+$ and $Q_{ik} = \mathbf{Tr}\{\overline{Y}_{ik} W\} \in \mathbb{R}_+$, then we can define new variables, $\hat{P}_{ik} \in \mathbb{R}$ and $\hat{Q}_{ik} \in \mathbb{R}$, replacing $\alpha_{ik} P_{ik}$ and $\alpha_{ik} Q_{ik}$ in (9), respectively, and impose the constraints

$$0 \le \hat{P}_{ik} \le P_{ik}, \quad 0 \le \hat{Q}_{ik} \le Q_{ik}. \tag{10}$$

This would eliminate the bilinear terms and the only remaining non-convexity is that the physical feasible solution should satisfy $\hat{P}_{ik} \in \{0, P_{ik}\}$ and $\hat{Q}_{ik} \in \{0, Q_{ik}\}$. In general, however, the direction of power flow of the lines $\{i,k\} \in \mathcal{E}_s$ is not known a priori and hence the trivial convex constraints (10) for the relaxation are no longer valid.

Our idea to approximate the bilinear terms builds on defining a virtual voltage for the terminal nodes. We impose constraints on the virtual voltage to make sure that it has physical sense. We make this precise next. Let $\hat{\mathcal{E}}_s$ be an arbitrary orientation of $\mathcal{E}_s$. To define the virtual voltages (and to keep with the SDP formulation), we introduce a two-by-two positive semi-definite matrix of virtual voltages $U_{ik} \in \mathbb{S}_+^2$, for each $\{i,k\} \in \hat{\mathcal{E}}_s$. This matrix encodes physically meaningful voltages of the terminal nodes if its rank is one, namely, $U_{ik} = u_{ik} u_{ik}^\top$, with $u_{ik}(1)$ and $u_{ik}(2)$ corresponding to the voltages of nodes i and k, respectively. We next impose the following constraints on U_{ik}

$$U_{ik}(1,1) \le \mathbf{Tr}\{M_i W\}, \tag{11a}$$
$$U_{ik}(2,2) \le \mathbf{Tr}\{M_k W\}, \tag{11b}$$
$$\mathbf{Tr}\{\hat{M}_{ik} U_{ik}\} \le \mathbf{Tr}\{M_{ik} W\}, \tag{11c}$$

where $\hat{M}_{ik}(1,1) = \hat{M}_{ik}(2,2) = 1$ and $\hat{M}_{ik}(1,2) = \hat{M}_{ik}(2,1) = -1$. Constraints (11a) and (11b) make the voltage magnitudes of i and k derived from U_{ik} is no bigger than the ones from W. Similarly, constraint (11c) ensures that the voltage difference between nodes i and k computed from U_{ik} is less than the corresponding difference from W. Therefore, by imposing constraint (11) on a rank-one matrix U_{ik}, we obtain a physically meaningful and feasible voltage value. Let $\hat{Y}_{ik} \in \mathbb{C}^{2\times 2}$ and $\mathbf{Tr}\{\hat{\overline{Y}}_{ik}\} \in \mathbb{C}^{2\times 2}$ being the principal submatrices which take entries associated with node i,k of Y_{ik} and $\overline{Y}_{ik}$, respectively. With this we replace $\alpha_{ik} P_{ik}$ and $\alpha_{ik} Q_{ik}$ in (9) by $\mathbf{Tr}\{\hat{Y}_{ik} U_{ik}\}$ and $\mathbf{Tr}\{\hat{\overline{Y}}_{ik} U_{ik}\}$, and convexify **(P2)** as follows

$$\textbf{(P3)} \quad \min_{\substack{W \succeq 0, \\ U_{ik} \succeq 0\ \forall \{i,k\} \in \hat{\mathcal{E}}_s}} \sum_{i \in \mathcal{N}_G} \Big(c_{i2} P_i^2 + c_{i1} P_i \Big),$$

subject to (8c)-(8f), (11), and

$$P_i = \mathbf{Tr}\{Y_i W\} + P_{D_i} + \sum_{k \in \mathcal{N}_{i,s}} \mathbf{Tr}\{\hat{Y}_{ik} U_{ik}\}, \tag{12a}$$
$$Q_i = \mathbf{Tr}\{\overline{Y}_i W\} + Q_{D_i} + \sum_{k \in \mathcal{N}_{i,s}} \mathbf{Tr}\{\hat{\overline{Y}}_{ik} U_{ik}\}. \tag{12b}$$

We next show the result that all non-trivial solutions U_{ik} of **(P3)** have $\mathrm{rank}(U_{ik}) \le 1$, as described in the following.

Proposition IV.1. ***(Optimal solutions have well-defined virtual voltages).*** *If Slater's condition holds for **(P3)**, then the optima of **(P3)** have $\mathrm{rank}(U_{ik}^{\mathrm{opt}}) \le 1$, for all $\{i,k\} \in \hat{\mathcal{E}}_s$.*

Recall that $\mathbf{Tr}\{Y_{ik} U_{ik}^{\mathrm{opt}}\}$ and $\mathbf{Tr}\{\overline{Y}_{ik} U_{ik}^{\mathrm{opt}}\}$ have the interpretation of the optimal line power flows from i to k on the edge $\{i,k\} \in \mathcal{E}_s$, that is, for all $\{i,k\} \in \mathcal{E}_s$

$$P_{ik}^{\mathrm{opt}} = \mathbf{Tr}\{Y_{ik} U_{ik}^{\mathrm{opt}}\}, \quad Q_{ik}^{\mathrm{opt}} = \mathbf{Tr}\{\overline{Y}_{ik} U_{ik}^{\mathrm{opt}}\}. \tag{13}$$

These power flows have a nice property regarding the sum of power loss on each edge. Let $W_{ik}^{\mathrm{opt}} \in \mathcal{H}^2$ be the principal sub-matrix of W^{opt} obtained by removing from W^{opt} the columns and rows different from i and k, respectively. The next result shows that the sum of the power losses is upper bounded by the one computed from $W^{\rm opt}$.

Lemma IV.2. ***(Bounds on the sums of line active and reactive powers).*** *If the line charging susceptance is zero for all $\{i,k\} \in \hat{\mathcal{E}}_s$, then the following inequalities hold for all $\{i,k\} \in \hat{\mathcal{E}}_s$*

$$0 \le P_{ik}^{\rm opt} + P_{ki}^{\rm opt} \le \mathbf{Tr}\{(Y_{ik}+Y_{ki})W_{ik}^{\rm opt}\}, \quad (14a)$$
$$0 \le Q_{ik}^{\rm opt} + Q_{ki}^{\rm opt} \le \mathbf{Tr}\{(\overline{Y}_{ik}+\overline{Y}_{ki})W_{ik}^{\rm opt}\}. \quad (14b)$$

The bounds shown in Lemma IV.2 are only on the optimal line power losses (or the sum of the directional active and reactive power flow). We seek the upper bound on each of $|P_{ik}|$, $|P_{ki}|$, $|Q_{ik}|$, and $|Q_{ki}|$, which are combined stricter than (14) because (14) is only on their sums, i.e., $P_{ik}^{\rm opt}+P_{ki}^{\rm opt}$ and $Q_{ik}^{\rm opt}+Q_{ki}^{\rm opt}$. We next show that under certain conditions for **(P3)**, stricter conditions on the voltages retrieved from the optimal solution $U_{ik}^{\rm opt}$ and $W^{\rm opt}$ hold. To state the result, we introduce $u_i, u_k, w_i, w_k \in \mathbb{C}$ such that $[u_i,u_k]^\top[u_i,u_k] = U_{ik}^{\rm opt}$ and $[w_i,w_k]^\top[w_i,w_k] = W_{ik}^{\rm opt}$.

Proposition IV.3. ***(Bounds on directional power flow).*** *Assume $\{i,k\} \in \hat{\mathcal{E}}_s$ is purely inductive and has zero charging susceptance, $|u_k| \in \{0,|w_k|\}$ and*

$$|w_i| \ge \frac{1}{2}|w_k|, \quad |w_k| \ge \frac{1}{2}|w_i|. \quad (15)$$

Then the following inequalities hold

$$|P_{ik}^{\rm opt}| \le |\mathbf{Tr}\{Y_{ik}W_{ik}^{\rm opt}\}|,\ |P_{ki}^{\rm opt}| \le |\mathbf{Tr}\{Y_{ki}W_{ik}^{\rm opt}\}|, \quad (16a)$$
$$|Q_{ik}^{\rm opt}| \le |\mathbf{Tr}\{(\overline{Y}_{ik}+\overline{Y}_{ki})W_{ik}^{\rm opt}\}|. \quad (16b)$$

The assumption (15) holds for most existing power systems [20]. An analogous result holds by restricting u_i as u_k, as shown in the next result.

Proposition IV.4. ***(Bounds on directional power flow. II).*** *If $\{i,k\} \in \hat{\mathcal{E}}_s$ is purely inductive and has zero charging susceptance, $|u_i| \in \{0,|w_i|\}$ and $\overline{V}_{ik}$ is sufficiently small such that* (15) *holds, then the following inequalities hold*

$$|P_{ik}^{\rm opt}| \le |\mathbf{Tr}\{Y_{ik}W_{ik}^{\rm opt}\}|,\ |P_{ki}^{\rm opt}| \le |\mathbf{Tr}\{Y_{ki}W_{ik}^{\rm opt}\}|,$$
$$|Q_{ki}^{\rm opt}| \le |\mathbf{Tr}\{(\overline{Y}_{ik}+\overline{Y}_{ki})W_{ik}^{\rm opt}\}|.$$

For the general impedance case, similar nonlinear inequalities hold, while the LHS of (16) becomes more complicated. We do not pursue those results here. Propositions IV.3 and IV.4 show that when the diagonal entries of U_{ik} are at the boundary points of their constraints, **(P3)** eliminates the bilinear terms on the active line power flow of **(P2)** in the same way as (10). Based on the intuition that in **(P3)**, the diagonal elements of U_{ik} scaled the "α", we approximate the value of α_{ik} in (9) by

$$\hat{\alpha}_{ik} = \mathbf{Tr}\{U_{ik}^{\rm opt}\}/\mathbf{Tr}\{W_{ik}^{\rm opt}\}. \quad (17)$$

Note that $\hat{\alpha}_{ik} \in [0,1]$ because of (11). One can always round $\hat{\alpha}$ to $\{0,1\}^{|\hat{\mathcal{E}}_s|}$ and derive a candidate possible solution for **(P2)**, $\hat{\alpha}_r$. The following result is straightforward, which draws the relation between **(P2)** and **(P3)** based on the rounded solution $\hat{\alpha}_r$.

Lemma IV.5. ***(Properties of the reconstructed solution).*** *Optimal values of **(P2)**, **(P3)** satisfy $p_2^{opt} \ge p_3^{opt}$. Moreover, if **(P1)**-$\hat{\alpha}_r$ has the optimal value $p_1^{opt} = p_3^{opt}$, then the optimal solution of **(P1)**-$\hat{\alpha}_r$, W_1^{opt}, combined with $\hat{\alpha}_r$ is an optimal solution of **(P2)**.*

Beyond providing candidate values for the binary variables, we rely on the optimization **(P3)** and its solution as a key component of our algorithmic solution to solve **(P2)** described in the next section.

Remark IV.6. ***(Comparison with the McCormick relaxation).*** For comparison, we briefly explain how we implement the McCormick relaxation on the problem **(P2)**. The bilinear terms in **(P2)** are on $\alpha_{ik}P_{ik}$ and $\alpha_{ik}Q_{ik}$. We define new variables $\hat{P}_{ik} \in \mathbb{R}$ and $\hat{Q}_{ik} \in \mathbb{R}$ for each $\{i,k\} \in \hat{\mathcal{E}}_s$. Next, we impose constraints (3) on $\hat{P}_{ik}$ and $\hat{Q}_{ik}$ based on $\alpha_{ik} \in \{0,1\}$ and the upper and lower bounds of line active/reactive power flow, $\overline{P}_{ik}, \overline{Q}_{ik} \in \mathbb{R}_+$, $\underline{P}_{ik} = -\overline{P}_{ik}, \underline{Q}_{ik} = -\overline{Q}_{ik}$. The upper bound of the active power is typically given by line specifications to prevent overheat. We impose the upper bound on the reactive power part for the purpose of the relaxation. We offer two advantages of the formulation proposed in **(P3)** over the McCormick approximation we just described. On the one hand, $\hat{P}_{ik}$ and $\hat{Q}_{ik}$ are loosely tied with the decision variable W in the McCormick relaxation, whereas **(P3)** introduces constraints (11a)-(11c) that enforce a stronger physical connection between the virtual voltages and W. Additionally, the upper and lower bounds on the line power flows may be far from the actual optimal line power flows, which affects the quality of the McCormick relaxation. In contrast, the proposed **(P3)** is not sensitive to those line power bounds as the virtual voltages are bounded by the power computed from W instead of $\overline{P}_{ik}, \overline{Q}_{ik}, \underline{P}_{ik}, \underline{Q}_{ik}$. One therefore needs to accurately estimate the bounds of the line power flow to make the McCormick relaxation effective. For example, knowing the power flow direction improves the estimates significantly, but such information is not available in general. □

Remark IV.7. ***(Optimal switching of transformer taps and capacitor banks).*** Our focus has been on the co-optimization of network topology and OPF, but in fact, the formulation **(P3)** can also be employed to relax the bilinear terms involved in transformer taps and capacitor banks. Both of them entail a physical variable x such that $x = \alpha_s\sqrt{W(i,i)}$, where α_s can take multiple positive integer values, which can always be equivalently represented by some binary variables. Without loss of generality, we assume $\alpha_s \in \{0,1\}$ and relax the bilinear term by introducing a two-by-two real-value positive semidefinite matrix U_s such that

$$U_s(1,2) + U_s(2,1) = x,$$
$$U_s(1,1) \le \mathbf{Tr}\{M_iW\},$$

$$U_s(2,2) \leq 1/4.$$

In this way, the last two inequalities being strict corresponds to $\alpha_s = 1$ and $\alpha_s = 0$ if $U_s = 0$. We can also show that $\text{rank}(U_s) = 1$. We do not incorporate transformer taps and capacitor banks in the OPF formulation only for simplicity of presentation. $\square$

V. Proposed Algorithm for Mixed-Integer OPF

This section focuses on further addressing the complexity associated with the binary variables. Note that the values obtained from the solution of **(P3)** by means of (17) may not in general belong to $\{0,1\}^{|\hat{\mathcal{E}}_s|}$. In such case, one may resort to branch-and-bound algorithm and use **(P3)** to generate the lower bounds in the algorithm. The approach can easily become intractable as $|\hat{\mathcal{E}}_s|$ grows because each lower bound computation from **(P3)** is expensive. We therefore propose to further manipulate the solution of **(P3)** with the objective of relieving the exponential complexity while finding a more accurate solution.

Our algorithm is based on partitioning the graph $\mathcal{G}$ to reduce the original problem into several others of smaller size. Solving the optimization problems associated with each subgraph can dramatically reduce the computational complexity. The proposed partition minimizes the correlation between the resulting subgraphs. If the correlation between the subgraphs (that is, the sensitivity of perturbing certain constraints) is small, then the reconstruction of the solution to **(P2)** from the ones of the subgraphs may result in a better approximation. Intuitively speaking, our procedure is based on the idea that, if the optimal solutions to the problems in each subgraph do not violate the constraints that connect them to other graphs' solutions, then we should be able to recover the full problem solution by putting them together.

1) Graph reduction: We view the disconnection of a line in $\mathcal{E}$ as a perturbation on the constraints (8c)-(8d). The viewpoint is natural in the sense that the disconnection changes the nodal active and reactive power injections of the terminal nodes, which in turn may cause (8c)-(8d) to be violated. Therefore, we seek to partition the graph so that the edge cut set, $\mathcal{E}_c$, induces the minimal perturbation on the optimal value p^{opt}. In particular, edge removal should not affect switchable lines and result into terminal nodes belonging to different subgraphs. This is because, if there exists a line $\{i,k\} \in \hat{\mathcal{E}}_s$ with i and k belonging to different subgraphs, then solving the OPF associated with each subgraph cannot capture how the switch in $\{i,k\}$ affects the optimal value.

We therefore aim to find a partition such that $\mathcal{E}_s \cap \mathcal{E}_c = \emptyset$. To do so, we 'hide' the nodes that are connected by $\mathcal{E}_s$ to the partitioning algorithm that finds $\mathcal{E}_c$. Let $\mathcal{N}_s := \{i \in \mathcal{N} \mid \{i,k\} \in \mathcal{E}_s\}$ and let $\mathcal{N}_{s,i}$ be the set of nodes that are connected to node $i \in \mathcal{N}_s$ through a line in $\mathcal{E}_s$. All nodes in $\mathcal{N}_{s,i}$ are clustered as one representative node and all the edges connected to one of $\mathcal{N}_{s,i}$ are considered being connected to the representative node. This results in a graph $\mathcal{G}_s = \big((\mathcal{N} \setminus \mathcal{N}_s) \cup \mathcal{V}, \mathcal{E}_v\big)$, where $\mathcal{V}$ is the collection of representative nodes. Notice that $\mathcal{E}_v \subseteq \mathcal{E}$ and $\mathcal{E}_v$ is a strict subset of $\mathcal{E}$ if there is $\{i,k\} \in \mathcal{E}$ such that a path connecting nodes i and k exists in the graph $(\mathcal{N}, \mathcal{E}_s)$. Figure 2 illustrates the construction of $\mathcal{G}_s$ and has $\mathcal{E}_v \subset \mathcal{E}$ as one edge of $\mathcal{E}$ is dropped in the process of graph reduction.

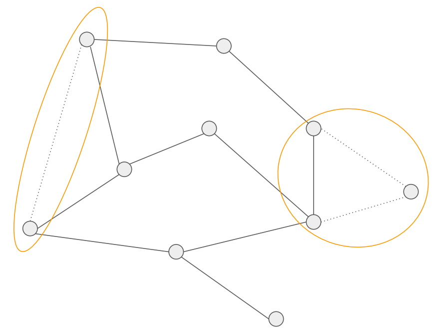


Fig. 2. Simplified graph with nodes connected by $\mathcal{E}_s$ are collapsed into one node. The dash lines denote the edges in $\mathcal{E}_s$; the solid lines denote the edges in $\mathcal{E}$

2) Graph partitioning: Our next step is to find an edge cut set $\mathcal{E}_c$ of the graph $\mathcal{G}_s$ with $\mathcal{E}_s \cap \mathcal{E}_c = \emptyset$. In order to minimally affect the optimal value p^{opt}, the graph partitioning is based on the optimal dual variables of **(P3)**. The optimum dual variables measure how the optimal value p_3^{opt} of **(P3)** changes with respect to the corresponding constraint. Formally, by taking the derivative of the Lagrangian of **(P3)**, we have for each $i \in \mathcal{N}$ that

$$\underline{\lambda}_i^{\text{opt}} = \frac{\partial p_3^{\text{opt}}}{\partial \underline{P}_i}, \quad \overline{\lambda}_i^{\text{opt}} = \frac{\partial p_3^{\text{opt}}}{\partial \overline{P}_i},$$
$$\underline{\gamma}_i^{\text{opt}} = \frac{\partial p_3^{\text{opt}}}{\partial \underline{Q}_i}, \quad \overline{\gamma}_i^{\text{opt}} = \frac{\partial p_3^{\text{opt}}}{\partial \overline{Q}_i},$$

With this interpretation, we define a weighted adjacency matrix $\mathcal{A}$ as follows

$$\begin{cases} \mathcal{A}(i,k) = \sum_{l \in \{i,k\}} \overline{\lambda}_l + \underline{\lambda}_l + \overline{\gamma}_l + \underline{\gamma}_l, & \{i,k\} \in \mathcal{E}_v, \\ \mathcal{A}(i,k) = 0, & \text{otherwise.} \end{cases} \tag{18}$$

Note that if $i \in \mathcal{V}$, then the computation on the entry $\mathcal{A}(i,k)$ uses $\overline{\lambda}_l, \underline{\lambda}_l, \overline{\gamma}_l, \underline{\gamma}_l$ of node $l \in \mathcal{N}_s$ which connects to k in the original graph. Given the weighted adjacency matrix associated with $\mathcal{G}_s$, we do an n-optimal partition on $\mathcal{G}_s$, which gives $\mathcal{G}_s[\mathcal{V}_1^0], \cdots, \mathcal{G}_s[\mathcal{V}_n^0]$ with $\cup_{i=1}^n \mathcal{V}_i^0 = (\mathcal{N} \setminus \mathcal{N}_s) \cup \mathcal{V}$. Since all the removed edges are in $\mathcal{E}$, we can use the same cut for the partition of $\mathcal{G}$: $\mathcal{G}[\mathcal{V}_1], \cdots, \mathcal{G}[\mathcal{V}_n]$ with $\cup_{i=1}^n \mathcal{V}_i = \mathcal{N}$. Such partition ensures $\mathcal{E}_c \cap \mathcal{E}_s = \emptyset$. The intuition is that the cut minimally perturbs p^{opt} because it select edges with minimal weight for the graph $\mathcal{G}$ with adjacency matrix $\mathcal{A}$. Though finding the cut set is NP-hard, there are algorithms that can find a cut set with small weights to determine an n-optimal partition of the graph in few seconds for graphs with the order of a thousand nodes, e.g., [21], [22].

3) Integer optimization on subgraphs: Given an n-partitioned graph $\mathcal{G}[\mathcal{V}_l]$, $l = 1, \cdots, n$, we define an optimization problem associated with each subgraph, which is a variant of **(P2)** that is convenient for the reconstruction of the solution of **(P2)** over the original $\mathcal{G}$. For subgraph l, let $\mathcal{E}_l$ be its set of edges, $W_l \in \mathbb{S}_+^{|\mathcal{V}_l|}$ the decision variable, $\hat{\mathcal{E}}_{s,l}$ the set of switchable lines, and $\mathcal{B}_l$ the set of nodes in $\mathcal{V}_l$ that connects to at least one node of another subgraph. We propose that each subgraph l solves the following

$$\textbf{(P4)} \quad \min_{\substack{W_l \succeq 0, \alpha_{ik} \in \{0,1\} \\ \forall \{i,k\} \in \hat{\mathcal{E}}_{s,i}}} \sum_{i \in \mathcal{N}_G \cap \mathcal{V}_l} \left(c_{i2} P_i^2 + c_{i1} P_i \right),$$

subject to

$$\underline{P}_i \leq P_i \leq \overline{P}_i, \ \forall i \in \mathcal{V}_l,$$
$$\underline{Q}_i \leq Q_i \leq \overline{Q}_i, \ \forall i \in \mathcal{V}_l,$$
$$\underline{V}_i^2 \leq \mathbf{Tr}\{M_i W_l\} \leq \overline{V}_i^2, \ \forall i \in \mathcal{V}_l,$$
$$\mathbf{Tr}\{M_{ik} W_l\} \leq \overline{V}_{ik}, \ \forall \{i,k\} \in \mathcal{E}_l \, .$$

For all $i \in \mathcal{V}_l \setminus \mathcal{B}_l$,

$$P_i = \mathbf{Tr}\{Y_i W_l\} + P_{D_i} + \sum_{k, \{i,k\} \in \mathcal{E}_{s,i}} \alpha_{ik} \, \mathbf{Tr}\{Y_{ik} W_{l,ik}\},$$
$$Q_i = \mathbf{Tr}\{\overline{Y}_i W_l\} + Q_{D_i} + \sum_{k, \{i,k\} \in \mathcal{E}_{s,i}} \alpha_{ik} \, \mathbf{Tr}\{\overline{Y}_{ik} W_{l,ik}\}.$$

For all $i \in \mathcal{B}_l$,

$$P_i = \mathbf{Tr}\{Y_i W_l\} + P_{D_i} + \mathcal{P}_{l,i} + \sum_{k, \{i,k\} \in \mathcal{E}_{s,i}} \alpha_{ik} \, \mathbf{Tr}\{Y_{ik} W_{l,ik}\},$$
$$Q_i = \mathbf{Tr}\{\overline{Y}_i W_l\} + Q_{D_i} + \mathcal{Q}_{l,i} + \sum_{k, \{i,k\} \in \mathcal{E}_{s,i}} \alpha_{ik} \, \mathbf{Tr}\{\overline{Y}_{ik} W_{l,ik}\},$$

where $\mathcal{P}_{l,i} = \sum_{k \in \mathcal{N} \setminus \mathcal{V}_l, \{i,k\} \in \mathcal{E}} P_{ik}^{\text{opt}}$ sums the active power flow from the solution of **(P3)**, $\mathcal{Q}_{l,i}$ is defined in a similar way, and with slight abuse of notation, all M_i, M_{ik}, Y_i, $\overline{Y}_i$, Y_{ik} take proper dimensions matching W_l. Adding $\mathcal{P}_{l,i}$ and $\mathcal{Q}_{l,i}$ in **(P4)** accounts for the "coupling" between $\mathcal{G}[\mathcal{V}_l]$ and the other subgraphs. Since $\mathcal{P}_{l,i}$ and $\mathcal{Q}_{l,i}$ are constants without considering their dependency on the terminal voltage determined by the solutions of the subgraphs, the $\mathcal{P}_{l,i}$ and $\mathcal{Q}_{l,i}$ approximation on the power exchanged (or coupling) between the subgraphs is not accurate. Therefore, the solutions from the previous problems may not result into a feasible solution of **(P2)**. Instead, their combination provides a solution to **(P2)** with a perturbation on (8c) and (8d), which further justifies the graph partitioning with respect to $\mathcal{A}$. Notice that **(P4)** is NP-hard due to α_{ik}, but it is more approachable as the number of switches in each partition, $|\mathcal{E}_{s,i}|$, is far less than $|\mathcal{E}_s|$ if the partition consists of a large number of subgraphs.

4) Full SDP optimization with fixed topology: In the last step, we define the candidate optimal switch $\alpha_c^{\text{opt}} \in [0,1]^{|\mathcal{E}_s|}$ from the solutions of **(P4)**. With this in place, we solve **(P1)**-α_c^{opt} to obtain the candidate optimal solution W_c^{opt}. We regard $(\alpha_c^{\text{opt}}, W_c^{\text{opt}})$ as the reconstructed solution of **(P2)**. We summarize this section by Algorithm 1.

Algorithm 1 Partition-Based MIP-OPF Algorithm

1: **Compute** the optimal solution W^{opt} of **(P3)**
2: **Compute** a graph reduction $\mathcal{G}_s$ (Section V-.1)
3: **Compute** the adjacency matrix for $\mathcal{G}_s$ (Section V-.2)
4: **Compute** a cut set $\mathcal{E}_c$ to partition $\mathcal{G}_s$ into n subgraphs, use it to partition $\mathcal{G}$ into n subgraphs (Section V-.2)
5: **Solve** an integer optimization problem **(P4)** on each subgraph to find α_c^{opt} (Section V-.3)
6: **Solve (P1)**-α_c^{opt} (Section V-.4)

VI. Simulation Studies

In this section, we present simulation studies on the IEEE 118 bus test system. The test system has 152 lines, 27 of which are switchable as described in Table I (we take this list from [23]). We solve two convexified versions of the optimizations of **(P2)**: **(P3)** and the optimization resulting from the McCormick relaxation. We include additional an cost function on the line power losses for both of them so that the optimal solution is more likely not to have all the edges connected.

TABLE I

List of the switchable lines in IEEE 118 test system.

Index	1	2	3	4	5	6
Line	(31,32)	(49,66)	(27,32)	(3,12)	(61,62)	(69,70)
Index	7	8	9	10	11	12
Line	(46,47)	(77,82)	(24,70)	(55,59)	(54,59)	(56,59)
Index	13	14	15	16	17	18
Line	(69,77)	(59,61)	(37,40)	(59,60)	(64,65)	(15,19)
Index	19	20	21	22	23	24
Line	(65,66)	(80,96)	(64,61)	(30,38)	(70,71)	(92,100)
Index	25	26	27			
Line	(11,12)	(94,100)	(23,25)			

We simulate **(P2)** with the McCormick relaxation as described in Remark IV.6 for two different estimates on the upper bounds of the line power flow. One places conservative bounds on $\overline{P}_{ik}$, $\underline{P}_{ik}$, $\overline{Q}_{ik}$, $\underline{Q}_{ik}$ with $\overline{P}_{ik} = \overline{Q}_{ik} = 5$(p.u.) and $\underline{P}_{ik} = \underline{Q}_{ik} = -5$(p.u.) for all $\{i,k\} \in \hat{\mathcal{E}}_s$. We impose those bounds by solving the nominal IEEE 118 example with all the switchable lines connected, and pick the value of the largest line active/reactive power flow. The other case sets $\overline{P}_{ik} = \overline{Q}_{ik} = 1$(p.u.) and $\underline{P}_{ik} = \underline{Q}_{ik} = -1$(p.u.) for all switchable lines, which bound the line power flow more aggressively but which results in a better McCormick approximation because the bounds are closer to the line power flow solution of the nominal IEEE test case. Their results on α are shown in the rows of McC and McC-t, respectively, of Table II. As shown in this table, all the α from **(P3)** are close to zero or one except line 19. In contrast, the α obtained from the McCormick relaxations provide far less information on what value α should take as they are mostly close to 0.5. The optimal values of **(P1)**-α after substituting the rounded values of α from the McCormick relaxations are 153660 and 151370, respectively. On the other hand, substituting α retrieved from **(P3)** in **(P1)**-α gives the optimal value 151040, which is noticeable better than the former two solutions.

TABLE II
SOLUTION OF α FROM **(P3)** AND THE OPTIMIZATION WITH THE MCCORMICK RELAXATION.

Index	1	2	3	4	5	6	7	8	9
(P3)	1	1	.99	1	1	1	1	1	.98
McC.	.5	.5	.49	.49	.51	.49	.49	.49	.49
McC.-t	.52	.87	.51	.5	.47	.73	.49	.49	.49
Index	10	11	12	13	14	15	16	17	18
(P3)	.99	.98	.98	.99	1	1	1	1	1
McC.	.49	.49	.49	.51	.51	.49	.49	.54	.5
McC.-t	.49	.5	.5	.78	.6	.5	.56	.98	.5
Index	19	20	21	22	23	24	25	26	27
(P3)	.39	1	0	1	1	1	1	1	1
McC.	.5	.49	.51	.49	.5	.49	.5	.49	.5
McC.-t	.42	.57	.52	.58	.14	.50	.51	.62	.8

We also have implemented Algorithm 1 to improve the solution obtained from **(P3)**. The graph partitioning step described in Section V-.2 is done by using the MATLAB script in [21]. Algorithm 1 gives the same rounded-up value of α as the one from **(P3)**. Algorithm 1 does not improve the solution from the one of **(P3)** in this case. A possible reason is that the solution of **(P3)** is already close to the true optimum of **(P2)**. The optimal value of **(P3)** is 150990, while the retrieved solution from **(P1)**-α is 151040. The difference is less than 0.1% and **(P3)** is likely to give the optimal α for **(P2)** already, and thus there is no room for Algorithm 1 to further improve it.

VII. CONCLUSIONS

This paper considers MIP-OPF problems that appear in the co-optimization of network topology and OPF. We have addressed the challenges posed by the non-convex bilinear terms and discrete variables. To handle the non-convex bilinear part, we have introduced auxiliary positive semidefinite matrices that convexify the bilinear terms. We show that those PSD matrices can be interpreted as physically meaningful virtual voltages for the original network. Such relationship between the auxiliary variables and the physical ones in the original problem is not captured by the purely geometric McCormick relaxation. To handle the non-convex discrete variable part, we have proposed a graph partitioning method that significantly reduces the computational complexity of the original problem. Future work will provide further analysis on the proposed relaxation methods and additional numerical tests on a larger class of power system testbeds.

ACKNOWLEDGMENTS

This research was supported by the ARPA-e Network Optimized Distributed Energy Systems (NODES) program, Cooperative Agreement DE-AR0000695.

REFERENCES

[1] K. W. Hedman, S. S. Oren, and R. P. O'Neill, "A review of transmission switching and network topology optimization," in *IEEE Power and Energy Society General Meeting*, 2011, pp. 1–7.
[2] J. G. Rolim and L. J. B. Machado, "A study of the use of corrective switching in transmission systems," *IEEE Transactions on Power Systems*, vol. 14, no. 1, pp. 336–341, 1999.
[3] J. D. Fuller, R. Ramasra, and A. Cha, "Fast heuristics for transmission-line switching," *IEEE Transactions on Power Systems*, vol. 27, no. 3, pp. 1377–1386, 2012.
[4] P. A. Ruiz, J. M. Foster, A. Rudkevich, and M. C. Caramanis, "Tractable transmission topology control using sensitivity analysis," *IEEE Transactions on Power Systems*, vol. 27, no. 3, pp. 1550–1559, 2012.
[5] T. Potluri and K. W. Hedman, "Impacts of topology control on the ACOPF," in *IEEE Power and Energy Society General Meeting*, 2012, pp. 1–7.
[6] M. Khanabadi, H. Ghasemi, and M. Doostizadeh, "Optimal transmission switching considering voltage security and $n-1$ contingency analysis," *IEEE Transactions on Power Systems*, vol. 28, no. 1, pp. 542–550, 2013.
[7] M. Soroush and J. D. Fuller, "Accuracies of optimal transmission switching heuristics based on DCOPF and ACOPF," *IEEE Transactions on Power Systems*, vol. 29, no. 2, pp. 924–932, 2014.
[8] H. Hijazi, C. Coffrin, and P. V. Hentenryck, "Convex quadratic relaxations for mixed-integer nonlinear programs in power systems," *Mathematical Programming Computation*, vol. 9, no. 3, pp. 321–367, 2017.
[9] J. Lavaei and S. H. Low, "Zero duality gap in optimal power flow problem," *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 92–107, 2012.
[10] R. Madani, M. Ashraphijuo, and J. Lavaei, "Promises of conic relaxation for contingency-constrained optimal power flow problem," in *Allerton Conf. on Communications, Control and Computing*, Monticello, IL, 2014, pp. 1064–1071.
[11] D. K. Molzahn, C. Josz, I. A. Hiskens, and P. Panciatici, "A Laplacian-based approach for finding near globally optimal solutions to OPF problems," *IEEE Transactions on Power Systems*, vol. 32, no. 1, pp. 305–315, 2017.
[12] J. Mareček, M. Mevissen, and J. C. Villumsen, "MINLP in transmission expansion planning," in *Power Systems Computation Conference (PSCC)*, 2016, pp. 1–8.
[13] E. Briglia, S. Alaggia, and F. Paganini, "Distribution network management based on optimal power flow: Integration of discrete decision variables," in *51st Annual Conference on Information Sciences and Systems (CISS)*, 2017, pp. 1–6.
[14] F. Bullo, J. Cortés, and S. Martínez, *Distributed Control of Robotic Networks*, ser. Applied Mathematics Series. Princeton University Press, 2009, electronically available at http://coordinationbook.info.
[15] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2009.
[16] G. P. McCormick, "Computability of global solutions to factorable nonconvex programs: Part I – convex underestimating problems," *Mathematical programming*, vol. 10, no. 1, pp. 147–175, 1976.
[17] J. Lavaei and S. H. Low, "Zero duality gap in optimal power flow problem," *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 92–107, 2012.
[18] S. H. Low, "Convex relaxation of optimal power flow – part i: Formulations and equivalence," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 1, pp. 15–27, 2014.
[19] R. A. Jabr, R. Singh, and B. C. Pal, "Minimum loss network reconfiguration using mixed-integer convex programming," *IEEE Transactions on Power Systems*, vol. 27, no. 2, pp. 1106–1115, 2012.
[20] A. J. Wood and B. F. Wollenberg, *Power generation, operation, and control*. John Wiley & Sons, 2012.
[21] J. P. Hespanha, "An efficient matlab algorithm for graph partitioning," *University of California*, pp. 1–8, 2004.
[22] A. Abou-Rjeili and G. Karypis, "Multilevel algorithms for partitioning power-law graphs," in *20th International Parallel and Distributed Processing Symposium (IPDPS)*, 2006, pp. 10–pp.
[23] P. A. Ruiz, A. Rudkevich, M. C. Caramanis, E. Goldis, E. Ntakou, and C. R. Philbrick, "Reduced MIP formulation for transmission topology control," in *Allerton Conf. on Communications, Control and Computing*, Monticello, IL, 2012, pp. 1073–1079.