Differentially Private Optimal Power Flow for Distribution Grids

Vladimir Dvorkin Jr., *Student member, IEEE,* Ferdinando Fioretto, Pascal Van Hentenryck, *Member, IEEE,*
Pierre Pinson, *Fellow, IEEE,* and Jalal Kazempour, *Senior Member, IEEE*

Abstract—**Although distribution grid customers are obliged to share their consumption data with distribution system operators (DSOs), a possible leakage of this data is often disregarded in operational routines of DSOs. This paper introduces a privacy-preserving optimal power flow (OPF) mechanism for distribution grids that secures customer privacy from unauthorised access to OPF solutions, e.g., current and voltage measurements. The mechanism is based on the framework of *differential privacy* that allows to control the participation risks of individuals in a dataset by applying a carefully calibrated noise to the output of a computation. Unlike existing private mechanisms, this mechanism does not apply the noise to the optimization parameters or its result. Instead, it optimizes OPF variables as affine functions of the random noise, which weakens the correlation between the grid loads and OPF variables. To ensure feasibility of the randomized OPF solution, the mechanism makes use of chance constraints enforced on the grid limits. The mechanism is further extended to control the optimality loss induced by the random noise, as well as the variance of OPF variables. The paper shows that the differentially private OPF solution does not leak customer loads up to specified parameters.**

Index Terms—**Data obfuscation, optimization methods, privacy**

I. INTRODUCTION

THE increasing observability of distribution grids enables advanced operational practices for distribution system operators (DSOs). In particular, high-resolution voltage and current measurements available to DSOs allow for continuously steering the system operation towards an optimal power flow (OPF) solution [1]–[3]. However, when collected, these measurements expose distribution grid customers to privacy breaches. Several studies have shown that the measurements of OPF variables can be used by an adversary to identify the type of appliances and load patterns of grid customers [4], [5]. The public response to these privacy risks has been demonstrated by the Dutch Parliament's decision to thwart the deployment of smart meters until the privacy concerns are resolved [6].

Although grid customers tend to entrust DSOs with their data in exchange for a reliable supply, their privacy rights are often disregarded in operational routines of DSOs. To resolve this issue, this paper augments the OPF computations with the preservation of customer privacy in the following sense.

V. Dvorkin Jr., P. Pinson, and J. Kazempour are with the Technical University of Denmark, Kgs. Lyngby, Denmark. F. Fioretto is with the Syracuse University, Syracuse, NY, USA. P. Van Hentenryck is with the Georgia Institute of Technology, Atlanta, GA, USA.

Definition 1 (Customer privacy)**.** *The right of grid customers to be secured from an unauthorized disclosure of sensitive information that can be inferred from the OPF solution.*

To ensure this right, privacy needs to be rigorously quantified and guaranteed. Differential privacy (DP) [7] is a strong privacy notion that quantifies and bounds privacy risks in computations involving sensitive datasets. By augmenting the computations with a carefully calibrated *random* noise, a DP *mechanism* guarantees that the noisy results do not disclose the attributes of individual items in a dataset. Chaudhuri *et al.* [8] and Hsu *et al.* [9] introduced several mechanisms to solve optimization models while preventing the recovery of the input data from optimization results. These mechanisms apply noise to either the parameters or the results of an optimization. The applied noise, however, fundamentally alters the optimization problem of interest. Therefore, the direct application of these mechanisms to OPF problems has been limited. First, they may fail to provide a *feasible* solution for constrained optimization problems. To restore feasibility, they require an additional level of complexity such as the post-processing steps proposed in [10], [11]. Second, although these mechanisms provide bounds on the worst-case performance, they do not consider the *optimality loss* as a control variable. As a result, they cannot provide appropriate trade-offs between the expected and the worst-case mechanism performances. Finally, the previously proposed mechanisms overlook the impact of the noise on the *variance* of the optimization results. Hence, their direct application to OPF problems may lead to undesired overloads of system components [12].

Contributions: To overcome these limitations, this paper proposes a novel differentially private OPF mechanism that does not add the noise to the optimization parameters or to the results. Instead, it obtains DP by optimizing OPF variables as *affine functions* of the noise, bypassing the above-mentioned theoretical drawbacks. More precisely, the paper makes the following contributions:

1. The proposed mechanism produces a random OPF solution that follows a Normal distribution and guarantees $(\varepsilon, \delta)-$differential privacy [7]. Parameters ε and δ, respectively, bound the multiplicative and additive differences between the probability distributions of OPF solutions obtained on adjacent datasets (i.e., differing in at most one load value). The mechanism is particularly suitable for protecting grid loads from unauthorized access to OPF solutions, as fine-tuned (ε, δ) values make randomized OPF solutions similar, irrespective of the used load dataset.

2. The mechanism enforces chance constraints on random OPF variables to guarantee solution feasibility for a given constraint satisfaction probability. This way, it does not require a post-processing step to restore OPF feasibility, as in [10] and [11]. Since the OPF variables are affine in the Gaussian noise, the chance constraints are reformulated into computationally efficient second-order cone constraints.

3. The mechanism enables the control of random OPF outcomes without weakening the DP guarantees. Using results from stochastic programming [13], the optimality loss induced by the noise is controlled using Conditional Value-at-Risk (CVaR) risk measure, enabling a trade-off between the expected and the worst-case performance. Furthermore, with a variance-aware control from [14], the mechanism attains DP with a smaller variance of OPF variables.

Broader Impact: Distribution OPF proposals have been around for at least a decade, though their adoption in real operations is complicated by the need of utilizing load datasets, which raises significant privacy concerns by many regulators worldwide. The adoption of the proposed mechanism, in turn, extends standard OPF models to enable a privacy-cognizant utilization[1] of this data, thus facilitating the digitization of the energy sector. The mechanism treats the DSOs as trustworthy parties and places them on the same ground with the digital service providers, e.g. Amazon, enabling the regulation and securing legal responsibility of the digitalized distribution grids under modern data protection and privacy standards, including the General Data Protection Rights (GDPR) in the European Union, the California Consumer Privacy Act (CCPA) and the New York Privacy Act (NYPA) in the United States. Moreover, the mechanism provides the means to hedge the financial risks of the DSOs by avoiding the cost incurred by privacy violations, such as legal costs, as it relies on a strong quantification of privacy and co-optimizes the joint cost of energy supply and privacy.

Related Work: Thanks to its strong privacy guarantees, DP has been recently applied to private OPF computations. In particular, the mechanism of Zhou *et al.* [15] releases aggregated OPF statistics, e.g., aggregated load, while ensuring the privacy for individual loads, even if all but one loads are compromised. The proposals by Fioretto *et al.* [10] and Mak *et al.* [11] provide a differentially private mechanism to release high-fidelity OPF datasets (e.g., load and network parameters) from real power systems while minimizing the risks of disclosing the actual system parameters. The mechanisms, however, are meant for the private release of aggregate statistics and input datasets and do not provide the OPF solution itself.

Private OPF computations have also been studied in a decentralized and distributed setting. Dvorkin *et al.* [16] designed a distributed OPF algorithm with a differentially private exchange of coordination signals, hence preventing the leakage of the sensitive information contained in the algorithm subproblems. Han *et al.* [17] proposed a privacy-aware distributed coordination scheme for electrical vehicle charging. The privacy frameworks in [16] and [17], however,

---

[1]Note that privacy concerns a safe utilization of the data, not its storage, which falls within the field of cyber-security.

are not suitable for centralized computations and solely focus on the privacy leakage through the exchange of coordination signals. Moreover, to negate the privacy loss induced at every iteration, they require scaling the parameters of the random perturbation, thus involving larger optimality losses and poorer convergence. The centralized mechanisms proposed in this work, however, allow obtaining the private OPF solution in a single computation run. In distribution systems, Zhang *et al.* [18], among other proposals reviewed in [6], designed a privacy-aware optimization of behind-the-meter energy storage systems to prevent the leakage of consumption data from the smart meter readings. However, they disregard OPF feasibility of distribution systems, which has to be preserved in all circumstances.

Paper Organization: Following the preliminaries in Section II, Section III formalizes the privacy goals and provides an overview of the proposed solution. Section IV provides the formulation of the proposed privacy-preserving OPF mechanism and its properties, whereas Section V presents its extensions. Section VI provides numerical experiments and Section VII concludes. The proofs are relegated to the appendix.

II. PRELIMINARIES

A. Optimal Power Flow Problem

The paper considers a low-voltage radial distribution grid with controllable distributed energy resources (DERs). A DSO is responsible for controlling the DERs and supplying power from the high-voltage grid while meeting the technical limits of the grid. The grid is modeled as a graph $\Gamma(\mathcal{N}, \mathcal{L})$, where $\mathcal{N} = \{0, 1, \ldots, n\}$ is the set of nodes and $\mathcal{L} = \mathcal{N} \setminus \{0\}$ is the set of lines connecting those nodes. The root node, indexed by 0, is a substation with a large capacity and fixed voltage magnitude $v_0 = 1$. The radial topology, depicted in Fig. 1, associates each node $i \in \mathcal{N}$ with the sets $\mathcal{U}_i$ and $\mathcal{D}_i$ of, respectively, upstream and downstream nodes, as well as with the set $\mathcal{R}_i$ of nodes on the path to the root node.

Each node i is characterized by its fixed active d_i^p and reactive d_i^q power load and by its voltage magnitude $v_i \in [\underline{v}_i, \overline{v}_i]$. For modeling convenience, the voltage variables are substituted by $u_i = v_i^2$, $\forall i \in \mathcal{N}$. A controllable DER sited at node i outputs an amount of active $g_i^p \in [\underline{g}_i^p, \overline{g}_i^p]$ and reactive $g_i^q \in [\underline{g}_i^q, \overline{g}_i^q]$ power. Its costs are linear with a cost coefficient c_i. To model the relation between the active and reactive DER power output, the constant power factor $\tan \phi_i$ is assumed for each node i. The active and reactive power flows, f_ℓ^p and f_ℓ^q, $\forall \ell \in \mathcal{L}$, respectively, are constrained by the apparent power limit $\overline{f}_\ell$, and each line is characterized by its resistance r_ℓ and reactance x_ℓ. The deterministic OPF model is formulated as:

$$\text{D-OPF:} \quad \min_{g^\dagger, f^\dagger, u} \sum_{i \in \mathcal{N}} c_i g_i^p \tag{1a}$$

$$\text{s.t.} \quad g_0^\dagger = \sum_{i \in \mathcal{D}_0} (d_i^\dagger - g_i^\dagger), \ u_0 = 1, \tag{1b}$$

$$f_\ell^\dagger = d_\ell^\dagger - g_\ell^\dagger + \sum_{i \in \mathcal{D}_\ell} (d_i^\dagger - g_i^\dagger), \ \forall \ell \in \mathcal{L}, \tag{1c}$$

$$u_i = u_0 - 2 \sum_{\ell \in \mathcal{R}_i} (f_\ell^p r_\ell + f_\ell^q x_\ell), \ \forall i \in \mathcal{L}, \tag{1d}$$

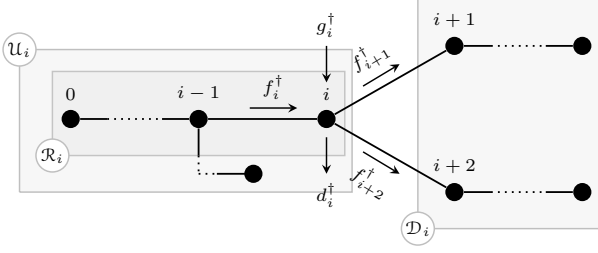$$(f_\ell^p)^2 + (f_\ell^q)^2 \leqslant \overline{f}_\ell^2, \ \forall \ell \in \mathcal{L}, \tag{1e}$$

Fig. 1. Topology of the distribution grid relative to node i, $\dagger = \{p, q\}$.

$$\underline{g}_i^\dagger \leqslant g_i^\dagger \leqslant \overline{g}_i^\dagger, \ \forall i \in \mathcal{N}, \tag{1f}$$

$$\underline{v}_i^2 \leqslant u_i \leqslant \overline{v}_i^2, \ \forall i \in \mathcal{N} \setminus \{0\}, \tag{1g}$$

where superscript $\dagger = \{p, q\}$ indexes active and reactive power. The objective is to minimize the total operational cost subject to the OPF equations (1b)–(1d), that balance the grid based on the *LinDistFlow* AC power flow equations [19, equations (9)] for distribution grids, and grid limits (1e)–(1g). Equation (1b) requires the total mismatch between power generation and loads in the distribution grid to be compensated for by the power from the substation at the root node. Equation (1c) requires the balance between the power flow along every edge ℓ, power mismatch at the in-flow node ℓ as well as that at the downstream nodes. The last term in (1c) can be also rewritten as the sum of power flows in the adjacent downstream lines, but kept as it is in the interest of the subsequent derivations. Last, equation (1d) models the voltage drop along the path from the root node to the node of interest.

Although OPF equations (1b)-(1d) establish the affine relation between the OPF variables, which is necessary for the subsequent chance-constrained formulation, they neglect distribution grid losses. The losses, however, can be included in an affine manner using various linearization techniques, such as in [20], [21] and [22] to mention but a few examples.

B. Differential Privacy

The paper uses the framework of *differential privacy* [7] to quantify and control the privacy risks of the customer loads. It considers datasets $D \in \mathbb{R}^n$ as n-dimensional vectors describing the *active* load values, denoted by d_i for each node i. To protect the participation of the load in the ith entry of the dataset, the following *adjacency relation* is introduced:

$$D \sim_\beta D' \Leftrightarrow \exists i \text{ s.t. } |d_i - d_i'| \leqslant \beta_i \text{ and } d_j = d_j', \forall j \neq i,$$

where D and D' are two adjacent datasets, $\beta \in \mathbb{R}^n$ is a vector of positive real values, and values d_i and d_i' are the load values corresponding to customer i in D and D', respectively. The adjacency relates two load vectors that differ in at most one item, at position i, by a value not greater than β_i.

If a mechanism satisfies the definition of differential privacy, it returns similar results on adjacent datasets in a probabilistic sense. This intuition is formalized in the following definition.

Definition 2 (Differential Privacy). *Given a value $\beta \in \mathbb{R}^n_+$, a randomized mechanism $\tilde{\mathcal{M}} : \mathcal{D} \to \mathcal{R}$ with domain $\mathcal{D}$ and range $\mathcal{R}$ is (ε, δ)-differential private if, for any output $s \subseteq \mathcal{R}$ and any two adjacent inputs $D \sim_\beta D' \in \mathbb{R}^n$*

$$\mathbb{P}[\tilde{\mathcal{M}}(D) \in s] \leqslant e^\varepsilon \mathbb{P}[\tilde{\mathcal{M}}(D') \in s] + \delta,$$

where $\mathbb{P}$ denotes the probability over runs of $\tilde{\mathcal{M}}$.

In the context of OPF problem (1), domain $\mathcal{D}$ includes all feasible load datasets, mechanism $\mathcal{M}$ denotes the OPF problem itself, and $\tilde{\mathcal{M}}$ is its randomized counterpart, and range $\mathcal{R}$ denotes the feasible region of the OPF problem.

The level of privacy is controlled by DP parameters (ε, δ). The former corresponds to the maximal multiplicative difference in distributions obtained by the mechanism on adjacent datasets, whereas the latter defines the maximal additive difference. Consequently, smaller values of ε and δ provide stronger privacy protection. Definition 2 extends the metric-based differential privacy introduced by Chatzikokolakis *et al.* [23] to control of *individual* privacy risks.

If a mechanism satisfies Definition 2, it features two important properties. First, by acting on adjacent datasets D and D', it provides privacy for each item i *irrespective* of the properties of all remaining items in a dataset. Second, it is immune to the so-called side attacks, i.e., it ensures that even if an attacker acquires the data of all other users but i, when accessing the output $\tilde{\mathcal{M}}(D)$ of the differential private mechanism, it will not be able to infer the load value of user i up to differential privacy bounds ε and δ [24].

The differentially private design of any mechanism is obtained by means of randomization using, among others, Laplace or Gaussian noise for numerical queries and exponential noise for the so-called non-numerical events [24]. The DP requirements for an optimization problem are achieved by introducing a calibrated noise to the input data [9] or to the output or objective function of the mechanism itself [8]. Regardless of the strategy adopted to attain DP, the amount of noise to inject depends on the mechanism *sensitivity*. In particular, the L_2−sensitivity of a deterministic mechanism $\mathcal{M}$ on β-adjacent datasets, denoted by Δ^β, is defined as:

$$\Delta^\beta = \max_{D \sim_\beta D'} \|\mathcal{M}(D) - \mathcal{M}(D')\|_2.$$

This work employs the Gaussian mechanism, because the Gaussian noise allows for the exact analytic reformulation of chance constraints into tractable second-order cone constraints.

Theorem 1 (Gaussian mechanism [24]). *Let $\mathcal{M}$ be a mechanism of interest that maps datasets D to $\mathbb{R}^n$, and let Δ^β be its L_2−sensitivity. For $\varepsilon \in (0, 1)$ and $\gamma^2 > 2\ln(\frac{1.25}{\delta})$, the Gaussian mechanism that outputs $\tilde{\mathcal{M}}(D) = \mathcal{M}(D) + \xi$, with ξ noise drawn from the Normal distribution with 0 mean and standard deviation $\sigma \geqslant \gamma \frac{\Delta^\beta}{\varepsilon}$ is (ε, δ)-differentially private.*

When the DP mechanism produces solutions to an optimization problem, it is also important to quantify the *optimality loss*, i.e., the distance between the optimal solutions of the original mechanism $\mathcal{M}(D)$ and its differentially private counterpart $\tilde{\mathcal{M}}(D)$ evaluated on the original dataset D.
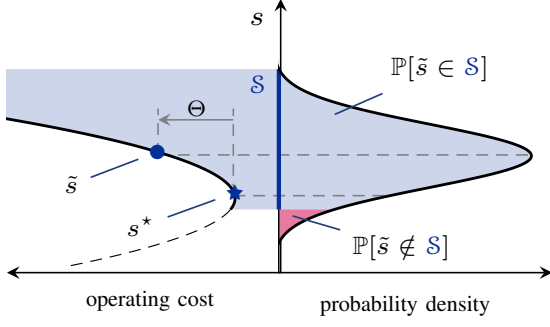
Fig. 2. Projections of OPF solutions onto operating cost and feasibility space.

III. PROBLEM STATEMENT

In the context of the underlying dispatch problem, the DSO collects a dataset $D = \{d_i^p\}_{i \in \mathcal{N}}$ of customer *sensitive* loads and dispatches the DER according to the solution of the OPF model (1). The OPF model acts as a mechanism $\mathcal{M} : D \mapsto s$ that maps the dataset D into an optimal OPF solution $s^\star$. The solution is a tuple comprising generator set points $\{g_i^p, g_i^q\}_{i \in \mathcal{N}}$, power flows $\{f_\ell^p, f_\ell^q\}_{\ell \in \mathcal{L}}$, and voltages $\{u_i\}_{i \in \mathcal{N}}$, as depicted on the left plane in Fig. 2. However, the release of $s^\star$ poses a privacy threat: an adversary with access to the items in $s^\star$ could *decode* the customers activities [4], [5]. For instance, the voltage sags at a node of interest discloses the activity of residential costumers (e.g., charging an electrical vehicle). Voltages and flows (currents) also encode information about the technology, production patterns, and other commercial data of industrial customers [10].

To minimize privacy risks, this work proposes a mechanism $\tilde{\mathcal{M}}$ for the DSO, which returns a feasible solution $\tilde{s}$ at the expense of an optimality loss, as shown in Fig. 2. A non-trivial benefit of choosing $\tilde{s}$ over $s^\star$ is that the former includes a particular perturbation of the optimal solution and thus carries less information about the real load data D. For instance, the sub-optimal solution can feature a more intensive deployment of expensive DERs instead of purchasing less expensive power from the high-voltage grid. To ensure that $\tilde{\mathcal{M}}$ returns a differentially private solution, $\tilde{s}$ has to follow a carefully calibrated noise distribution, as depicted on the right plane in Fig. 2. In other words, the mechanism must satisfy Definition 2, i.e., on adjacent load datasets, it must output distributions that differ by at most ε and δ in multiplicative and additive terms, respectively. However, with smaller ε and δ, the variance of the OPF solutions and hence the probability of producing an infeasible solution increases. The mechanism thus needs to address this feasibility issue. Finally, the last desired property is the ability to control the induced optimality loss Θ in order to ensure cost-effective grid operations.

To provide differentially private OPF solutions, the work focuses on the randomization of active power flows as their sensitivities to grid loads can be directly upper-bounded by the load magnitudes in radial grids. Since the OPF equations (1b)–(1d) couple OPF variables, the randomization of power flows will also induce the randomization of reactive power flows and voltages. Therefore, the randomized OPF mechanism $\tilde{\mathcal{M}}$ is now seen as a mapping from a dataset D to the active power

flow solution f^p. Let $F^p \in \mathbb{R}^{|\mathcal{L}|}$ be a particular realization of the randomized active power flows. The privacy goal of this work is to ensure that $\tilde{\mathcal{M}}$ satisfies

$$D \sim_\beta D' : \mathbb{P}[\tilde{\mathcal{M}}(D) \in F^p] \leqslant e^\varepsilon \mathbb{P}[\tilde{\mathcal{M}}(D') \in F^p] + \delta, \text{ i.e.,}$$

the definition of (ε, δ)−DP on β−adjacent load datasets.

IV. DIFFERENTIALLY PRIVATE OPF MECHANISM

This section provides a mathematical description of mechanism $\tilde{\mathcal{M}}$ and details its application. Section IV-A describes the perturbation of generator outputs to attain the randomization of power flows, Section IV-B details the chance-constrained program that accommodates the perturbation in a feasible manner, and Section IV-C explains the mechanism application as well as its feasibility and privacy guarantees.

A. Random Perturbation of OPF Solutions

Consider a random perturbation $\xi \in \mathbb{R}^{|\mathcal{L}|}$ which obeys a Gaussian distribution $\mathcal{N}(0, \Sigma)$ with covariance matrix

$$\Sigma = \text{diag}([\sigma_1^2, \ldots, \sigma_{|\mathcal{L}|}^2]) = \text{diag}(\sigma^2) \in \mathbb{R}^{|\mathcal{L}| \times |\mathcal{L}|}.$$

Throughout the paper, Σ, σ^2, and σ are used interchangeably to discuss the perturbation parameters. The power flows are conditioned on perturbation ξ when the following affine policies are imposed on DERs and substation supplies:

$$\tilde{g}_i^p(\xi) = g_i^p + \sum_{\ell \in \mathcal{D}_i} \alpha_{i\ell}\xi_\ell - \sum_{\ell \in \mathcal{U}_i} \alpha_{i\ell}\xi_\ell, \ \forall i \in \mathcal{N}, \quad (2a)$$

where $\tilde{g}_i^p(\xi)$ and g_i^p are, respectively, the random and nominal (mean) active power outputs, and $\alpha_{i\ell}$ is the portion of random perturbation ξ_ℓ provided by the supplier at node i, modeled as a free variable. The policies in (2a) are viable when the following balancing conditions are enforced:

$$\sum_{i \in \mathcal{U}_\ell} \alpha_{i\ell} = 1, \ \sum_{i \in \mathcal{D}_\ell} \alpha_{i\ell} = 1, \ \forall \ell \in \mathcal{L}, \quad (2b)$$

such that for each line ℓ, the upstream suppliers adjust their aggregated output by ξ_ℓ and the downstream DERs counter-balance this perturbation by ξ_ℓ, thus satisfying power balance.

The policies in (2) differ from those in stochastic dispatch models in [2], [3], [14], [25], [26], where the overall generator recourse compensates for the mismatch between grid loads and renewable forecast error realizations. However, since the affine nature of generator response remains similar, the proposed policy directly extends to balance renewable forecast errors.

To provide a succinct representation of the randomized OPF variables, consider a topology matrix $T \in \mathbb{R}^{|\mathcal{N}| \times |\mathcal{L}|}$ whose elements are such that:

$$T_{i\ell} = \begin{cases} 1, & \text{if line } \ell \text{ is downstream w.r.t. node } i, \\ -1, & \text{if line } \ell \text{ is upstream w.r.t. node } i \\ 0, & \text{otherwise.} \end{cases}$$

Consider also an auxiliary row vector $\rho_i^p = T_i \circ \alpha_i$ that returns a Schur product of ith row of T and i^{th} row of α, and set $\rho_i^q = \rho_i^p \tan \phi_i$. If the grid DERs allow, the later can be relaxed to model variable DER power factors. Using this notation, the

perturbed OPF solution is modeled as the following set of random variables:

$$\tilde{g}_i^\dagger(\xi) = g_i^\dagger + \rho_i^\dagger \xi, \ \forall i \in \mathcal{N}, \tag{3a}$$

$$\tilde{f}_\ell^\dagger(\xi) = f_\ell^\dagger - \Big[\rho_\ell^\dagger + \sum_{j \in \mathcal{D}_\ell} \rho_j^\dagger\Big]\xi, \ \forall \ell \in \mathcal{L}, \tag{3b}$$

$$\tilde{u}_i(\xi) = u_i + 2 \sum_{j \in \mathcal{R}_i} \Big[r_j\big(\rho_j^p + \sum_{k \in \mathcal{D}_j} \rho_k^p\big) +$$
$$x_j\big(\rho_j^q + \sum_{k \in \mathcal{D}_j} \rho_k^q\big)\Big]\xi, \ \forall i \in \mathcal{L}, \tag{3c}$$

where the randomized power flows $\tilde{f}_i^\dagger$ are obtained by substituting generator policy (2a) into (1c), and randomized voltage magnitudes $\tilde{u}_i$ are expressed by substituting $\tilde{f}_i^\dagger$ into (1d), refer to Appendix A for details. Each variable is thus represented by its nominal component and its random component whose realization depends on ξ. Furthermore, the random components of power flows in (3b) and voltages (3c) also depend on the generator dispatch decisions $\rho^\dagger$. Therefore, by properly calibrating the parameters of ξ and finding the optimal dispatch decisions, the randomized OPF solution in (3) provides the required privacy guarantees (see Section IV-C, Theorem 2). However, there is yet no guarantee that the randomized OPF solution in (3) is feasible.

B. The Chance-Constrained Optimization Program

To obtain a feasible dispatch, the proposed mechanism uses a chance-constrained program which optimizes the affine functions in (3) to make OPF solution feasible for any realization of random variable ξ with a high probability. The chance-constrained program is obtained by substituting variables (3) into the base OPF model (1) and enforcing chance constraints on the grid limits. Its tractable formulation is provided in (5), which is obtained considering the following reformulations.

1) Objective Function Reformulation: The chance-constrained program minimizes the *expected* cost, which is reformulated as follows:

$$\mathbb{E}_\xi\Big[\sum_{i \in \mathcal{N}} c_i \tilde{g}_i^p\Big] = \mathbb{E}_\xi\Big[\sum_{i \in \mathcal{N}} c_i\big(g_i^\dagger + \rho_i^\dagger \xi\big)\Big] = \sum_{i \in \mathcal{N}} c_i g_i^p, \ \forall i \in \mathcal{N},$$

due to the zero-mean distribution of ξ.

2) Inner Polygon Approximation of the Quadratic Power Flow Constraints: The substitution of the random power flow variables in (3b) into the apparent power flow limit constraints (1e) results in the following expression

$$(\tilde{f}_\ell^p)^2 + (\tilde{f}_\ell^p)^2 \le \bar{f}_\ell^2, \ \forall \ell \in \mathcal{L},$$

which exhibits a quadratic dependency on random variable ξ, for which no tractable chance-constrained reformulation is known. To resolve this issue, the above quadratic constraint is replaced by the inner polygon [22], [27], which writes as

$$\gamma_c^p \tilde{f}_i^p + \gamma_c^q \tilde{f}_i^q + \gamma_c^s \bar{f}_i \le 0, \ \forall i \in \mathcal{L}, \forall c \in \mathcal{C}, \tag{4}$$

where $\gamma_c^p, \gamma_c^q, \gamma_c^s$ are the coefficients for each side c of the polygon. The cardinality $|\mathcal{C}|$ is arbitrary, but a higher cardinality brings a better accuracy. Equation (4) is not a relaxation but an inner convex approximation: if the power flow solution is feasible for (4), it is also feasible for the original quadratic constraint (1e).

Algorithm 1: DP CC-OPF mechanism $\tilde{\mathcal{M}}$

1 **Input:** $D, \varepsilon, \delta, \beta, \eta^g, \eta^u, \eta^f$
2 Define covariance $\Sigma = f(\varepsilon, \delta, \beta)$ as per Theorem 2
3 Solve $\overset{\star}{\mathcal{V}} \leftarrow$ argmin problem (5) using D and Σ
4 Sample random perturbation $\hat{\xi} \sim \mathcal{N}(0, \Sigma)$
5 Obtain final OPF solution from (3) using $\overset{\star}{\mathcal{V}}$ and $\hat{\xi}$
6 **Release:** $\tilde{f}^\dagger(\hat{\xi}), \tilde{u}(\hat{\xi}), \tilde{g}^\dagger(\hat{\xi})$

---

3) Conic Reformulation of Linear Chance Constraints: For the normally distributed variable ξ with known moments, the chance constraint of the form $\mathbb{P}_\xi[\xi^\top x \le b] \ge 1 - \eta$ is translated into a second-order cone constraint as [28, Chapter 4.2.2]:

$$z_\eta \|\text{std}(\xi^\top x)\|_2 \le b - \mathbb{E}_\xi[\xi^\top x],$$

where $z_\eta = \Phi^{-1}(1 - \eta)$ is the inverse cumulative distribution function of the standard Gaussian distribution at the $(1 - \eta)$ quantile, and η is the constraint violation probability. Therefore, the individual chance constraints on the generation, voltage, and power flow variables are formulated in a conic form in (5c)–(5g), respectively. The resulting tractable formulation of the chance-constrained OPF program is as follows:

$$\text{CC-OPF:} \quad \min_{\mathcal{V} = \{g^\dagger, f^\dagger, u, \rho^\dagger\}} \sum_{i \in \mathcal{N}} c_i g_i^p \tag{5a}$$

$$\text{s.t.} \quad \text{Equations (1b)} - \text{(1d)}, \text{(2b)}, \tag{5b}$$

$$z_{\eta^g} \big\|\rho_i^\dagger \sigma\big\|_2 \le \overline{g}_i^\dagger - g_i^\dagger, \ \forall i \in \mathcal{N}, \tag{5c}$$

$$z_{\eta^g} \big\|\rho_i^\dagger \sigma\big\|_2 \le g_i^\dagger - \underline{g}_i^\dagger, \ \forall i \in \mathcal{N}, \tag{5d}$$

$$z_{\eta^u} \Big\|\Big[\sum_{j \in \mathcal{R}_i} \big[r_j\big(\rho_j^p + \sum_{k \in \mathcal{D}_j} \rho_k^p\big) + x_j\big(\rho_j^q + \sum_{k \in \mathcal{D}_j} \rho_k^q\big)\big]\Big]\sigma\Big\|_2$$
$$\le \tfrac{1}{2}\big(\overline{u}_i - u_i\big), \ \forall i \in \mathcal{L}, \tag{5e}$$

$$z_{\eta^u} \Big\|\Big[\sum_{j \in \mathcal{R}_i} \big[r_j\big(\rho_j^p + \sum_{k \in \mathcal{D}_j} \rho_k^p\big) + x_j\big(\rho_j^q + \sum_{k \in \mathcal{D}_j} \rho_k^q\big)\big]\Big]\sigma\Big\|_2$$
$$\le \tfrac{1}{2}\big(u_i - \underline{u}_i\big), \ \forall i \in \mathcal{L}, \tag{5f}$$

$$z_{\eta^f} \Big\|\big(\gamma_c^p\big[\rho_\ell^p + \sum_{i \in \mathcal{D}_\ell} \rho_i^p\big] + \gamma_c^q\big[\rho_\ell^q + \sum_{i \in \mathcal{D}_\ell} \rho_i^q\big]\big)\sigma\Big\|_2 \le$$
$$- \gamma_c^p f_\ell^p - \gamma_c^q f_\ell^q - \gamma_c^s \overline{f}_\ell, \ \forall \ell \in \mathcal{L}, \forall c \in \mathcal{C}. \tag{5g}$$

C. The Privacy-Preserving Mechanism and Guarantees

The functioning of the privacy-preserving mechanism $\tilde{\mathcal{M}}$ is explained in Algorithm 1. The Algorithm first computes the covariance matrix Σ that encodes the DP parameters (ε, δ), and adjacency parameter β. The mechanism then solves the optimization problem in (5) to obtain an optimal chance-constrained solution $\overset{\star}{\mathcal{V}}$. Last, the mechanism samples the random perturbation and obtains the final OPF solution using equations (3). By design of problem (5), the sampled OPF solution is guaranteed to satisfy grid limits and customer loads up to specified violation probabilities η^g, η^u and η^f, of the generator, voltage, and power flow constraints.

The privacy guarantees, in turn, depend on the specification of DP parameters (ε, δ) and the vector of adjacency coefficients β. For simplicity, the DP parameters are assumed to be uniform for all customers and specified by the DSO,

whereas customer privacy preferences are expressed in the submitted adjacency coefficients. In this setting, the load of every customer i is guaranteed to be indistinguishable from any other load in the range $[d_i^p - \beta_i, d_i^p + \beta_i]$ in the release of OPF solution related to node i up to DP parameters (ε, δ). This guarantee is formalized by the following result.

Theorem 2 (Privacy Guarantees). *Let $(\varepsilon, \delta) \in (0,1)$ and $\sigma_i \geqslant \beta_i \sqrt{2ln(1.25/\delta)}/\varepsilon, \; \forall i \in \mathcal{L}$. Then, if problem (5) returns an optimal solution, mechanism $\tilde{\mathcal{M}}$ is (ε, δ)-differentially private for β-adjacent load datasets. That is, the probabilities of returning a power flow solution in set F^p on any two β-adjacent datasets D and D' are such that*

$$\mathbb{P}[\tilde{\mathcal{M}}(D) \in F^p] \leqslant e^\varepsilon \mathbb{P}[\tilde{\mathcal{M}}(D') \in F^p] + \delta,$$

where $\mathbb{P}$ denotes the probability over runs of $\tilde{\mathcal{M}}$.

Proof. The full proof is available in Appendix B and relies on two intermediate results summarized in Lemmas 1 and 2. The first lemma shows that the standard deviation of power flow related to customer i is at least as much as σ_i. Therefore, by specifying σ_i, the DSO attains the desired degree of randomization. The second lemma shows that $\beta_i \geqslant \Delta_i^\beta$, i.e., if σ_i is parameterized by β_i, then σ_i is also parameterized by sensitivity Δ_i^β, required by the Gaussian mechanism in Theorem 1. □

V. MECHANISM EXTENSIONS

A. OPF Variance Control

Due to the radial topology of distribution grids, the flow perturbations along the same radial branch induce larger flow variances than those intended by the covariance matrix Σ. This section extends the mechanism $\tilde{\mathcal{M}}$ to reduce the overall flow variance while still preserving privacy guarantees. Two strategies are proposed to achieve this goal.

1) Total Variance Minimization: The flow standard deviation, obtained from (3b), depends on the DER participation variables $\rho^\dagger$. Therefore, the variance of power flows can be controlled by optimizing the DER dispatch. This variance control strategy is enabled by replacing problem (5) at the core of mechanism $\tilde{\mathcal{M}}$ by the following optimization:

$$\text{ToV-CC-OPF:} \quad \min_{\mathcal{V} \cup \{t\}} \sum_{i \in \mathcal{N}} c_i g_i^p + \sum_{\ell \in \mathcal{L}} \psi_\ell t_\ell \tag{6a}$$

$$\text{s.t.} \quad \left\| \left[\rho_\ell^p + \sum_{i \in \mathcal{D}_\ell} \rho_i^p\right]\sigma \right\|_2 \leqslant t_\ell, \; \forall \ell \in \mathcal{L}, \tag{6b}$$

$$\text{Equations (5b)} - \text{(5g)}, \tag{6c}$$

where the decision variable t_ℓ represents the standard deviation of the active power flow in line ℓ, which is penalized in the objective function by a non-negative parameter ψ_ℓ. By choosing $\psi_\ell, \forall \ell \in \mathcal{L}$, the DSO minimizes the total variance at the expense of operational cost. Note that, by Lemma 1, optimization (6) does not violate the privacy guarantees.

2) Pursuing Target Variance: This strategy solely perturbs the flow in the selected line of the radial branch (e.g., adjacent to the customer with the strongest privacy requirement) and constrains the DERs to maintain the flow variance in each line as required by the original matrix Σ. It specifies a new

matrix $\hat{\Sigma} = \text{diag}([\hat{\sigma}_1^2, \dots, \hat{\sigma}_{|\mathcal{L}|}^2]), \; \mathbb{1}^\top \hat{\sigma}^2 \leqslant \mathbb{1}^\top \sigma^2$, that contains a smaller number of perturbations. This control is enabled by replacing problem (5) by the following optimization:

$$\text{TaV-CC-OPF:} \quad \min_{\mathcal{V} \cup \{t, \tau\}} \sum_{i \in \mathcal{N}} c_i g_i^p + \sum_{\ell \in \mathcal{L}} \psi_\ell \tau_\ell \tag{7a}$$

$$\text{s.t.} \quad \left\| \left[\rho_\ell^p + \sum_{i \in \mathcal{D}_\ell} \rho_i^p\right]\hat{\sigma} \right\|_2 \leqslant t_\ell, \; \forall \ell \in \mathcal{L}, \tag{7b}$$

$$\left\| t_\ell - \sigma_\ell \right\|_2 \leqslant \tau_\ell, \; \forall \ell \in \mathcal{L}, \tag{7c}$$

$$\text{Equations (5b)} - \text{(5g) with } \hat{\sigma}, \tag{7d}$$

where, t_ℓ returns the resulting flow standard deviation, while constraint (7c) yields the distance τ_ℓ between the resulting standard deviation and original one $\sigma_\ell = \Sigma_{\ell,\ell}^{1/2}$ required to provide customer at node ℓ with differential privacy. By penalizing this distance in the objective function, the DSO attains privacy at a smaller amount of random perturbations. Note, as optimization (7) acts on covariance matrix $\hat{\Sigma}$ instead of Σ, the DSO needs to verify a posteriori that $t_\ell \geqslant \sigma_\ell, \; \forall \ell \in \mathcal{L}$.

B. Optimality Loss Control

The application of mechanism $\tilde{\mathcal{M}}$ necessarily leads to an optimality loss compared to the solution of non-private mechanism $\mathcal{M}$. This section slightly abuses the notation and denotes the cost of the non-private OPF solution and that of the proposed DP mechanism when evaluated on a dataset D by $\mathcal{M}(D)$ and $\tilde{\mathcal{M}}(D)$, respectively. The optimality loss Θ is measured in expectation as the L_2 distance, i.e.,

$$\mathbb{E}[\Theta] = \|\mathcal{M}(D) - \mathbb{E}[\tilde{\mathcal{M}}(D)]\|_2,$$

as $\mathcal{M}(D)$ always provides a deterministic solution. However, the worst-case realization of $\tilde{\mathcal{M}}(D)$ may significantly exceed the expected value and lead to a larger optimality loss. To this end, this section introduces the optimality loss control strategy using the Conditional Value-at-Risk (CVaR) measure [13].

Consider $\varrho\%$ of the worst-case realizations of the optimally loss. The expected value of these worst-case realizations can be modeled as a decision variable using the CVaR measure as

$$\text{CVaR}_\varrho = \mu_c + \sigma_c \phi \left(\Phi^{-1}(1 - \varrho) \right) / \varrho, \tag{8}$$

where μ_c and σ_c represent the expected value and standard deviation of operational costs, while ϕ and $\Phi^{-1}(1 - \varrho)$ denote the probability density function and the inverse cumulative distribution function at the $(1 - \varrho)$ quantile of the standard Normal distribution. From Section (IV-B), it follows that

$$\mu_c = \mathbb{E}[c^\top \tilde{g}^p] = \mathbb{E}[c^\top (g^p + \rho\xi)] = c^\top g^p,$$

for zero-mean ξ, and the standard deviation finds as

$$\sigma_c = \text{std}[c^\top (g^p + \rho\xi)] = \text{std}[c^\top (\rho\xi)] = \|c^\top (\rho\sigma)\|_2,$$

providing a convex reformulation of the CVaR in (8). Therefore, for some trade-off parameter $\theta \in [0,1]$, the DSO can trade off the mean and CVaR_ϱ of the optimality loss by substituting problem (5) at the core of mechanism $\tilde{\mathcal{M}}$ by the following optimization CVaR-CC-OPF:

$$\min_{\mathcal{V} \cup \sigma_c} (1 - \theta)c^\top g^p + \theta \left[c^\top g^p + \sigma_c \phi \left(\Phi^{-1}(1 - \varrho) \right) / \varrho \right] \tag{9a}$$

s.t. $\quad \|c^\top(\rho\sigma)\|_2 \leqslant \sigma_c,$ (9b)

$\quad$ Equations (5b) $-$ (5g), (9c)

where the standard deviation σ_c is modeled as a decision variable. Notice, the optimality loss control by means of (9) does not violate the privacy guarantees as per Lemma 1.

VI. NUMERICAL EXPERIMENTS

The experiments consider a modified 15-node radial grid from [29], which includes network parameters taken from [30], nodal loads as given in Table I, and nodal DERs with cost coefficients drawn from Normal distribution $c_i \sim \mathcal{N}(10, 2)$ \$/MWh, generation limits $\underline{g}_i^p = 0$ MW, $\overline{g}_i^p = 8$ MW, and power factor $\tan\phi_i = 0.5$, $\forall i \in \mathcal{N}$. The constraint violation probabilities are set as $\eta^g = 1\%, \eta^u = 2\%$ and $\eta^f = 10\%$. The DP parameters are set to $\varepsilon \to 1, \delta = 1/n = 0.071$ with n being a number of grid customers, while the adjacency parameters $\beta_i, \forall i \in \mathcal{L}$, vary across the experiments. All models are implemented in Julia using the JuMP package [31], and all data and codes are relegated to the e-companion [32].

A. Illustrative Example

The purpose of the illustrative example is to simulate and obfuscate periodic components of the load profile in power flow and voltage measurements. Assume that the customer at node 7 has an atypical load pattern representing its production technology. Her pattern is obtained by multiplying the maximum load by $k(t)$, a multiplier with the following three periodic components:

$$k(t) = \max\left\{\sin\tfrac{5}{10^2}t, \tfrac{7}{10}\right\} + \tfrac{5}{10^2}\sin\tfrac{5}{10^2}t + \tfrac{25}{10^3}\sin\tfrac{75}{10^2}t$$

where t is the time step. The parameters of multiplier $k(t)$ are selected such that the load components have different magnitudes and frequencies. The non-private OPF solution provided by the D-OPF model leaks the information about this pattern through the power flow f_7^p and voltage v_7 readings, as displayed on the left plots in Fig. 3. To obfuscate the load pattern in the OPF solution, the customer submits the privacy preference β_7, which is accommodated by the DSO using mechanism $\tilde{\mathcal{M}}$. Figure 3 shows that by setting $\beta_7 \to 0.07$ MW, the presence of the smallest periodic component is obfuscated through randomization, while the presence of the two remaining components is still distinguished. With an increasing β_7, the mechanism further obfuscates the medium and largest periodic components.

B. Privacy Guarantees

To illustrate the privacy guarantees of Theorem 2, consider the same grid customer at node 7 with the load of 2.35 MW. For β_7, consider two adjacent load datasets D' and D'', containing $d_7'^p = d_7^p - \beta_7$ and $d_7''^p = d_7^p + \beta_7$, respectively. The non-private OPF mechanism returns the following power flows

$$\mathcal{M}(D') = 2.05\text{MW}, \ \mathcal{M}(D) = 2.35\text{MW}, \ \mathcal{M}(D'') = 2.65\text{MW},$$

for $\beta_7 = 0.3$ MW, clearly distinguishing the differences in datasets through power flow readings. The differentially

private mechanism $\tilde{\mathcal{M}}$ in Algorithm 1, in turn, obfuscates the load value used in the computation. Figure 4 shows that the mechanism makes the OPF solutions on the three datasets similar in the probabilistic sense, thus providing privacy guarantees for the original load dataset D. The maximal difference between the distributions of power flow solutions is bounded by the parameters ε and δ. Observe that the larger specification $\delta = 0.75$ results in weaker guarantees, as the distributions slightly stand out from one another. On the other hand, $\delta = 0.07$ yields a larger noise magnitude overlapping the support of the three distributions. The OPF solution to be implemented is obtained from a single sample drawn from the blue distribution. By observing a single sample, an adversary cannot distinguish the distribution, and thus the dataset, it was sampled from. Finally, Fig. 4 shows randomized OPF solutions obtained on a given load dataset. The parameters of the noise, however, are independent from load dataset (see Theorem 2), and the privacy guarantee for the customer at node 7 is independent from the loads and their variations at other grid nodes.

C. OPF Variance Control

Consider the application of the mechanism when all grid customers have their adjacency coefficients set to 10% of their loads. The non-private D-OPF and private OPF solutions, obtained with the variance-agnostic CC-OPF and variance-aware ToV- and TaV-CC-OPF models, are summarized in Table I: Each row i presents the power flow and voltage solutions related to customer i, and the bottom rows report the expected operational cost, optimality loss $\mathbb{E}[\Theta]$ in %, sum of power flow standard deviations, percentage $\hat{\eta}$ of infeasible instances on 5000 noise samples, and CPU times.

The table shows that, unlike non-private, deterministic D-OPF model, the DP mechanisms return OPF variables as probability densities with given means and standard deviations. For all three DP mechanisms, the flow standard deviations are at least as much as those required by Theorem 2, thus providing differential privacy. However, due to the noise applied to each network flow, the flow standard deviations provided by the CC-OPF model exceed the intended quantities, e.g., by 458% for the first customer close to the substation. To minimize the OPF variance, the ToV-CC-OPF and TaV-CC-OPF models are used with the uniform variance penalty factor $\psi_\ell = 10^5, \forall \ell \in \mathcal{L}$. The ToV-CC-OPF model also perturbs each flow in the network but it alters the optimal DER dispatch to reduce the sum of flow standard deviations by 50%. The TaV-CC-OPF model, in turn, introduces a limited number of perturbations to lines $\{1, 5-7, 9, 11-13\}$ and constrains the DERs to maintain the intended standard deviation σ across the entire network, reducing the flow standard deviation by 63%. As ToV-CC-OPF and TaV-CC-OPF prioritize the flow variance over expected cost, the models provide larger optimality loss than the CC-OPF model.

The OPF solution to be implemented by the DSO is a sample drawn from the probability densities reported in Table I. The empirical probability of the joint constraint violation $\hat{\eta}$ demonstrates an appropriate out-of-sample performance.
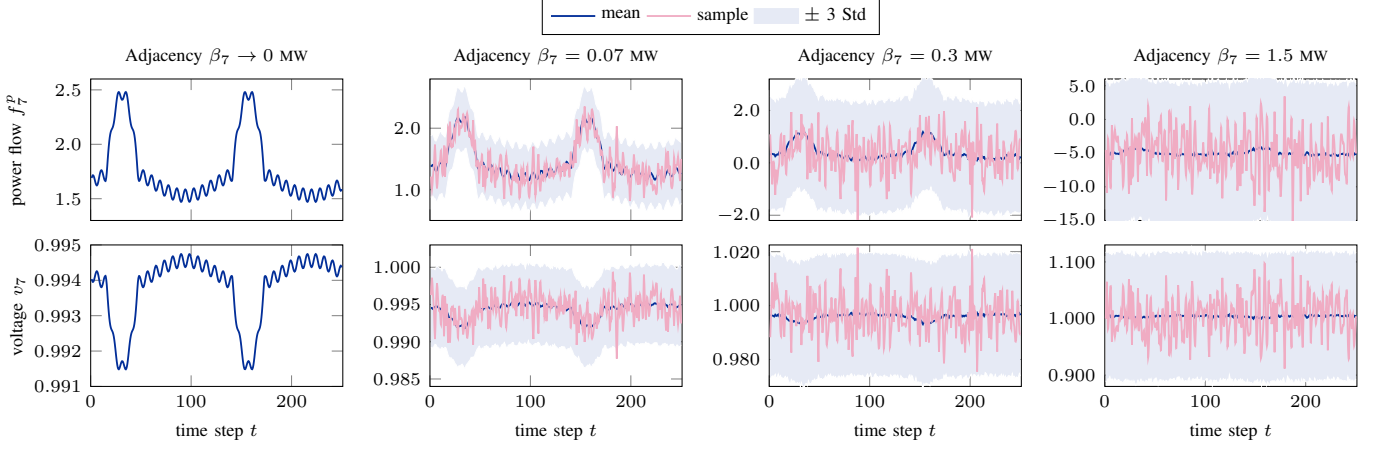
Fig. 3. Power flow and voltage magnitude at node 7 as functions of adjacency coefficient β_7. The flow and voltage solutions are given by their mean value (blue) and the range of ± 3 standard deviations (light blue). The OPF solution implemented by the DSO is given by sample trajectories (red).

TABLE I
SOLUTION SUMMARY FOR THE NON-PRIVATE AND DIFFERENTIALLY PRIVATE OPF MECHANISMS.

| i	d_i^p	σ_i	D-OPF, Eq. (1)		CC-OPF, Eq. (5)				ToV-CC-OPF, Eq. (6)				TaV-CC-OPF, Eq. (7)			
			f_i^p	v_i	f_i^p		v_i		f_i^p		v_i		f_i^p		v_i	
					mean	std	mean	std	mean	std	mean	std	mean	std	mean	std
0	0	–	–	1.00	–	–	1.00	–	–	–	1.00	–	–	–	1.00	–
1	2.01	**0.48**	8.5	1.00	11.3	**2.68**	1.00	0.0016	12.6	**0.69**	1.00	0.0004	13.0	**0.48**	1.00	0.0003
2	2.01	**0.48**	6.5	1.00	9.3	**2.68**	0.99	0.0057	11.4	**0.71**	0.99	0.0015	11.0	**0.48**	0.99	0.0010
3	2.01	**0.48**	4.4	1.00	7.3	**2.68**	0.99	0.0123	10.2	**0.78**	0.97	0.0033	9.0	**0.48**	0.98	0.0022
4	1.73	**0.41**	-8.0	1.00	-1.4	**1.72**	0.99	0.0128	3.6	**0.69**	0.97	0.0034	1.7	**0.41**	0.98	0.0023
5	2.91	**0.70**	5.1	1.00	3.1	**0.87**	0.99	0.0128	2.5	**0.82**	0.97	0.0035	1.9	**0.70**	0.98	0.0024
6	2.19	**0.52**	2.2	1.00	0.1	**0.87**	0.99	0.0128	0.7	**0.63**	0.97	0.0038	1.0	**0.52**	0.98	0.0024
7	2.35	**0.56**	2.3	0.99	0.9	**0.63**	0.98	0.0134	0.9	**0.61**	0.97	0.0039	1.0	**0.56**	0.98	0.0024
8	2.35	**0.56**	10.5	0.99	6.7	**1.18**	0.98	0.0130	5.8	**0.78**	0.97	0.0036	6.4	**0.56**	0.98	0.0023
9	2.29	**0.55**	5.8	0.99	3.5	**0.88**	0.98	0.0132	3.1	**0.70**	0.97	0.0037	3.6	**0.55**	0.98	0.0023
10	2.17	**0.52**	3.5	0.99	1.2	**0.88**	0.98	0.0135	1.6	**0.65**	0.97	0.0038	1.3	**0.52**	0.97	0.0023
11	1.32	**0.32**	1.3	0.99	0.4	**0.39**	0.98	0.0135	0.4	**0.40**	0.97	0.0038	0.6	**0.32**	0.97	0.0023
12	2.01	**0.48**	6.5	1.00	3.6	**1.23**	1.00	0.0008	3.3	**0.73**	1.00	0.0004	3.6	**0.48**	1.00	0.0003
13	2.24	**0.54**	4.5	0.99	1.6	**1.23**	1.00	0.0034	2.1	**0.72**	1.00	0.0019	3.2	**0.54**	0.99	0.0012
14	2.24	**0.54**	2.2	0.99	-0.6	**1.23**	1.00	0.0050	0.8	**0.64**	0.99	0.0027	1.0	**0.54**	0.99	0.0018
cost ($\mathbb{E}[\Theta]$)			$396.0 (0%)		$428.0 (8.1%)				$463.5 (17.1%)				$459.3 (16.0%)			
$\sum_i \text{std}[f_i^p]$			0 MW		19.1 MW				9.5 MW				7.1 MW			
infeas. $\hat{\eta}$			0%		3.3%				6.9%				5.5%			
CPU time			0.016s		0.037s				0.043s				0.052s			
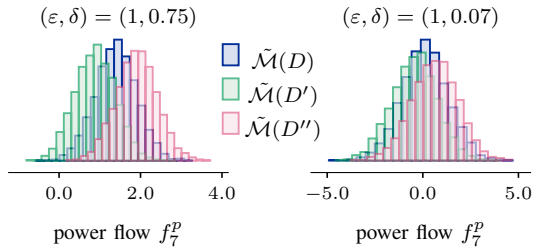
Fig. 4. The overlay of power flow probability densities obtained on the three β_7−adjacent load datasets for $\delta = 0.75$ and $\delta = 0.07$ (5000 samples).
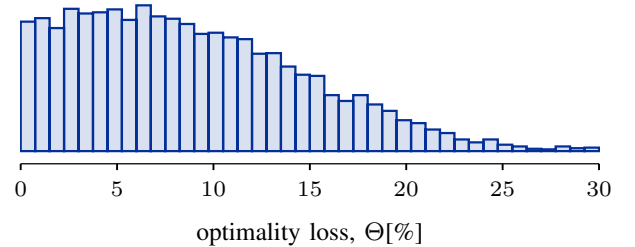

Fig. 5. CVaR-CC-OPF: Empirical out-of-sample density of the optimality loss for a trade-off parameter $\theta = 0$ (5000 samples)

However, if the DP mechanism returns an infeasible sample, the DSO may re-sample the OPF solution, yet it comes at the expense of the relaxation of privacy guarantees: every re-sampling round increases the privacy loss linearly, by ε, as per composition of DP [24, Theorem 3.14]. Finally, the privacy-preserving mechanisms keep the CPU time within the same time-frame as the standard, non-private D-OPF mechanism.

D. Optimality Loss Control

The application of mechanism $\tilde{\mathcal{M}}$ yields an optimality loss with respect to the solution of the standard, non-private OPF

TABLE II
TRADE-OFFS OF THE EXPECTED AND CVaR$_{10\%}$ PERFORMANCE

θ	exp. value		CVaR$_{10\%}$		$\sum_i \text{std}[f_i^p]$, MW
	cost, \$	Θ,%	cost, \$	Θ,%	
0.0	428.0	8.1	478.1	20.7	19.1
0.1	428.0	8.1	476.3	20.3	19.4
0.2	428.3	8.2	475.0	19.9	19.6
0.3	428.9	8.3	473.3	19.5	19.8
0.4	431.9	9.1	467.8	18.1	17.3
0.5	434.5	9.7	464.4	17.3	15.7
0.6	438.2	10.7	461.7	16.6	14.6
0.7	452.9	14.4	452.9	14.4	13.0

mechanism. For the same setting as in the previous experiment, the out-of-sample empirical distribution of the cost is depicted in Fig. VI-D. Observe that the probability mass is centered at 8.1% of optimality loss and that the distribution is biased towards smaller optimality losses. Although the probability of the worst-case outcomes is shown very small on out-of-sample, the expected value over 10% of the worst-case outcomes is at relatively large 20.7%. The DSO, however, is capable to trade off between the expected and the worst-case optimality loss by substituting the CC-OPF model in mechanism $\mathcal{M}$ by the CVaR-CC-OPF model in (9). For a trade-off parameter $\theta \in [0,1]$, the expected optimality loss in $\varrho = 10\%$ of the worst-case scenarios is contrasted with the expected loss in Table II. For $\theta = 0$, the CVaR$_{10\%}$ significantly exceeds the expected value. However, by increasing θ, the DSO alters the DER dispatch to reduce the worst-case optimality loss at the expense of increasing the expected value. For $\theta \geqslant 0.7$, the expected value corresponds to CVaR$_{10\%}$, thus providing differential privacy at a fixed cost. Eventually, the choice of θ is driven by the DSO's risk preference. Finally, the optimality loss can be further reduced by relaxing the feasibility guarantee with larger probabilities η, though it may result in an increasing out-of-sample violation probability $\hat{\eta}$.

E. Comparison with the Output Perturbation Mechanism

It remains to compare the proposed privacy-preserving OPF mechanism with the standard, non-adapted to the specifics of OPF problems, output perturbation (OP) mechanism [24]. The functioning of this mechanism is depicted in Algorithm 2: it solves the deterministic OPF problem, perturbs the optimal power flow solution with the random noise, and then finds the feasible DER and substation dispatch that satisfies AC-OPF equations for the perturbed values of power flows. If exists, the mechanism returns an $(\varepsilon, \delta)-$differentially private OPF solution on $\beta-$adjacent datasets, or reports infeasibility otherwise. Observe, unlike the proposed mechanism in Algorithm 1, the output perturbation mechanism does not offer feasibility guarantees, because the perturbation step is done independently from the OPF computations.

To compare the two mechanisms, consider the provision of differential privacy for sets of nodes $1 : n$, i.e., from 1 to n, for which $\beta_i \to 10\%$. By increasing n, the amount of noise that the DSO needs to accommodate in the grid increases. Table III summarizes the feasibility statistics for

---

Algorithm 2: Output Perturbation (OP) Mechanism

1 **Input:** $D, \varepsilon, \delta, \beta$
2 Solve $\{\overset{\star}{f_\ell^p}\}_{\forall \ell \in \mathcal{L}} \leftarrow$ argmin problem (1) using D
3 Sample $\hat{\xi}_\ell \sim \mathcal{N}(0, \beta_\ell \sqrt{2\ln(1.25/\delta)}/\varepsilon), \forall \ell \in \mathcal{L}$
4 Perturb power flows $\hat{f}_\ell^p = \overset{\star}{f_\ell^p} + \hat{\xi}_\ell, \forall \ell \in \mathcal{L}$
5 Re-solve problem (1) for the fixed $f_\ell^p = \hat{f}_\ell^p, \forall \ell \in \mathcal{L}$
6 **if** problem (1) feasible **then**
7 | Implement perturbed solution
8 **else**
9 | Report infeasibility
10 **end**

TABLE III
PERCENTAGE OF INFEASIBLE OPF INSTANCES (5000 SAMPLES) [%]

Mechanism	Node set $1 : n$ (with non-zero adjacency β_i)						
	1	1:2	1:3	1:4	1:5	...	1:14
OP	52.1	87.0	97.9	99.7	100	...	100
CC-OPF	0.1	0.9	0.9	1.0	1.1	...	3.3

the two algorithms. Observe, even for a single customer, the output perturbation mechanism produces infeasible solutions in 52.1% of instances, and its performance further reduces in n. By optimizing affine functions in (3), the proposed mechanism, instead, produces feasible OPF instances with a high probability, e.g. 96.7% for the entire set of customers.

VII. CONCLUSIONS AND FUTURE WORK

This paper introduced a differentially private OPF mechanism for distribution grids, which provides formal privacy guarantees for grid customer loads. The mechanism parametrizes OPF variables as affine functions of a carefully calibrated noise to weaken the correlations between grid loads and OPF variables, thus preventing the recovery of customer loads from the voltage and power flow measurements. Furthermore, the mechanism was extended to enable the DSO to control the OPF variance induced by the noise in the computations, providing better practices for systems with more emphasis on component overloads than on operational costs. Finally, the optimality loss induced by the mechanism translates into privacy costs. To minimize the risk of large privacy costs, the mechanism was extended to enable the trade-off between the expected and worst-case performances.

There are several avenues for future work. To understand the impacts of the privacy preservation on distribution electricity pricing, one can explore the connection between DP parameters and distribution locational marginal prices following price decomposition approach from [22] and [29]. Alternatively, the coalition game theory can be used to find an adequate privacy cost allocation among customers, similar to game-theoretic reserve cost allocation in [33]. Finally, the private OPF mechanism has been developed in a technically suitable ecosystem: it builds upon LinDistFlow OPF equations neglecting the effects of power losses, adopts a constant DER power factor, and does not include the control of stochastic DERs. Although these three factors are well-studied in the context of the chance-constrained OPF problems under uncertainty, it remains valid,

if not crucial, for future work to explore their effects on the limits of the differential privacy provision in distribution grids.

<div align="center">APPENDIX</div>

A. System Response to the Random Perturbation

The affine dependency of power flows on the random perturbation is obtain by substituting generator response policy (2a) into OPF equations (1c), that is:

$$
\begin{aligned}
\tilde{f}_\ell^\dagger(\xi) &= d_\ell^\dagger - g_\ell^\dagger(\xi) + \sum_{i\in\mathcal{D}_\ell}(d_i^\dagger - g_i^\dagger(\xi)) \\
&= d_\ell^\dagger - g_\ell^\dagger - \rho_\ell^\dagger\xi + \sum_{i\in\mathcal{D}_\ell}(d_i^\dagger - g_i^\dagger - \rho_i^\dagger\xi) \\
&= d_\ell^\dagger - g_\ell^\dagger + \sum_{i\in\mathcal{D}_\ell}(d_i^\dagger - g_i^\dagger) - [\rho_\ell^\dagger\xi + \sum_{i\in\mathcal{D}_\ell}\rho_i^\dagger\xi] \\
&\overset{\text{due to (1c)}}{=} f_\ell^\dagger - \Big[\rho_\ell^\dagger + \sum_{j\in\mathcal{D}_\ell}\rho_j^\dagger\Big]\xi, \quad \forall\ell\in\mathcal{L},
\end{aligned}
\tag{10a}
$$

where $f_\ell^\dagger$ is the nominal (average) component and the last term is the random flow component. The affine dependency of voltages on the random perturbation is obtain by substituting (10a) into voltage drop equation (1d), that is:

$$
\begin{aligned}
\tilde{u}_i(\xi) &= u_0 - 2\sum_{\ell\in\mathcal{R}_i}(\tilde{f}_\ell^p(\xi)r_\ell + \tilde{f}_\ell^q(\xi)x_\ell) \\
&= u_0 - 2\sum_{\ell\in\mathcal{R}_i}(f_\ell^p r_\ell + f_\ell^q x_\ell) \\
&\quad + 2\sum_{\ell\in\mathcal{R}_i}\Big[r_\ell\big(\rho_\ell^p + \sum_{k\in\mathcal{D}_\ell}\rho_k^p\big) + x_\ell\big(\rho_\ell^q + \sum_{k\in\mathcal{D}_\ell}\rho_k^q\big)\Big]\xi \\
&\overset{\text{due to (1d)}}{=} u_i + 2\sum_{\ell\in\mathcal{R}_i}\Big[r_\ell\big(\rho_\ell^p + \sum_{k\in\mathcal{D}_\ell}\rho_k^p\big) + x_\ell\big(\rho_\ell^q + \sum_{k\in\mathcal{D}_\ell}\rho_k^q\big)\Big]\xi,
\end{aligned}
\tag{10b}
$$

where u_i is the nominal (average) component and the last term is the random voltage component. Together with the response policy in (2a), equations (10) constitute the model of system response to the random perturbation, given in equations (3).

B. Proof of Theorem 2

The proof of Theorem 2 relies on Lemmas 1 and 2. The first lemma shows that the standard deviation of power flow related to customer i is at least as much as σ_i. Therefore, by specifying σ_i, the DSO attains the desired degree of randomization.

Lemma 1. *If OPF mechanism (5) returns optimal solution, then σ_ℓ is the lower bound on $std[\tilde{f}_\ell^p]$.*

Proof. Consider a single flow perturbation with $\xi_\ell \sim \mathcal{N}(0,\sigma_\ell^2)$ and $\xi_j = 0,\ \forall j\in\mathcal{L}\backslash\ell$. The standards deviation of active power flow (3b) in optimum finds as

$$
\begin{aligned}
&\mathrm{std}\Big[\overset{\star}{f}_\ell^p - \big[\overset{\star}{\rho}_\ell^p + \sum_{j\in\mathcal{D}_\ell}\overset{\star}{\rho}_j^p\big]\xi\Big] = \mathrm{std}\Big[\big[\overset{\star}{\rho}_\ell^p + \sum_{j\in\mathcal{D}_\ell}\overset{\star}{\rho}_j^p\big]\xi\Big] \\
&= \mathrm{std}\Big[\sum_{j\in\mathcal{D}_\ell}\overset{\star}{\alpha}_{j\ell}\xi_\ell\Big] \overset{(2b)}{=} \mathrm{std}\big[\xi_\ell\big] = \sigma_\ell,
\end{aligned}
\tag{11}
$$

where the second to the last equality follows from balancing conditions (2b). As for any pair $(\ell,j)\in\mathcal{L}$ the covariance

matrix returns $\Sigma_{\ell,j} = 0$, σ_ℓ is a lower bound on $std[\tilde{f}_\ell^p]$ in the optimum for any additional perturbation in the network. $\square$

Remark 1. *The result of Lemma 1 holds independently from the choice of objective function and is solely driven by the feasibility conditions.*

The second lemma shows that $\beta_i \geqslant \Delta_i^\beta$, i.e., if σ_i is parameterized by β_i, then σ_i is also parameterized by sensitivity Δ_i^β.

Lemma 2. *Let D and D' be two adjacent datasets differing in at most one load d_i^p by at most $\beta_i > 0$. Then,*

$$
\Delta_i^\beta = \max_{\ell\in\mathcal{L}}\|\mathcal{M}(D)_{|f_\ell^p} - \mathcal{M}(D')_{|f_\ell^p}\|_2 \leqslant \beta_i,
$$

where the notation $\mathcal{M}(\cdot)_{|f_\ell^p}$ denotes the value of the optimal active power flow on line ℓ returned by the computation $\mathcal{M}(\cdot)$.

Proof. Let $\overset{\star}{f}_\ell^p$ be the optimal solution for the active power flow in line ℓ obtained on input dataset $D = (d_1^p,\ldots,d_n^p)$. From OPF equation (1c), it can be written as

$$
\overset{\star}{f}_\ell^p = d_\ell^p - \overset{\star}{g}_\ell^p + \sum_{i\in\mathcal{D}_\ell}(d_i^p - \overset{\star}{g}_i^p),
$$

which expresses the flow as a function of the downstream loads and the optimal DER dispatch. A change in the active load d_ℓ^p translates into a change of power flow as

$$
\begin{aligned}
\frac{\partial \overset{\star}{f}_\ell^p}{\partial d_\ell^p} &= \underbrace{\frac{\partial d_\ell^p}{\partial d_\ell^p}}_{1} - \frac{\partial \overset{\star}{g}_\ell^p}{\partial d_\ell^p} + \sum_{i\in\mathcal{D}_\ell}\Big(\underbrace{\frac{\partial d_i^p}{\partial d_\ell^p}}_{0} - \frac{\partial \overset{\star}{g}_i^p}{\partial d_\ell^p}\Big) \\
&= 1 - \frac{\partial \overset{\star}{g}_\ell^p}{\partial d_\ell^p} - \sum_{i\in\mathcal{D}_\ell}\frac{\partial \overset{\star}{g}_i^p}{\partial d_\ell^p},
\end{aligned}
\tag{12}
$$

where the last two terms are always non-negative due to convexity of model (1). The value of (12) attains maximum when

$$
\overset{\star}{g}_k^p = \overline{g}_k^p \mapsto \frac{\partial \overset{\star}{g}_k^p}{\partial d_\ell^p} = 0, \quad \forall k\in\{\ell\}\cup\mathcal{D}_\ell.
\tag{13}
$$

Therefore, by combining (12) with (13) we obtain the maximal change of power flows as

$$
\frac{\partial \overset{\star}{f}_\ell^p}{\partial d_\ell^p} = 1.
$$

Since the dataset adjacency relation considers loads d_ℓ^p that differ by at most β_ℓ, it suffices to multiply the above by β_ℓ to attain the result. It finds similarly that for a β_i change of any load $i\in\mathcal{N}$, all network flows change by at most β_i. $\square$

Proof of Theorem 2. Consider a customer at non-root node i. Mechanism $\tilde{\mathcal{M}}$ induces a perturbation on the active power flow f_i^p by a random variable $\xi_i \sim \mathcal{N}(0,\sigma_i^2)$. The randomized active power flow f_i^p is then given as follows:

$$
\tilde{f}_i^p = \overset{\star}{f}_i^p - \big[\overset{\star}{\rho}_i^p + \sum_{j\in\mathcal{D}_i}\overset{\star}{\rho}_j^p\big]\xi,
$$

where $\star$ denotes optimal solution for optimization variables. For privacy parameters (ε, δ), the mechanism specifies

$$\sigma_i \geqslant \beta_i \sqrt{2\ln(1.25/\delta)}/\varepsilon, \; \forall i \in \mathcal{L}.$$

As per Lemma 1, we know that σ_i is the lower bound on the standard deviation of power flow f_ℓ^p. From Lemma 2 we also know that the sensitivity Δ_i^β of power flow in line i to load d_i^p is upper-bounded by β_i, so we have

$$\text{std}[\tilde{f}_i^p] \geqslant \sigma_i \geqslant \Delta_i^\beta \sqrt{2\ln(1.25/\delta)}/\varepsilon.$$

Since the randomized power flow follow is now given by a Normal distribution with the standard deviation $\text{std}[\tilde{f}_i^p]$ as above, by Theorem 1, mechanism $\tilde{\mathcal{M}}$ satisfies (ε, δ)-differential privacy for each grid customer up to adjacency parameter β. $\qquad\square$

REFERENCES

[1] E. Dall'Anese and A. Simonetto, "Optimal power flow pursuit," *IEEE Trans. Smart Grid*, vol. 9, no. 2, pp. 942–952, 2016.

[2] S. Bolognani, E. Arcari, and F. Dörfler, "A fast method for real-time chance-constrained decision with application to power systems," *IEEE Contr. Syst. Lett.*, vol. 1, no. 1, pp. 152–157, 2017.

[3] R. Mieth and Y. Dvorkin, "Data-driven distributionally robust optimal power flow for distribution systems," *IEEE Contr. Syst. Lett.*, vol. 2, no. 3, pp. 363–368, 2018.

[4] C. Duarte *et al.*, "Non-intrusive load monitoring based on switching voltage transients and wavelet transforms," in *2012 Future of Instrumentation International Workshop Proceedings*. IEEE, 2012, pp. 1–4.

[5] A. I. Cole and A. Albicki, "Data extraction for effective non-intrusive identification of residential power loads," in *IMTC/98 Conference Proceedings*, vol. 2. IEEE, 1998, pp. 812–815.

[6] Z. Erkin *et al.*, "Privacy-preserving data aggregation in smart metering systems: An overview," *IEEE Signal Processing Magazine*, vol. 30, no. 2, pp. 75–86, 2013.

[7] C. Dwork *et al.*, "Calibrating noise to sensitivity in private data analysis," in *Theory of Cryptography Conference*. Springer, 2006, pp. 265–284.

[8] K. Chaudhuri, C. Monteleoni, and A. D. Sarwate, "Differentially private empirical risk minimization," *Journal of Machine Learning Research*, vol. 12, no. Mar, pp. 1069–1109, 2011.

[9] J. Hsu *et al.*, "Privately solving linear programs," in *International Colloquium on Automata, Languages, and Programming*. Springer, 2014, pp. 612–624.

[10] F. Fioretto, T. W. Mak, and P. Van Hentenryck, "Differential privacy for power grid obfuscation," *IEEE Trans. Smart Grid*, vol. 11, no. 2, pp. 1356–1366, 2019.

[11] T. W. Mak *et al.*, "Privacy-preserving power system obfuscation: A bilevel optimization approach," *IEEE Trans. Power Syst.*, vol. 35, no. 2, pp. 1627–1637, 2019.

[12] S. S. Baghsorkhi and I. A. Hiskens, "Impact of wind power variability on sub-transmission networks," in *2012 IEEE Power and Energy Society General Meeting*. IEEE, 2012, pp. 1–7.

[13] A. Shapiro, D. Dentcheva, and A. Ruszczyński, *Lectures on stochastic programming: modeling and theory*. SIAM, 2009.

[14] D. Bienstock and A. Shukla, "Variance-aware optimal power flow: Addressing the tradeoff between cost, security, and variability," *IEEE Control Netw. Syst.*, vol. 6, no. 3, pp. 1185–1196, 2019.

[15] F. Zhou, J. Anderson, and S. H. Low, "Differential privacy of aggregated DC optimal power flow data," *arXiv preprint arXiv:1903.11237*, 2019.

[16] V. Dvorkin *et al.*, "Differentially private distributed optimal power flow," *arXiv preprint arXiv:1910.10136*, 2019.

[17] S. Han, U. Topcu, and G. J. Pappas, "Differentially private distributed constrained optimization," *IEEE Trans. Autom. Control*, vol. 62, no. 1, pp. 50–64, 2016.

[18] Z. Zhang *et al.*, "Cost-friendly differential privacy for smart meters: Exploiting the dual roles of the noise," *IEEE Trans. Smart Grid*, vol. 8, no. 2, pp. 619–626, March 2017.

[19] M. Baran and F. F. Wu, "Optimal sizing of capacitors placed on a radial distribution system," *IEEE Trans. Power Deliv.*, vol. 4, no. 1, pp. 735–743, 1989.

[20] L. A. Roald, "Optimization methods to manage uncertainty and risk in power systems operation," Ph.D. dissertation, ETH Zurich, 2016.

[21] A. Bernstein, C. Wang, E. Dall'Anese, J.-Y. Le Boudec, and C. Zhao, "Load flow in multiphase distribution networks: Existence, uniqueness, non-singularity and linear models," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 5832–5843, 2018.

[22] R. Mieth and Y. Dvorkin, "Distribution electricity pricing under uncertainty," *arXiv preprint arXiv:1905.07526*, 2019.

[23] K. Chatzikokolakis *et al.*, "Broadening the scope of differential privacy using metrics," in *International Symposium on Privacy Enhancing Technologies Symposium*. Springer, 2013, pp. 82–102.

[24] C. Dwork and A. Roth, "The algorithmic foundations of differential privacy," *Foundations and Trends® in Theoretical Computer Science*, vol. 9, no. 3–4, pp. 211–407, 2014.

[25] L. Roald, S. Misra, T. Krause, and G. Andersson, "Corrective control to handle forecast uncertainty: A chance constrained optimal power flow," *IEEE Trans. Power Syst.*, vol. 32, no. 2, pp. 1626–1637, 2016.

[26] V. Dvorkin, J. Kazempour, and P. Pinson, "Chance-constrained equilibrium in electricity markets with asymmetric forecasts," *arXiv preprint arXiv:2005.11749*, 2020.

[27] T. Akbari and M. T. Bina, "Linear approximated formulation of AC optimal power flow using binary discretisation," *IET Generation, Transmission & Distribution*, vol. 10, no. 5, pp. 1117–1123, 2016.

[28] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[29] A. Papavasiliou, "Analysis of distribution locational marginal prices," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 4872–4882, 2017.

[30] R. Mieth and Y. Dvorkin. Code supplement - DLMPs under uncertainty. [Online]. Available: https://github.com/korpuskel91/DLMP_uncertainty_CodeSupplement

[31] M. Lubin and I. Dunning, "Computing in operations research using julia," *INFORMS J Comput*, vol. 27, no. 2, pp. 238–248, 2015.

[32] V. Dvorkin *et al.* Online Appendix to DP-CC-OPF. [Online]. Available: https://github.com/wdvorkin/DP_CC_OPF

[33] O. Karaca, S. Delikaraoglou, G. Hug, and M. Kamgarpour, "Enabling inter-area reserve exchange through stable benefit allocation mechanisms," *arXiv preprint arXiv:1912.09933*, 2019.