LA-UR-24-23063
Accepted Manuscript

A data-driven sensor placement approach for detecting voltage violations in distribution systems

Buason, Paprapee
Misra, Sidhant
Talkington, Samuel
Molzahn, Daniel

Provided by the author(s) and the Los Alamos National Laboratory (1930-01-01).

To be published in: Electric Power Systems Research

DOI to publisher's version: 10.1016/j.epsr.2024.110387

Permalink to record:
https://permalink.lanl.gov/object/view?what=info:lanl-repo/lareport/LA-UR-24-23063

Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by Triad National Security, LLC for the National Nuclear Security Administration of U.S. Department of Energy under contract 89233218CNA000001. By approving this article, the publisher recognizes that the U.S. Government retains nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes. Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy. Los Alamos National Laboratory strongly supports academic freedom and a researcher's right to publish; as an institution, however, the Laboratory does not endorse the viewpoint of a publication or guarantee its technical correctness.

A Data-Driven Sensor Placement Approach for Detecting Voltage Violations in Distribution Systems

Paprapee Buason[a,*], Sidhant Misra[b], Samuel Talkington[a], Daniel K. Molzahn[a]

[a] *Georgia Institute of Technology*
[b] *Los Alamos National Laboratory*

Abstract

Stochastic fluctuations in power injections from distributed energy resources (DERs) combined with load variability can cause constraint violations (e.g., exceeded voltage limits) in electric distribution systems. To monitor grid operations, sensors are placed to measure important quantities such as the voltage magnitudes. In this paper, we consider a sensor placement problem which seeks to identify locations for installing sensors that can capture all possible violations of voltage magnitude limits. We formulate a bilevel optimization problem that minimizes the number of sensors and avoids false sensor alarms in the upper level while ensuring detection of any voltage violations in the lower level. This problem is challenging due to the nonlinearity of the power flow equations and the presence of binary variables. Accordingly, we employ recently developed conservative linear approximations of the

*Corresponding author

Email addresses: pbuason6@gatech.edu (Paprapee Buason), sidhant@lanl.gov (Sidhant Misra), talkington@gatech.edu (Samuel Talkington), molzahn@gatech.edu (Daniel K. Molzahn)

power flow equations that overestimate or underestimate the voltage magnitudes. By replacing the nonlinear power flow equations with conservative linear approximations, we can ensure that the resulting sensor locations and thresholds are sufficient to identify any constraint violations. Additionally, we apply various problem reformulations to significantly improve computational tractability while simultaneously ensuring an appropriate placement of sensors. Lastly, we improve the quality of the results via an approximate gradient descent method that adjusts the sensor thresholds. We demonstrate the effectiveness of our proposed method for several test cases, including a system with multiple switching configurations.
Keywords: Keywords, Sensor placement, voltage violations, bilevel optimization, approximate gradient descent.

---

1. Introduction

Distributed energy resources (DERs) are being rapidly deployed in distribution systems. Fluctuations in DER power outputs and varying load demands can potentially cause violations of voltage limits, i.e., voltages outside the bounds imposed in the ANSI C84.1 standard. These violations can cause equipment malfunctions, failures of electrical components, and, in severe situations, power outages.

To mitigate the impacts of violations, distribution system operators (DSOs) must identify when power injection fluctuations lead to voltages exceeding their limits. To do so, sensors are placed within the distribution system to

measure and communicate the voltage magnitudes at their locations. Due to the cost of sensor hardware and communication infrastructure and the structure of distribution systems, sensors are not placed at all buses. The question arises whether a voltage violation at a location without a sensor can nevertheless be detected.

1.1. Related work

Various studies have proposed sensor siting methods to capture constraint violations and outages in power systems. For instance, [1] and [2] focus on a cost minimization problem that aims to capture all node (e.g., voltage magnitude) and line (e.g., power flow) outages. However, these references assume that a power source/generation is only located at the root node, which is not always the case, especially in the distribution systems where DERs can be located further down a feeder. Additional research efforts such as [3, 4, 5, 6] seek to locate the minimum number of sensors to achieve full observability for the system. Alternatively, instead of considering full observability for the entire system, [7] considers satisfying observability requirements given a probability of observability at each bus. This is related to adjacent research efforts, such as [8, 9, 10, 11], which focus on model-based voltage control schemes that prevent voltage violations. We refer the reader to [12] for a survey. In parallel, there is also research on siting phasor measurement units (PMUs) and micro phasor measurement units (μ-PMUs), which are utilized as sensors in power grids [13, 14, 15]. This research complements

the discussion on sensor placement strategies by exploring alternative sensor technologies and their deployment strategies within power systems. Existing smart meters installed at customer locations measure power consumption over a long period of time (e.g., a day to a month). Reference [16] discusses the potential benefits of using smart meters to report voltages and currents at higher resolutions. However, handling large amounts of data remains challenging, necessitating methods to selectively site high-resolution sensors.

Growing numbers of sensing devices in electric distribution networks have fueled a broad range of applications. Exemplifying use cases of the measurements from these sensors are in control [12, 17, 18], state estimation [19, 20, 21], topology identification [22, 23], and non-intrusive load monitoring (NILM) [24]. Our work contributes to this rich literature. Some measurement-based voltage control approaches, such as [17, 18] utilize approximate representations of the power flow equations recovered from data; like our work, this approach is an abstract way to model an unknown distribution network. In contrast, network identification works such as [22, 25] seek to recover the network topology itself, in tandem with admittance [22] or impedance [25] parameters. We refer the reader to [23] for a comprehensive review of this topic.

1.2. Overview of approach

In this paper, we consider a *sensor placement* problem which seeks to locate the minimum number of sensors and determine corresponding sensor

alarm thresholds in order to reliably identify all possible violations of voltage magnitude limits in a distribution system. We formulate this sensor placement problem as a bilevel optimization with an upper level that minimizes the number of sensors and chooses sensor alarm thresholds and a lower level that computes the most extreme voltage magnitudes within given ranges of power injection variability. This problem additionally aims to reduce the number of false positive alarms, i.e., violations of the sensors' alarm thresholds that do not correspond to an actual voltage limit violation.

In contrast to previous work, this problem *does not* attempt to ensure full observability of the distribution system. Rather, we seek to locate (a potentially smaller number of) sensors that can nevertheless identify all voltage limit violations for any power injections within a specified range of power injection variability. With a small number of sensors, the proposed formulation also provides a simple means to design corrective actions if voltage violations are encountered in real-time operations. By restoring voltages at these few critical locations to within their alarm thresholds, the system operator can guarantee feasibility of the voltage limits for the full system. This guarantee is obtained by our sensor placement method purely by analyzing the geometric properties of the feasible set. We do not consider the design details of the feedback control protocol and thus dynamic properties of the sensors such as latency are not relevant in our approach.

Due to the nonlinear nature of the AC power flow equations, computing a globally optimal solution is challenging. We utilize *conservative linear*

approximations of the power flow equations to convert the lower-level problem to a linear program [26]. This bilevel problem can be reformulated to a single-level problem using the Karush-Kuhn-Tucker (KKT) conditions with binary variables via a big-M formulation [27, 28]. In this paper, we consider a duality-based approach, which has substantial computational advantages over traditional KKT-based approaches to solving the bilevel problem. The conservative linear approximations can incorporate the behavior of more complex components such as tap-changing transformers, smart inverters, etc., as long has we have access to a power flow solver for the system. By using these linear approximations as the first step, we are able to treat the power flow solver as a black-box. Consequently, all complexities of component behavior and power flow physics are absorbed by the power flow solver and the complexity of the resulting sensor placement formulation remains unaffected.

Note that conservativeness from the conservative linear approximations may increase the number of false positive alarms. We therefore propose an approximate gradient descent method as a post-processing step to further improve the quality of the results. This method iteratively adjusts the sensor thresholds while ensuring that all violations are still detected.

1.3. Contributions and article outline

In summary, our main contributions are:

(i) A bilevel optimization formulation for a sensor placement problem that minimizes the number of sensors needed to capture all possible viola-

tions of voltage limits while minimizing the number of false positive alarms.

(*ii*) Reformulations that substantially improve the computational tractability of this bilevel problem.

(*iii*) An approximate gradient descent method to improve solution quality.

(*iv*) Numerical demonstration of our proposed problem formulations for a variety of test cases, including networks with multiple switching configurations.

This paper is organized as follows. Section 2 formulates the sensor placement problem using bilevel optimization. Section 3 proposes different techniques to reformulate the optimization problem. Section 4 provides our numerical tests. Section 5 concludes the paper and discusses future work.

2. Sensor Placement Problem

This section describes the sensor placement problem by introducing notation, presenting the bilevel programming formulation that is the focus of this paper, and detailing the objective function that simultaneously minimizes the number of sensors and reduces the number of false positive alarms.

2.1. Notation

Consider an n-bus power system. The sets of buses and lines are denoted as $\mathcal{N} = \{1, \ldots, n\}$ and $\mathcal{L}$, respectively. One bus in the system is specified as

the slack bus where the voltage is $1\angle 0^\circ$ per unit. For the sake of simplicity, the remaining buses are modeled as PQ buses with given values for their active (P) and reactive (Q) power injections. Extensions to consider PV buses, which have given values for the active power (P) and the voltage magnitude (V), are straightforward. (The controlled voltage magnitudes at PV buses imply that voltage violations cannot occur at these buses so long as the voltage magnitude setpoints are within the voltage limits.) The set of all nonslack buses is denoted as $\mathcal{N}_{PQ}$. The set of neighboring buses to bus i is defined as $\mathcal{N}_i := \{k \,|\, (i,k) \in \mathcal{L}\}$. Subscript $(\cdot)_i$ denotes a quantity at bus i, and subscript $(\cdot)_{ik}$ denotes a quantity associated with the line from bus i to bus k, unless otherwise stated. Conductance (susceptance) is denoted as G (B) as the real (imaginary) part of the admittance.

To illustrate the main concepts in this paper, we consider a balanced single-phase equivalent network representation rather than introducing the additional notation and complexity needed to model an unbalanced three-phase network. Our work does not require assumptions regarding a radial network structure, and we are able to handle multiple network configurations. Extensions to consider other limits such as restrictions on line flows, budget uncertainty sets, and unbalanced three-phase network models impose limited additional complexity.

2.2. Bilevel optimization formulation

The main goal of this problem is to find sensor location(s) such that sensor(s) can capture all possible voltage violations. This paper assumes that the voltages read by the sensors are accurate and noise-free measurements. We formulate this problem as a bilevel optimization with the following upper-level and lower-level problems.

- **Upper level:** Determine sensor locations and alarm thresholds such that when the voltages at the sensors are within the chosen thresholds, the voltages at all other buses are within pre-specified safety limits.

- **Lower level:** Find the extreme achievable voltages at all buses given the sensor locations, sensor alarm thresholds, and the specified range of power injection variability.

The sensor locations and alarm thresholds output from the upper-level problem are input to the lower-level problem, and the extreme achievable voltage magnitudes output from the lower-level problem are used to evaluate the bounds in the upper-level problem. We first introduce notation for various quantities associated with the voltage at bus i:

$\underset{\sim}{V}_i$ ($\widetilde{V}_i$)	: Lower (Upper) sensor alarm threshold.
$\underline{V}_i$ ($\overline{V}_i$)	: Lowest (Highest) achievable voltage obtained from the lower-level problem.
$\underset{\sim}{U}_i$ ($\widetilde{U}_i$)	: Translation of lower (upper) sensor threshold via a big-M formulation; see (1c).

$V_i^{\min}$ ($V_i^{\max}$) : Specified lower (upper) voltage limit.

We formulate the following bilevel optimization formulation:

$$\min_{\mathbf{s},\widetilde{\mathbf{V}},\underset{\sim}{\mathbf{V}}} \; c(\mathbf{s},\widetilde{\mathbf{V}},\underset{\sim}{\mathbf{V}}) \tag{1a}$$

$$\text{s.t. } (\forall i \in \mathcal{N}_{PQ})$$

$$\underline{V}_i \geq V_i^{\min},\ \overline{V}_i \leq V_i^{\max}, \tag{1b}$$

$$\underset{\sim}{U}_i = \underset{\sim}{V}_i s_i,\ \widetilde{U}_i = \widetilde{V}_i s_i + M(1-s_i), \tag{1c}$$

$$\underline{V}_i = \mathbb{L}_i(\mathbf{s},\widetilde{\mathbf{U}},\underset{\sim}{\mathbf{U}}),\ \overline{V}_i = \mathbb{U}_i(\mathbf{s},\widetilde{\mathbf{U}},\underset{\sim}{\mathbf{U}}), \tag{1d}$$

where c is the cost function associated with the placement of sensors including costs for the hardware, installation, communication network, etc. and $\mathbf{s}$ is a vector of sensor locations modeled as binary variables (1 if a sensor is placed, 0 otherwise). All bold quantities are vectors. The quantities $\mathbb{L}_i(\mathbf{s},\widetilde{\mathbf{U}},\underset{\sim}{\mathbf{U}})$ and $\mathbb{U}_i(\mathbf{s},\widetilde{\mathbf{U}},\underset{\sim}{\mathbf{U}})$ are the solutions to the lower-level problems which, for each $i \in \mathcal{N}_{PQ}$, are given by

$$\mathbb{L}_i(\mathbf{s},\widetilde{\mathbf{U}},\underset{\sim}{\mathbf{U}}) = \min_{V_i} \; V_i \quad (\mathbb{U}_i(\mathbf{s},\widetilde{\mathbf{U}},\underset{\sim}{\mathbf{U}}) = \max_{V_i} \; V_i) \tag{2a}$$

$$\text{s.t. } (\forall j \in \mathcal{N}_{PQ})$$

$$P_j = V_j \sum_{k\in\mathcal{N}_j} V_k(G_{jk}\cos(\theta_{jk}) + B_{jk}\sin(\theta_{jk})), \tag{2b}$$

$$Q_j = V_j \sum_{k\in\mathcal{N}_j} V_k(G_{jk}\sin(\theta_{jk}) - B_{jk}\cos(\theta_{jk})), \tag{2c}$$

$$\underset{\sim}{U}_j \leq V_j \leq \widetilde{U}_j, \tag{2d}$$

$$P_j^{\min} \leq P_j \leq P_j^{\max}, \tag{2e}$$

$$Q_j^{\min} \leq Q_j \leq Q_j^{\max}, \tag{2f}$$

where P_j and Q_j denote the active and reactive power injections at bus j within a particular lower-level problem, $\theta_{jk} := \theta_j - \theta_k$ denotes the voltage angle difference between buses j and k, and superscripts max and min denote upper and lower limits, respectively, on the corresponding quantity. The quantities $\mathbb{L}_i$ and $\mathbb{U}_i$ are functions of $\mathbf{s}$, $\widetilde{\mathbf{U}}$, and $\underset{\sim}{\mathbf{U}}$ as shown in (1d), but these dependencies are hereafter omitted for the sake of notational brevity. For the upper-level problem, the objective function in (1a) minimizes a cost function $c(\mathbf{s}, \widetilde{\mathbf{V}}, \underset{\sim}{\mathbf{V}})$ associated with the sensor locations $\mathbf{s}$ and alarm thresholds $\widetilde{\mathbf{V}}$, $\underset{\sim}{\mathbf{V}}$ while ensuring that the extreme achievable voltage magnitudes calculated in the lower-level problem, $\underline{V}_i, \overline{V}_i$ are within safety limits as shown in (1b). The cost function $c(\mathbf{s}, \widetilde{\mathbf{V}}, \underset{\sim}{\mathbf{V}})$ will be detailed in the following subsection. In the lower-level problem, the objective function (2a) computes the maximum or minimum voltage magnitude for each PQ bus $i \in \mathcal{N}_{PQ}$. For each lower-level problem, constraints (2b)–(2c) are the power flow equations at each bus j, constraint (2d) forces the voltage magnitudes to be within voltage alarm thresholds if a sensor is placed at the corresponding bus, and constraints (2e)–(2f) model the range of variability in the net power injections. We typically set $\theta_1 = 0$ as the angle reference.

2.3. Cost function

Overly restrictive sensor thresholds can potentially trigger an alarm even when there are no voltage violations actually occurring in the system, thus resulting in a *false positive*. To reduce both the number of sensors and the number of false positive alarms due to unnecessarily restrictive alarm thresholds, our cost function, $c(\mathbf{s}, \underset{\sim}{\mathbf{V}}, \widetilde{\mathbf{V}})$, is:

$$c(\mathbf{s}, \underset{\sim}{\mathbf{V}}, \widetilde{\mathbf{V}}) = \sum_{i \in \mathcal{N}} c_i(s_i, \underset{\sim}{V}_i, \widetilde{V}_i) \tag{3}$$

where

$$c_i(s_i, \underset{\sim}{V}_i, \widetilde{V}_i) = \begin{cases} (\underset{\sim}{V}_i - V_i^{\min}) + (V_i^{\max} - \widetilde{V}_i) + \delta; & s_i = 1, \\ 0; & s_i = 0, \end{cases} \tag{4}$$

where δ is a specified cost of placing a sensor. When $s_i = 1$, the objective $c(\mathbf{s}, \underset{\sim}{\mathbf{V}}, \widetilde{\mathbf{V}})$ in (4) seeks to reduce the restrictiveness of the sensor alarm thresholds to have fewer false positives. Changing the value of δ in (4) models the tradeoff between placing an additional sensor and making the sensor range more restrictive. This is a crucial part of our formulation since our main goal is to identify a small number of critical locations that carry sufficient information about the feasibility of the entire network. Beyond the clear financial benefit of having to place fewer sensors, this also provides a simple and practical mechanism for deploying corrective actions in real-time. Indeed, when the system operator encounters a voltage violation, a reactive

power compensation protocol that brings the voltages at these few critical locations to within the alarm thresholds will guarantee feasibility of the voltage limits for the entire network.

3. Reformulations of the Sensor Location Problem

The bilevel problem (1) is computationally challenging due to the nonconvexity in the lower-level problem induced by the AC power flow equations in (2b)–(2c) and the presence of two levels. In this section, we provide methods for obtaining a tractable version of the bilevel sensor placement problem. We first use the *conservative linear approximations* of the power flow equations to convert the lower-level problem to a more tractable linear programming formulation that nevertheless retains characteristics from the nonlinear AC power flow equations. This bilevel problem can be reformulated to a single-level problem using the Karush-Kuhn-Tucker (KKT) conditions with binary variables via a big-M formulation [27, 28]. However, as we will show numerically in Section 4, traditional methods for reformulating the bilevel problem into a single-level problem suitable for standard solvers using the KKT conditions turn out to yield computationally burdensome problems. (The full problem setup using the KKT conditions is shown in Appendix A.) We then use various additional reformulation techniques that yield significantly more tractable problems than standard KKT-based reformulations. These reformulations first yield a (single-level) mixed-integer bilinear programming formulation that can be solved using commercial mixed-integer

programming solvers like Gurobi. We further discretize the sensor alarm thresholds and transform the bilinear terms to a mixed-integer linear program (MILP).

3.1. Conservative linear power flow approximations

To address challenges associated with power flow nonlinearities, we employ a linear approximation of the power flow equations that is adaptive (i.e., tailored to a specific system and a range of load variability) and conservative (i.e., intend to over- or under-estimate a quantity of interest to avoid constraint violations). These linear approximations are called *conservative linear approximations* (CLAs) and were first proposed in [26]. As a sample-based approach, the CLAs are computed using the solution to a constrained regression problem across all samples within the range of power injection variability. They linearly relate the voltage magnitudes at a particular bus to the power injections at all PQ buses. These linear approximations can also effectively incorporate the characteristics of more complex components (e.g., tap-changing transformers, smart inverters, etc.), only requiring the ability to apply a power flow solver to the system. Additionally, in the context of long-term planning, the CLAs can be readily computed with knowledge of expected DER locations and their potential power injection ranges. The accuracy and conservativeness of our proposed method is based on the information of the location of DERs and their power injections variability. As inputs, our method uses the net load profiles including the size of PVs when

computing the CLAs. In practice, this data can be obtained by leveraging the extensive existing research on load modeling and monitoring to identify the locations and capabilities of behind-the-meter devices (refer to, e.g., [29, 30]).

An example of an *overestimating* CLA of the voltage magnitude at bus i is the linear expression

$$a_{i,0} + \boldsymbol{a}_{i,1}^T \begin{pmatrix} \mathbf{P} \\ \mathbf{Q} \end{pmatrix}$$

such that the following relationship is satisfied for some specified range of power injections $\mathbf{P}$ and $\mathbf{Q}$:

$$V_i - \left(a_{i,0} + \boldsymbol{a}_{i,1}^T \begin{pmatrix} \mathbf{P} \\ \mathbf{Q} \end{pmatrix} \right) \leq 0, \tag{5}$$

where a_0 and $\boldsymbol{a}_1$ are the coefficients of the CLA computed by solving a constrained regression problem (refer to [26]) and superscript T denotes the transpose. We replace the AC power flow equations in (2b)–(2c) with a CLA as in (5) for all $i \in \mathcal{N}_{PQ}$. The quantities $\mathbb{L}_i$ and $\mathbb{U}_i$ in (2a) become:

$$\mathbb{L}_i = \min_{\mathbf{P},\mathbf{Q}} \ \underline{a}_{i,0} + \underline{\boldsymbol{a}}_{i,1}^T \begin{pmatrix} \mathbf{P} \\ \mathbf{Q} \end{pmatrix}^i \quad \left(\mathbb{U}_i = \max_{\mathbf{P},\mathbf{Q}} \ \overline{a}_{i,0} + \overline{\boldsymbol{a}}_{i,1}^T \begin{pmatrix} \mathbf{P} \\ \mathbf{Q} \end{pmatrix}^i \right) \tag{6}$$

where superscripts i denote quantities associated with the i^{th} lower-level problem. Using conservative linear approximations yields a linear programming formulation for the lower-level problem rather than the non-convex

lower-level problem in (2). By assuming that the conservative linear approximations do indeed reliably over- or under- estimate the voltage magnitudes, it is sufficient to ensure satisfaction of (1b). As a result, solving the reformulationwill compute sensor locations and thresholds where alarms will always be raised if there are indeed violations of the voltage limits.

3.2. Duality of the lower-level problem

One can reformulate a bilevel problem into a single-level problem by *dualizing* the lower-level problem. This technique can only be usefully applied to problems with specific structure where the optimal objective value of the lower-level problem is constrained in the upper-level problem in the appropriate sense ($\max \leq \cdot$ or $\min \geq \cdot$). In this special case, we can significantly improve tractability compared to the KKT formulation.

Let $\underset{\sim}{\mathbf{y}}^i$ be the vector of all dual variables associated with the lower-level problem $\mathbb{L}_i$ and $\widetilde{\mathbf{y}}^i$ be the vector of all dual variables associated with lower-level problem $\mathbb{U}_i$. Let I be the identity matrix of appropriate dimension. By dualizing the lower-level problem (6) with conservative linear power flow approximations as constraints, we obtain the following:

$$\mathbb{L}_i = \max_{\underset{\sim}{\mathbf{y}}^i} \ \mathbf{b}^T \underset{\sim}{\mathbf{y}}^i + \underline{a}_{i,0} \qquad (7a)$$

$$\text{s.t. } A\underset{\sim}{\mathbf{y}}^i = \underline{\boldsymbol{a}}_{i,1}, \qquad (7b)$$

$$\underset{\sim}{\mathbf{y}}^i \geq \mathbf{0}, \qquad (7c)$$

$$\mathbb{U}_i = \min_{\widetilde{\mathbf{y}}^i} \ \mathbf{b}^T \widetilde{\mathbf{y}}^i + \overline{a}_{i,0} \qquad (8a)$$

$$\text{s.t. } A\widetilde{\mathbf{y}}^i = \overline{\boldsymbol{a}}_{i,1}, \qquad (8b)$$

$$\widetilde{\mathbf{y}}^i \leq \mathbf{0}, \qquad (8c)$$

where

$$A = \left[-I, I, \overline{\boldsymbol{a}}_{1,1}, \ldots, \overline{\boldsymbol{a}}_{n,1}, -\underline{\boldsymbol{a}}_{1,1}, \ldots, -\underline{\boldsymbol{a}}_{n,1}\right],$$

$$\mathbf{b} = [(-\mathbf{P}^{\max})^T, (-\mathbf{Q}^{\max})^T, (\mathbf{P}^{\min})^T, (\mathbf{Q}^{\min})^T, \underset{\sim}{U}_1 - \overline{a}_{1,0}, \ldots, \underset{\sim}{U}_n - \overline{a}_{n,0},$$
$$- \widetilde{U}_1 + \underline{a}_{1,0}, \ldots, -\widetilde{U}_n + \underline{a}_{n,0}]^T.$$

Due to strong duality of the linear lower-level problem, the dual (7a) (and (8a)) has the same objective value as (6) and does not directly provide any advantages. However, the problem has a specific structure where objectives from each lower-level problem (7a) and (8a) only appear in a single inequality constraint (1b). Hence, we *only need to show that there exists some choice of duals* $\underset{\sim}{\mathbf{y}}^i$ *and* $\widetilde{\mathbf{y}}^i$ *for which* (1b) *is feasible.* This allows us to obtain a single-level formulation by defining the lower-level coupling quantities via the following set of constraints:

$$\mathbb{L}_i = \mathbf{b}^T \underset{\sim}{\mathbf{y}}^i + \underline{a}_{i,0}, \qquad (9a) \qquad\qquad \mathbb{U}_i = \mathbf{b}^T \widetilde{\mathbf{y}}^i + \overline{a}_{i,0}, \qquad (10a)$$

$$A\,\underset{\sim}{\mathbf{y}}^i = \underline{\boldsymbol{a}}_{i,1}, \qquad (9b) \qquad\qquad A\,\widetilde{\mathbf{y}}^i = \overline{\boldsymbol{a}}_{i,1}, \qquad (10b)$$

$$\underset{\sim}{\mathbf{y}}^i \geq \mathbf{0}. \qquad (9c) \qquad\qquad \widetilde{\mathbf{y}}^i \leq \mathbf{0}. \qquad (10c)$$

We refer to the formulation using (9) and (10) as the "bilinear formulation" due to the bilinear product of the sensor threshold variables ($\underset{\sim}{\mathbf{U}}$ and $\widetilde{\mathbf{U}}$) and the dual variables $\underset{\sim}{\mathbf{y}}^i$ and $\widetilde{\mathbf{y}}^i$ in (9a) and (10a). Using (9) and (10) leads to a single-level optimization problem. However, the latter has the major advantage that no additional binary variables are required (beyond those associated with the sensor locations in the upper-level problem) since there are no analogous equations to the complementarity condition as in the KKT

reformulation. Our bilinear formulation can be further converted into an MILP by discretizing the continuous-valued sensor thresholds. The details for this MILP reformulation and the removal of unnecessary binary variables from discretizing the sensor thresholds (referred to as binary variable removal (BVR)) are described in Appendix B. Further details about the comparison of each problem formulation provided in this paper, including the use of the KKT conditions, are in Appendix C.

3.3. Approximate gradient descent

Solving any of the reformulated bilevel optimization problems may lead to *false positives*. This is both due to the limited number of sensors and the conservative nature of the linear power flow approximations used in the lower-level problem. To reduce the number of false positives, we propose a post-processing step that iteratively adjusts the sensor thresholds that result from the reformulated bilevel optimization problems. We refer to this post-processing step as the Approximate Gradient Descent (AGD) method.

Let superscript k denote the k^{th} iteration of the AGD method. Let ϵ_{AGD} be a step size for adjusting the sensor thresholds and $\mathbf{f}^k$ be the vector of the number of false positives from the sampled power injections. Using the sampled power injections, this method computes an "approximate gradient" indicating how small changes to the sensor alarm thresholds affect the number of false positives. The approximate gradient at iteration k is given by $\mathbf{g}^k$. We denote the set of buses with sensors as $\mathcal{N}_s$. Subscripts give the bus number.

Let Δf_i^k represent the change in the number of false positives among the sampled power injections using the sensor thresholds in the k^{th} iteration when the sensor alarm threshold i is changed by ϵ_{AGD} (leaving all other sensor thresholds unchanged). We then compute an approximate gradient $\mathbf{g}^k$ by comparing the values of Δf_i^k across different buses i:

$$\mathbf{g}^k = \frac{\Delta \mathbf{f}^k}{\sqrt{\sum_{i \in \mathcal{N}_s} (\Delta f_i^k)^2}}. \tag{11}$$

In each iteration, we update the sensor thresholds as follows:

$$\underset{\sim}{\mathbf{V}}^{k+1} = \underset{\sim}{\mathbf{V}}^k + \epsilon_{AGD} \cdot \mathbf{g}^k \quad (\widetilde{\mathbf{V}}^{k+1} = \widetilde{\mathbf{V}}^k + \epsilon_{AGD} \cdot \mathbf{g}^k). \tag{12}$$

The AGD method stops when taking an additional step would result in the appearance of false *negatives*, i.e., undetected violations of voltage limits.

4. Numerical Tests

This section describes numerical experiments on a number of test cases to analyze the sensor locations and thresholds, demonstrate the advantages of our problem reformulations and the post-processing step, and compare results and computational efficiency from different problem formulations.

The test cases we use in these experiments are the 10-bus system *case10ba*, the 33-bus system *case33bw*, and the 141-bus system *case141* from MATPOWER [31]. For the CLAs, we minimize the ℓ_1 error with 1000 samples in the first it-

eration and 4000 additional samples in a sample selection step, and we choose a quadratic output function of voltage magnitude. (See [26] for a discussion on computationally efficient iterative methods for computing CLAs and variants of CLAs that approximate different quantities in order to improve their accuracy.) All power injections vary within 50% to 150% of the load demand values given in the MATPOWER files except for *case33bw* where we consider a variant with solar panels at buses 18 and 33. The active power demands at these two buses vary within -200% to 150% of the nominal values. Note that a manufacturer can provide actual data regarding the range of power injections from DERs like solar PV.

We implement the single-level reformulations of the sensor placement problem in MATLAB using YALMIP [32] and use Gurobi as a solver with a MIP gap tolerance of 0.5%. The AGD step size is $\epsilon_{AGD} = 2 \times 10^{-4}$ per unit. The value of δ (i.e., cost of placing a sensor) in the objective (4) is 0.02. In *case10ba*, *case33bw*, and *case141*, the lower voltage limits are 0.90, 0.91, and 0.92, respectively, and the computation times for the CLAs are 58, 198, and 1415 seconds, respectively.

4.1. Sensor locations

We compare the quality of results and the computation time from the following reformulations: (i) the KKT formulation, (ii) the duality-based bilinear formulation (9) and (10), and (iii) the MILP formulation with the BVR pre-processing.

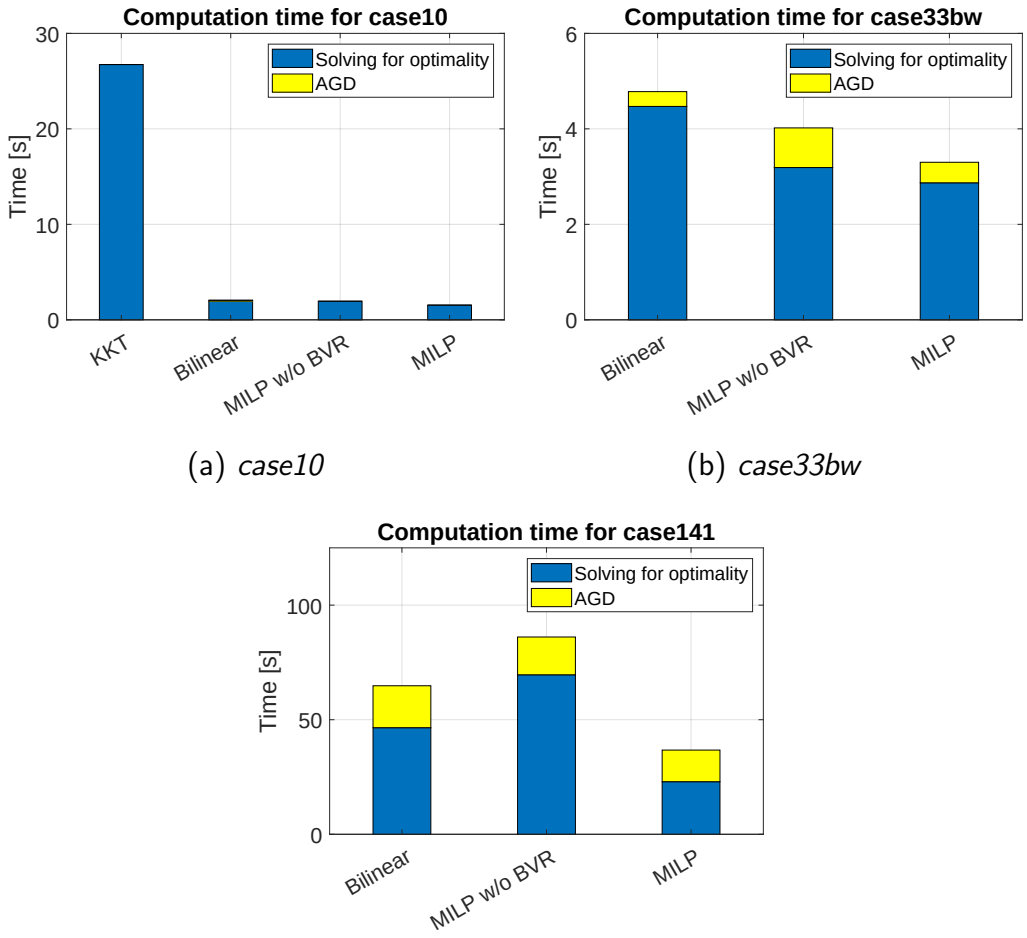
Table 1: Results showing sensor alarm thresholds and number of false positives from KKT, bilinear, and MILP formulations.

		KKT[§]	**Bilinear**			**MILP**		
		case10ba	*case10ba*	*case33bw*	*case141*	*case10ba*	*case33bw*	*case141*
Computation time [s]	Optimality	26.7	1.96	4.47	46.52	1.54	2.87	22.95
	AGD	—	0.11	0.31	18.3	—	0.43	13.8
Sensor location(s)		10	10	14, 15, 17, 31	79, 80, 82, 85	10	14, 30	80, 86
Sensor threshold(s)		0.9	0.9017	0.91, 0.91, 0.9118, 0.9126	0.92, 0.9213, 0.93, 0.93	0.9	0.9195 0.9185	0.929 0.9295
with AGD		—	0.9	0.91, 0.91, 0.9107, 0.9112	0.92, 0.9212, 0.9218, 0.9201	—	0.9163 0.9161	0.9213 0.9201
# feasible points		7317	7317	9753	9955	7317	9753	9955
% false positive(s)		0%	4.01%	1.91%	72.07%	0%	7.64%	66.64%
with AGD		—	0%	0.24%	0.03%	—	1.34%	0.01%
% false negatives		0%	0%	0%	0%	0%	0%	0%

[§]The solver does not find a solution to *case33bw* and *case141* within 55000 seconds.

The first test case is the 10-bus system *case10ba*, a simple single-branch network. We consider a variant where the nominal loads are 60% of the values in the MATPOWER file. The results from each formulation place a sensor at the end of the branch (furthest bus from the substation) with an alarm threshold of 0.9 per unit (at the voltage limit). Fig. 1a compares computation times from the three formulations. The the KKT formulation takes 26.7 seconds while the bilinear and MILP formulations take 1.96 and 1.54 seconds, respectively. Since the sensor threshold for the KKT and MILP formulations is at the voltage limit, AGD is not needed. Conversely, the bilinear formulation gives a higher alarm threshold. As a result, the AGD method is applied as a post-processing step to achieve the lowest possible threshold without introducing false alarms. The number of false positives reduces from 5.48% to 0%. Executing the AGD method takes 0.11 seconds.

The second test case is the 33-bus system *case33bw*, which has multiple branches. In this example, we demonstrate the efficacy of our approach in



(a) *case10* (b) *case33bw*

(c) *case141*

Figure 1: Bar graphs showing the computation times for solving various problem formulations and executing the approximate gradient descent method.

handling a system with complex components through the implementation of volt-VAR control, which represents smarter inverter behavior (whose characteristics are described in [33]). To incorporate the behavior of volt-VAR control, we enhance the power flow solver used to compute the CLAs by integrating an additional fixed-point iterative method. Table 1 shows the

computation times for the bilinear and the two MILP formulations. We exclude the computation time for the KKT formulation since the solver fails to find even a feasible (but potentially suboptimal) point within 55000 seconds (15 hours). Our final test case is the 141-bus system *case141*. Similar to the 33-bus system, the solver could not find the optimal solution for the KKT formulation within a time limit of 15 hours. It is evident the KKT formulation is intractable. Table 1 again shows the results for this test case, and Figs. 1b and 1c compare the computation times for the bilinear and MILP formulations.

Table 1 shows both the computation times and the results of randomly drawing sampled power injections within the specified range of variability, computing the associated voltages by solving the power flow equations, and finding the number of false positive alarms (i.e., the voltage at a bus with a sensor is outside the sensor's threshold but there are no voltage violations in the system). The results for the 33-bus and 141-bus test cases given in Table 1 illustrate the performance of the proposed reformulations. Whereas the KKT formulation is computationally intractable, our proposed reformulations find solutions within approximately one minute, where the MILP formulation with the BVR method typically exhibits the fastest performance. The solutions to the reformulated problems place a small number of sensors (two to four sensors in systems with an order of magnitude or more buses). No solutions suffer from false negatives since all samples where there is a voltage violation trigger an alarm. There are a number of false alarms

prior to applying the AGD that after its application decrease dramatically to a small fraction of the total number of samples (1.34% and 0.01% in the 33-bus and the 141-bus systems, respectively). These observations suggest that our sensor placement formulations provide a computationally efficient method for identifying a small number of sensor locations and associated alarm thresholds that reliably identify voltage constraint violations with no false negatives (missed alarms) and few false positives (spurious alarms).

4.2. Multiple configurations

The previous results described the sensor placements for the *case10ba*, *case33bw*, and *case141* systems in their nominal network topologies. We next demonstrate the effectiveness of our problem reformulations for variants of these systems with multiple network configurations. We consider a variant of the *case33bw* system with three distinct network configurations and two solar PV generators installed at buses 18 and 33, as an illustrative example. Other network configurations are not included from this study, as they either exhibit no violations or yield identical optimal solutions. The first configuration is the nominal topology given in the MATPOWER version of the test case. In the second configuration, the line connecting buses 6 and 7 is removed and a new line connecting buses 4 and 18 is added. The third configuration removes the line connecting buses 6 and 26 and adds a new line connecting buses 25 and 33. All network configurations are displayed in Fig. 2.

Table 2 shows the results from using the bilinear and MILP formulations

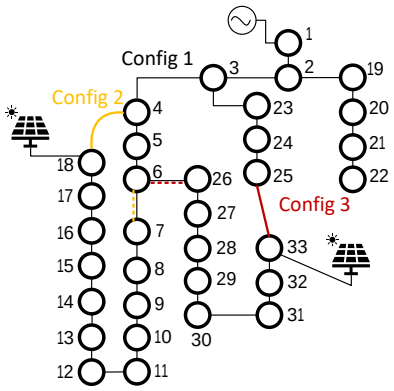


Figure 2: The *case33bw* system with three different network configurations. A dash line means that line is removed when a solid line with the same color is added such that the system has a radial topology in each configuration.

Table 2: Results for the *case33bw* system with three network configurations.

		Bilinear			**MILP**		
		Config 1	**Config 2**	**Config 3**	**Config 1**	**Config 2**	**Config 3**
Computation time [s]	Optimality	20.1			7.93		
	AGD	0.53	0.55	0.10	0.86	0.93	0.40
Sensor location(s)		8, 14, 26, 33			9, 14, 30		
Sensor thresholds		0.91, 0.919, 0.91, 0.9113	0.9126, 0.91, 0.91, 0.9111	0.91, 0.91, 0.9106, 0.9137	0.91, 0.919, 0.919	0.9185, 0.91, 0.919	0.91, 0.91, 0.9165
with AGD		0.91, 0.9167, 0.91, 0.9104	0.9117, 0.91, 0.91, 0.9109	0.91, 0.91, 0.9101, 0.9136	0.91, 0.9167, 0.9168	0.9166, 0.91, 0.9189	0.91, 0.91, 0.9151
# feasible points		9560	8292	9112	9560	8292	9112
% false positives		6.19%	4.86%	1.28%	9.19%	12.38%	6.55%
with AGD		1.86%	1.21%	0.05%	2.87%	3.38%	2.38%
% false negatives		0%	0%	0%	0%	0%	0%

to solve the multiple-configuration problem for this case. The results generally mirror those from the single-network-configuration test cases shown earlier in that computation times are still reasonable (approximately a factor of four larger) and there are no false negatives and a small number of false positives after applying the AGD method.

We note that some configurations may not need to utilize all available sensors. To show this, we describe an experiment that considers each config-

uration separately. In this experiment, we compare the number of sensors and the locations of the sensors with those in the previous experiment. As Table 3 shows, configurations 1 and 2 require only two sensors while configuration 3 requires only one sensor as opposed to three-sensor solution obtained from the multiple-configuration problem. This demonstrates the need to jointly consider network topologies in one problem for such situations.

Table 3: Results for *case33bw* with three network configurations where sensor locations are not necessary the same.

	Config 1	**Config 2**	**Config 3**
Sensor location(s)	14, 30	9, 31	30
Sensor threshold(s)	0.9195 0.9190	0.9185 0.9185	0.9185
with AGD	0.9167 0.9168	0.9164 0.9186	0.9151
# feasible points	9560	8292	9112
% false positives	10.48%	12.45%	14.53%
with AGD	2.89%	3.06%	2.38%
% false negatives	0%	0%	0%

5. Conclusion

This paper has formulated a bilevel optimization problem that seeks to minimize the number of sensors needed to detect violations of voltage magnitude limits in an electric distribution system. We first addressed the power flow nonlinearities in the lower-level problem via previously developed conservative linear approximations of the power flow equations. To handle computational challenges from the bilevel nature, we exploited the structure of the problem to obtain single-level mixed-integer programming formulations,

thereby avoiding the introduction of unnecessary additional discrete variables. We also developed a mixed-integer-linear programming (MILP) formulation by discretizing the sensor thresholds. Our MILP reformulation proves tractable across all test cases, a notable advantage over the traditional KKT reformulation, which only remains tractable for the 10-bus system. Furthermore, the MILP formulation yields a reduced number of sensors compared to the bilinear formulation. Additionally, we extended these reformulations to accommodate multiple network topologies, wherein several lines can be opened or closed without any nodes being isolated. Our proposed sensor placement reformulations significantly reduce computation time compared to standard techniques.

Furthermore, we developed a post-processing technique to minimize false alarms using an approximate gradient descent method. The integration of this technique with our bilevel problem reformulation ensures the computation of sensor locations and alarm thresholds that yield minimal false alarms and no missed alarms. This was validated through numerical testing conducted via out-of-sample testing.

For enhanced computational efficiency, we may further expedite computations by considering sensor placements solely at locations where voltage violations occurred, utilizing insights from sample-based CLAs. Additionally, employing a constraint generation approach enables us to explicitly formulate constraints associated only with lower-level problems where violations were observed in the samples used for CLA computation. Essentially, lever-

aging samples from CLA computation minimizes the necessity for explicit modeling of variables and constraints in the bilevel sensor placement problem. Subsequently, conducting out-of-sample testing, as demonstrated in our numerical analyses, can validate the efficacy of these approaches in further enhancing tractability.

In our future work, we seek to identify where the violations occur using the information obtained from CLAs and solutions from the sensor placement problem. We intend to use the sensor locations and thresholds resulting from the proposed formulations to design corrective control actions, including voltage controls, that ensure all voltages remain within safety limits. These actions may also involve adjustments to generation levels, switching devices, or control of reactive power sources.

Expanding our research scope, we aim to work on real-time operations, focusing particularly on optimizing the utilization of the limited communication bandwidth accessible at smart meters. Our objective is to devise strategies to efficiently extract the most relevant real-time information from a network of sensors, therefore enhancing the overall operational effectiveness. Additionally, we intend to address this challenge while considering potential cyber threats, such as false data injection attacks [34].

Additionally, it is worth noting that prior research has demonstrated the tractability of extending the CLA framework to unbalanced three-phase systems [35]. We intend to explore the extension of our work to three-phase networks in future work.

Acknowledgement

Support from NSF #2023140 and PSERC #T-64 (P. Buason and D.K. Molzahn), NSF Graduate Research Fellowship Program under Grant No. DGE-1650044 (S. Talkington), and DOE Office of Electricity Advanced Grid Modeling Program (S. Misra).

References

[1] A. N. Samudrala, M. H. Amini, S. Kar, and R. S. Blum, "Optimal sensor placement for topology identification in smart power grids," in *53rd Annual Conference on Information Sciences and Systems (CISS)*, 2019.

[2] ——, "Sensor placement for outage identifiability in power distribution networks," *IEEE Transactions on Smart Grid*, vol. 11, no. 3, pp. 1996–2013, 2020.

[3] T. Baldwin, L. Mili, M. Boisen, and R. Adapa, "Power system observability with minimal phasor measurement placement," *IEEE Transactions on Power Systems*, vol. 8, no. 2, pp. 707–715, 1993.

[4] B. Gou, "Optimal placement of PMUs by integer linear programming," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 1525–1526, 2008.

[5] R. J. Albuquerque and V. L. Paucar, "Evaluation of the PMUs measurement channels availability for observability analysis," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 2536–2544, 2013.

[6] M. S. Thomas, S. Ranjan, and N. Bhaskar, "Optimization of PMU placement by performing observability analysis," in *IEEE 6th India International Conference on Power Electronics (IICPE)*, 2014.

[7] F. Aminifar, M. Fotuhi-Firuzabad, M. Shahidehpour, and A. Khodaei, "Probabilistic multistage PMU placement in electric power systems," *IEEE Transactions on Power Delivery*, vol. 26, no. 2, pp. 841–849, 2011.

[8] H. Mehrjerdi, S. Lefebvre, D. Asber, and M. Saad, "Eliminating voltage violations in power systems using secondary voltage control and decentralized neural network," in *IEEE Power & Energy Society General Meeting*, 2013.

[9] A. M. Nour, A. Y. Hatata, A. A. Helal, and M. M. El-Saadawi, "Review on voltage-violation mitigation techniques of distribution networks with distributed rooftop PV systems," *IET Generation, Transmission & Distribution*, vol. 14, no. 3, pp. 349–361, 2020.

[10] N. M. Dehkordi and S. Z. Moussavi, "Distributed resilient adaptive control of islanded microgrids under sensor/actuator faults," *IEEE Transactions on Smart Grid*, vol. 11, no. 3, pp. 2699–2708, 2020.

[11] H. M. Soliman, E. H. E. Bayoumi, F. A. El-Sheikhi, and M. De Santis, "Decentralized sensor fault-tolerant control of dc microgrids using the attracting ellipsoid method," *Sensors*, vol. 23, no. 16, 2023.

[12] W. Murray, M. Adonis, and A. Raji, "Voltage control in future electrical distribution networks," *Renewable and Sustainable Energy Reviews*, vol. 146, p. 111100, 2021.

[13] D. K. Mohanta, C. Murthy, and D. S. Roy, "A brief review of phasor measurement units as sensors for smart grid," *Electric Power Components and Systems*, vol. 44, no. 4, pp. 411–425, 2016.

[14] M. Jamei, A. Scaglione, C. Roberts, E. Stewart, S. Peisert, C. McParland, and A. McEachern, "Anomaly detection using optimally placed μPMU sensors in distribution grids," *IEEE Transactions on Power Systems*, vol. 33, no. 4, pp. 3611–3623, 2018.

[15] E. Dusabimana and S.-G. Yoon, "A survey on the micro-phasor measurement unit in distribution networks," *Electronics*, vol. 9, no. 2, 2020. [Online]. Available: https://www.mdpi.com/2079-9292/9/2/305

[16] K. McKenna, P. Gotseff, M. Chee, and E. Ifuku, "Advanced metering infrastructure for distribution planning and operation: Closing the loop on grid-edge visibility," *IEEE Electrification Magazine*, vol. 10, no. 4, pp. 58–65, 2022.

[17] X. Wang, C. Wang, T. Xu, H. Meng, P. Li, and L. Yu, "Distributed voltage control for active distribution networks based on distribution phasor measurement units," *Applied energy*, vol. 229, pp. 804–813, 2018.

[18] R. Gupta, F. Sossan, and M. Paolone, "Model-less robust voltage control in active distribution networks using sensitivity coefficients estimated from measurements," *Electric Power Systems Research*, vol. 212, p. 108547, 2022.

[19] M. Baran and A. Kelley, "State estimation for real-time monitoring of distribution systems," *IEEE Transactions on Power Systems*, vol. 9, no. 3, pp. 1601–1609, 1994.

[20] A. Primadianto and C.-N. Lu, "A review on distribution system state estimation," *IEEE Transactions on Power Systems*, vol. 32, no. 5, pp. 3875–3883, 2017.

[21] I. D. Melo, M. O. N. Teixeira, and J. S. Mingorança, "Neutral-to-Earth Voltage (NEV) and state estimation for unbalanced multiphase distribution systems based on an optimization model," *Electric Power Systems Research*, vol. 217, p. 109123, Apr. 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378779623000123

[22] O. Stanojev, L. Werner, S. Low, and G. Hug, "Tractable identification of electric distribution networks," in *2023 62nd IEEE Conference on Decision and Control (CDC)*, 2023, pp. 549–555.

[23] D. Deka, V. Kekatos, and G. Cavraro, "Learning distribution grid topologies: A tutorial," *IEEE Transactions on Smart Grid*, vol. 15, no. 1, pp. 999–1013, 2024.

[24] P. A. Schirmer and I. Mporas, "Non-intrusive load monitoring: A review," *IEEE Transactions on Smart Grid*, vol. 14, no. 1, pp. 769–784, 2023.

[25] M. Vanin, F. Geth, R. D'hulst, and D. Van Hertem, "Combined unbalanced distribution system state and line impedance matrix estimation," *International Journal of Electrical Power & Energy Systems*, vol. 151, p. 109155, 2023.

[26] P. Buason, S. Misra, and D. K. Molzahn, "A sample-based approach for computing conservative linear power flow approximations," *Electric Power Systems Research*, vol. 212, p. 108579, 2022, presented at the *22nd Power Systems Computation Conference (PSCC 2022)*.

[27] U.-P. Wen and S.-T. Hsu, "Linear bi-level programming problems – A review," *The Journal of the Operational Research Society*, vol. 42, no. 2, pp. 125–133, 1991.

[28] S. Dempe and A. Zemkoho, "On the Karush–Kuhn–Tucker reformulation of the bilevel optimization problem," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 75, no. 3, pp. 1202–1218, 2012.

[29] S. Grijalva, A. U. Khan, J. S. Mbeleg, C. Gomez-Peces, M. J. Reno, and L. Blakely, "Estimation of PV location in distribution systems based on voltage sensitivities," in *2020 52nd North American Power Symposium (NAPS)*, 2021, pp. 1–6.

[30] P. A. Schirmer and I. Mporas, "Non-intrusive load monitoring: A review," *IEEE Transactions on Smart Grid*, vol. 14, no. 1, pp. 769–784, 2023.

[31] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 12–19, Feb. 2011.

[32] J. Löfberg, "YALMIP: A toolbox for modeling and optimization in MATLAB," in *IEEE International Symposium on Computer Aided Control Systems Design (CACSD)*, September 2004, pp. 284–289.

[33] F. Ding, A. Nagarajan, S. Chakraborty, M. Baggu, A. Nguyen, S. Walinga, M. McCarty, and F. Bell, "Photovoltaic impact assessment of smart inverter Volt-VAR control on distribution system conservation voltage reduction and power quality," Dec. 2016. [Online]. Available: https://www.osti.gov/biblio/1337541

[34] G. Liang, J. Zhao, F. Luo, S. R. Weller, and Z. Y. Dong, "A review

of false data injection attacks against modern power systems," *IEEE Transactions on Smart Grid*, vol. 8, no. 4, pp. 1630–1638, 2017.

[35] A. D. O. Aquino, S. Talkington, and D. K. Molzahn, "Managing vehicle charging during emergencies via conservative distribution system modeling," 2023, to Appear in the 2024 Texas Power and Energy Conference (TPEC 2024).

[36] G. P. McCormick, "Computability of global solutions to factorable nonconvex programs: Part I–Convex underestimating problems," *Mathematical Programming*, vol. 10, no. 1, pp. 147–175, 1976.

[37] S. Pineda and J. M. Morales, "Solving linear bilevel problems using big-Ms: Not all that glitters is gold," *IEEE Transactions on Power Systems*, vol. 34, no. 3, pp. 2469–2471, 2019.

Appendix A. Reformulation using KKT constraints

With a linear lower-level problem, we can apply standard techniques for reformulating the bilevel problem (1) with CLAs as a (single-level) mixed-integer linear program (MILP). These techniques replace the lower-level problem (6) with its KKT conditions that are both necessary and sufficient for optimality of this problem [28] and also apply McCormick envelopes [36] to convert the bilinear product of the continuous and discrete variables in (1c) to an equivalent linear form. The resulting single-level problem still involves bilinear constraints associated with the complementarity conditions. These bilinear constraints are traditionally addressed using binary variables in a "Big-M" formulation. Commercial MILP solvers are applicable to this traditional reformulation, which we denote throughout the paper as the "KKT formulation". This formulation is obtained by defining the lower-level coupling quantities $\mathbb{L}_i$ and $\mathbb{U}_i$ using the KKT conditions given below:

$$\mathbb{L}_i = \underline{a}_{i,0} + \underline{\boldsymbol{a}}_{i,1}^T \begin{pmatrix} \mathbf{P} \\ \mathbf{Q} \end{pmatrix}^i \quad \left(\mathbb{U}_i = \ \overline{a}_{i,0} + \overline{\boldsymbol{a}}_{i,1}^T \begin{pmatrix} \mathbf{P} \\ \mathbf{Q} \end{pmatrix}^i \right), \tag{A.1a}$$

$$(\forall j \in N_{PQ} \setminus \{i\})$$

$$\underline{a}_{i,1} - \sum_{k \in \mathcal{N}_{PQ}} \lambda_k^i \overline{\boldsymbol{a}}_{k,1} + \sum_{k \in \mathcal{N}_{PQ}} \mu_k^i \underline{\boldsymbol{a}}_{k,1}$$

$$- \sum_{k \in \mathcal{N}_{PQ}} \gamma_k^i e_k + \sum_{k \in \mathcal{N}_{PQ}} \pi_k^i e_k = 0$$

$$(\text{for } \mathbb{U}_i : \overline{a}_{i,1} - \sum_{k \in \mathcal{N}_{PQ}} \lambda_k^i \overline{\boldsymbol{a}}_{k,1} + \sum_{k \in \mathcal{N}_{PQ}} \mu_k^i \underline{\boldsymbol{a}}_{k,1}$$

$$- \sum_{k \in \mathcal{N}_{PQ}} \gamma_k^i e_k + \sum_{k \in \mathcal{N}_{PQ}} \pi_k^i e_k = 0), \tag{A.1b}$$

$$\overline{a}_{j,0} + \overline{\boldsymbol{a}}_{j,1}^T \begin{pmatrix} \mathbf{P} \\ \mathbf{Q} \end{pmatrix}^i \geq \underset{\sim}{U}_j, \tag{A.1c}$$

$$\underline{a}_{j,0} + \underline{\boldsymbol{a}}_{j,1}^T \begin{pmatrix} \mathbf{P} \\ \mathbf{Q} \end{pmatrix}^i \leq \widetilde{U}_j, \tag{A.1d}$$

$$\begin{pmatrix} \mathbf{P} \\ \mathbf{Q} \end{pmatrix}^{\min} \leq \begin{pmatrix} \mathbf{P} \\ \mathbf{Q} \end{pmatrix}^i \leq \begin{pmatrix} \mathbf{P} \\ \mathbf{Q} \end{pmatrix}^{\max}, \tag{A.1e}$$

$$\lambda_j^i \cdot \left(\underset{\sim}{U}_j - \overline{a}_{j,0} - \overline{\boldsymbol{a}}_{j,1}^T \begin{pmatrix} \mathbf{P} \\ \mathbf{Q} \end{pmatrix}^i \right) = 0, \tag{A.1f}$$

$$\mu_j^i \cdot \left(-\widetilde{U}_j + \underline{a}_{j,0} + \underline{\boldsymbol{a}}_{j,1}^T \begin{pmatrix} \mathbf{P} \\ \mathbf{Q} \end{pmatrix}^i \right) = 0, \tag{A.1g}$$

$$\boldsymbol{\gamma}^i \odot \left(\begin{pmatrix} \mathbf{P} \\ \mathbf{Q} \end{pmatrix}^{\min} - \begin{pmatrix} \mathbf{P} \\ \mathbf{Q} \end{pmatrix}^i \right) = \mathbf{0}, \tag{A.1h}$$

$$\boldsymbol{\pi}^i \odot \left(\begin{pmatrix} \mathbf{P} \\ \mathbf{Q} \end{pmatrix}^i - \begin{pmatrix} \mathbf{P} \\ \mathbf{Q} \end{pmatrix}^{\max} \right) = \mathbf{0}, \tag{A.1i}$$

$$\boldsymbol{\gamma}^i, \boldsymbol{\pi}^i \geq \mathbf{0}; \lambda_j^i, \mu_j^i \geq 0, \ \forall j \in N_{PQ} \setminus \{i\}, \tag{A.1j}$$

where the operator $\odot$ is the element-wise multiplication; e_i is the i^{th} column of the identity matrix; $\boldsymbol{\lambda}$, $\boldsymbol{\mu}$, $\boldsymbol{\gamma} := (\gamma_1, \gamma_2, \ldots, \gamma_{2m})^T$, and $\boldsymbol{\pi} := (\pi_1, \pi_2, \ldots, \pi_{2m})^T$ are dual variables associated with the voltage and power

injection bounds. Note that the solution to the set of equations in $\mathbb{L}_i$ is completely decoupled from that in $\mathbb{U}_i$. Equations (A.1b)–(A.1i) are the KKT conditions of the lower-level problem. Equation (A.1b) is the stationarity condition. Equations (A.1c)–(A.1e) are the primal feasibility conditions. The complementary slackness conditions are (A.1f)–(A.1i) and the dual feasibility condition is (A.1j). Observe that the complementary slackness conditions give rise to nonlinear functions due to the multiplication of the dual variables $\boldsymbol{\lambda}$, $\boldsymbol{\mu}$, $\boldsymbol{\gamma}$, and $\boldsymbol{\pi}$ with the primal variables $\mathbf{P}$ and $\mathbf{Q}$. To handle these nonlinearities, traditional methods for bilevel optimization replace these products using additional binary variables and a big-M reformulation. This requires bounds on the dual variables that are difficult to determine, and bad choices for these bounds can result in either infeasibility or poor computational performance [37].

Appendix B. Bilinear to mixed-integer linear programming

The bilinear formulation can be further converted into an MILP by discretizing the continuous-valued sensor thresholds. This formulation has the advantage of being within the scope of a larger range of mixed-integer programming solvers since not all of them can handle billinear forms. We partition the sensor threshold ranges into d discrete steps with size ϵ and define the vectors of threshold variables, $\underset{\sim}{\mathbf{U}}$ and $\widetilde{\mathbf{U}}$, as

$$\underset{\sim}{\mathbf{U}} = \eta^T \underset{\sim}{v}_t \quad (\widetilde{\mathbf{U}} = \eta^T \widetilde{v}_t), \tag{B.1}$$

where

$$\eta = \begin{bmatrix} \eta_{0,1} & \eta_{0,2} & \cdots & \eta_{0,n} \\ \eta_{1,1} & \eta_{1,2} & \cdots & \eta_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ \eta_{d,1} & \eta_{d,2} & \cdots & \eta_{d,n} \end{bmatrix}, \tag{B.2a}$$

$$\underset{\sim}{v}_t = \left[0, V^{\min}, V^{\min} + \epsilon, \ldots, V^{\min} + (d-1)\epsilon\right]^T$$

$$(\widetilde{v}_t = \left[0, V^{\max} - \epsilon, V^{\max} - 2\epsilon, \ldots, V^{\max}\right]^T), \tag{B.2b}$$

$$\sum_{i=1}^{d} \eta_{i,k} = 1, \quad \forall k \in \mathcal{N}_{PQ}. \tag{B.2c}$$

Note that this discretization exploits the fact that any sensor threshold will necessarily be above the lower voltage limit $V_i^{\min}$ and below the upper voltage limit $V_i^{\max}$. Equations (B.2a)–(B.2c) imply that when $\eta_{0,i} = 0$, no sensor is placed (i.e., $s_i = 0$). Using this discretization, the constraints (9a) and (10a) now contain bilinear products of binary variables. These products can be equivalently transformed into a mixed-integer *linear* (as opposed to bilinear) programming formulation using McCormick Envelopes [36]. With McCormick Envelopes and discrete sensor thresholds, the problems (9) and (10) become a MILP that can be computed using any MILP solver. We refer to the reformulation of the lower-level problems (9) and (10) using the discretization (B.1) as the "MILP formulation".

To further improve tractability, we can remove unnecessary binary vari-

ables by inspecting data from the samples of power injections used to compute the conservative linear approximations of the power flow equations. Let b be a bus where the voltage magnitude never reaches the highest discretized sensor threshold value (i.e., $V^{\min} + (d-1)\epsilon$) in any of the sampled power injections. Given a sufficiently comprehensive sampling of the range of possible power injections, we can then simplify the discretized representation of the sensor alarm threshold as:

$$\underset{\sim}{U}_b = 0 \cdot \eta_{0,b} + V^{\min} \cdot \eta_{1,b}, \tag{B.3a}$$

$$\eta_{0,b} + \eta_{1,b} = 1. \tag{B.3b}$$

A similar simplification can be used for the upper sensor thresholds. This pre-screening thus eliminates binary variables associated with sensor thresholds at buses that will never violate their voltage limits. We henceforth call this data-driven simplification technique "binary variable removal" (BVR).

Appendix C. Comparisons of each formulation

The previous subsections present several problem reformulations that convert the bilevel sensor placement problem (1) into various single-level problems that can be solved with mixed-integer solvers like Gurobi. Each reformulation has different computational characteristics and yields solutions with differing accuracy. We next compare the KKT formulation (A.1) described in Appendix A with the duality-based bilinear formulation (9) and

(10) described in Section 3.2 according to the numbers of decision variables and constraints.

Both formulations involve bilinear terms but the bilinear formulation is more compact. Consider a system with b PQ buses, of which there are r buses where the voltage magnitudes may violate their limits after the pre-screening described in Appendix B. The total number of decision variables in the KKT formulation is $5 \cdot b \cdot r + 2 \cdot b$ ($2 \cdot b \cdot r$ from power injections, $3 \cdot b \cdot r$ from dual variables, b from the voltage thresholds, and b from the sensor locations). Our proposed duality-based bilinear formulation involves only $3 \cdot b \cdot r + 2 \cdot b$ decision variables. The reduction happens because the variables corresponding to power injections are entirely removed by duality.

Regarding the number of constraints, the duality-based bilinear formulation does not have the stationarity conditions, primal feasibility, or power injections directly involved, resulting in a reduction of $2 \cdot b \cdot r + b \cdot r + 2 \cdot b \cdot r = 5 \cdot b \cdot r$ constraints. The implications of these differences on tractability is assessed via the solution times presented in Section 4.