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

Anonymous Author(s)

Abstract

With increasingly frequent and 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. Since many distributed energy resources are located in distribution systems, we investigate coordinating transmission and distribution systems when optimizing transmission line de-energization decisions. We consider a coordinated transmission-distribution optimization problem that balances tradeoffs between wildfire risk mitigation and load shedding. We model distribution systems that include battery energy storage systems which provide power to reduce load shedding 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. We demonstrate this approach on a large-scale test case geo-located in California that contains a transmission system with actual wildfire risk data coupled with multiple realistic distribution system models. Our results demonstrate that applying distributed optimization enables solving large-scale multi-period switching problems that are intractable using state-of-the-art centralized solvers.

CCS Concepts

• **Hardware → Power networks**; • **Applied computing → Multi-criterion optimization and decision-making**.

Keywords

integrated transmission and distribution (ITD) systems, distributed optimization, wildfire risk, multi-period switching

Conference'25, June 2025, Rotterdam, Netherlands

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM
https://doi.org/10.1145/nnnnnnn.nnnnnnn

ACM Reference Format:
Anonymous Author(s). 2024. Integrated Transmission-Distribution Multi-Period Switching for Wildfire Risk Mitigation: Improving Speed and Scalability with Distributed Optimization. In . ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 Introduction

Wildfires are expected to become more frequent and more severe as the climate changes, and the risk of electric power line faults igniting wildfires is a growing concern [11]. During times when wildfire risk is high, electric power utilities may proactively de-energize power lines to mitigate wildfire ignition risk, a practice commonly known as "public safety power shutoffs" [8].

De-energizing lines to reduce wildfire ignition risk may cause power outages to consumers. To mitigate wildfire risk while also minimizing such power outages, recent research has investigated optimal switching problems which balance the wildfire risk reductions with load shedding. One of the first of these papers formulates a transmission operation problem which finds the optimal line de-energization decisions for an objective which balances the wildfire ignition risk of energized lines with the harm of load shedding [19]. The authors compare their results with heuristics which simply use risk thresholds to de-energize power lines. A subsequent study also considered equity, formulating the problem with a rolling horizon to make power shut-offs more fair over time [13]. A survey paper reviewed utility practices and research directions for de-energizing power equipment to manage wildfire risk [11].

While the papers listed above all focused on transmission network operation, we next review papers which formulated wildfire risk mitigation problems at the distribution level. One recent paper studied optimal distribution network reconfiguration and microgrid formation for mitigating wildfire risk [23]. They tested their method by solving this problem for a single time period on a reduced version of several distribution networks from the NREL SMART-DS [14] dataset containing 3,122 buses. Another study solved a similar network reconfiguration problem, but incorporated equity considerations when making load shedding decisions [26]. Another study presented a stochastic model which accounts for uncertainty in wildfire ignition risks and incorporates mobile power sources to reduce load shedding [20].

To the best of our knowledge, this paper is the first to consider the optimal switching problem for wildfire risk mitigation as a transmission-distribution co-optimization. Distributed energy resources (DERs) can help support the power grid during severe wildfire conditions. To leverage the flexibility of these distributed resources, often located in distribution systems, we must jointly model and optimize over transmission and distribution systems. As in previous literature [13, 19, 26], we use a linear approximation of the power flow equations to facilitate solving a problem with

discrete variables representing energization statuses. We also use realistic transmission and distribution models which result in a large-scale multi-period problem. This large mixed-integer linear problem requires prohibitive amounts of time and memory to solve for state-of-the-art mixed-integer solvers. To scale to large systems, we decompose the problem and use distributed optimization techniques to reach consensus at the optimal solution.

We carry out a brief review of past work on solving coordinated transmission-distribution dispatch problems in a distributed manner. Since the past work is extensive, we focus on papers which consider multi-period dispatch or problems with integer variables. One study solved a multi-period economic dispatch problem by decomposing it over both space and time and applying an algorithm based on the alternating direction method of multipliers (ADMM) to find the solution [25]. Considering the uncertainty in renewable resources, another group presented a robust multiperiod economic dispatch, which they solved in a distributed manner using an accelerated augmented Lagrangian method [2]. Another paper used an enhanced ADMM algorithm to solve a coordinated transmission-distribution reserve scheduling problem [5], where the authors improve the convergence by adding an inner loop to the ADMM algorithm 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 which model both medium- and low-voltage buses so that we can leverage distributed resources at every level. These distribution networks each contain thousands of buses, and the full test case contains over ten thousand buses. Therefore, our results show that the proposed method can effectively be used to make optimal power shut-off decisions for realistically sized systems.

We note that ADMM is not guaranteed to converge for non-convex problems. Some papers have proposed using ADMM as a heuristic to solve mixed-integer programming problems [1, 12, 21]. They note that although ADMM is not guaranteed to converge to the optimal solution, 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. Our results demonstrate that using ADMM to solve the decomposed problem allows us to solve our large multi-period problem in a reasonable amount of time, while using a state-of-the-art centralized solver takes much longer and can be intractable for sufficiently large problems.

Our contributions are as follows:

(1) We formulate and solve a transmission-distribution co-optimization problem which optimizes transmission line de-energizations to minimize both wildfire risk and load shedding. Our paper is the first to consider coordinating transmission and distribution networks to optimize power shut-offs under wildfire risk.
(2) To make the large wildfire mitigation problem scalable, we decompose the problem over both space and time. Then we apply an ADMM-based distributed algorithm to the decomposed problem to find the optimal solution. After the decomposition, individual subproblems can be solved relatively quickly at each ADMM iteration. To the best of our knowledge, this paper is the first to apply an ADMM algorithm to solve optimal switching problems for wildfire risk mitigation.
(3) We present numerical results from test cases which are significantly larger in scale compared to previous literature. We demonstrate that applying our proposed decomposition and distributed algorithm solves the large-scale mixed-integer problem much faster than state-of-the-art centralized solvers.

We organize the remainder of the paper as follows. In Section 2, we formulate the mixed-integer linear optimal switching problem, which balances wildfire risk reduction with load shedding. In Section 3, we decompose the optimal switching problem across networks and time periods. We also describe how we use the ADMM algorithm to solve the decomposed problem. In Section 4, we then present results from a large-scale transmission-distribution test case to illustrate the computational advantages of our decomposition and distributed solution method. Finally, we conclude and discuss future work in Section 5.

2 Problem Formulation

We present the optimal switching for wildfire risk mitigation problem in this section. The problem aims to balance wildfire risk with load shedding by finding the optimal system operating point which minimizes a weighted sum of transmission line wildfire ignition risk and load shedding across all time periods. The constraints consist of the power flow physics and engineering limits which are typical of optimal power flow, augmented so that they can support component shut-offs. The control variables consist of transmission line energization statuses, transmission generator setpoints, distribution bus load shed, and distribution storage system charge/discharge setpoints. The most natural way to formulate this problem would contain non-linear AC power flow equations and discrete variables to represent transmission line statuses. However, solving mixed-integer nonlinear optimization problems is not desirable. The linearized LinDistFlow power flow approximation [7] 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. However, using linearized power flow equations may result in a solution which is not AC feasible. In particular, using generator setpoints from a LinDistFlow approximation may be problematic, since LinDistFlow neglects all losses along power lines. To make the mixed-integer problem tractable, and to produce an AC feasible solution, we formulate the problem in two stages:

(1) Solve the multi-period optimal switching for wildfire risk mitigation problem as a mixed-integer linear program (MILP) using the LinDistFlow approximation to model power flow. Save the transmission line switching decisions, distribution bus load shed, and distribution storage system setpoints from each time period.
(2) At each time period, solve a single-period AC OPF given the transmission line switching decisions, distribution bus

load shed, and distribution storage system setpoints. Save the generator setpoints selected by the AC OPF at each time period.

For the single-period AC OPF solved in Step 2, the formulation is standard. We note that rather than allowing the storage device setpoints to be free variables, we fix them to the setpoints decided by the solution of the multi-period problem in Step 1. In addition, we use the current-voltage rectangular formulation for the power flow equations, since this seems to provide the best convergence performance on the distribution networks. See [17] for additional details on the integrated transmission-distribution AC OPF solved in Step 2.

The reminder of this section details the formulation for the multi-period linearized optimal switching problem solved in Step 1. Throughout this section, we denote components of the transmission network with the superscript $\mathcal{H}$ and components of distribution networks with the superscript $\mathcal{D}$. The problem formulation contains over multiple time periods, $\mathcal{T} = \{1, 2, ..., T\}$. Variables are indexed by t to denote which time period they represent.

For the optimal switching problem, we model power flow in the transmission system with a balanced single-phase equivalent LinDistFlow approximation, and we model distribution system power flow with an unbalanced three-phase LinDistFlow approximation. Following the structure in [17], we also include constraints that balance the power flow and ensure that voltage magnitudes are equal at the transmission-distribution boundary.

2.1 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 power flow and reactive power flow 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}. The voltage drop across line (i,k) is given as

$$\begin{aligned} &\forall (i,k) \in \mathcal{L}^{\mathcal{H}},\ \forall t \in \mathcal{T}: \\ &2(r_{ik}p_{ik,t} + x_{ik}q_{ik,t}) + (1-\ell_{ik,t})\underline{M} \le w_{i,t} - w_{k,t} \\ &w_{i,t} - w_{k,t} \le 2(r_{ik}p_{ik,t} + x_{ik}q_{ik,t}) + (1-\ell_{ik,t})\overline{M} \end{aligned} \tag{2}$$

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.

We apply the following power balance constraints:

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

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 have bounds on the voltage magnitudes at all buses, given as

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

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 [16]. 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 by allowing artificial commodity to flow across 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 [10] and could be applied in our formulation.

2.2 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,

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

Using the same notation as in Section 2.1, we denote the squared voltage magnitudes at bus i at time t as $\boldsymbol{w}_{i,t}$. Also, the active and reactive power flows across line (i,k) are denoted by $\boldsymbol{p}_{ik,t}$ and $\boldsymbol{q}_{ik,t}$. The resistance is denoted as $\boldsymbol{r}_{ik}$ and the reactance is $\boldsymbol{x}_{ik}$. While in the transmission modeling these quantities were scalar values, when modeling all phases in the distribution network we have that $\boldsymbol{r}_{ik}$ and $\boldsymbol{x}_{ik}$ are matrices with terms that reflect self-impedance as well as mutual impedances between phases. The difference in squared voltage magnitudes between two buses (i,k) is

$$\boldsymbol{w}_{k,t} = \boldsymbol{w}_{i,t} - \boldsymbol{M}^P_{ik,t}\boldsymbol{p}_{ik,t} - \boldsymbol{M}^Q\boldsymbol{q}_{ik,t} \quad \forall (i,k) \in \mathcal{L}^{\mathcal{D}}, \quad \forall t \in \mathcal{T} \tag{8}$$

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

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

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.

Power balance is given by

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

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 due to the transmission line de-energizations. Previous papers on transmission line switching generally model a continuous load shed at bulk transmission loads [13, 19]. 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

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

where $\underline{\boldsymbol{V}}^2_i, \overline{\boldsymbol{V}}^2_i$ 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. For the purposes of this paper, we use an ideal model for the storage system in which charging and discharging is perfectly efficient. The variable C_i is positive when the storage system is charging and negative when it is discharging. The energy storage system 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 charging power. We assume that the charging/discharging power C_i is fixed over one time period $t \in \mathcal{T}$.

$$\boldsymbol{E}_{i,t} - \boldsymbol{E}_{i,t-1} = \boldsymbol{C}_i \qquad \forall i \in \mathcal{N}^{\mathcal{D}},\ \forall t \in \mathcal{T} \tag{12}$$
$$0 \le \boldsymbol{E}_{i,t} \le \overline{\boldsymbol{E}_i} \qquad \forall i \in \mathcal{N}^{\mathcal{D}},\ \forall t \in \mathcal{T} \tag{13}$$
$$\underline{C}_i \le \boldsymbol{C}_i \le \overline{C}_i \qquad \forall i \in \mathcal{N}^{\mathcal{D}},\ \forall t \in \mathcal{T} \tag{14}$$

2.3 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 optimization software we use for our experiments [17].

2.4 Objective Function

To these constraints, we add an objective function inspired by [13, 19], which balances wildfire risk with load shedding:

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

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

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

This is a mixed-integer linear program.

3 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. An illustration of how the problem is decomposed and where the coupling constraints are imposed for a small example system is shown in Figure 1. In this figure, each block represents a subproblem solved by its own computing agent at each ADMM iteration. The dashed lines represent power flow coupling constraints between the central transmission network and its attached distribution networks at each time period.

The transmission network subproblem is computationally intensive due to the large number of binary transmission switching variables. Therefore, we solve separate problems for the transmission network at each timestep. The objective and constraints for each distribution network across all time periods also form individual subproblems. We do not decompose the distribution network subproblems across time periods because the distribution networks contain many small storage systems with energy state constraints that couple time periods. Decomposing these distribution storage

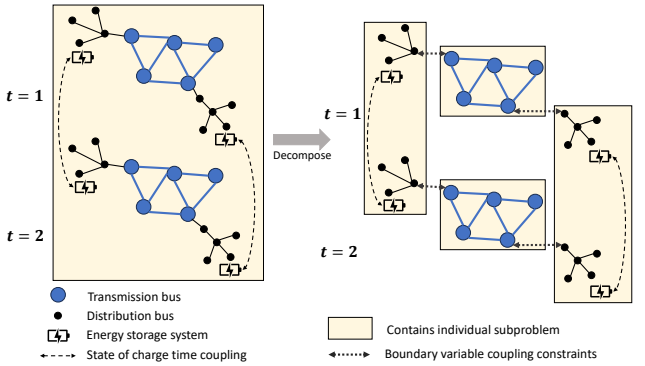

Figure 1: This figure shows the decomposition of the multi-period problem. Dashed lines show the coupling between energy storage systems' states of charge across time periods. After decomposition, each transmission network at each time period forms a subproblem. Each distribution network across all time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

system constraints across time periods would result in thousands of coupling constraints between subproblems which would slow the ADMM algorithm's convergence such that decomposing these subproblems over time would be counterproductive.

We now specify the exact nature of the coupling constraints. We add to the transmission subproblem a copy of each three-phase distribution boundary bus. Each phase of the distribution boundary bus copies contain a virtual generator which can inject or consume unconstrained amounts of active and reactive power. This generator thus accounts for power flow between the transmission and distribution networks. For each distribution network subproblem, we add a copy of the transmission boundary bus equipped with a virtual generator to allow for power flow into or out of the transmission network. Then, we add simple coupling constraints between the transmission subproblem and each of the distribution network subproblems, which ensure that the subproblems agree on the values of the squared voltage magnitudes and active and reactive power flows at the boundary. Figure 2 illustrates how copies of the boundary buses are added to subproblems and shows the coupling constraints.

Consider a transmission-distribution boundary between transmission bus m and distribution bus n. The transmission subproblem contains a virtual bus representing a copy of the distribution bus denoted n' and the distribution subproblem contains a virtual bus representing a copy of the transmission bus denoted m'. The coupling constraints ensure that voltage magnitudes are equal at the transmission bus m and its copy m', and that voltage magnitudes are equal for each phase ϕ of the distribution bus n_ϕ and its copy n'_ϕ:

$$\begin{aligned} w_m &= w_{m'} \\ w_{n_\phi} &= w_{n'_\phi} \quad \forall \phi \in \Phi_n \end{aligned} \tag{19}$$

Also, the coupling constraints ensure that the power flow variables are consistent across both subproblems. Therefore, the power flow from the transmission bus m to the copy of the distribution bus

n' must be equal to the power flow from the transmission bus copy to the distribution bus. In addition, for each phase ϕ on the distribution side, the power flow from the distribution bus copy n'_ϕ to the transmission bus m_ϕ must be equal to the power flow from the distribution bus n_ϕ to the transmission bus copy m'_ϕ. The coupling constraints for power flows between transmission bus m and distribution bus n are as follows:

$$\begin{aligned} &p_{mn'} = p_{m'n}, \quad q_{mn'} = q_{m'n} \\ &p_{n'_\phi m_\phi} = p_{n_\phi m'_\phi}, \quad q_{n'_\phi m_\phi} = q_{n_\phi m'_\phi} \quad \forall \phi \in \Phi_n \end{aligned} \tag{20}$$

Some readers may wonder why we include the transmission-distribution boundary and a virtual distribution bus in the transmission subproblems, and why we include the boundary and a virtual transmission bus in the distribution subprblems. Since the boundary constraints (15), (16) are linear, we could instead impose these boundary constraints as coupling constraints, removing the need for virtual buses. We choose to add the virtual buses and keep the transmission-distribution boundary constraints internal to subproblems for two reasons. First, it makes our subproblems compatible with the integrated transmission-distribution optimization framework of [17]. In addition, we want to remain flexible to using new power flow formulations with nonlinear boundary constraints in the future. See, for example, the quadratic constraints which ensure that voltage magnitudes are equal at the transmission-distribution boundary for a rectangular current-voltage power flow formulation considering a three-phase distribution model in [17]. Such nonlinear boundary constraints could not be used as coupling constraints for a distributed solution method.

We use the ADMM algorithm to solve the decomposed problem [4]. The ADMM algorithm solves the general problem

$$\begin{aligned} \min_{\boldsymbol{x}\in\mathcal{X},\boldsymbol{z}\in\mathcal{Z}} \quad & f(\boldsymbol{x}) + g(\boldsymbol{z}) \\ \text{s.t.} \quad & A\boldsymbol{x} + B\boldsymbol{z} = \boldsymbol{c} \end{aligned} \tag{21}$$

Now, the optimal switching problem for wildfire risk mitigation has been decomposed into subproblems. Each subproblem is in the form of (18), but not all variables and constraints are included in every subproblem. 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

$$\begin{aligned} \min_{\substack{\boldsymbol{\ell}^{\mathcal{H}},\boldsymbol{p}^{\mathcal{H}},\boldsymbol{q}^{\mathcal{H}},\boldsymbol{w}^{\mathcal{H}} \\ \boldsymbol{p}^{\mathcal{D}},\boldsymbol{q}^{\mathcal{D}}}} \quad & \gamma \sum_{t\in\mathcal{T}} \sum_{(i,k)\in\mathcal{L}^{\mathcal{H}}} \rho_{ik,t}\ell_{ik,t} \\ \text{s.t.} \quad & (1)-(7) \end{aligned} \tag{22}$$

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

$$\begin{aligned} \min_{\substack{\boldsymbol{p}^{\mathcal{H}},\boldsymbol{q}^{\mathcal{H}},\boldsymbol{s}^{\mathcal{D}}, \\ \boldsymbol{p}^{\mathcal{D}},\boldsymbol{q}^{\mathcal{D}},\boldsymbol{w}^{\mathcal{D}},\boldsymbol{C}^{\mathcal{D}},\boldsymbol{E}^{\mathcal{D}}}} \quad & (1-\gamma) \sum_{t\in\mathcal{T}} \Big(\sum_{i\in\mathcal{N}^{\mathcal{D}}} s_i \sum_{\phi\in\Phi_i} -P^{d,\phi}_{i,t} \Big) \\ \text{s.t.} \quad & (8)-(14) \end{aligned} \tag{23}$$

To write the decomposed optimal power shut-off problem in the form of (21), 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' 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' 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 variables at *virtual* buses in the subproblems. Note that power flow variables $\boldsymbol{p}^H$, $\boldsymbol{q}^H$, $\boldsymbol{p}^{\mathcal{D}}$, and $\boldsymbol{q}^{\mathcal{D}}$ include both physical flows and virtual flows associated with the virtual buses used to decompose the subproblems at the transmission/distribution system boundary. Similarly, the squared voltage magnitude variables $\boldsymbol{w}^{\mathcal{H}}$, $\boldsymbol{w}^{\mathcal{D}}$ include physical voltage magnitudes as well as voltage magnitudes at virtual buses in the decomposed problem.

Next, we introduce a "central" copy of the variables at the virtual buses, and gather these central variables into a vector $\boldsymbol{z}$. When we introduce the "central" variables, we modify the coupling constraints described in (19), (20) as follows to fit the ADMM structure. Instead of constraining virtual bus variables to be directly equal to their copies in neighboring subproblems, we set virtual bus variable values equal to the central variable copy. We illustrate how these coupling constraints are modified below. Let the tilde denote 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:

$$\begin{aligned} &w_m = \tilde{w}_m \\ &w_{n'_\phi} = \tilde{w}_{n_\phi} \quad \forall \phi \in \Phi_n \\ &p_{mn'} = \tilde{p}_{mn}, \quad q_{mn'} = \tilde{q}_{mn} \\ &p_{n'_\phi m_\phi} = \tilde{p}_{n_\phi m_\phi}, \quad q_{n'_\phi m_\phi} = \tilde{q}_{n_\phi m_\phi} \quad \forall \phi \in \Phi_n \end{aligned} \tag{24}$$

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

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

This formulation 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.

The ADMM algorithm augments the Lagrangian function for (21) 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 when solving MILPs over distribution network models, using the squared ℓ_2-norm may result in numerical issues. Distribution network subproblem numerics are much improved by penalizing the ℓ_1-norm of the coupling constraint violations instead. This is likely because the ℓ_1-norm is formulated as a linear objective function, with additional linear constraints. Quadratic programs may experience more numerical problems compared to linear programs during execution of the

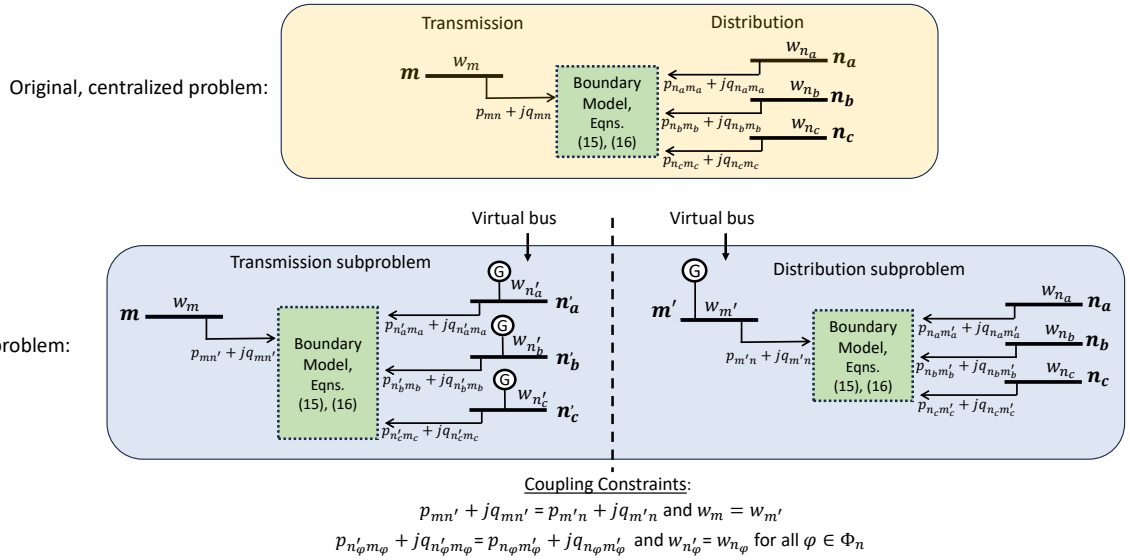

Figure 2: Coupling constraints between the transmission and distribution subproblems for a single timestep. To the transmission subproblem, we add a copy of the three-phase distribution boundary bus equipped with virtual generators on each phase which can inject or absorb unlimited power. Similarly, the distribution subproblem also contains a copy of the transmission boundary bus, with only one phase modeled, with a virtual generator. The coupling constraints then ensure that the transmission and distribution subproblems agree on the voltage magnitudes and power flows at the boundary.

simplex or interior-point algorithms implemented by Gurobi, the solver we choose for our experiments [15]. Therefore, we run the ADMM algorithm on the ℓ_1-norm augmented Lagrangian:

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

Here, the vector $\boldsymbol{\lambda}$ contains the dual variables for each coupling constraint in (21). The penalty parameter α is user-selected.

At each iteration k, the ADMM algorithm minimizes over the $\boldsymbol{x}$ and z variables separately, while holding all other variables fixed, and then updates the dual variables. The variable updates at iteration k are as follows:

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

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

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

The $\boldsymbol{x}$-update step corresponds to solving all transmission and distribution subproblems in parallel. Note that the transmission and distribution subproblems are given in (22) and (23), 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.

For transmission subproblems, the number of continuous variables is $3|\mathcal{L}^{\mathcal{H}}| + |\mathcal{N}^{\mathcal{H}}|$ and the number of binary variables is $|\mathcal{L}^{\mathcal{H}}|$. The number of inequality constraints is $8|\mathcal{L}^{\mathcal{H}}| + 2|\mathcal{N}^{\mathcal{H}}|$ and the number of equality constraints is $2|\mathcal{L}^{\mathcal{H}}| + |\mathcal{N}^{\mathcal{H}}| + 5|\beta|$. For distribution network n, let $\mathcal{N}_\phi^{\mathcal{D},n}$ be the set containing each phase of each bus, let $\mathcal{L}_\phi^{\mathcal{D},n}$ be the set containing each phase of each line, let $\mathcal{N}^{\mathcal{D},n}$ be the set containing each bus, and let $\mathcal{S}^{\mathcal{D},n}$ be the set containing battery energy storage systems. Then the number of continuous variables for distribution subproblems is $|\mathcal{T}|\big(2|\mathcal{L}_\phi^{\mathcal{D},n}| + |\mathcal{N}_\phi^{\mathcal{D},n}| + |\mathcal{N}^{\mathcal{D},n}| + |\mathcal{S}^{\mathcal{D},n}|\big)$. Also, the number of inequality constraints is $|\mathcal{T}|\big(2|\mathcal{N}_\phi^{\mathcal{D},n}| + 4|\mathcal{S}^{\mathcal{D},n}|\big)$ and the number of equality constraints is $|\mathcal{T}|\big(3|\mathcal{L}_\phi^{\mathcal{D},n}| + 2|\mathcal{N}_\phi^{\mathcal{D},n}||\mathcal{S}^{\mathcal{D},n}| + 5\big)$. Note that the term $5|\beta|$ in the transmission subproblem equality constraint count and the 5 in the distribution subproblem equality constraint count come from the transmission-distribution boundary constraints. To summarize, 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.

4 Case Study

We construct a test case using the CATS (California Test System) [22] and the SMART-DS synthetic distribution networks [14]. 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 20% of the low-voltage nodes in the distribution networks, each with energy capacity randomly selected between 30-50 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 3. In Tables 1 and 2, we show the number of components in the transmission and distribution networks, respectively. 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 leverage the work in [18]. This study extracts the wind-enhanced fire potential index (WFPI) as computed by the United States Geological Survey [24] for regions across the power network. The authors of [18] then identify which regions each transmission line crosses and use the corresponding WFPI values to compute the total wildfire ignition risk for the transmission line. We selected the "high-risk cumulative metric" out of the various metrics proposed in the paper for our wildfire ignition risk values; interested readers can find further details in [18].

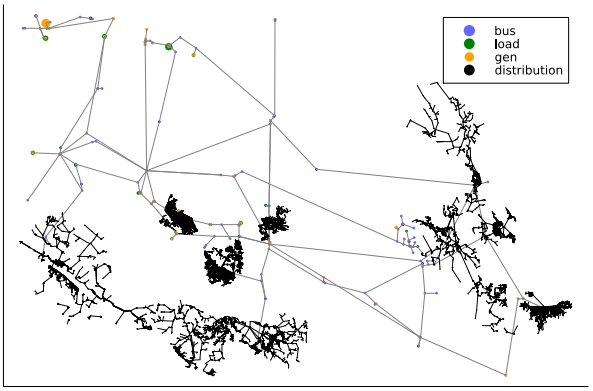

Figure 3: Synthetic test case near San Francisco, California for integrated transmission-distribution optimization.

Table 1: Transmission network components

Buses	Generators	Branches
155	15	171

4.1 Experiment Setup: Multi-Period Linearized Optimal Switching for Wildfire Risk Mitigation

In this section, we provide results on the solution of the multi-period linearized optimal switching problem for wildfire risk mitigation. We compare the performance of our distributed solution method with the performance of a state-of-the-art solver, Gurobi. All experiments ran on the Georgia Institute of Technology PACE high-performance computing cluster. For each experiment, which consisted of solving the optimal switching problem for wildfire risk mitigation over some number of time periods $\mathcal{T}$, we used one PACE CPU compute node equipped with two 12-core 2.7 GHz processors and 192 GB RAM. We set a time limit of 48 hours and terminate any algorithm which has not converged at this point.

Table 2: Distribution system components across all distribution networks

Buses	Lines	Storage Devices	Indiv. Phases, Buses	Indiv. Phases, Lines
15083	15456	1203	24823	25942

We note that our distributed solution method also uses the Gurobi solver on individual sub-problems. The method described in our results as "central" solves the full MILP in one shot with Gurobi. Our method, described as "distributed," decomposes the problem as described in Section 3 and uses the ADMM algorithm to reach the solution, where individual subproblems are solved with Gurobi. We terminate the "central" solution method when the MIP gap reaches 1%. When solving the distributed algorithm, we also solve individual subproblems to a 1% MIP gap. We terminate the distributed algorithm when the norm of the mismatches between subproblems' shared variables falls below $\epsilon = 10^{-4}$.

4.2 Comparing Central and Distributed Computation Time

Figure 4 shows the runtime for each solution method as we increase the number of time-steps in the multi-period problem. When the number of time-steps is four or less, the centralized solution 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. In fact, the centralized method fails to solve the 12-, 16-, and 24-timestep cases within 48 hours, at which point the MIP gap is 24.2% for the 12-timestep case, 27.3% for the 16-timestep case, and 30.1% for the 24-timestep case.

We also show how the individual subproblems for the distributed method scale as the number of timesteps increases. Note that the transmission subproblems cover one time period only, and they consistently solve in less than one second. The distribution subproblems cover one distribution network across all time periods, so the solve time for distribution subproblems increases with the number of timesteps. Every iteration of the distributed algorithm solves $|\mathcal{T}|$ transmission subproblems, where $\mathcal{T}$ is the set of all time periods, and N distribution subproblems, where N is equal to the number of distribution networks across the system. Note that at each iteration of the distributed algorithm, we solve subproblems in parallel across six processors. These subproblems could be solved with further parallelization across more processors if desired. Across all timesteps, the distributed algorithm takes between 11 and 17 iterations to converge, where all subproblems are solved at each iteration.

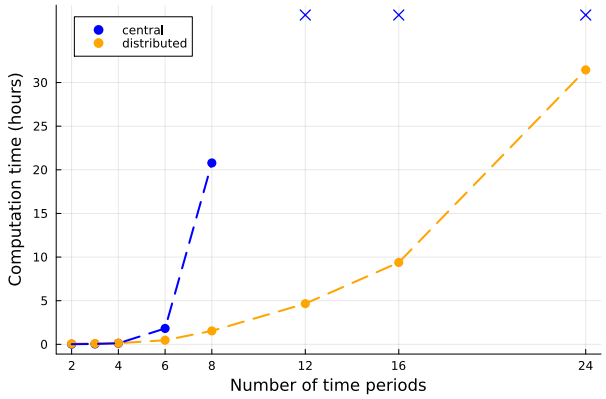

Figure 4: Comparing computation time for the central solution method using the Gurobi solver vs. our distributed solution method. Note that the mark × indicates that the solver could not reach a solution within 48 hours for that number of time periods.

Table 3: Average subproblem solve time, where times are given in minutes : seconds

Timesteps	Transmission	Distribution
4	00:01	00:07
6	00:01	00:20
8	00:01	01:14
12	00:01	03:51
16	00:01	08:58
24	00:01	20:58

4.3 Discussion on Optimality

Recall that the ADMM distributed algorithm does not provide convergence guarantees for non-convex problems. Despite this, previous work on solving mixed-integer problems with ADMM suggests that ADMM can often find very good solutions to such problems [1, 12, 21]. For large-scale mixed-integer linear problems which pose significant computational challenges to centralized branch-and-bound solvers, using distributed algorithms may be an effective alternative. In addition, 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.

Despite the potential for suboptimality, we found that our distributed algorithm obtained the same solution as the centralized solver for all cases where we could perform a comparison. For every $|\mathcal{T}|$-timestep case for which the centralized solution method solved within 48 hours, we compare the central solution to the distributed solution. Our metric is the cost percent difference between the central objective cost f^c and the distributed objective cost f^d:

$$(f^d - f^c)/f^c.$$

The cost percent difference is less than 1% (the MIP gap tolerance) for the 2-, 3-, 4-, 6-, and 8-timestep cases. For the 12- and 16-timestep cases, the centralized solver did not converge within 48 hours, so we could not compare to the central solution. In most cases, the cost percent difference is less than 0.0001%, but for the 4-timestep case the cost percent difference is negative, with the distributed objective cost 0.1% less than the central objective cost. Within the 1% MIP gap tolerance, the distributed objective cost may be less than the central objective cost.

4.4 Final Operating Points from Single-Period AC OPF

The operating point produced by the mixed-integer linear optimal switching problem may not be AC feasible. See, for example, references [3, 6] which investigate AC infeasibility of transmission switching decisions made using the DC power flow approximation as well as reference [9] for a study on power flow modeling for wildfire switching problems. To assess AC feasibility for solutions to our formulation, we solve AC OPF problems for each time period to find the generator setpoints in the transmission system. We take the transmission line switching decisions, distribution bus load sheds, and distribution storage system setpoints as fixed when solving the AC OPF. The time to solve an instance of single-period AC OPF on the test case is 5 minutes and 58 seconds on average. We find that for the all time steps we considered (2, 3, 4, 6, 8, 12, 16, and 24), the switching decisions with corresponding load sheds and storage setpoints result in AC feasible operating points. This empirically demonstrates that our proposed mixed-integer formulation of the optimal switching problem for wildfire risk mitigation can often provide AC feasible network topologies. Our future work aims to further explore the AC feasibility of transmission switching decisions made using the LinDistFlow power flow approximation.

5 Conclusion

We show that decomposition and distributed optimization methods can solve large-scale coordinated transmission-distribution problems which may be intractable for centralized solvers. Specifically, we investigate solving an integrated transmission-distribution problem which finds optimal switching decisions to reduce wildfire ignition risk. The problem coordinates line switching at the transmission level with load shedding and battery energy storage system charging/discharging at the distribution level. We make switching and storage setpoint decisions by solving a multi-period mixed-integer linear optimization problem. Next, we solve AC OPF problem over single time periods to obtain generator setpoints and ensure that the operating point is AC feasible.

We solve this problem with realistic, large-scale distribution networks. We show that when enough time periods are modeled, state-of-the-art solvers struggle to solve this mixed-integer optimization problem. We propose a method of decomposing the problem over both space and time and then using a distributed algorithm to find the solution. With the same computing resources, our distributed method significantly outperforms the centralized solver on sufficiently large problems. The distributed method successfully solves problems with up to 24 time-steps, while the centralized solver becomes much slower and then fails to solve as the number of time-steps increases. Although the distributed algorithm does not provide optimality guarantees for mixed-integer linear programs,

our experiments demonstrated that the distributed algorithm successfully reached an optimal solution for all problems for which the central solver's solution was available for comparison. We also solved AC OPF over single time periods to obtain generator setpoints and found that for all experiments, the switching topology selected by the mixed-integer linear program was AC feasible.

This paper shows the significant computational benefit of using distributed algorithms to solve large-scale coordinated transmission-distribution optimization problems. For our future work, we intend to make the load shed model more realistic by making switching decisions at the distribution level to de-energize blocks of loads. In addition, we plan to increase the network's capacity to adapt and reduce load shedding while mitigating wildfire risk by extending the problem to include distribution system reconfiguration and islanded microgrids. We also plan to investigate ways to accelerate the distributed algorithm through improved initialization and penalty parameter tuning.

References

[1] Alborz Alavian and Michael C. Rotkowitz. 2017. Improving ADMM-based optimization of Mixed Integer Objectives. In *51st Annual Conference on Information Sciences and Systems (CISS)*.

[2] Moossa Khodadadi Arpanahi, Mohammad E. Hamedani Golshan, and Pierluigi Siano. 2021. A Comprehensive and Efficient Decentralized Framework for Coordinated Multiperiod Economic Dispatch of Transmission and Distribution Systems. *IEEE Systems Journal* 15, 2 (2021), 2583–2594.

[3] Clayton Barrows, Seth Blumsack, and Paul Hines. 2014. Correcting Optimal Transmission Switching for AC Power Flows. In *47th Hawaii International Conference on System Sciences (HICSS)*. 2374–2379.

[4] Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, and Jonathan Eckstein. 2011. Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. *Foundations and Trends® in Machine Learning* 3, 1 (2011), 1–122.

[5] Zhe Chen, Zhengshuo Li, Chuangxin Guo, Jianhui Wang, and Yi Ding. 2021. Fully Distributed Robust Reserve Scheduling for Coupled Transmission and Distribution Systems. *IEEE Transactions on Power Systems* 36, 1 (2021), 169–182.

[6] Carleton Coffrin, Hassan L. Hijazi, Karsten Lehmann, and Pascal Van Hentenryck. 2014. Primal and Dual Bounds for Optimal Transmission Switching. In *18th Power Systems Computation Conference (PSCC)*.

[7] Lingwen Gan and Steven H. Low. 2014. Convex Relaxations and Linear Approximation for Optimal Power Flow in Multiphase Radial Networks. In *18th Power Systems Computation Conference (PSCC)*.

[8] Pacific Gas and Electric Company (PG&E). 2022. 2022 Wildfire Mitigation Plan Update. https://www.pge.com/en/outages-and-safety/safety/community-wildfire-safety-program.html?WT.mc_id=Vanity_wildfiremitigationplan

[9] Eric Haag, Noah Rhodes, and Line Roald. 2024. Long Solution Times or Low Solution Quality: On Trade-offs in Choosing a Power Flow Formulation for the Optimal Power Shutoff Problem. *Electric Power Systems Research* 234 (2024), 110713. Presented at the 23rd Power Systems Computation Conference (PSCC).

[10] Tong Han, Yue Song, and David J. Hill. 2021. Ensuring Network Connectedness in Optimal Transmission Switching Problems. *IEEE Transactions on Circuits and Systems II: Express Briefs* 68, 7 (2021), 2603–2607.

[11] Can Huang, Qinran Hu, Linwei Sang, Donald D. Lucas, Robin Wong, Bin Wang, Wanshi Hong, Mengqi Yao, and Vaibhav Donde. 2023. A Review of Public Safety Power Shutoffs (PSPS) for Wildfire Mitigation: Policies, Practices, Models and Data Sources. *IEEE Transactions on Energy Markets, Policy and Regulation* 1, 3 (2023), 187–197.

[12] Yoshihiro Kanno and Satoshi Kitayama. 2018. Alternating Direction Method of Multipliers as a Simple Effective Heuristic for Mixed-Integer Nonlinear Optimization. *Structural and Multidisciplinary Optimization* 58, 3 (September 2018), 1291–1295.

[13] Alyssa Kody, Amanda West, and Daniel K. Molzahn. 2022. 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)*. 5705–5712.

[14] Venkat K. Krishnan, Bryan S. Palmintier, Bri-Mathias S. Hodge, Elaine T. Hale, Tarek Elgindy, Bruce Bugbee, Michael N. Rossol, Anthony J. Lopez, Dheepak Krishnamurthy, and Claudio Vergara. 2017. *SMART-DS: Synthetic Models for Advanced, Realistic Testing: Distribution Systems and Scenarios.* Technical Report. National Renewable Energy Laboratory (NREL).

[15] Robert Luce. 2022. *Quadratic Optimization.* Technical Report. Gurobi Optimization. https://cdn.gurobi.com/wp-content/uploads/quadratic_optimization.pdf

[16] Muhammad Numan, Muhammad Farasat Abbas, Muhammad Yousif, Sherif S. M. Ghoneim, Alsharef Mohammad, and Abdulfattah Noorwali. 2023. The Role of Optimal Transmission Switching in Enhancing Grid Flexibility: A Review. *IEEE Access* 11 (2023), 32437–32463.

[17] Juan Ospina, David M. Fobes, Russell Bent, and Andreas Wächter. 2024. Modeling and Rapid Prototyping of Integrated Transmission-Distribution OPF Formulations With PowerModelsITD.jl. *IEEE Transactions on Power Systems* 39, 1 (2024), 172–185.

[18] Ryan Piansky, Sofia Taylor, Noah Rhodes, Daniel K. Molzahn, Line A. Roald, and Jean-Paul Watson. 2025. Quantifying Metrics for Wildfire Ignition Risk from Geographic Data in Power Shutoff Decision-Making. To appear in *58th Hawaii International Conference on System Sciences (HICSS)* (January 2025).

[19] Noah Rhodes, Lewis Ntaimo, and Line Roald. 2021. Balancing Wildfire Risk and Power Outages Through Optimized Power Shut-Offs. *IEEE Transactions on Power Systems* 36, 4 (2021), 3118–3128.

[20] Jinshun Su, Saharnaz Mehrani, Payman Dehghanian, and Miguel A. Lejeune. 2024. Quasi Second-Order Stochastic Dominance Model for Balancing Wildfire Risks and Power Outages due to Proactive Public Safety De-Energizations. *IEEE Transactions on Power Systems* 39, 2 (2024), 2528–2542.

[21] Reza Takapoui, Nicholas Moehle, Stephen Boyd, and Alberto Bemporad. 2016. A Simple Effective Heuristic for Embedded Mixed-Integer Quadratic Programming. In *American Control Conference (ACC)*. 5619–5625.

[22] Sofia Taylor, Aditya Rangarajan, Noah Rhodes, Jonathan Snodgrass, Bernie Lesieutre, and Line A. Roald. 2023. California Test System (CATS): A Geographically Accurate Test System based on the California Grid. (June 2023). arXiv:2210.04351.

[23] Sofia Taylor, Gabriela Setyawan, Bai Cui, Ahmed Zamzam, and Line A. Roald. 2023. Managing Wildfire Risk and Promoting Equity through Optimal Configuration of Networked Microgrids *(e-Energy '23)*. Association for Computing Machinery, New York, NY, USA, 189–199.

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

[25] Qi Wang, Wenchuan Wu, Chenhui Lin, Yue Yang, and Bin Wang. 2024. A Spatio-Temporal Decomposition Method for the Coordinated Economic Dispatch of Integrated Transmission and Distribution Grids. *IEEE Transactions on Power Systems* 39, 3 (2024), 4835–4851. https://doi.org/10.1109/TPWRS.2023.3329820

[26] Yuzi Zhou, Ahmed Zamzam, and Andrey Bernstein. 2024. Equitable Networked Microgrid Topology Reconfiguration for Wildfire Risk Mitigation. (Feb 2024). arXiv:2402.04444.