Integrated Transmission-Distribution Multi-Period Switching for Wildfire Risk Mitigation: Improving Speed and Scalability with Distributed Optimization

Rachel Harris
Georgia Institute of Technology
Atlanta, GA, USA
rharris94@gatech.edu

Chin-Yao Chang
National Renewable Energy Laboratory
Golden, CO, USA
chinyao.chang@nrel.gov

Daniel K. Molzahn
Georgia Institute of Technology
Atlanta, GA, USA
molzahn@gatech.edu

***Abstract*—With increasingly severe wildfire conditions driven by climate change, utilities must manage the risk of wildfire ignitions from electric power lines. During "public safety power shutoff" events, utilities de-energize power lines to reduce wildfire ignition risk, which may result in load shedding. Distributed energy resources provide flexibility that can help support the system to reduce load shedding when lines are de-energized. We investigate a coordinated transmission-distribution optimization problem that balances wildfire risk mitigation and load shedding. We model distribution systems that include battery energy storage systems which may support loads when transmission lines are de-energized. This multi-period integrated transmission-distribution optimal switching problem jointly optimizes line switching decisions, the generators' setpoints, load shedding, and the batteries' states of charge, resulting in significant computational challenges. To improve scalability, we decompose the problem over both space and time and apply a distributed optimization algorithm. Using a large-scale synthetic California test case with realistic distribution models and real wildfire risk data, we show that distributed optimization can solve large-scale multi-period switching problems that are otherwise intractable for centralized solvers. We also discuss challenges and future directions for improving the distributed algorithm's convergence performance as the number of time periods increases.**

***Index Terms*—Integrated transmission-distribution, distributed optimization, wildfire mitigation, mixed-integer linear program**

I. Introduction

Wildfires are expected to become more frequent and severe as the climate changes. To reduce wildfire ignition risk from electric power line faults, utilities may de-energize power lines during high-risk periods, a practice known as "public safety power shutoffs" [1]. De-energizing lines may cause power outages, motivating solutions that balance wildfire risk reduction while minimizing power outages.

This work was authored in part by NREL, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by DOE Office of Electricity, Advanced Grid Modeling Program, through agreement NO. 33652. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work or allow others to do so, for the U.S. Government purposes.

This work was supported by the NSF AI Institute for Advances in Optimization (AI4OPT), #211253.

Prior work has explored optimal transmission switching to balance wildfire risk reductions with load shedding [2]. Subsequent work also considered equity to make power shut-offs more fair over time [3]. Recent work investigated mitigating wildfire risk in distribution networks through optimal network reconfiguration and microgrid formation [4]. Researchers have also accounted for equity while performing distribution network reconfiguration for wildfire risk mitigation [5], and introduced formulations which are robust to uncerainty [6].

To the best of our knowledge, this paper is the first to solve a transmission-distribution co-optimization problem for wildfire risk mitigation. Specifically, we optimize transmission line de-energizations while accounting for the flexibility provided by distributed energy resources (DERs) in distribution systems. The problem is most naturally formulated as a large-scale, multi-time-period mixed-integer linear program (MILP), which requires prohibitive amounts of time and memory to solve for state-of-the-art mixed-integer solvers. To improve scalability, we decompose the problem over both space and time and apply an ADMM-based distributed optimization algorithm.

We next provide a brief review on distributed algorithms for coordinated transmission-distribution optimization problems which consider multi-period dispatch or problems with integer variables. Several papers used the alternating direction method of multipliers (ADMM) algorithm to solve multi-period economic dispatch [7], [8]. Another paper used an enhanced ADMM algorithm to solve a coordinated transmission-distribution reserve scheduling problem [9], where the authors improve the convergence with an ADMM inner loop which operates with all integer variables fixed. One of the gaps seen in these studies is a lack of high-fidelity, realistic distribution system models. These papers used synthetic distribution networks which are assumed to be balanced so that only one phase is modeled, each containing less than 150 buses. The largest test case used for numerical results contained 3892 buses in total. In contrast, we test our methods on multi-phase distribution networks with thousands of buses each, and the full test case contains over 10,000 buses.

Although ADMM is not guaranteed to converge for non-convex problems, it may be used as a heuristic to solve mixed-

integer programming problems [10]–[12]. Instances of ADMM on mixed-integer programs may provide significant computational advantages compared to global solution methods like branch-and-bound, which suffer from exponential worst-case time complexity. We find that ADMM may solve large mixed-integer switching problems much more quickly than centralized solvers, yielding significant scalability advantages. However, when the number of time periods in our multi-period problem is sufficiently large, the distributed algorithm can fail to converge or may reach consensus at a sub-optimal solution. We provide associated discussion in Section IV.

Our contributions are as follows:

1) We formulate and solve a transmission-distribution co-optimization problem to select transmission line de-energizations which minimize wildfire risk and load shedding. Our paper is the first to coordinate transmission and distribution networks to optimize power shut-offs under wildfire risk.
2) We develop a scalable solution by decomposing the problem over space and time and applying an ADMM-based distributed algorithm. To the best of our knowledge, this paper is the first to use ADMM to solve optimal switching problems for wildfire risk mitigation.
3) We present numerical results from large-scale, realistic systems. We show that coordinated transmission-distribution optimization produces switching decisions that reduce wildfire ignition risk compared to solving a transmission-level switching problem alone. Our proposed distributed approach solves this optimal switching problem much more quickly than a centralized solver when we consider up to eight time periods, but both the centralized and distributed algorithms fail to converge to the optimal solution for problems with more than eight time periods. We discuss future directions for improving the distributed algorithm's convergence performance.

We organize the remainder of the paper as follows. Section II formulates the optimal switching problem. Section III details the decomposition and distributed solution methods. Section IV presents results showing the improved performance from transmission-distribution coordination and computational advantages of the distributed approach. Finally, Section V concludes the paper and discusses future work.

II. Problem Formulation

We formulate the optimal switching for wildfire risk mitigation problem to minimize a weighted sum of wildfire ignition risk and load shedding over multiple time periods. The constraints consist of the power flow physics and engineering limits which are typical of optimal power flow, augmented to support component shut-offs. The control variables are transmission line energization statuses, transmission generator setpoints, distribution bus load shed, and distribution storage system charge/discharge setpoints. The problem, ideally formulated with nonlinear AC power flow equations and discrete variables for energization statuses, is computationally intractable as a mixed-integer nonlinear program. The linearized LinDistFlow power flow approximation [13] is widely used to model power flow in distribution systems and, while less common, can also be used to approximate transmission system power flow models as well.

Throughout this section, transmission components are denoted with the superscript $\mathcal{H}$ and distribution components with the superscript $\mathcal{D}$. Time periods are indexed by $t \in \mathcal{T} = \{1, 2, ..., T\}$. Power flow in the transmission network is modeled with a balanced single-phase LinDistFlow approximation, while we use an unbalanced three-phase LinDistFlow model for the distribution network. Following [14], we include constraints that balance the power flow and match voltage magnitudes at the transmission-distribution boundary.

A. Transmission Constraints

Consider a transmission network with a set of buses $\mathcal{N}^{\mathcal{H}}$ and a set of lines $\mathcal{L}^{\mathcal{H}}$. The active and reactive power flows through line (i,k) at time t are $p_{ik,t}$ and $q_{ik,t}$, respectively. The binary variables $\ell_{ik,t}$ represent the energization status at time $t \in \mathcal{T}$ of the transmission line $(i,k) \in \mathcal{L}^{\mathcal{H}}$ connecting bus i to bus k. Operational limits restrict the amount of power flow across lines. The lower bounds on active and reactive power flows are $\underline{p}_{ik}$ and $\underline{q}_{ik}$, while the upper bounds on active and reactive power flows are $\overline{p}_{ik}$ and $\overline{q}_{ik}$, respectively. The line flow is then defined as

$$\begin{aligned} \underline{p}_{ik}\ell_{ik,t} \le p_{ik,t} \le \overline{p}_{ik}\ell_{ik,t} \quad \forall (i,k) \in \mathcal{L}^{\mathcal{H}},\ \forall t \in \mathcal{T} \\ \underline{q}_{ik}\ell_{ik,t} \le q_{ik,t} \le \overline{q}_{ik}\ell_{ik,t} \quad \forall (i,k) \in \mathcal{L}^{\mathcal{H}},\ \forall t \in \mathcal{T} \end{aligned} \tag{1}$$

so that the active and reactive power flows $p_{ik,t}$ and $q_{ik,t}$ must be within their operational limits if line (i,k) is energized, and must be 0 if line (i,k) is de-energized.

Next, we introduce the notation $w_{i,t}$ to represent the squared voltage magnitude at bus i at time t. Also note that the line resistance is r_{ik} and the line reactance is x_{ik}. For all $(i,k) \in \mathcal{L}^{\mathcal{H}}$ and $t \in \mathcal{T}$, the voltage drop across line (i,k) is given as

$$\begin{aligned} 2(r_{ik}p_{ik,t} + x_{ik}q_{ik,t}) + (1-\ell_{ik,t})\underline{M} \le w_{i,t} - w_{k,t} \\ w_{i,t} - w_{k,t} \le 2(r_{ik}p_{ik,t} + x_{ik}q_{ik,t}) + (1-\ell_{ik,t})\overline{M} \end{aligned} \tag{2}$$

where $\underline{M}$ and $\overline{M}$ are big-M constants computed from voltage magnitude limits. This constraint ensures that if line (i,k) is energized, the voltage magnitude difference between buses i and k follows the LinDistFlow equations, but if line (i,k) is de-energized, then there is no prescribed relationship between the voltage magnitudes at buses i and k. To see how to compute the big-M constants, note that if the line is energized, i.e., $\ell_{ik,t} = 1$, then the terms with big-M constants vanish from both inequalities and the voltage magnitude difference follows the LinDistFlow equations. However, if the line is de-energized and thus $\ell_{ik,t} = 0$, then we must allow $w_{k,t} - w_{i,t}$ to take on any possible values. Since voltage magnitudes at every bus are constrained by (5), we can easily compute $\underline{M} = \underline{V}_i^2 - \overline{V}_k^2$ and $\overline{M} = \overline{V}_i^2 - \underline{V}_k^2$. Note that by (1) we know that if $\ell_{ik,t} = 0$ then $p_{ik,t} = q_{ik,t} = 0$, so we need not consider the term $2(r_{ik}p_{ik,t} + x_{ik}q_{ik,t})$ when computing the big-M constants.

For all $t \in \mathcal{T}$, we apply power balance constraints:

$$\begin{aligned} p_{ik,t} &= -P_{k,t} + \sum_{m:k\to m} p_{km,t} \quad \forall (i,k) \in \mathcal{L}^{\mathcal{H}} \\ q_{ik,t} &= -Q_{k,t} + \sum_{m:k\to m} q_{km,t} \quad \forall (i,k) \in \mathcal{L}^{\mathcal{H}} \end{aligned} \tag{3}$$

Here, $P_{i,t}$ and $Q_{i,t}$ represent the active and reactive power injections, respectively, at bus i.

At each bus $i \in \mathcal{N}^{\mathcal{H}}$, the injected power is given by

$$\begin{aligned} P_{i,t} &= \sum_{m\in\mathcal{G}_i} P^g_{m,t} - \sum_{m\in\mathcal{D}_i} P^d_{m,t} \\ Q_{i,t} &= \sum_{m\in\mathcal{G}_i} Q^g_{m,t} - \sum_{k\in\mathcal{D}_i} Q^d_{m,t} \end{aligned} \tag{4}$$

where $P^g_{m,t}$ and $Q^g_{m,t}$ denote the active and reactive power generated at generator m at time t, and $P^d_{m,t}$ and $Q^d_{m,t}$ denote the active and reactive power consumed at load m at time t. Also, $\mathcal{D}_i$ denotes the set of loads at bus i and $\mathcal{G}_i$ denotes the set of generators at bus i.

We also bound the voltage magnitudes at all buses:

$$\underline{V}_i^2 \le w_{i,t} \le \overline{V}_i^2 \quad \forall i \in \mathcal{N}^{\mathcal{H}},\ \forall t \in \mathcal{T} \tag{5}$$

We also want to ensure that the transmission network remains connected after switching. To do so, we use a network flow formulation that introduces an artificial commodity as in [15]. We supply the reference bus with $\mathcal{N}^{\mathcal{H}} - 1$ units of the commodity, and set a constraint that all buses other than the reference bus must consume one unit of the commodity. Let $\mathcal{S}$ denote the one-element set containing the reference bus. We also introduce artificial flow variables $f_{ik,t}$ for each branch (i,k). The constraint for artificial commodity flow balance at each node other than the reference bus is

$$\sum_{k:i\to k} f_{ik,t} - \sum_{k:k\to i} f_{ik,t} = 1 \quad \forall i \in \mathcal{N}^{\mathcal{H}} \setminus \mathcal{S}, \quad \forall t \in \mathcal{T} \tag{6}$$

In addition, we ensure that no artificial commodity can flow across a de-energized line:

$$\begin{aligned} &\forall t \in \mathcal{T}: \\ &-(|\mathcal{N}^{\mathcal{H}}|-1)\ell_{ik,t} \le f_{ik,t} \le (|\mathcal{N}^{\mathcal{H}}|-1)\ell_{ik,t} \quad \forall (i,k) \in \mathcal{L}^{\mathcal{H}} \end{aligned} \tag{7}$$

By requiring each non-reference bus to consume one unit of this artificial commodity and restricting flows of the artificial commodity to the energized lines only, these constraints ensure that there exists some path across energized lines from the reference bus to every other bus in the network. Therefore, the network will remain connected after switching if these artificial flow constraints are imposed. Note that these artificial flows are a simple way of ensuring connectivity, but more sophisticated methods which minimize the number of constraints needed to maintain network connectedness during optimal transmission switching have been explored in, e.g., [16] and could be applied in our formulation.

B. Distribution Constraints

To account for the unbalanced nature of distribution systems, we use a three-phase model for power flow. We consider distribution systems with battery storage systems, which we model for simplicity as ideal batteries with perfect efficiency. We also consider load shedding, which we model as a continuous load shed at every bus containing loads.

Note that we use a bold notation to indicate a vector of variables which contains values for all phases of the bus or line. For example, for a three-phase bus,

$$\boldsymbol{w}_{i,t} = \begin{bmatrix} w^a_{i,t} & w^b_{i,t} & w^c_{i,t} \end{bmatrix}^T$$

Using the same notation as in Section II-A, we denote the squared voltage magnitudes at bus i at time t as $\boldsymbol{w}_{i,t}$. The active and reactive power flows across line (i,k) are $\boldsymbol{p}_{ik,t}$ and $\boldsymbol{q}_{ik,t}$. The resistance is $\boldsymbol{r}_{ik}$ and the reactance is $\boldsymbol{x}_{ik}$. When modeling all phases in the distribution network, $\boldsymbol{r}_{ik}$ and $\boldsymbol{x}_{ik}$ are matrices with terms that reflect self-impedance as well as mutual impedances between phases. For all $(i,k) \in \mathcal{L}^{\mathcal{D}}$ and $t \in \mathcal{T}$, the difference in squared voltage magnitudes between buses i and k is

$$\boldsymbol{w}_{k,t} = \boldsymbol{w}_{i,t} - \boldsymbol{M}^P_{ik,t}\boldsymbol{p}_{ik,t} - \boldsymbol{M}^Q\boldsymbol{q}_{ik,t} \tag{8}$$

where we have that

$$\boldsymbol{\Gamma} = \begin{bmatrix} 1 & \alpha^2 & \alpha \\ \alpha & 1 & \alpha^2 \\ \alpha^2 & \alpha & 1 \end{bmatrix}$$

for $\alpha = \exp(-j\frac{2\pi}{3})$ and

$$\begin{aligned} \boldsymbol{M}^P &= 2\big(\Re(\boldsymbol{\Gamma}) \odot \boldsymbol{r}_{ik} + \Im(\boldsymbol{\Gamma}) \odot \boldsymbol{x}_{ik}\big) \\ \boldsymbol{M}^Q &= 2\big(\Re(\boldsymbol{\Gamma}) \odot \boldsymbol{x}_{ik} - \Im(\boldsymbol{\Gamma}) \odot \boldsymbol{r}_{ik}\big) \end{aligned}$$

where we denote $\boldsymbol{A} \odot \boldsymbol{B}$ as the element-wise product of $\boldsymbol{A}$ and $\boldsymbol{B}$. Also note that $j = \sqrt{-1}$, and $\Re$ and $\Im$ are the real and imaginary part operators, respectively.

For all $t \in \mathcal{T}$, power balance is given by

$$\begin{aligned} \boldsymbol{p}_{ik,t} &= -\boldsymbol{P}_{k,t} + \sum_{m:k\to m} \boldsymbol{p}_{km,t} \quad \forall (i,k) \in \mathcal{L}^{\mathcal{D}} \\ \boldsymbol{q}_{ik,t} &= -\boldsymbol{Q}_{k,t} + \sum_{m:k\to m} \boldsymbol{q}_{km,t} \quad \forall (i,k) \in \mathcal{L}^{\mathcal{D}} \end{aligned} \tag{9}$$

where $\boldsymbol{P}_{k,t}$ and $\boldsymbol{Q}_{k,t}$ represent active and reactive power injections at bus k.

The formulation also allows load shedding, which may be necessary given transmission line de-energizations. Previous papers on transmission line switching generally model a continuous load shed at bulk transmission loads [2], [3]. We model a continuous load shedding at individual distribution loads. In practice, load shedding would be actuated by opening switches in the distribution networks. Our future work includes more accurately modeling load shedding by making binary switching decisions to de-energize blocks of loads in the distribution network. For the purposes of this paper, we use a continuous variable $s_i \in [0,1]$ to denote the proportion of load served at bus i. For example, if $s_i = 0.8$, then 20% of

the load at bus i is shed. The power injections at each bus account for storage devices and load shedding at that bus:

$$\begin{aligned} \boldsymbol{P}_{i,t} &= \sum_{m \in \mathcal{S}_i} \boldsymbol{P}^s_{m,t} - s_i \sum_{m \in \mathcal{D}_i} \boldsymbol{P}^d_{m,t} \\ \boldsymbol{Q}_{i,t} &= \sum_{m \in \mathcal{S}_i} \boldsymbol{Q}^s_{m,t} - s_i \sum_{m \in \mathcal{D}_i} \boldsymbol{Q}^d_{m,t} \end{aligned} \tag{10}$$

Here, $\boldsymbol{P}^s_{m,t}$ and $\boldsymbol{Q}^s_{m,t}$ represent active and reactive power injections from storage device m at time t. Similarly, $\boldsymbol{P}^d_{m,t}$ and $\boldsymbol{Q}^d_{m,t}$ represent the active and reactive power demands from load m at time t.

We also limit the voltage magnitudes at bus i as

$$\underline{\boldsymbol{V}}_i^2 \leq \boldsymbol{w}_i \leq \overline{\boldsymbol{V}}_i^2 \quad \forall i \in \mathcal{N}^{\mathcal{D}}, \quad \forall t \in \mathcal{T} \tag{11}$$

where $\underline{\boldsymbol{V}}_i^2$, $\overline{\boldsymbol{V}}_i^2$ are vectors of the lower and upper bounds, respectively, on the squared voltage magnitudes of each phase at bus i.

Each distribution network may contain battery energy storage systems. We use an ideal storage model with perfectly efficient charging and discharging. The variable C_i is positive when the storage system is charging and negative when it is discharging. The state of charge E_i is bounded by its energy capacity $\overline{E}_i$. There are also bounds on the charging/discharging power, where $\underline{C}_i$ denotes the lower bound that limits discharging power and $\overline{C}_i$ is the upper bound which limits its charging power. We assume that the charging/discharging power C_i is fixed during each time period $t \in \mathcal{T}$:

$$\boldsymbol{E}_{i,t} - \boldsymbol{E}_{i,t-1} = \boldsymbol{C}_i \qquad \forall i \in \mathcal{N}^{\mathcal{D}} \tag{12}$$

$$0 \leq \boldsymbol{E}_{i,t} \leq \overline{\boldsymbol{E}_i} \qquad \forall i \in \mathcal{N}^{\mathcal{D}} \tag{13}$$

$$\underline{\boldsymbol{C}}_i \leq \boldsymbol{C}_i \leq \overline{\boldsymbol{C}}_i \qquad \forall i \in \mathcal{N}^{\mathcal{D}} \tag{14}$$

C. Transmission-Distribution Boundary

We collect the buses at transmission-distribution boundaries into a set β containing tuples $(b^{\mathcal{H}}, b^{\mathcal{D}}) \in \beta$ where $b^{\mathcal{H}}$ is the transmission bus and $b^{\mathcal{D}}$ is the distribution bus. On the transmission side, we have modeled single-phase equivalent power flow from each transmission boundary bus $b^{\mathcal{H}}$ to the corresponding distribution boundary bus $b^{\mathcal{D}}$. On the distribution side, we have modeled three-phase power flow from the distribution boundary bus to its corresponding transmission bus. To ensure that these power flows are consistent, we impose the following constraints:

$$\begin{aligned} &\forall (b^{\mathcal{H}}, b^{\mathcal{D}}) \in \beta, \ \forall t \in \mathcal{T}: \\ &p^{\mathcal{H}}_{b^{\mathcal{H}} b^{\mathcal{D}},t} + \sum_{\phi \in \Phi_{b^{\mathcal{D}}}} p^{\mathcal{D},\phi}_{b^{\mathcal{D}} b^{\mathcal{H}},t} = 0 \\ &q^{\mathcal{H}}_{b^{\mathcal{H}} b^{\mathcal{D}},t} + \sum_{\phi \in \Phi_{b^{\mathcal{D}}}} q^{\mathcal{D},\phi}_{b^{\mathcal{D}} b^{\mathcal{H}},t} = 0, \end{aligned} \tag{15}$$

where Φ_i is the set of phases at a given bus i. The constraints in (15) ensure that the transmission model's power flow across the boundary toward the distribution side is equal and opposite to the sum across all phases in Φ_i of the distribution model's power flow across the boundary toward the transmission side. We also require that voltage magnitudes on either side of the boundary bus are equal:

$$w^{\mathcal{H}}_{b^{\mathcal{H}},t} = w^{\mathcal{D},\phi}_{b^{\mathcal{D}},t}, \quad \forall \phi \in \Phi_{b^{\mathcal{D}}}, \ \forall (b^{\mathcal{H}}, b^{\mathcal{D}}) \in \beta, \ \forall t \in \mathcal{T} \tag{16}$$

This formulation results in balanced voltage magnitudes at distribution substations, an assumption made by the integrated transmission-distribution software in our experiments [14].

D. Objective Function

To these constraints, we add an objective function inspired by [2], [3], which balances wildfire risk with load shedding:

$$\begin{aligned} C(\boldsymbol{\ell}, \boldsymbol{s}) = \sum_{t \in \mathcal{T}} \Big[&\gamma \sum_{(i,k) \in \mathcal{L}^{\mathcal{H}}} \rho_{ik,t} \ell_{ik,t} \\ &+ (1-\gamma) \big(\sum_{i \in \mathcal{N}^{\mathcal{D}}} s_i \sum_{\phi \in \Phi_i} -P^{d,\phi}_{i,t} \big) \Big] \end{aligned} \tag{17}$$

Here, ρ_{ik} is the wildfire ignition risk for line (i,k) at time t, and $P^{d,\phi}_{i,t}$ is the active power demand at phase ϕ of bus i during time period t. Also, the parameter γ allows controlling the tradeoff between load shedding and wildfire risk. The cost function depends on both $\boldsymbol{\ell}$, the vector of all transmission line energization statuses, and $\boldsymbol{s}$, the vector of all bus load sheds.

To summarize, the optimal switching problem for wildfire risk mitigation is the following mixed-integer linear program:

$$\begin{aligned} \min_{\substack{\boldsymbol{\ell}^{\mathcal{H}}, \boldsymbol{p}^{\mathcal{H}}, \boldsymbol{q}^{\mathcal{H}}, \boldsymbol{w}^{\mathcal{H}}, \\ \boldsymbol{s}^{\mathcal{D}}, \boldsymbol{p}^{\mathcal{D}}, \boldsymbol{q}^{\mathcal{D}}, \boldsymbol{w}^{\mathcal{D}}, \boldsymbol{C}^{\mathcal{D}}, \boldsymbol{E}^{\mathcal{D}}}} \quad & (17) \\ \text{s.t.} \quad & (1) - (16) \end{aligned} \tag{18}$$

III. Decomposition and Distributed Optimization

To solve the large-scale optimal switching problem for wildfire risk mitigation, we decompose the problem across both time and space using an ADMM-based distributed optimization algorithm. In Figure 1, we show how a small example system is decomposed, where each block represents a separate subproblem. The dashed lines represent coupling constraints between the central transmission network and its attached distribution networks at each time period.

We solve separate problems for the transmission network at each time period to manage complexity from the many binary transmission switching variables. We do not decompose distribution network subproblems over time because distribution energy storage systems introduce state-of-charge constraints that couple time periods. Decomposing distribution networks over time would thus result in thousands of coupling constraints between subproblems, slowing ADMM convergence such that decomposing these subproblems over time would be counterproductive. Instead, the objective and constraints for each distribution network across all time periods form individual subproblems.

The ADMM algorithm [17] solves the general problem

$$\begin{aligned} \min_{\boldsymbol{x} \in \mathcal{X}, \boldsymbol{z} \in \mathcal{Z}} \quad & f(\boldsymbol{x}) + g(\boldsymbol{z}) \\ \text{s.t.} \quad & A\boldsymbol{x} + B\boldsymbol{z} = \boldsymbol{c} \end{aligned} \tag{19}$$

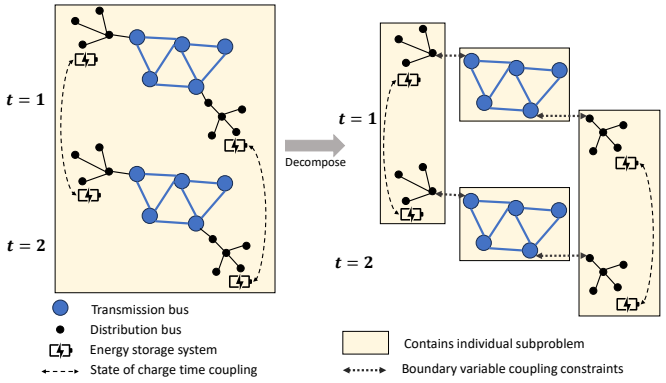

Fig. 1. This figure shows the decomposition of the multi-period problem. Dashed lines show the coupling between energy storage systems' states of charge across time periods. After decomposition, each transmission network at each time period forms a subproblem. Each distribution network across all time periods forms a subproblem. Dotted lines show the coupling constraints between subproblems' copies of boundary variables.

Now, the optimal switching problem for wildfire risk mitigation has been decomposed into subproblems. For transmission subproblems, the set of time periods $\mathcal{T}$ contains only one element, so there are $|\mathcal{T}|$ transmission subproblems. These transmission subproblems take the form

$$\begin{aligned} \min_{\substack{\boldsymbol{\ell}^{\mathcal{H}},\boldsymbol{p}^{\mathcal{H}},\boldsymbol{q}^{\mathcal{H}},\boldsymbol{w}^{\mathcal{H}},\\ \boldsymbol{p}^{\mathcal{D}},\boldsymbol{q}^{\mathcal{D}}}} \quad & \gamma \sum_{t\in\mathcal{T}} \sum_{(i,k)\in\mathcal{L}^{\mathcal{H}}} \rho_{ik,t}\ell_{ik,t} \\ \text{s.t.} \quad & (1)-(7) \end{aligned} \tag{20}$$

The distribution subproblems cover all time periods in $\mathcal{T}$, and the number of distribution subproblems is equal to the number of distinct distribution networks. These distribution subproblems take the form

$$\begin{aligned} \min_{\substack{\boldsymbol{p}^{\mathcal{H}},\boldsymbol{q}^{\mathcal{H}},\boldsymbol{s}^{\mathcal{D}},\\ \boldsymbol{p}^{\mathcal{D}},\boldsymbol{q}^{\mathcal{D}},\boldsymbol{w}^{\mathcal{D}},\boldsymbol{C}^{\mathcal{D}},\boldsymbol{E}^{\mathcal{D}}}} \quad & (1-\gamma)\sum_{t\in\mathcal{T}}\Big(\sum_{i\in\mathcal{N}^{\mathcal{D}}} s_i \sum_{\phi\in\Phi_i} -P_{i,t}^{d,\phi}\Big) \\ \text{s.t.} \quad & (8)-(14) \end{aligned} \tag{21}$$

To model the interaction between transmission and distribution subproblems, we introduce virtual buses at their boundaries. For transmission subproblems, the distribution network is represented by a three-phase virtual bus with a generator that can inject or consume unlimited active and reactive power. Similarly, each distribution subproblem includes a copy of the transmission boundary bus with a virtual generator.

To write the decomposed optimal power shut-off problem in ADMM form (19), we collect the variables belonging to every subproblem in a vector $\boldsymbol{x}$. That is, $\boldsymbol{x}$ contains variables from transmission subproblems representing the lines' power flows $\boldsymbol{p}^{\mathcal{H}}, \boldsymbol{q}^{\mathcal{H}}$, the lines' energization statuses $\boldsymbol{\ell}$, the buses' squared voltages $\boldsymbol{w}^{\mathcal{H}}$, the buses' power injections $\boldsymbol{P}^{\mathcal{H}}, \boldsymbol{Q}^{\mathcal{H}}$, and artificial flows for connectivity $\boldsymbol{f}$. From the distribution subproblems, $\boldsymbol{x}$ also contains variables representing the lines' power flows $\boldsymbol{p}^{\mathcal{D}}, \boldsymbol{q}^{\mathcal{D}}$, the buses' squared voltages $\boldsymbol{w}^{\mathcal{D}}$, the buses' power injections $\boldsymbol{P}^{\mathcal{D}}, \boldsymbol{Q}^{\mathcal{D}}$, the buses' load sheds $\boldsymbol{s}^{\mathcal{D}}$, the battery energy storage systems' states of charge $\boldsymbol{E}^{\mathcal{D}}$, and the batteries' charging/discharging power $\boldsymbol{C}^{\mathcal{D}}$. Here we have indexed variables with $\mathcal{H}$ and $\mathcal{D}$ to indicate whether they belong to the transmission or distribution systems, respectively. We must also include in $\boldsymbol{x}$ the voltage magnitude and power injection variables at *virtual* buses in the subproblems.

Next, we introduce a "central" copy of the variables at the virtual buses, and gather these central variables into a vector $\boldsymbol{z}$. This formulation with "central" variables allows us to use the ADMM algorithm, which is designed to optimize over two sets of variables with simple coupling constraints, to solve a problem with parallel subproblems over many regions of the power system. Let the tilde notation indicate central variables in the vector $\boldsymbol{z}$, so that, for example, $\tilde{w}_m$ is the central copy of the squared voltage magnitude at bus m. For a transmission subproblem with boundary bus m and virtual distribution boundary bus n', the constraints are as follows:

$$\begin{aligned} & w_m = \tilde{w}_m \\ & w_{n'_\phi} = \tilde{w}_{n_\phi} \quad \forall \phi \in \Phi_n \end{aligned} \tag{22}$$

$$\begin{aligned} & p_{mn'} = \tilde{p}_{mn}, \quad q_{mn'} = \tilde{q}_{mn} \\ & p_{n'_\phi m_\phi} = \tilde{p}_{n_\phi m_\phi}, \quad q_{n'_\phi m_\phi} = \tilde{q}_{n_\phi m_\phi} \quad \forall \phi \in \Phi_n \end{aligned} \tag{23}$$

Similarly, for a distribution subproblem with boundary bus n and virtual transmission boundary bus m', the constraints are:

$$\begin{aligned} & w_{m'} = \tilde{w}_m \\ & w_{n_\phi} = \tilde{w}_{n_\phi} \quad \forall \phi \in \Phi_n \end{aligned} \tag{24}$$

$$\begin{aligned} & p_{m'n} = \tilde{p}_{mn}, \quad q_{m'n} = \tilde{q}_{mn} \\ & p_{n_\phi m'_\phi} = \tilde{p}_{n_\phi m_\phi}, \quad q_{n_\phi m'_\phi} = \tilde{q}_{n_\phi m_\phi} \quad \forall \phi \in \Phi_n \end{aligned} \tag{25}$$

The constraints (22) and (24) ensure that the transmission and distribution subproblems agree on voltage magnitudes at the boundary, while the constraints (23) and (25) ensure that the transmission and distribution subproblems agree on power flows across the boundary.

The ADMM algorithm augments the Lagrangian function for (19) with a penalty term. The augmented term typically penalizes the squared ℓ_2-norm of the coupling constraint violations, $||A\boldsymbol{x}+B\boldsymbol{z}-\boldsymbol{c}||_2^2$. However, we find that subproblem numerics are improved by penalizing the ℓ_1-norm of the coupling constraint violations instead. Quadratic programs may experience more numerical problems compared to linear programs during execution of Gurobi's simplex or interior-point algorithms [18]. Therefore, we run the ADMM algorithm on the ℓ_1-norm augmented Lagrangian:

$$\begin{aligned} L_\alpha(\boldsymbol{x},\boldsymbol{z},\boldsymbol{\lambda}) = f(\boldsymbol{x}) + g(\boldsymbol{z}) + \boldsymbol{\lambda}^T(A\boldsymbol{x}+B\boldsymbol{z}-\boldsymbol{c}) \\ + \alpha||A\boldsymbol{x}+B\boldsymbol{z}-\boldsymbol{c}||_1 \end{aligned} \tag{26}$$

Here, $\boldsymbol{\lambda}$ contains the dual variables for each coupling constraint in (19). The penalty parameter α is user-selected.

At each iteration k, the ADMM algorithm minimizes over the $\boldsymbol{x}$ and $\boldsymbol{z}$ variables separately, while holding all other

variables fixed, and then updates the dual variables. The variable updates at iteration k are as follows:

$$\boldsymbol{x}^{k+1} = \arg\min_{\boldsymbol{x}} L_\alpha(\boldsymbol{x}, \boldsymbol{z}^k, \boldsymbol{\lambda}^k) \tag{27}$$

$$\boldsymbol{z}^{k+1} = \arg\min_{\boldsymbol{z}} L_\alpha(\boldsymbol{x}^{k+1}, \boldsymbol{z}, \boldsymbol{\lambda}^k) \tag{28}$$

$$\boldsymbol{\lambda}^{k+1} = \boldsymbol{\lambda}^k + \alpha(A\boldsymbol{x} + B\boldsymbol{z} - \boldsymbol{c}) \tag{29}$$

The $\boldsymbol{x}$-update step corresponds to solving all transmission and distribution subproblems in parallel. The transmission and distribution subproblems are in (20) and (21), respectively, but the relaxed coupling constraints augmented with the ℓ_1-norm penalty as shown in (26) are added to the objective. The transmission subproblems are mixed-integer linear programs, while the distribution subproblems are linear programs. Transmission subproblems cover single time periods and scale according to the number of buses and branches in the system. Distribution subproblems cover all time periods and scale according to the number of buses, lines, switches, and storage systems in the system, where separate variables and constraints are needed for each phase of distribution system components.

IV. Case Study

We construct a test case using the CATS (California Test System) [19] and the SMART-DS synthetic distribution networks [20]. We select eight distribution networks from the SMART-DS San Francisco dataset, and then extract a portion of the California test system in the San Francisco region which contains sufficient generation capacity to supply loads in the distribution networks. We place battery energy storage systems randomly at 30% of the low-voltage nodes in the distribution networks, each with energy capacity randomly selected between 45 and 100 kilowatt-hours. Next, we modify the distribution networks by aggregating the loads at voltage levels below 7 kV, so that the full test case contains buses and lines at voltage levels from 7 kV through 230 kV. When reducing the distribution networks, we also aggregate the energy capacity of storage devices at low-voltage nodes and place the aggregate energy storage system at the primary side of the distribution transformer. We show a plot of the test case in Figure 2. The transmission system consists of 155 buses, 15 generators, and 171 branches. The distribution system, across all networks, contains 15,083 buses, 15,496 lines and 1203 battery energy storage systems. Since optimizing over the unbalanced distribution system requires modeling each phase separately, we also note that there are 24,823 individual phases across all buses and 25,942 phases across all lines.

To obtain the values for the wildfire ignition risk for each transmission line, we select the "high-risk cumulative" metric from [21]. This study extracts the wind-enhanced fire potential index (WFPI) as computed by the United States Geological Survey [22] for regions across the power network and identifies which regions each transmission line crosses to compute the risk. Further details are available in [21].

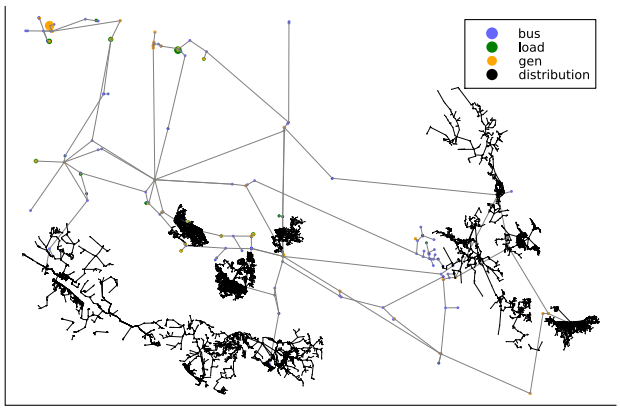

Fig. 2. Synthetic test case near San Francisco, California for integrated transmission-distribution optimization.

A. Comparing Central and Distributed Computation Time

We compare the performance of our distributed solution method with a state-of-the-art solver, Gurobi. All experiments ran on the Georgia Institute of Technology PACE high-performance computing cluster. Each experiment consisted of solving the optimal switching problem for wildfire risk mitigation over some number of time periods $\mathcal{T}$, and used one PACE CPU compute node equipped with two 12-core 2.7 GHz processors and 192 GB RAM. Any algorithm that did not converge within a 48-hour time period was terminated.

The central method solves the full integrated transmission-distribution (ITD) MILP directly with Gurobi. Our distributed method decomposes the problem as described in Section III and uses ADMM to reach the solution, where individual subproblems are solved with Gurobi. We terminate the central solution method when the MIP gap reaches 0.1%, and also solve distributed subproblems to a 0.1% MIP gap. We terminate the distributed algorithm when the norm of the mismatches between subproblems' shared variables falls below $\epsilon = 10^{-4}$.

Figure 3 shows the runtimes for each method as we increase the number of periods in the multi-period problem. With four or fewer time periods, the centralized method outperforms the distributed method. However, when the problem is sufficiently large, we see that the distributed method performs much better than the centralized method (e.g., approximately an order-of-magnitude faster with eight periods). In the future, we could potentially achieve further computational speedups by distributing suproblems across multiple computing nodes in addition to the parallelization across computing cores as in this paper.

For nine or more time periods, the central method fails to converge within 48 hours, at which point the MIP gap is 54.1% or higher. The distributed method also fails to find the optimal solution for nine or more time periods, converging to a sub-optimal solution which may be more costly than simply solving a transmission switching problem without accounting for distribution network flexibility. We discuss the problem of sub-optimality further below in Section IV-C.

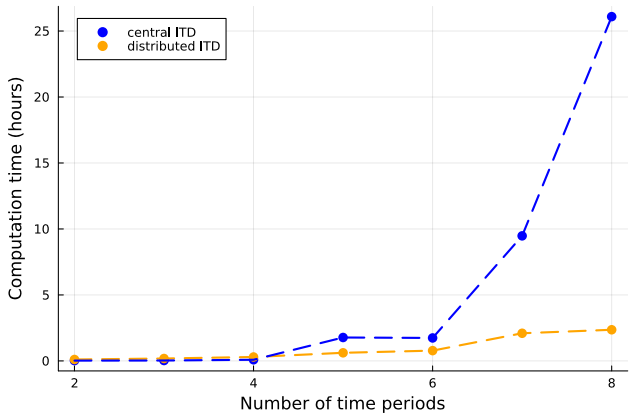

Fig. 3. Comparing computation time for the central solution method using the Gurobi solver vs. our distributed solution method.

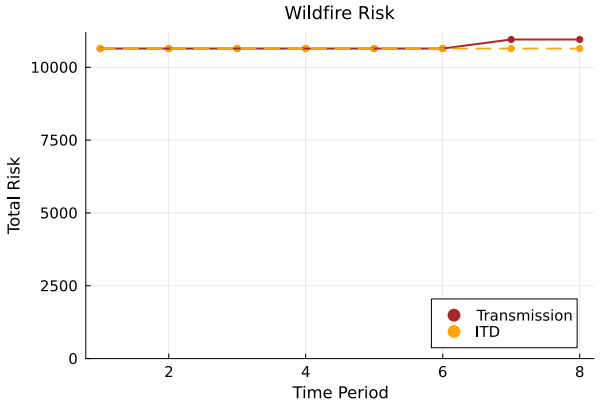

Fig. 4. Comparing the total wildfire risk, i.e., the sum of risk values on all energized lines, between the transmission-only problem and the ITD problem.

B. Effect of Coordinated Transmission-Distribution Switching

We next show how solving a coordinated transmission-distribution optimal switching problem produces transmission switching decisions that reduce wildfire ignition risk more than solving only a transmission-level problem. The "transmission-only" problem is formulated by neglecting all energy storage systems in the distribution networks. For each distribution network, the loads are aggregated and placed at the distribution boundary bus. We solve (18) over this simplified network and compare the objective cost with that of the ITD problem. We present results for the optimal switching problem solved over eight time periods, the highest number of periods for which the distributed algorithm reached the optimal solution. Figure 4 shows the total wildfire risk, i.e., the sum of the risk values on all energized lines, at each period. Observe that the ITD problem selects different switching decisions than the transmission-only problem, which slightly reduces the wildfire ignition risk for the last two time periods. Although risk reduction was relatively small for this test case, it could be more significant for scenarios with more high-risk transmission lines. Figure 5 shows the total load shed in MW at each time period t. Since the ITD problem includes distribution battery energy storage systems, it schedules less load shedding when selecting line de-energizations compared to the transmission-only problem.

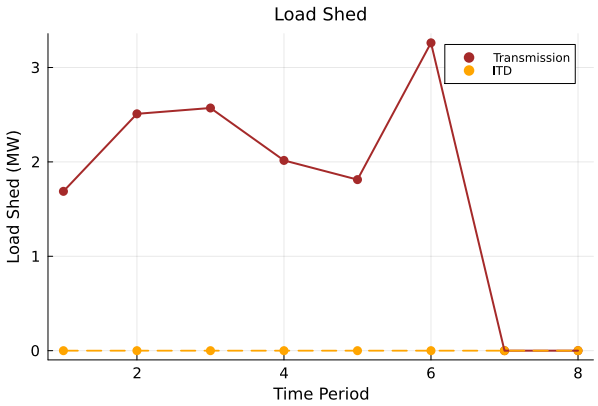

Fig. 5. Comparing the total load shed in MW between the transmission-only problem and the ITD problem.

C. Discussion on Optimality

Although ADMM does not provide convergence guarantees for non-convex problems, previous work suggests that ADMM can find good solutions to mixed-integer problems [10]–[12]. For our problem, the centralized solver converged within 48 hours for the 2-, 3-, 4-, 5-, 6-, 7- and 8-period cases, enabling a comparison between the centralized and distributed solutions. Our metric is the cost percent difference between the central objective cost f^c and the distributed objective cost f^d, computed as $(f^d - f^c)/f^c$. The cost percent difference is less than 0.1% (the MIP gap tolerance) for all these cases.

However, with nine or more time periods, the distributed approach converged to a sub-optimal solution. Although we could not compare to the centralized solution for these time periods, we did compare to the "transmission-only" solution described in Section IV-B. The transmission-only problem minimizes the same objective as the full ITD problem (17), but models each distribution network as an aggregate load, neglecting the energy storage systems. Clearly, the optimal solution to the ITD problem, which leverages energy storage to support loads, should be no more costly than the optimal solution to the transmission-only problem. For cases with nine or more periods, the distributed algorithm converged to an operating point with higher cost than the transmission-only" solution, and thus the distributed solution is sub-optimal.

Our future work will investigate improvements to the distributed algorithm to achieve better solutions for large multi-period problems. The authors of [12] suggest using ADMM as a heuristic, running the distributed algorithm several times with different penalty parameters and initializations and selecting the best result. We may also explore better methods for initializing the primal and dual variables associated with the coupling constraints, perhaps by solving some transmission-level problem with reduced models of the distribution networks. We may also leverage new machine learning techniques to perform data-driven penalty parameter tuning, such as the reinforcement learning approach in [23].

V. Conclusion

We show that decomposition and distributed optimization may provide significant computational benefits for solv-

ing large-scale integrated transmission-distribution switching problems. Specifically, we investigate coordinated transmission-distribution optimal switching to reduce wildfire ignition risk while minimizing load shedding. The resulting problem is a multi-period mixed-integer linear program, which we test using realistic, large-scale distribution networks. Our results show how the coordinated transmission-distribution problem, which leverages battery energy storage systems in the distribution networks to support loads, selects different line de-energizations which reduce wildfire risk compared to solving only a transmission-level problem.

We propose decomposing the problem over space and time into smaller subproblems. Then, we use an ADMM-based distributed algorithm to solve the decomposed problem. Given the same computing resources, our distributed method significantly outperforms the centralized solver on some problems. However, the distributed algorithm does not provide optimality guarantees for mixed-integer programs, and reached sub-optimal solutions for sufficiently large problems. Our experiments demonstrated that the distributed approach successfully reached the optimal solution for problems with up to eight time periods. For larger problems with more time periods, the centralized solver failed to converge and the distributed algorithm converged to a sub-optimal solution. Our future work includes improving the distributed algorithm's convergence performance. Potential directions include finding better initializations and using machine learning for data-driven penalty parameter tuning. We also intend to make the load shed model more realistic by making switching decisions at the distribution level to de-energize blocks of loads.

REFERENCES

[1] Pacific Gas and Electric Company (PG&E), "2023-2025 Wildfire Mitigation Plan," 2024. [Online]. Available: https://www.pge.com/assets/pge/docs/outages-and-safety/outage-preparedness-and-support/2024-06-07-PGE-2023-WMP-R4-1.pdf

[2] N. Rhodes, L. Ntaimo, and L. Roald, "Balancing Wildfire Risk and Power Outages Through Optimized Power Shut-Offs," *IEEE Trans. Power Syst.*, vol. 36, no. 4, pp. 3118–3128, 2021.

[3] A. Kody, A. West, and D. K. Molzahn, "Sharing the Load: Considering Fairness in De-energization Scheduling to Mitigate Wildfire Ignition Risk using Rolling Optimization," in *61st IEEE Conference on Decision and Control (CDC)*, 2022, pp. 5705–5712.

[4] S. Taylor, G. Setyawan, B. Cui, A. Zamzam, and L. A. Roald, "Managing Wildfire Risk and Promoting Equity through Optimal Configuration of Networked Microgrids," in *ACM e-Energy*, 2023, p. 189–199.

[5] Y. Zhou, A. Zamzam, and A. Bernstein, "Equitable Networked Microgrid Topology Reconfiguration for Wildfire Risk Mitigation," Feb. 2024, arXiv:2402.04444.

[6] J. Su, S. Mehrani, P. Dehghanian, and M. A. Lejeune, "Quasi Second-Order Stochastic Dominance Model for Balancing Wildfire Risks and Power Outages due to Proactive Public Safety De-Energizations," *IEEE Trans. Power Syst.*, vol. 39, no. 2, pp. 2528–2542, 2024.

[7] Q. Wang, W. Wu, C. Lin, Y. Yang, and B. Wang, "A Spatio-Temporal Decomposition Method for the Coordinated Economic Dispatch of Integrated Transmission and Distribution Grids," *IEEE Trans. Power Syst.*, vol. 39, no. 3, pp. 4835–4851, 2024.

[8] M. K. Arpanahi, M. E. H. Golshan, and P. Siano, "A Comprehensive and Efficient Decentralized Framework for Coordinated Multiperiod Economic Dispatch of Transmission and Distribution Systems," *IEEE Systems Journal*, vol. 15, no. 2, pp. 2583–2594, 2021.

[9] Z. Chen, Z. Li, C. Guo, J. Wang, and Y. Ding, "Fully Distributed Robust Reserve Scheduling for Coupled Transmission and Distribution Systems," *IEEE Trans. Power Syst.*, vol. 36, no. 1, pp. 169–182, 2021.

[10] R. Takapoui, N. Moehle, S. Boyd, and A. Bemporad, "A Simple Effective Heuristic for Embedded Mixed-Integer Quadratic Programming," in *American Control Conference (ACC)*, 2016, pp. 5619–5625.

[11] A. Alavian and M. C. Rotkowitz, "Improving ADMM-based optimization of Mixed Integer Objectives," in *51st Annual Conference on Information Sciences and Systems (CISS)*, 2017.

[12] Y. Kanno and S. Kitayama, "Alternating Direction Method of Multipliers as a Simple Effective Heuristic for Mixed-Integer Nonlinear Optimization," *Structural and Multidisciplinary Optimization*, vol. 58, no. 3, p. 1291–1295, September 2018.

[13] L. Gan and S. H. Low, "Convex Relaxations and Linear Approximation for Optimal Power Flow in Multiphase Radial Networks," in *18th Power Systems Computation Conference (PSCC)*, 2014.

[14] J. Ospina, D. M. Fobes, R. Bent, and A. Wächter, "Modeling and Rapid Prototyping of Integrated Transmission-Distribution OPF Formulations With PowerModelsITD.jl," *IEEE Trans. Power Syst.*, vol. 39, no. 1, pp. 172–185, 2024.

[15] M. Numan, M. F. Abbas, M. Yousif, S. S. M. Ghoneim, A. Mohammad, and A. Noorwali, "The Role of Optimal Transmission Switching in Enhancing Grid Flexibility: A Review," *IEEE Access*, vol. 11, pp. 32 437–32 463, 2023.

[16] T. Han, Y. Song, and D. J. Hill, "Ensuring Network Connectedness in Optimal Transmission Switching Problems," *IEEE Trans. Circuits and Syst. II: Express Briefs*, vol. 68, no. 7, pp. 2603–2607, 2021.

[17] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[18] R. Luce, "Quadratic Optimization," Gurobi Optimization, Tech. Rep., 2022. [Online]. Available: https://cdn.gurobi.com/wp-content/uploads/quadratic_optimization.pdf

[19] S. Taylor, A. Rangarajan, N. Rhodes, J. Snodgrass, B. C. Lesieutre, and L. A. Roald, "California Test System (CATS): A Geographically Accurate Test System Based on the California Grid," *IEEE Trans. Energy Markets, Policy and Regulation*, vol. 2, no. 1, pp. 107–118, 2024.

[20] V. K. Krishnan *et al.*, "SMART-DS: Synthetic Models for Advanced, Realistic Testing: Distribution Systems and Scenarios," National Renewable Energy Laboratory (NREL), Tech. Rep., 2017.

[21] R. Piansky, S. Taylor, N. Rhodes, D. K. Molzahn, L. A. Roald, and J.-P. Watson, "Quantifying Metrics for Wildfire Ignition Risk from Geographic Data in Power Shutoff Decision-Making," in *58th Hawaii International Conference on System Sciences (HICSS)*, January 2025.

[22] United States Geological Survey (USGS), "Wildland Fire Potential Index," 2024. [Online]. Available: https://www.usgs.gov/fire-danger-forecast/wildland-fire-potential-index-wfpi

[23] S. Zeng, A. Kody, Y. Kim, K. Kim, and D. K. Molzahn, "A Reinforcement Learning Approach to Parameter Selection for Distributed Optimization in Power Systems," *Electric Power Systems Research*, vol. 212, p. 108546, 2022, presented at the *22nd Power Systems Computation Conference (PSCC 2022)*.