Virtual-Voltage Partition-Based Approach to Optimal Transmission Switching

Chin-Yao Chang, Sonia Martínez, *Fellow, IEEE*, Jorge Cortés, *Fellow, IEEE*

***Abstract*— This paper deals with optimal transmission switching (OTS) problems involving binary decisions about network topology and non-convex power flow constraints. We adopt a semidefinite programming formulation for the OPF problem which, however, remains nonconvex due to the presence of discrete variables and bilinear products between the decision variables. To tackle the latter, we introduce a physically-inspired, virtual-voltage approximation that leads to provable lower and upper bounds on the solution of the original problem. To deal with the exponential complexity caused by the discrete variables, we introduce a graph partition-based algorithm which breaks the problem into several parallel mixed-integer subproblems of smaller size. Simulations on the IEEE bus test cases demonstrate the high degree of accuracy and affordable computational requirements of our approach.**

I. Introduction

Optimal transmission switching (OTS) is concerned with the identification of topologies that minimize generation cost while maintaining the secure operation of the grid. This forms a class of non-convex optimization problems with discrete decision variables and bilinear constraints. Motivated by the emphasis [2] of the Federal Energy Regulatory Commission on research addressing the optimization of flexible assets (including transmission switching) in power system for better efficiency, we aim to develop a computationally efficient approach to solve OTS problems.

Literature review: Transmission line switching or network topology reconfiguration commonly serve as corrective mechanisms in response to system contingencies, see e.g., [3], [4] and references therein. The work [5] introduces the notion of optimal transmission line switching for power dispatch. This paper initiated a series of works [6], [7], [8] that put transmission switching in the context of optimal power flow (OPF) by considering mixed-integer OPF, which is also referred as OTS. Inspired by the potential benefits brought by OTS, a number of recent studies have explored methods to solve it. In [9], [10], [11], linearized OPF, also known as DCOPF, is deployed to solve OTS efficiently. Despite its relatively low complexity, DCOPF may lead, especially in congested systems, to poor solutions that can even result in voltage collapse [12], [13], [14]. The work [15] proposes quadratic convex (QC) relaxations for the OTS problem, which provides more accurate results than DCOPF, while still retaining a fast computation time. Recent studies [15], [16] show that methods based on semidefinite programming (SDP) convex relaxations of ACOPF may lead to better solutions than DCOPF and QC. However, how to handle variables for transmission switching in the context of SDP is challenging and not fully understood. The challenges stem from not only the integer-valued nature of these variables, but also from the presence of bilinear terms involving the product of discrete decision variables with continuous ones, reflecting the impact on the physical modeling of the line being connected. The paper [16] uses a lift-and-branch-and-bound procedure to deal with the SDP formulation of mixed-integer OPF (OTS as a special case), but has still exponential complexity in the worst case. The work [17] also uses SDP to solve OTS problems, where bilinear terms associated to line connections are addressed by assuming certain nominal network topology and bilinear terms of other discrete decision variables are dealt with using the McCormick relaxation [18].

Statement of contributions: We consider the SDP convex relaxation of the OTS problem and make two major contributions. First, we introduce a novel way of dealing with the nonconvexity coming from the presence of bilinear terms, which we term virtual-voltage approximation. Our approach is based on introducing virtual-voltage variables for the terminal nodes of each switchable line and impose physically-meaningful constraints on them. We show that this approach leads to sound bounds on individual flows of the switchable lines, and provides lower and upper bounds on the optimal value of the original problem.

Our second contribution deals with the nonconvexity coming from the discrete variables. We build on the virtual-voltage approximation to propose a graph partition-based algorithm that significantly reduces the computational complexity of solving the original problem. This algorithm uses the values of the optimal dual variables from the virtual-voltage method to define a weighted network graph, which is then partitioned with a minimum weight edge-cut set. The algorithm breaks the original network into sub-networks so as to minimize the correlation between the solutions to the optimization problem on each sub-network. Finally, the algorithm solves the OTS problem on each sub-network in parallel and combines them to reconstruct the solution of the original problem. We implement the proposed algorithms on various IEEE standard

A preliminary version of this work appeared as [1] at the 2017 Allerton Conference on Communications, Control, and Computing.

C.-Y. Chang is with the National Renewable Energy Laboratory at Golden, Colorado, chinyao.chang@nrel.gov, and S. Martínez and J. Cortés are with the Department of Mechanical and Aerospace Engineering, UC San Diego, {soniamd,cortes}@ucsd.edu

test cases, and compare them with available approaches from the literature to illustrate their superior performance regarding convergence to the optima and computation time.

II. Preliminaries

This section introduces basic concepts used in the paper.

1) Notation: We denote by $\mathbb{N}$, $\mathbb{R}$, $\mathbb{R}_+$, and $\mathbb{C}$ the sets of positive integer, real, positive real, and complex numbers, resp. We denote by $|\mathcal{N}|$ the cardinality of $\mathcal{N}$. For $a \in \mathbb{C}$, we let $|a|$ and $\angle a$ be the complex modulus and angle of a, and its real and imaginary parts are $\Re(a)$ and $\Im(a)$. We let $\|V\|$ denote the 2-norm of $V \in \mathbb{C}^{n\times m}$. Let $\mathbb{S}^n_+ \subset \mathbb{C}^{n\times n}$ and $\mathcal{H}^n$ be the set of positive semidefinite and Hermitian matrices, resp. For $A \in \mathbb{C}^{n\times n}$, we let A^* and $\mathbf{Tr}\{A\}$ denote its conjugate transpose and trace, resp. Given $\epsilon \in \mathbb{R}_+$ and $A \in \mathbb{C}^{n\times n}$, we define $\mathcal{B}_\epsilon(A) = \{C \mid |C - A| \le \epsilon\}$.

2) Graph Theory: Following [19], a graph is a pair $\mathcal{G} = (\mathcal{N}, \mathcal{E})$, where $\mathcal{N} \subseteq \mathbb{N}$ is the set of nodes and $\mathcal{E} \subseteq \mathcal{N} \times \mathcal{N}$ is the set of edges. A *self-loop* is an edge that connects a node to itself. The graph is *undirected* if $\{i, k\} = \{k, i\} \in \mathcal{E}$. A *path* is a sequence of nodes such that any two consecutive nodes correspond to an edge. The graph is *connected* if there exists a path between any two nodes. An *orientation* of an undirected graph is an assignment of exactly one direction to each of its edges. A graph is *simple* if it does not have self-loops or multiple edges connecting any pair of nodes. We limit our discussion to undirected, simple graphs. A *vertex-induced subgraph* of $\mathcal{G}$, written $\mathcal{G}[\mathcal{N}_s] = (\mathcal{N}_s, \mathcal{E}_s)$, satisfies $\mathcal{N}_s \subseteq \mathcal{N}$ and $\mathcal{E}_s = \mathcal{E} \cap (\mathcal{N}_s \times \mathcal{N}_s)$. An *edge cut set* is a subset of edges which, if removed, disconnects the graph. A *weighted graph* is a graph where each branch $\{i, k\}$ has a weight, $w_{ik} \in \mathbb{R}_+$. Given $w \in \mathbb{R}_+^{|\mathcal{E}|}$, the *adjacency matrix* A has $A(i, k) = A(k, i) = w_{ik}$ if $\{i, k\} \in \mathcal{E}$, and $A(i, k) = 0$ otherwise. The *degree matrix* D is the diagonal matrix with $D(i, i) = \sum_{k,\{i,k\}\in\mathcal{E}} w_{ik}$. The *normalized adjacency matrix* is $A_n = \sqrt{D}^{-1} A \sqrt{D}^{-1}$. The *Laplacian matrix* is $L = D - A$. The *Fiedler vector* is the eigenvector associated with the second smallest eigenvalue of L. An n-*partition* of a connected $\mathcal{G} = (\mathcal{N}, \mathcal{E}, A)$ divides $\mathcal{G}$ into a number of n connected vertex-induced subgraphs, $\mathcal{G}[\mathcal{V}_i]$, such that $\cup_{i=1}^n \mathcal{V}_i = \mathcal{N}$ and $\mathcal{V}_i \cap \mathcal{V}_k = \emptyset$ for all $i \neq k$. An n-*optimal partition* of $\mathcal{G} = (\mathcal{N}, \mathcal{E}, A)$ is an n-partition of $\mathcal{G}$ with $\sum_{\{i,k\}\in\mathcal{E}_c} w_{ik}$ minimized, where $\mathcal{E}_c = \mathcal{E} \setminus (\cup_{i=1}^n \mathcal{V}_i \times \mathcal{V}_i)$. *Spectral graph partitioning* partitions a connected graph $\mathcal{G}$ into two vertex-induced subgraphs, $\mathcal{G}[\mathcal{N}_1]$ and $\mathcal{G}[\mathcal{N}_2]$, where $\mathcal{N}_1$ and $\mathcal{N}_2$ are the nodes corresponding to the positive and non-positive entries of the Fiedler vector, respectively.

3) McCormick Relaxation of Bilinear Terms: The McCormick envelopes [18] provide linear relaxations for optimization problems that involve bilinear terms. Consider a bilinear term xy on the variables $x, y \in \mathbb{R}$, for which there exist upper and lower bounds, $\underline{x} \le x \le \overline{x}$, $\underline{y} \le y \le \overline{y}$. The McCormick relaxation consists of substituting in the optimization problem the term xy by its surrogate $v \in \mathbb{R}$ and adding the following McCormick envelopes on v,

$$v \ge \underline{x}y + x\underline{y} - \underline{x}\underline{y}, \quad v \ge \overline{x}y + x\overline{y} - \overline{x}\overline{y}, \tag{1a}$$

$$v \le \overline{x}y + x\underline{y} - \overline{x}\underline{y}, \quad v \le x\overline{y} + \underline{x}y - \underline{x}\overline{y}. \tag{1b}$$

Constraints (1) are tight, in the sense that each plane in (1) is tangent to the bilinear-constraint manifold at two boundary lines. The convex polyhedron in the variables (x, y, v) encloses the actual bilinear-constraint manifold.

III. Problem Statement

We begin with the formulation of the OPF problem over an electrical network and its SDP convex relaxation following [20]. Then, we introduce binary variables leading to the OTS problem formulation of interest in this paper.

Consider an electrical network with generation buses $\mathcal{N}_G$, load buses $\mathcal{N}_L$, and electrical interconnections described by an undirected edge set $\mathcal{E}_0$. 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. For convenience, $V = \{V_i \mid i \in \mathcal{N}\}$ denotes the collection of voltages at all buses. The active and reactive power injections at bus i are given by the power flow equations, cf. [21],

$$P_i = \mathbf{Tr}\{Y_i V V^*\} + P_{D_i}, \quad Q_i = \mathbf{Tr}\{\overline{Y}_i V V^*\} + Q_{D_i}, \tag{2}$$

where $P_{D_i}, Q_{D_i} \in \mathbb{R}$ are the active and reactive power demands 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

$$Y_i = \frac{(e_i e_i^\top \mathbf{Y})^* + e_i e_i^\top \mathbf{Y}}{2}, \quad \overline{Y}_i = \frac{(e_i e_i^\top \mathbf{Y})^* - e_i e_i^\top \mathbf{Y}}{2j}.$$

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

$$\begin{aligned} &\underline{V}_i^2 \le |V_i|^2 \le \overline{V}_i^2, \ \forall i \in \mathcal{N}, \\ &\underline{P}_i \le P_i \le \overline{P}_i, \ \ \underline{Q}_i \le Q_i \le \overline{Q}_i, \ \forall i \in \mathcal{N}, \\ &|V_i - V_k|^2 \le \overline{V}_{ik}, \ \forall \{i, k\} \in \mathcal{E}_0, \end{aligned} \tag{3}$$

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. All $\underline{P}_i, \underline{Q}_i, \overline{P}_i, \overline{Q}_i$, are defined similarly. Notice that the upper bound of the voltage difference is equivalent to line thermal constraints. Detailed derivations of $\overline{V}_{ik}$ and its relationship with line thermal limit can be found in [22, Theorem 1]. The voltage difference constraints prevent overheating of transmission lines. The objective function is a quadratic function of the active power,

$$\sum\nolimits_{i\in\mathcal{N}_G} c_{i2} P_i^2 + c_{i1} P_i, \tag{4}$$

where $c_{i2} \ge 0$, and $c_{i1} \in \mathbb{R}$. The OPF problem is the minimization over (4) subject to (2) and (3). Such optimization is non-convex 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. Dropping the rank constraint on W makes the OPF problem convex, giving rise to the SDP convex relaxation,

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

subject to

$$P_i = \mathbf{Tr}\{Y_i W\} + P_{D_i},\ \forall i \in \mathcal{N}, \tag{5a}$$
$$Q_i = \mathbf{Tr}\{\overline{Y}_i W\} + Q_{D_i},\ \forall i \in \mathcal{N}, \tag{5b}$$
$$\underline{P}_i \le P_i \le \overline{P}_i,\ \underline{Q}_i \le Q_i \le \overline{Q}_i,\ \ \forall i \in \mathcal{N}, \tag{5c}$$
$$\underline{V}_i^2 \le \mathbf{Tr}\{M_i W\} \le \overline{V}_i^2,\ \forall i \in \mathcal{N}, \tag{5d}$$
$$\mathbf{Tr}\{M_{ik} W\} \le \overline{V}_{ik},\ \forall \{i,k\} \in \mathcal{E}_0, \tag{5e}$$

where $M_i, M_{ik} \in \mathcal{H}^N$ are defined as follows

$$M_i(l,m) = \begin{cases} 1, & \text{if } l = m = i, \\ 0, & \text{otherwise,} \end{cases}$$
$$M_{ik}(l,m) = \begin{cases} 1, & \text{if } l = m = i \text{ or } l = m = k, \\ -1, & \text{if } (l,m) = (i,k) \text{ or } (l,m) = (k,i), \\ 0, & \text{otherwise.} \end{cases}$$

The definition of M_i above makes $\mathbf{Tr}\{M_i W\}$ equal $|V_i|^2$ if $\text{rank}(W) = 1$. Similarly, the definition of M_{ik} makes $\mathbf{Tr}\{M_{ik} W\}$ equal $|V_i - V_k|^2$ if $\text{rank}(W) = 1$.

Remark III.1. ***(Low-rank SDP convex relaxation).*** There a number of known conditions under which the SDP convex relaxation (P1) has a rank-one optimal solution, cf. [23]. Even if those conditions are not met, one can still observe that the rank of the SDP optimal solutions is usually close to one for many existing test cases, see e.g., [24]. There is a simple intuition behind this fact. Since we employ per-unit expressions throughout the paper, the bounds of the voltage magnitude for all buses, $\underline{V}_i$ and $\overline{V}_i$, are close to one. Therefore, a feasible solution should have per-unit voltage magnitude of all buses close to one. Accordingly, we expect that the optimal solution of **(P1)**, W^{opt_1}, has a simple eigenvalue with the associated eigenvector having its entries close to one. In addition, this eigenvalue is much bigger than all the other eigenvalues. The observation partially explains why it is widely observed that the SDP solution W^{opt_1} is approximately rank-one if $\mathcal{G}$ is connected. □

In the OTS problem, the set of transmission lines $\mathcal{E}_0$ is divided into a set of switchable $\mathcal{E}_s$ and non-switchable $\mathcal{E}$ lines such that $\mathcal{E}_0 = \mathcal{E}_s \cup \mathcal{E}$ (note that $\mathcal{E}_s = \mathcal{E}_0$ and $\mathcal{E} = \emptyset$ is possible). Choosing which ones are active among the switchable lines affects the nodal active and reactive power injections in (5c). The question is then to determine what the optimal choice of switching lines is. We formalize this problem next. For each line $\{i,k\} \in \mathcal{E}_s$, we define a binary variable $\alpha_{ik} \in \{0,1\}$, and we say the line is connected if $\alpha_{ik} = 1$ and disconnected otherwise. If $\alpha_{ik} = 1$, then the power flow from node i to k through edge $\{i,k\} \in \mathcal{E}_s$ is

$$P_{ik} = \mathbf{Tr}\{Y_{ik} W\}, \quad Q_{ik} = \mathbf{Tr}\{\overline{Y}_{ik} W\}, \tag{6}$$

where $Y_{ik}, \overline{Y}_{ik} \in \mathbb{C}^{N \times N}$ are defined as follows: all entries are prescribed to be zero except the ones defined by

$$Y_{ik}(i,i) = \Re(y_{ik}), \quad Y_{ik}(i,k) = Y_{ik}^*(i,k) = -y_{ik}/2,$$
$$\overline{Y}_{ik}(i,i) = -\Im(y_{ik}), \quad \overline{Y}_{ik}(i,k) = \overline{Y}_{ik}^*(i,k) = j \cdot y_{ik}/2.$$

Here $y_{ik} \in \mathbb{C}$ is the admittance of line $\{i,k\}$. Taking (6) into account, the active and reactive power of each node become

$$P_i = \mathbf{Tr}\{Y_i W\} + P_{D_i} + \sum\nolimits_{k \in \mathcal{S}_i} \alpha_{ik} P_{ik}, \tag{7}$$
$$Q_i = \mathbf{Tr}\{\overline{Y}_i W\} + Q_{D_i} + \sum\nolimits_{k \in \mathcal{S}_i} \alpha_{ik} Q_{ik},$$

where $\mathcal{S}_i := \{k \mid \{i,k\} \in \mathcal{E}_s\}$. [1]

We are interested in solving what we call **(P2)**, which is the optimization **(P1)** with constraints (5a) and (5b) replaced by (6) and (7). If we fix a particular $\alpha \in \{0,1\}^{|\mathcal{E}_s|}$ so that this binary variable is no longer a decision variable, we refer to the resulting optimization problem instead as **(P1)**-α. Note that the problem **(P1)**-α is convex, whereas 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 issue can be addressed using existing integer-programming solvers [26], [17]. The McCormick relaxation (cf. Section II-.3) is the standard way to deal with the second issue. Here, we instead provide alternative routes to address each of these issues for **(P2)**.

Remark III.2. ***(Networks where all lines are switchable).*** General formulations of the OTS problem assume all transmission lines are switchable, see e.g. [27], [15]. Some works also employ heuristics [27] to select a pool of switchable lines. Our approach is applicable to both types of scenarios. □

IV. Virtual-Voltage Approximation of Bilinear Terms

We introduce the proposed *virtual-voltage approximation* to deal with the bilinear terms in **(P2)**. Let us start by illustrating the challenges in using the ad-hoc convex relaxation of the OTS problem, which serves as a motivatation for the need of a better approach. We start by noting that every binary variable α only appears in the bilinear products in (7) together with another continuous variable W. If we convexify the binary variables by having them take values in $[0,1]$, then we can interpret each bilinear term corresponding to $\{i,k\} \in \mathcal{E}_s$ as a line power flow from i to k, with the magnitude bounded by what W indicates. Following this reasoning, 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 could 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 (7), respectively, and impose

$$0 \le \hat{P}_{ik} \le P_{ik}, \quad 0 \le \hat{Q}_{ik} \le Q_{ik}. \tag{8}$$

This would eliminate the bilinear terms and the only remaining non-convexity would be that the 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 $\{i,k\} \in \mathcal{E}_s$ is not known a priori and, hence, the trivial convex constraints (8) for the relaxation

[1] In this reformulation, Y_i and $\overline{Y}_i$ are derived from the admittance matrix associated with the edge set $\mathcal{E}$ instead of $\mathcal{E}_0$. We also omit charging susceptance, tap ratio and phase shift of transformers for P_{ik} and Q_{ik} for simplicity. The interested reader can find derivations for matrices Y_i, $\overline{Y}_i$, Y_{ik}, and $\overline{Y}_{ik}$ incorporating all aspects in [25].

are no longer valid. As we show next, the proposed virtual-voltage method provides an alternative convex relaxation that does not require knowledge of the directions of line power flows.

A. Convex Relaxation Via Virtual Voltages

Our idea to approximate each bilinear term builds on defining a virtual voltage for the terminal nodes of the line and impose constraints on them to make sure they have physical meaning. Let $\vec{\mathcal{E}}_s$ be an arbitrary orientation of $\mathcal{E}_s$. To define the virtual voltages, and in keeping with the SDP approach, for each $\{i,k\} \in \vec{\mathcal{E}}_s$ we introduce a positive semidefinite matrix $U_{ik} \in \mathbb{S}^2_+$. This matrix encodes physically meaningful voltages at the terminal nodes if its rank is one, namely, $U_{ik} = u_{ik}u_{ik}^\top$, with $u_{ik}(1)$ and $u_{ik}(2)$ being the voltages of nodes i and k, respectively. For convenience, we introduce $\hat{M} = \left[\begin{smallmatrix} 1 & -1 \\ -1 & 1 \end{smallmatrix}\right]$ and impose the following constraints on U_{ik}

$$U_{ik}(1,1) \leq \mathbf{Tr}\{M_i W\}, \tag{9a}$$
$$U_{ik}(2,2) \leq \mathbf{Tr}\{M_k W\}, \tag{9b}$$
$$\mathbf{Tr}\{\hat{M}U_{ik}\} \leq \mathbf{Tr}\{M_{ik} W\}. \tag{9c}$$

Constraints (9a) and (9b) ensure that the voltage magnitudes of nodes i and k derived from U_{ik} are no bigger than the ones from W. Constraint (9c) ensures that the voltage difference between nodes i and k computed from U_{ik} is less than the corresponding difference from W. Therefore, if the matrix U_{ik} has rank one, constraints (9) ensure that we obtain feasible voltage values.

Let $\hat{Y}_{ik} \in \mathbb{C}^{2\times 2}$ be the principal sub-matrix of $Y_{ik} \in \mathbb{C}^{N\times N}$ constructed by only keeping the rows and columns associated with nodes i and k. We define $\hat{\overline{Y}}_{ik}$ similarly. We replace $\alpha_{ik}P_{ik}$ and $\alpha_{ik}Q_{ik}$ in (7) by $\mathbf{Tr}\{\hat{Y}_{ik}U_{ik}\}$ and $\mathbf{Tr}\{\hat{\overline{Y}}_{ik}U_{ik}\}$, respectively. We now have all the elements necessary to convexify **(P2)** as follows

$$\textbf{(P3)} \qquad \min_{W\succeq 0, U_{ik}\succeq 0\ \forall\{i,k\}\in\vec{\mathcal{E}}_s} \sum\nolimits_{i\in\mathcal{N}_G} \left(c_{i2}P_i^2 + c_{i1}P_i\right),$$

subject to (5c)-(5e), (9), and $\forall i \in \mathcal{N}$

$$P_i = \mathbf{Tr}\{Y_i W\} + P_{D_i} + \sum\nolimits_{k\in\mathcal{S}_i} \mathbf{Tr}\{\hat{Y}_{ik}U_{ik}\}, \tag{10a}$$
$$Q_i = \mathbf{Tr}\{\overline{Y}_i W\} + Q_{D_i} + \sum\nolimits_{k\in\mathcal{S}_i} \mathbf{Tr}\{\hat{\overline{Y}}_{ik}U_{ik}\}. \tag{10b}$$

At the optimal solution of **(P3)**, every two-dimensional matrix $U_{ik}^{\text{opt}_3} \in \mathbb{C}^{2\times 2}$ has a dominant eigenvalue, namely, one eigenvalue is much larger than the other one. To formally state the result, we use the spectral decomposition to rewrite $U_{ik}^{\text{opt}_3}$ as

$$U_{ik}^{\text{opt}_3} = a_{ik}\begin{bmatrix} u_i \\ u_k^* \end{bmatrix}\begin{bmatrix} u_i^* \\ u_k \end{bmatrix}^\top + \begin{bmatrix} u_k \\ -u_i^* \end{bmatrix}\begin{bmatrix} u_k^* \\ -u_i \end{bmatrix}^\top,$$

where $u_i \in \mathbb{C}$, $u_k \in \mathbb{C}$, and $a_{ik} \geq 1$ is the condition number of $U_{ik}^{\text{opt}_3}$. Lemma IV.1 establishes a useful lower bound on a_{ik}.

Lemma IV.1. ***(Lower bound on the condition number).*** *For all $\{i,k\} \in \vec{\mathcal{E}}_s$,*

$$a_{ik} \geq \frac{|u_i|^2 + |u_k|^2 + 2\Re(u_i u_k)}{\overline{V}_{ik} - (|u_i|^2 + |u_k|^2 - 2\Re(u_i u_k))}. \tag{11}$$

Proof. By constraint (9c) and (5e), we have

$$(a_{ik}+1)(|u_i|^2 + |u_k|^2) - 2(a_{ik}-1)\Re(u_i u_k) \leq a_{ik}\,\mathbf{Tr}\{M_{ik}W^{\text{opt}_3}\} \leq a_{ik}\overline{V}_{ik}$$

Lemma IV.1 follows by rearranging the terms of the first line and the RHS of the second line in the inequality above. □

Lemma IV.1 implies that $U_{ik}^{\text{opt}_3}$ specifies well-defined virtual voltages at the terminal nodes, as we explain next.

Remark IV.2. ***(Optimal solutions have well-defined virtual voltages).*** Using Lemma IV.1, we justify that the optimal solution $U_{ik}^{\text{opt}_3}$ has a dominant eigenvalue. The denominator of (11) is always non-negative due to (9c). The order of the denominator of (11) is at most 10^{-2} as $\overline{V}_{ik} \approx 10^{-2}$ in most test cases (involving optimal transmission switching). On the other hand, when the virtual voltage satisfies $\mathbf{Tr}\{\hat{M}U_{ik}^{\text{opt}_3}\} \approx \mathbf{Tr}\{M_i W^{\text{opt}_3}\}$ or $\mathbf{Tr}\{\hat{M}U_{ik}^{\text{opt}_3}\} \approx \mathbf{Tr}\{M_k W^{\text{opt}_3}\}$, then the numerator is lower bounded by a scalar of the order of one, as $\underline{V}_i \approx 1$. As a consequence, the fraction in (11) is usually bigger than 10^2. Our simulations on IEEE 118 and 300 bus text cases confirm that a_{ik} is at least 100. □

B. Physical Properties of the Convex Relaxation

The active and reactive power flows in **(P3)** on a switchable line $\{i,k\} \in \mathcal{E}_s$ are determined by U_{ik} according to

$$P_{ik}^{\text{opt}_3} = \mathbf{Tr}\{\hat{Y}_{ik}U_{ik}^{\text{opt}_3}\}, \quad Q_{ik}^{\text{opt}_3} = \mathbf{Tr}\{\hat{\overline{Y}}_{ik}U_{ik}^{\text{opt}_3}\}. \tag{12}$$

The next result shows that the optimal power losses on each edge are bounded by the ones computed from W^{opt_3}.

Lemma IV.3. ***(Bounds on the sums of line active and reactive powers).*** *The following inequalities hold*

$$0 \leq P_{ik}^{\text{opt}_3} + P_{ki}^{\text{opt}_3} \leq \mathbf{Tr}\{(Y_{ik}+Y_{ki})W^{\text{opt}_3}\}, \tag{13a}$$
$$0 \leq Q_{ik}^{\text{opt}_3} + Q_{ki}^{\text{opt}_3} \leq \mathbf{Tr}\{(\overline{Y}_{ik}+\overline{Y}_{ki})W^{\text{opt}_3}\}. \tag{13b}$$

Proof. The sums, $Y_{ik}+Y_{ki}$ and $\overline{Y}_{ik}+\overline{Y}_{ki}$, take the following form

$$Y_{ik}+Y_{ki} = M_{ik}\Re(y_{ik}),\ \overline{Y}_{ik}+\overline{Y}_{ki} = M_{ik}\Im(-y_{ik}). \tag{14}$$

Since both $\Re(y_{ik})$ and $\Im(-y_{ik})$ are non-negative, the result of Lemma IV.3 follows from (9c) and the equalities (12). □

Remark IV.4. ***(Bounds with complete expression of the line power flow).*** As we mentioned in Section III, the expressions for $Y_{ik} \in \mathbb{C}^{N\times N}$ and $\overline{Y}_{ik} \in \mathbb{C}^{N\times N}$ only take into account series admittances, without including other modeling elements such as charging susceptance, tap ratio, and phase shift. If these were included, tap ratio and phase shift would change their off-diagonal values. In addition, charging susceptance and tap ratio would change the diagonal entries. As a result, $Y_{ik}+Y_{ki}$ would no longer take the simple form shown in (14), and (13a) may not hold. However, the changes induced by these omitted factors are usually small relative to the values corresponding to the series admittance, so even if (13a) no longer holds exactly, the violation is small relative to $\mathbf{Tr}\{(Y_{ik}+Y_{ki})W_{ik}^{\text{opt}_3}\}$. A similar argument holds regarding (13b). □

Note that $P_{ik}^{\text{opt}_3} + P_{ki}^{\text{opt}_3}$ and $Q_{ik}^{\text{opt}_3} + Q_{ki}^{\text{opt}_3}$ respectively have the meaning of active and reactive line power losses. Lemma IV.3 and Remark IV.4 then imply that the line power flow computed from the virtual voltage has less power loss than the one computed from W^{opt_3}. Even though the virtual-voltage approach exhibits this nice property regarding the bounded line power losses, we observe in numerical examples that, for many switchable lines, $P_{ik}^{\text{opt}_3}$ is quite different from $\mathbf{Tr}\{Y_{ik}W^{\text{opt}_3}\}$ (we have made similar observations for $P_{ki}^{\text{opt}_3}$, $Q_{ik}^{\text{opt}_3}$, and $Q_{ki}^{\text{opt}_3}$). Our explanation for this is that, since **(P3)** does not explicitly constrain the line power flows, the virtual-voltage approach gives flexibility to choose the line power flows of the switchable lines advantageously as long as the virtual voltages satisfy the relationship (9) with the actual voltages.

Given the discussion above, analytical results that characterize how $P_{ik}^{\text{opt}_3}$ and $\mathbf{Tr}\{Y_{ik}W^{\text{opt}_3}\}$ are related could be restrictive. Our next result shows that under appropriate assumptions, the line power flows derived by the virtual voltage are bounded by $\|\mathbf{Tr}\{Y_{ik}W^{\text{opt}_3}\}\|$.

Proposition IV.5. ***(Bounds on directional power flow).*** *Assume $\{i,k\} \in \vec{\mathcal{E}}_s$ is purely inductive and*

$$\text{rank}(U_{ik}^{\text{opt}_3}) = 1, \quad \text{rank}(W_{ik}^{\text{opt}_3}) = 1, \tag{15}$$

where $W_{ik}^{\text{opt}_3} = \left[\begin{smallmatrix} W^{\text{opt}_3}(i,i) & W^{\text{opt}_3}(i,k) \\ W^{\text{opt}_3}(k,i) & W^{\text{opt}_3}(k,k) \end{smallmatrix}\right]$. Then, there exists $\epsilon > 0$ such that for all $U_{ik}^{\text{opt}_3} \in \mathcal{B}_\epsilon(W_{ik}^{\text{opt}_3})$,

$$|P_{ik}^{\text{opt}_3}| \leq |\mathbf{Tr}\{Y_{ik}W^{\text{opt}_3}\}|, \; |P_{ki}^{\text{opt}_3}| \leq |\mathbf{Tr}\{Y_{ki}W^{\text{opt}_3}\}|. \tag{16}$$

Proof. If $\{i,k\}$ is purely inductive, then

$$\hat{Y}_{ik} = \frac{1}{2}\begin{bmatrix} 0 & y_{ik}^* \\ y_{ik} & 0 \end{bmatrix}, \hat{Y}_{ki} = \frac{1}{2}\begin{bmatrix} 0 & y_{ik} \\ y_{ik}^* & 0 \end{bmatrix}. \tag{17}$$

Equation (17) implies that $P_{ik}^{\text{opt}_3} = -P_{ki}^{\text{opt}_3}$ and $|\mathbf{Tr}\{Y_{ik}W^{\text{opt}_3}\}| = |\mathbf{Tr}\{Y_{ki}W^{\text{opt}_3}\}|$. It is then sufficient to prove one of the inequalities of (16). Because of the rank-1 assumption, we rewrite $U_{ik}^{\text{opt}_3}$ and $W_{ik}^{\text{opt}_3}$ as

$$U_{ik}^{\text{opt}_3} = \begin{bmatrix} u_i \\ u_k^* \end{bmatrix}\begin{bmatrix} u_i^* \\ u_k \end{bmatrix}^\top, \quad W_{ik}^{\text{opt}_3} = \begin{bmatrix} w_i \\ w_k^* \end{bmatrix}\begin{bmatrix} w_i^* \\ w_k \end{bmatrix}^\top.$$

Define $\xi_i, \xi_k \in [0,1]$ such that $\xi_i|w_i| = |u_i|$ and $\xi_k|w_k| = |u_k|$. Let θ_{ik}^u be the angle differences between u_i and u_k. Similarly, θ_{ik}^w is defined as the angle differences between w_i and w_k. Assuming $|P_{ik}^{\text{opt}_3}| > |\mathbf{Tr}\{Y_{ik}W^{\text{opt}_3}\}|$, then

$$\begin{aligned} &|u_i||u_k||\sin(\theta_{ik}^u)| > |w_i||w_k||\sin(\theta_{ik}^w)| \\ \implies &\xi_i^2\xi_k^2(1-\cos^2(\theta_{ik}^u)) > (1-\cos^2(\theta_{ik}^w)) \\ \implies &\cos(\theta_{ik}^w) > \sqrt{(1-\xi_i^2\xi_k^2)+\xi_i^2\xi_k^2\cos^2(\theta_{ik}^u)}. \end{aligned}$$

We next rewrite (9c) with $w_i, w_k, \xi_i, \xi_k, \theta_{ik}^u, \theta_{ik}^w$, and then substitute $\cos(\theta_{ik}^w)$ by using the inequality above, shown in the following.

$$\begin{aligned} &(1-\xi_i^2)|w_i|^2 + (1-\xi_k^2)|w_k|^2 + 2\xi_i\xi_k|w_i||w_k|\cos(\theta_{ik}^u) \geq \\ &\qquad 2|w_i||w_k|\cos(\theta_{ik}^w) \\ \implies &(1-\xi_i^2)|w_i|^2 + (1-\xi_k^2)|w_k|^2 + 2\xi_i\xi_k|w_i||w_k|\cos(\theta_{ik}^u) > \\ &\qquad 2|w_i||w_k|\sqrt{(1-\xi_i^2\xi_k^2)+\xi_i^2\xi_k^2\cos^2(\theta_{ik}^u)} \end{aligned}$$

If $\xi_i = \xi_k = 1$, then there is a contradiction for the inequality above as the both LHS and RHS are zero. This implies that there exists $\epsilon > 0$ such that the contradiction follows for any $\forall U_{ik}^{\text{opt}_3} \in \mathcal{B}_\epsilon(W_{ik}^{\text{opt}_3})$, which completes the proof. □

Note that the value of ϵ in Proposition IV.5 could be small, which makes the result difficult to hold true. This is reflected in the simulations of the IEEE 118 and 300 bus systems, which only have around 70% of switchable lines satisfy (16). The assumption of purely inductive lines is also restrictive. For lines with general impedance, the arguments in the proof of Proposition IV.5 only hold for certain instances of θ_{ik}^u and θ_{ik}^w have.

C. Reconstructed Solution to the OTS Problem

We note that the ratio of the voltage magnitudes derived from $U_{ik}^{\text{opt}_3}$ and W^{opt_3} provides an approximation of the discrete variables α_{ik} in (7) as

$$\hat{\alpha}_{ik} = \mathbf{Tr}\{U_{ik}^{\text{opt}_3}\}/\,\mathbf{Tr}\{M_{ik}W^{\text{opt}_3}\}. \tag{18}$$

Note that $\hat{\alpha} \in [0,1]^{|\vec{\mathcal{E}}_s|}$ because of (9). If we round the entries of $\hat{\alpha}$ to the closest number in $\{0,1\}$, we obtain a candidate solution $\hat{\alpha}_r \in \{0,1\}^{|\vec{\mathcal{E}}_s|}$ to **(P2)**. The following result, whose proof is straightforward, states the relationship between **(P2)** and **(P3)** based on the rounded solution $\hat{\alpha}_r$.

Proposition IV.6. ***(Properties of the reconstructed solution).*** *The optimal values of **(P1)**-$\hat{\alpha}_r$, **(P2)**, and **(P3)** satisfy $p_1^{\text{opt}} \geq p_2^{\text{opt}} \geq p_3^{\text{opt}}$. Moreover, if $p_1^{\text{opt}} = p_3^{\text{opt}}$, then the optimal solution of **(P1)**-$\hat{\alpha}_r$, W_1^{opt}, combined with $\hat{\alpha}_r$, is an optimal solution of **(P2)**.*

Note that even if $\hat{\alpha} = \hat{\alpha}_r \in \{0,1\}^{|\vec{\mathcal{E}}_s|}$, p_3^{opt} does not necessary equal p_2^{opt}. The reason is that (18) computes $\hat{\alpha}_{ik}$ from the diagonal of $U_{ik}^{\text{opt}_3}$, and hence we can not conclude any equality for the off-diagonal elements of $U_{ik}^{\text{opt}_3}$ and $M_{ik}W^{\text{opt}_3}$. Hence, even if $\hat{\alpha} \in \{0,1\}^{|\vec{\mathcal{E}}_s|}$, the optimal solution of **(P3)** does not necessarily lie in the feasible region of **(P2)**.

Remark IV.7. ***(Comparison with the McCormick relaxation).*** We explain how we implement the McCormick relaxation on the problem **(P2)** for comparison purposes. For each $\{i,k\} \in \vec{\mathcal{E}}_s$, we define new variables $\hat{P}_{ik}, \hat{Q}_{ik} \in \mathbb{R}$ to substitute the bilinear terms $\alpha_{ik}P_{ik}, \alpha_{ik}Q_{ik}$, respectively. Then, we impose constraints of the form (1) on the new variables based on $\alpha_{ik} \in \{0,1\}$ and upper and lower bounds of active/reactive line power flow, $\overline{P}_{ik}, \overline{Q}_{ik} \in \mathbb{R}_+$, $\underline{P}_{ik} = -\overline{P}_{ik}, \underline{Q}_{ik} = -\overline{Q}_{ik}$. If these bounds are far from the actual optimal line power flows, this can significantly affect the quality of the solution obtained by the McCormick relaxation, a point that we illustrate later in our simulations, along with rationale for how to select them. In contrast, the proposed relaxation **(P3)** is not sensitive to those line power bounds, as the virtual voltages are bounded by the power computed from W. Additionally, the variables $\hat{P}_{ik}$ and $\hat{Q}_{ik}$ in the McCormick relaxation are loosely tied to the decision variable W, whereas **(P3)** introduces constraints (9a)-(9c) enforcing a stronger physical connection between the virtual voltages and W. □

D. N-1 Security Constraints

The virtual-voltage approximation approach described above can also be applied to scenarios with security constraints. N-1 security formulations are widely considered in the literature [8], [13] and require that, under any single component outage (failures of a generator or a transmission line), power flow constraints remain satisfied. This prevents cascading failures and benefits post-contingency controls. The proposed virtual-voltage approach can accommodate N-1 security constraints as we explain next. Let $\mathcal{C} := \{0, 1, \cdots, c\}$ be the set that enumerates contingency scenarios, with index $t = 0$ corresponding to the nominal operating scenario. We formulate the virtual-voltage convexifed OTS with N-1 security constraints

$$\min_{\substack{W^{[t]} \succeq 0, t \in \mathcal{C}, \\ U_{ik} \succeq 0 \ \forall \{i,k\} \in \vec{\mathcal{E}}_s}} \sum\nolimits_{i \in \mathcal{N}_G} \Big(c_{i2} P_{i0}^2 + c_{i1} P_{i0} \Big),$$

subject to the following for all $t \in \mathcal{C}$

$$\begin{aligned}
&U_{ik}(1,1) \leq \mathbf{Tr}\{M_i W^{[t]}\}, \ \forall \{i,k\} \in \mathcal{E}_s \\
&U_{ik}(2,2) \leq \mathbf{Tr}\{M_k W^{[t]}\}, \ \forall \{i,k\} \in \mathcal{E}_s \\
&\mathbf{Tr}\{\hat{M} U_{ik}\} \leq \mathbf{Tr}\{M_{ik} W^{[t]}\} \ \forall \{i,k\} \in \mathcal{E}_s \\
&\underline{P}_i \leq P_i^{[t]} \leq \overline{P}_i, \ \ \underline{Q}_i \leq Q_i^{[t]} \leq \overline{Q}_i, \ \forall i \in \mathcal{N}^{[t]}, \\
&\underline{V}_i \leq \mathbf{Tr}\{M_i W^{[t]}\} \leq \overline{V}_i, \ \forall i \in \mathcal{N}^{[t]}, \\
&\mathbf{Tr}\{M_{ik} W^{[t]}\} \leq \overline{V}_{ik}, \ \forall \{i,k\} \in \mathcal{E}^{[t]}, \\
&P_i^{[t]} = \mathbf{Tr}\{Y_i^{[t]} W^{[t]}\} + P_{D_i} + \sum\nolimits_{k \in \mathcal{N}_{i,s}^{[t]}} \mathbf{Tr}\{\hat{Y}_{ik} U_{ik}\}, \\
&Q_i^{[t]} = \mathbf{Tr}\{\overline{Y}_i^{[t]} W^{[t]}\} + Q_{D_i} + \sum\nolimits_{k \in \mathcal{N}_{i,s}^{[t]}} \mathbf{Tr}\{\hat{\overline{Y}}_{ik} U_{ik}\},
\end{aligned}$$

where $\mathcal{E}^{[t]}$ is the set of connected lines for contingency $t \in \mathcal{C}$. Each $W^{[t]}$ corresponds to the solution of the contingency $t \in \mathcal{C}$. All the constraints involved in **(P3)** are imposed on every $W^{[t]}$. Since there is no coupling between $W^{[t_1]}$ and $W^{[t_2]}$ for different $t_1, t_2 \in \mathcal{C}$, all the properties of the virtual-voltage approximation characterized above remain valid. The problem size of the security constrained OTS grows linearly with respect to $|\,\mathcal{C}\,|$.

V. Partition-Based OTS Algorithm

The virtual voltage approach described in Section IV finds a candidate switching $\hat{\alpha}_r$, cf. (18), for the OTS problem. This, together with the solution to **(P3)**, provides upper and lower bounds on the optimal value of the OTS problem, cf. Proposition IV.6. In this section, we discuss how the partition-based OTS algorithm described in Algorithm 1 may find a better solution than solving **(P1)**-$\hat{\alpha}_r$. The idea is to use the solution of **(P3)** to break the original OTS problem into a number of smaller problems so that each sub-problem can be solved exactly. Our numerical studies show that the reconstructed solution obtained by combining the optimal switchings of the sub-problems is usually better than **(P1)**-$\hat{\alpha}_r$. Note that the method relies on graph partitioning and is not directly applicable to security constrained OTS formulations, cf. Section IV-D.

An informal description of Algorithm 1 is as follows:

Algorithm 1 Partition-Based OTS Algorithm

1: **Compute** the optimal solution W^{opt} of **(P3)**
2: **Construct** graph reduction $\mathcal{G}_r$ (Section V-.1)
▷ *Clusters together nodes connected by switchable lines*
3: **Assign** adjacency matrix to $\mathcal{G}_r$ (Section V-.2)
▷ *Sets weights according to edge influence on optimal solution*
4: **Compute** cut set $\mathcal{E}_c$ to partition $\mathcal{G}_r$ into n subgraphs (Section V-.2)
▷ *Partitions network graph into smaller components accounting for impact on optimal solution*
5: **Solve** integer optimization problem **(P4)** on each subgraph to find α_p^{opt} (Section V-.3)
▷ *Solves subproblems of smaller size*
6: **Solve (P1)**-α_p^{opt} (Section V-.4)
▷ *Reconstructs solution of original problem*

[Informal description:] The algorithm partitions the network graph into smaller components to break the original problem into subproblems of smaller size. To do this, we view the disconnection of a line as a perturbation on the constraint (5c). This is natural, in the sense that the disconnection changes the active and reactive power injections of the terminal nodes, which in turn may cause the constraints to be violated. Hence, we partition the graph so that the correlation between the solutions to the optimization problem on each of the resulting subgraphs is minimized. The idea is that, if the optimal solution in each subgraph minimally violates the constraints that connect them to other graphs' solutions, then, when put together, the reconstructed solution to **(P2)** would be of better quality.

In the following, we provide a detailed explanation of each step of the algorithm.

1) Graph reduction: This is a step prior to graph partitioning which is motivated by the following observation. The graph partitioning should not result in nodes connected by a switchable line belonging to different subgraphs. This is because, if that were the case, then solving the OPF associated with each subgraph cannot capture how the switch in that specific line affects the optimal value of the original OTS problem. To address this, we 'hide' the nodes that are connected by $\mathcal{E}_s$ to the partitioning algorithm that finds the edge cut $\mathcal{E}_c$ so as to ensure $\mathcal{E}_s \cap \mathcal{E}_c = \emptyset$. Let $\mathcal{N}_s := \{i \in \mathcal{N} \mid \exists k \in \mathcal{N} \text{ s.t. } \{i,k\} \in \mathcal{E}_s\}$ 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 $\mathcal{G}_r = \big((\mathcal{N} \setminus \mathcal{N}_s) \cup \mathcal{V}, \mathcal{E}_r\big)$, where $\mathcal{V}$ is the collection of representative nodes. Notice that $\mathcal{E}_r \subseteq \mathcal{E}$ and $\mathcal{E}_r$ 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 1 illustrates the construction of $\mathcal{G}_r$ and has $\mathcal{E}_r \subset \mathcal{E}$ as one edge of $\mathcal{E}$ is dropped in the process of graph reduction.

The graph reduction step described above only makes sense

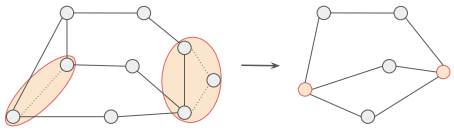

Fig. 1: Graph reduction. Nodes connected by $\mathcal{E}_s$ are collapsed into one node. The dash lines denote edges in $\mathcal{E}_s$; the solid lines denote the edges in $\mathcal{E}$.

when not all lines are switchable because otherwise, it results in a graph with a single node. For OTS scenarios where all lines are switchable, one can either skip the graph reduction step, or pre-select a set of lines that should remain non-switchable.

2) Graph partitioning: Our next step is to find an edge cut set $\mathcal{E}_c$ of the graph $\mathcal{G}_r$. 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 the following for $i \in \mathcal{N} \setminus \mathcal{N}_s$,

$$\underline{\lambda}_i^{\text{opt}_3} = \frac{\partial p_3^{\text{opt}}}{\partial \underline{P}_i},\ \overline{\lambda}_i^{\text{opt}_3} = \frac{\partial p_3^{\text{opt}}}{\partial \overline{P}_i},\ \underline{\gamma}_i^{\text{opt}_3} = \frac{\partial p_3^{\text{opt}}}{\partial \underline{Q}_i},\ \overline{\gamma}_i^{\text{opt}_3} = \frac{\partial p_3^{\text{opt}}}{\partial \overline{Q}_i},$$

and for $i \in \mathcal{V}$,

$$\underline{\lambda}_i^{\text{opt}_3} = \sum_{k \in \mathcal{N}_{s,i}} \frac{\partial p_3^{\text{opt}}}{\partial \underline{P}_k},\ \overline{\lambda}_i^{\text{opt}_3} = \sum_{k \in \mathcal{N}_{s,i}} \frac{\partial p_3^{\text{opt}}}{\partial \overline{P}_k},$$

$$\underline{\gamma}_i^{\text{opt}_3} = \sum_{k \in \mathcal{N}_{s,i}} \frac{\partial p_3^{\text{opt}}}{\partial \underline{Q}_k},\ \overline{\gamma}_i^{\text{opt}_3} = \sum_{k \in \mathcal{N}_{s,i}} \frac{\partial p_3^{\text{opt}}}{\partial \overline{Q}_k},$$

With this interpretation, we define edge weights as follows

$$A(i,k) = \begin{cases} \sum\limits_{l \in \{i,k\}} \overline{\lambda}_l^{\text{opt}_3} + \underline{\lambda}_l^{\text{opt}_3} + \overline{\gamma}_l^{\text{opt}_3} + \underline{\gamma}_l^{\text{opt}_3}, & \{i,k\} \in \mathcal{E}_r, \\ 0, & \text{otherwise.} \end{cases}$$

The intuition behind defining the adjacency matrix A in this way is as follows: the power injections of nodes i and k change when line $\{i,k\}$ is removed, which in turn may cause violations of the nodal power injection constraints. The sum of the λ's and γ's in the definition of $A(i,k)$ then approximates the change of p_3^{opt} as a consequence of removing $\{i,k\}$. Note that this is only a crude way to estimate how p_3^{opt} would change when a line is removed. The exact value would require solving **(P3)** with the line removed, which would be impractical.

Given the adjacency matrix A, we do an n-optimal partition on $\mathcal{G}_r$, which gives $\mathcal{G}_r[\mathcal{V}_1^0], \cdots, \mathcal{G}_r[\mathcal{V}_n^0]$ 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 weighted graph $(\mathcal{G}, A)$. The intuition for using a minimum cut is to have minimum power flows between different components so that solving the optimization on each sub-graph and combining the solutions together has minimum error.

Finding the optimal cut set is NP-hard. There are algorithms [28], [29] that can approximate it in a few seconds for graphs of the order of a thousand nodes. However, they do not guarantee that the resulting subgraphs are connected. To ensure this property, we resort to spectral graph partitioning.

Theorem V.1. ***(Fiedler's theorem of connectivity of spectral graph partitions).*** *The two subgraphs resulting from spectral graph partitioning on a connected graph are also connected.*

The proof is available in [30, Corollary 2.9]. To derive a n-partition, one can implement spectral graph partitioning recursively n times. Since we aim for subgraphs with similar size, each iteration applies spectral graph partitioning on the subgraph with the largest number of nodes. The most computationally expensive step in this process is the eigenvector computation, which only takes linear time, or $\mathcal{O}(n)$. This low complexity is reflected on the negligible computational time in our simulations. Even though this recursive spectral partitioning does not in general lead to an n-optimal partition, we can characterize a lower bound for the sum of weights for each iteration l of the recursive partitioning by

$$\frac{1}{2} C_{\text{opt},l}^2 \leq \sum_{\{i,k\} \in \mathcal{E}_c(l)} A(i,k), \tag{20}$$

where $C_{\text{opt},l}$ is the optimal value for 2-optimal partitioning and $\mathcal{E}_c(l)$ is the edge cut set in iteration l (w.l.o.g., we assume normalized A). Inequality (20) follows directly from Cheeger inequalities and the Courant-Fischer Theorem [31].

3) Integer optimization on subgraphs: Given a n-partition $\{\mathcal{G}[\mathcal{V}_l]\}_{l=1}^n$, we define an optimization problem associated with each subgraph. This problem 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, $\mathcal{E}_{s,l}$ 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. Each subgraph l solves

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

subject to

$$\underline{P}_i \leq P_i \leq \overline{P}_i, \quad \underline{Q}_i \leq Q_i \leq \overline{Q}_i,\ \forall i \in \mathcal{V}_l,$$

$$\underline{V}_i^2 \leq \mathbf{Tr}\{M_i W_l\} \leq \overline{V}_i^2,\ \forall i \in \mathcal{V}_l,$$

$$\mathbf{Tr}\{M_{ik} W_l\} \leq \overline{V}_{ik},\ \forall \{i,k\} \in \mathcal{E}_l\,.$$

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

$$P_i = \mathbf{Tr}\{Y_i W_l\} + P_{D_i} + \sum\nolimits_{k,\{i,k\} \in \mathcal{E}_{s,i}} \alpha_{ik}\, \mathbf{Tr}\{Y_{ik} W_l\},$$

$$Q_i = \mathbf{Tr}\{\overline{Y}_i W_l\} + Q_{D_i} + \sum\nolimits_{k,\{i,k\} \in \mathcal{E}_{s,i}} \alpha_{ik}\, \mathbf{Tr}\{\overline{Y}_{ik} W_l\}.$$

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

$$P_i = \mathbf{Tr}\{Y_i W_l\} + P_{D_i} + \mathcal{P}_{l,i} + \sum_{k,\{i,k\} \in \mathcal{E}_{s,i}} \alpha_{ik}\, \mathbf{Tr}\{Y_{ik} W_l\},$$

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

where $\mathcal{P}_{l,i} = \sum_{k \in \mathcal{N} \setminus \mathcal{V}_l, \{i,k\} \in \mathcal{E}} P_{ik}^{\text{opt}_3}$ sums the active power flow from the solution of **(P3)**, $\mathcal{Q}_{l,i}$ is defined similarly, and with a slight abuse of notation, all M_i, M_{ik}, Y_i, $\overline{Y}_i$, Y_{ik}, and $\overline{Y}_{ik}$ take proper dimensions matching W_l. Notice that

(P4) is still NP-hard due to α_{ik}, but the number of switches $|\mathcal{E}_{s,i}|$ in each partition is less than $|\mathcal{E}_s|$, and decreases with n. The addition of $\mathcal{P}_{l,i}$ and $\mathcal{Q}_{l,i}$ in **(P4)** accounts for the coupling between $\mathcal{G}[\mathcal{V}_l]$ and the other subgraphs. For each subgraph, these terms are constant and do not provide an exact approximation of the power exchanged between the subgraphs – since they do not take into account the dependency of the terminal voltage on the solutions determined on the other subgraphs. Therefore, putting together the solutions obtained for each subgraph may not result in a feasible solution of **(P2)**, but rather a solution to **(P2)** with a perturbation on (5c). We address this next.

4) Full SDP optimization with fixed topology: In the last step, we define the candidate optimal switch $\alpha_p^{\text{opt}} \in \{0,1\}^{|\mathcal{E}_s|}$ from the solutions of **(P4)** obtained in the previous step. With this in place, we solve **(P1)**-α_p^{opt} to obtain the candidate optimal solution W_p^{opt} and output $(\alpha_p^{\text{opt}}, W_p^{\text{opt}})$ as the reconstructed solution of **(P2)**.

VI. Simulation Studies

We illustrate the performance of the virtual-voltage approximation on standard IEEE test systems. All simulations are done on a desktop with 3.5GHz CPU and 16GB RAM, using MATLAB and its CVX toolbox [32] for convex optimization problems. In all tables except the last row in Table III, "lower bound" refers to the optimal value of **(P3)** and "upper bound" refers to the optimal value of **(P1)**-α (with α coming from the corresponding method).

A. Comparison with McCormick Relaxation

We implement the virtual-voltage approximation and compare its performance against the solution obtained from **(P2)** with the McCormick approximation (cf. Remark IV.7). For the latter, we use two different estimates on the upper bounds of the line power flows. In one case, we use the conservative bounds $\overline{P}_{ik} = -\underline{P}_{ik} = 5$(p.u.) and $\overline{Q}_{ik} = -\underline{Q}_{ik} = 5$(p.u.) for all $\{i,k\} \in \vec{\mathcal{E}}_s$. These bounds come from the heuristic estimation on the largest line active/reactive power flow. In the other case, we set the tighter bounds $\overline{P}_{ik} = -\underline{P}_{ik} = 1$(p.u.) and $\overline{Q}_{ik} = -\underline{Q}_{ik} = 0.5$(p.u.) for all $\{i,k\} \in \vec{\mathcal{E}}_s$, based on our knowledge of the solution of the IEEE test nominal case. We include an additional cost function on the line power losses to promote optimal solutions with some edges disconnected. For each test case, a set of switchable lines are selected by the following policy: given a design parameter $p \in \mathbb{N}$, we rank all lines by the norm of their admittance in ascending order and select the first p as the set of switchable lines. The rationale for this selection is that each line with small admittance places a small correlation between its terminal nodes and, as a result, they are more likely to be disconnected. The selection of p is based on network size.

The virtual-voltage approach yields discrete variables close to $\{0,1\}$ for all the test cases and, in contrast, the McCormick relaxation has most of them around 0.5. Table I shows the values obtained by both approximations, and confirms that the virtual-voltage approach gives better solutions than the McCormick relaxation. It is worthwhile to note that every virtual-voltage matrix U has its condition number comfortably more than 1000 for all the test cases. The smallest condition number is 1492. This validates the statements of Lemma IV.1 and Remark IV.2 for a physically meaningful virtual voltage.

In both the McCormick relaxation and virtual-voltage methods, the number of decision variables grows linearly with respect to p. Though p is relatively small compared to the dimension of W and is not the main factor for optimization complexity, we observe in Table II that for a given p, the McCormick relaxation returns a solution noticeably faster than the virtual-voltage method. The reason is that the McCormick relaxation does not introduce LMIs for the switchable lines. In this regard, the virtual-voltage approach trades time complexity for better-quality solutions.

B. Comparison with BARON

We compare the virtual-voltage approximation with the general purpose mixed-integer nonlinear programming solver BARON [33] in several small-scale IEEE examples. With the default settings and maximal computational time at 1000 seconds, BARON does not always return a feasible solution, see Table III.

We observe that BARON gives a feasible solution only when it finds one in the preprocessing stage, which happens for the IEEE 30 bus test case and the IEEE 39 bus test case with one switchable line. The branch-and-bound process made in BARON only slowly increases the lower bound, and barely refines the upper bound given in the preprocessing stage. On top of discrete binary variables for switching, OTS involves nonlinearities coming from bilinear products. All these factors can prevent general purpose solvers such as BARON from finding the optimal solution efficiently. On the other hand, the virtual-voltage approximation is built on provably accurate SDP relaxations of OPF and makes use of the specific structure of the OTS problem. Table III clearly shows the difference in performance between both methods.

C. Partition-based OTS algorithm

We examine here the performance of the partition-based algorithm. Table IV shows the result of implementing on the IEEE 118 and IEEE 300 bus test cases the following methods: (i) the virtual-voltage approximation with 40 switchable lines; and (ii) the partition-based OTS algorithm with 40 switchable lines; and (iii) the virtual-voltage approximation with all the lines switchable. For the case with 40 switchable lines, one can see that the partition-based OTS algorithm refines the virtual-voltage approximation. The solutions obtained by the virtual-voltage approximation and the partition-based OTS algorithm with 40 switchable lines are both close to the lower bound. In addition, SDP returns a rank-1 solution for every case with 40 switchable lines. We quantify their accuracy by upper bounding the percentage error with the true optimal solution using $(p_1^{\text{opt}} - p_{3,\text{all}}^{\text{opt}})/p_{3,\text{all}}^{\text{opt}} \cdot 100\%$ (this is justified by the fact that, if **(P1)**-α gives a rank-one solution, then the true optimal value belongs to the interval $[p_{3,\text{all}}^{\text{opt}}, p_1^{\text{opt}}]$).

We simulate case (iii) where all the lines are switchable to compare the results with the QC relaxation method [15].

		IEEE 30	IEEE 39	IEEE 57	IEEE 89	IEEE 118	IEEE 300
	# of switches	5	5	5	30	40	40
Virtual-voltage approximation	lower bound	1265 (1)	135003 (1)	50912 (1)	179325 (1)	151594 (3)	1086369 (1)
	upper bound	1270 (1)	135303 (1)	52267 (2)	188896 (1)	152707 (1)	1096117 (1)
McCormick relaxation w/ 5(p.u.) bounds	lower bound	1046 (1)	133591 (4)	50209 (1)	177643 (1)	146466 (1)	1054130 (2)
	upper bound	4654 (3)	135660 (1)	52267 (2)	230355 (1)	153477 (1)	NaN
McCormick relaxation w/ 1(p.u.) bounds	lower bound	1046 (1)	133591 (1)	50209 (1)	NaN	147919 (1)	NaN
	upper bound	NaN	135303 (1)	67667 (2)	NaN	152484 (1)	NaN

TABLE I: Performance of the virtual-voltage and the McCormick approximations. The integer value in parentheses is the rank of the SDP solutions. "NaN" means that CVX cannot find a feasible solution (even though the problem may still be feasible, see [32] for details).

		IEEE 30	IEEE 39	IEEE 57	IEEE 89	IEEE 118	IEEE 300
	# of switches	5	5	5	30	40	300
Virtual-voltage approximation	lower bound	7.78	9.55	18.72	59.01	149.37	2668
	upper bound	5.52	6.71	12.63	39.05	88.55	2356
McC. relaxation w/ 5(p.u.) bounds	lower bound	4.93	7.69	12.92	57.19	99.19	2394
	upper bound	5.1	5.8	12.62	44.48	88.28	2250
McC. relaxation w/ 1(p.u.) bounds	lower bound	6.24	5.8	13.34	94.51	104.18	2338
	upper bound	5.82	6.61	9.87	42.25	87.88	2311

TABLE II: Comparison of the computational time between McCormick and virtual-voltage approximations.

		IEEE 30		IEEE 39		IEEE 57	
# of switchable lines		5	10	1	5	1	5
Virtual-voltage approx.	lower bound	1265	1265	135295	135003	50997	50912
	upper bound	1269.9	1269.9	135303	135303	52267	52267
BARON	lower bound	$-2.2 \cdot 10^7$	$-2.3 \cdot 10^7$	$-2.7 \cdot 10^8$	$-2.6 \cdot 10^8$	$-6.8 \cdot 10^7$	$-6.9 \cdot 10^7$
	upper bound	1269.9	1269.9	135303	$3.4 \cdot 10^8$	$7.6 \cdot 10^8$	$7.6 \cdot 10^8$

TABLE III: Comparison between solutions of BARON and virtual-voltage approximation. We pre-select 5 switchable lines for all the test cases. The upper and lower bounds for BARON are, respectively, the best feasible solution found and the largest lower bound of the branches in the branch-and-bound method.

The QC relaxation method takes 10 hours to solve the OTS problem for both IEEE 118 and 300 bus test cases employing servers with 4334 CPUs and 64GB memory. The QC relaxation method converges to a near optimum with around 1% error for the IEEE 118 test case, while it is unable to provide a feasible solution for the IEEE 300 bus case. In contrast, solving **(P3)** with all lines switchable takes significantly less time for both cases (around 2min30sec and 1h30min, respectively) as shown in Table IV. Though we only retrieve a rank-2 solution from **(P3)** for IEEE 118 bus case, the condition number of the decision variable W is larger than 250, which translates to an error (or violation of constraints) less than 1%. Therefore, the virtual-voltage approximation leads to a solution of a similar quality as QC. The reason of not retrieving a rank-1 solution with the virtual-voltage approximation is that it provides high-rank solutions when the U matrix is introduced on every line (this is one of the reasons for pre-selecting a subset of lines).

D. OTS with Security Constraints

We also validate the virtual-voltage approximation method on OTS with security constraints. We simulate both IEEE 118 and 300 bus test systems with one contingency scenario and 40 pre-selected switchable lines. Table V shows the results for the IEEE 118 bus test system with security constraint on line $\{106, 107\}$. The combined computational time in computing the lower and upper bounds raises to 834 (395+438) seconds, which is longer than the one without security constraint (149.37 seconds in Table IV). For the IEEE 300 bus test case, the default SDPT3 algorithm of CVX cannot find a feasible solution even when we increase the maximum number of iterations to 600 (the default is 100), which took around 3.7 hours. In fact, the diagnosis report of the SDPT3 algorithm indicated that the algorithm diverged. We therefore tried the SeDuMi algorithm, also supported by CVX. Though [32] states that SDPT3 is more reliable then SeDuMi, in this case SeDuMi finds a solution while SDPT3 cannot. Table V shows the results for the IEEE 300 bus test system with security constraint on line $\{157, 158\}$. As in the IEEE 118 bus test case, we also observe that the computational times roughly doubled (9811.8 seconds in total) when a contingency is added.

	IEEE 118		IEEE 300	
	Optimal Value	Time (s)	Optimal Value	Time (s)
lower bound	156896 (5)	396	1731182 ($\overline{10}$)	5015.2
upper bound	158660 (1)	438	1854889 (4)	4796.6

TABLE V: Performance of the virtual-voltage approximation for OTS with security constraint on line $\{106, 107\}$ for IEEE 118 system and $\{157, 158\}$ for IEEE 300 system.

VII. Conclusions

We have considered OTS problems with security constraints. For these scenarios, the standard SDP relaxation of the problem remains non-convex because of the presence of bilinear terms and the discrete variables. We have proposed an approximation based on the introduction of virtual voltages to convexify the bilinear terms. We have also characterized several of its properties regarding physical interpretation and lower and upper bounds on the optimal value of the original problem. To handle the presence of the discrete vari-

		Optimal values		Bound on errors		Computation times (sec)	
		IEEE 118	IEEE 300	IEEE 118	IEEE 300	IEEE 118	IEEE 300
Virtual-voltage approximation w/ all lines switchable	lower bound	150791 ($\overline{10}$)	1090397 ($\overline{10}$)	0.36%	N/A	142.57	4859.62
	upper bound	151340 (2)	NaN				
Virtual-voltage approximation w/ 40 switchable lines	lower bound	151594 (1)	1086369 (1)	1.27%	0.52%	149.37	5023.90
	upper bound	152707 (1)	1096117 (1)				
Partition-based OTS w/ 40 switchable lines		152505 (1)	1092967 (1)	1.14%	0.24%	173	5055.88

TABLE IV: Performance of the virtual-voltage approximation and the partition-based OTS algorithm on the IEEE 118 and IEEE 300 bus test cases. The rank of the decision variable is shown in parenthesis. ($\overline{10}$) indicates that the rank of the decision variable is larger than 10.

ables, we have built on the virtual-voltage approximation to propose a graph partition-based algorithm that significantly reduces the computational complexity of solving the original problem. The high degree of accuracy and the reduction in computational complexity observed in simulations makes the proposed algorithms promising for OTS applications. We note here that all the simulations in the paper inherit common challenges of the SDP-based approach, including scalability and feasibility. Many works have attempted to address these issues (see for example [34] and references therein), but fully accommodating them remain an open problem. Our focus here is on incorporating discrete variables for OTS in the SDP formulation in a computationally efficient way. We expect that offline tree decomposition or chordal extensions can potentially address the complexity issues of SDP-based approaches, and combining those reduction techniques with the solution proposed here will be part of our future work.

Acknowledgments

This research was supported by the ARPA-e NODES program through Cooperative Agreement DE-AR0000695.

References

[1] C.-Y. Chang, S. Martinez, and J. Cortés, "Convex relaxation for mixed-integer optimal power flow problems," in *Allerton Conf. on Communications, Control and Computing*, Monticello, IL, 2017, pp. 307–314.

[2] F. E. R. Commission, "FY 2009-2014 strategic plan," 2013, https://www.ferc.gov/about/strat-docs/FY-09-14-strat-plan-print.pdf.

[3] 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*, Detroit, MI, 2011, electronic proceedings.

[4] 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.

[5] R. P. O'Neill, R. Baldick, U. Helman, M. H. Rothkopf, and W. Stewart, "Dispatchable transmission in RTO markets," *IEEE Transactions on Power Systems*, vol. 20, no. 1, pp. 171–179, 2005.

[6] E. B. Fisher, R. P. O'Neill, and M. C. Ferris, "Optimal transmission switching," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 1346–1355, 2008.

[7] K. W. Hedman, R. P. O'Neill, E. B. Fisher, and S. S. Oren, "Optimal transmission switchingsensitivity analysis and extensions," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 1469–1479, 2008.

[8] ——, "Optimal transmission switching with contingency analysis," *IEEE Transactions on Power Systems*, vol. 24, no. 3, pp. 1577–1586, 2009.

[9] 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.

[10] 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.

[11] S. Fattahi, J. Lavaei, and A. Atamtürk, "A bound strengthening method for optimal transmission switching in power systems," *arXiv preprint arXiv:1711.10428*, 2017.

[12] T. Potluri and K. W. Hedman, "Impacts of topology control on the ACOPF," in *IEEE Power and Energy Society General Meeting*, San Diego, CA, 2012, electronic proceedings.

[13] 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.

[14] 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.

[15] H. Hijazi, C. Coffrin, and P. Van Hentenryck, "Convex quadratic relaxations for mixed-integer nonlinear programs in power systems," *Mathematical Programming Computation*, vol. 9, no. 3, pp. 321–367, 2017.

[16] J. Mareček, M. Mevissen, and J. C. Villumsen, "MINLP in transmission expansion planning," in *Power Systems Computation Conference*, Genoa, Italy, 2016, electronic proceedings.

[17] E. Briglia, S. Alaggia, and F. Paganini, "Distribution network management based on optimal power flow: Integration of discrete decision variables," in *Annual Conference on Information Systems and Sciences*, Baltimore, MD, 2017, electronic proceedings.

[18] 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.

[19] F. Bullo, J. Cortés, and S. Martinez, *Distributed Control of Robotic Networks*, ser. Applied Mathematics Series. Princeton University Press, 2009.

[20] 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.

[21] S. Bose, D. F. Gayme, S. H. Low, and K. M. Chandy, "Optimal power flow over tree networks," in *Allerton Conf. on Communications, Control and Computing*, Monticello, IL, 2011, pp. 1342–1348.

[22] R. Madani, S. Sojoudi, and J. Lavaei, "Convex relaxation for optimal power flow problem: Mesh networks," *IEEE Transactions on Power Systems*, vol. 30, no. 1, pp. 199–211, 2015.

[23] 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.

[24] 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.

[25] D. Molzahn, J. T. Holzer, B. C. Lesieutre, and C. L. DeMarco, "Implementation of a large-scale optimal power flow solver based on semidefinite programming," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 3987–3998, 2013.

[26] 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.

[27] 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.

[28] J. P. Hespanha, "An efficient Matlab algorithm for graph partitioning," University of California, Santa Barbara, Tech. Rep., 2004.

[29] A. Abou-Rjeili and G. Karypis, "Multilevel algorithms for partitioning power-law graphs," in *International Parallel and Distributed Processing Symposium (IPDPS)*, Rhodes Island, Greece, 2006, electronic proceedings.

[30] M. Fiedler, "A property of eigenvectors of nonnegative symmetric matrices and its application to graph theory," *Czechoslovak Mathematical Journal*, vol. 25, no. 4, pp. 619–633, 1975.

[31] B. Mohar, Y. Alavi, G. Chartrand, and O. Oellermann, "The Laplacian

spectrum of graphs," *Graph theory, combinatorics, and applications*, vol. 2, no. 871-898, p. 12, 1991.
[32] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," Mar. 2014, available at http://cvxr.com/cvx.
[33] The Optimization Firm, "BARON," https://minlp.com/baron.
[34] D. K. Molzahn and I. A. Hiskens, "A survey of relaxations and approximations of the power flow equations," *Foundations and Trends in Electric Energy Systems*, vol. 4, no. 1-2, pp. 1–221, 2019.

Chin-Yao Chang received the B.S. degree in Physics from National Taiwan University, Taipei, Taiwan, R.O.C. He received the Ph.D. degree in Electrical and Computer Engineering at the Ohio State University, Columbus, OH, USA, in July 2016. He was a postdoctoral researcher with the Mechanical and Aerospace Engineering at the University of California, San Diego, CA, USA, from 2016 to 2018. He is now a Research Engineer of National Renewable Energy Laboratory (NREL), Golden, CO, USA. His research interests include demand response, distributed optimization, cyber physical systems, optimal power flow, microgrid control, and quantum computing.

Sonia Martínez (M'02-SM'07-F'18) is a Professor of Mechanical and Aerospace Engineering at the University of California, San Diego, CA, USA. She received the Ph.D. degree in Engineering Mathematics from the Universidad Carlos III de Madrid, Spain, in May 2002. Following a year as a Visiting Assistant Professor of Applied Mathematics at the Technical University of Catalonia, Spain, she obtained a Postdoctoral Fulbright Fellowship and held appointments at the Coordinated Science Laboratory of the University of Illinois, Urbana-Champaign during 2004, and at the Center for Control, Dynamical systems and Computation of the University of California, Santa Barbara during 2005. In a broad sense, her main research interests include the control of network systems, multi-agent systems, nonlinear control theory, and robotics. For her work on the control of underactuated mechanical systems she received the Best Student Paper award at the 2002 IEEE Conference on Decision and Control. She was the recipient of a NSF CAREER Award in 2007. For the paper "Motion coordination with Distributed Information," co-authored with Jorge Cortés and Francesco Bullo, she received the 2008 Control Systems Magazine Outstanding Paper Award. She has served on the editorial board of the European Journal of Control (2011-2013) and the Journal of Geometric Mechanics (2009-present), and currently serves as a Senior Editor of the IEEE Transactions on Control of Network Systems.

Jorge Cortés (M'02, SM'06, F'14) received the Licenciatura degree in mathematics from Universidad de Zaragoza, Zaragoza, Spain, in 1997, and the Ph.D. degree in engineering mathematics from Universidad Carlos III de Madrid, Madrid, Spain, in 2001. He held postdoctoral positions with the University of Twente, Twente, The Netherlands, and the University of Illinois at Urbana-Champaign, Urbana, IL, USA. He was an Assistant Professor with the Department of Applied Mathematics and Statistics, University of California, Santa Cruz, CA, USA, from 2004 to 2007. He is currently a Professor in the Department of Mechanical and Aerospace Engineering, University of California, San Diego, CA, USA. He is the author of Geometric, Control and Numerical Aspects of Nonholonomic Systems (Springer-Verlag, 2002) and co-author (together with F. Bullo and S. Martínez) of Distributed Control of Robotic Networks (Princeton University Press, 2009). He is a Fellow of IEEE and SIAM. At the IEEE Control Systems Society, he has been a Distinguished Lecturer (2010-2014), and is currently its Director of Operations and an elected member (2018-2020) of its Board of Governors. His current research interests include distributed control and optimization, network science, resource-aware control, nonsmooth analysis, reasoning and decision making under uncertainty, network neuroscience, and multi-agent coordination in robotic, power, and transportation networks.