Studying Utility Metrics for Differentially Private Low-Voltage Grid Monitoring

Christine Schäler*, **Hans-Peter Schwefel***,**

*GridData GmbH, Anger, Germany.

**Aalbourg University, Denmark.

E-mail: {schaeler,schwefel}@griddata.eu

Received 19 February 2022; received in revised form 24 June 2024; accepted 3 July 2024

Abstract. The low-voltage energy distribution grid carries power to industrial and residential customers. To ensure its correct operation, distribution grid operators aim to monitor the grid with grid monitoring systems continuously. The system processes the stream of active and reactive power measurements at customer connections. Since this imposes major privacy concerns, data gateways typically sanitize the streams by adding noise to each measurement to achieve differential privacy. This however reduces the utility of the grid monitoring system. Due to missing studies, the utility of grid monitoring is not known. This leads to two research questions investigated in this study by means of a realistic case study. The first question is how to measure the utility appropriately. The second question is to give an intuition on whether one can achieve reasonable privacy and utility at the same time. Studying these questions is challenging for two reasons: The plurality of (1) grid analyses a grid monitoring system conducts, and (2) privacy requirements customers can have. To tackle the challenges, we identify a set of candidate utility metrics and use a differential privacy mechanism that unpacks multiple privacy requirements into one scaling parameter. Our experiments on a real-world grid and realistic measurements indicate the following. First, the utility of grid monitoring decreases faster than the sanitization error, that is frequently used in related work on differential privacy as utility metric. Second, already under weak privacy requirements, the utility is lower than under measurement errors.

Keywords. Differential Privacy, Electricity Distribution Grid, Utility Metrics.

1 Introduction

The smart meter roll-out in Europe comes with availability of quarter-hourly power measurement streams of customers. Distribution system operators (DSOs) aim to turn these data streams into value by using them to monitor the low-voltage grid continuously [5, 28]. To this end, DSOs use a plurality of automated grid analyses, like voltage and line loading analysis, that link measurements with additional data, like the grid topology from a geographic information system [28, 39, 45] (see Figure 1). However, it is well-known that active as well as reactive power measurements facilitate everyone who has access to this data to infer daily habits of the customers [41, 20]. Consequently, customers may have privacy requirements on their measurements. A prominent example is hiding certain power

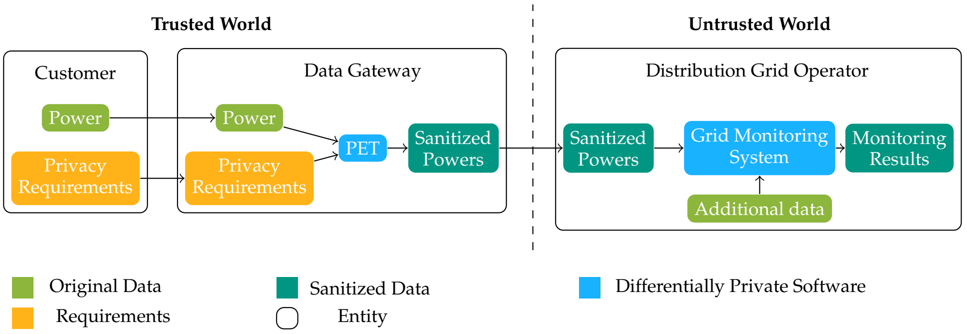


Figure 1: Differentially private grid monitoring. *Original* data refers to data that is measured and not sanitized.

patters like appliance usage cycles [55, 31]. As Figure 1 shows, to tackle these requirements, a trusted data gateway, that is locally deployed at each customer connection, sanitizes the measurements with respect to the privacy requirements. To this end, the gateway uses a privacy-enhancing technology (PET). After that, the gateway transmits the sanitized measurements to the DSO running the grid monitoring system.

The current gold-standard in research are PETs based on the w-event differential privacy framework [13, 30] and related definitions [55, 26, 32]. The w-event DP framework guarantees that anyone who inspects the measurements is not able to distinguish whether a power pattern, like an appliance usage cycle, of a maximum length of w time stamps is present or not. This guarantee is perfectly suitable to sanitize measurements before transmitting them to DSOs. The reason is that, in contrast to, e.g., k-anonymity [54], ad-hoc noise adding [11, 48] or temporal aggregation approaches [16], this guarantee features post-processing immunity [13]. This means that grid monitoring performed on the sanitized data still features differential privacy. However, a differentially private PET sanitizes a stream typically by adding a well-defined amount of noise to the measurements. As this noise may falsify the grid monitoring results, usually, the goal is to design a mechanism such that the resulting sanitized measurements feature *high utility*, i.e., the grid monitoring yields useful results for the DSO. Selecting a PET yielding high utility for grid monitoring is currently challenging because of two limitations of respective related work: First, most related work proposing, e.g., novel PETs evaluate the utility of PETs for use cases like location monitoring only [51, 18]. Since measurement streams have other properties, these results can hardly be transferred. Second, related work focuses on the sanitization error of the measurements as utility metric. But it is not known whether this metric is appropriate to measure the utility of grid monitoring, i.e., whether the error of grid monitoring behaves in the same way.

Consequently, in this paper, we study the utility of differentially private PETs for grid monitoring. Specifically, we are interested to answers the following research questions:

(RQ1) Is the sanitization error an appropriate utility metric for grid monitoring?

(RQ2) Having selected an utility metric, how behaves the utility of differentially private grid monitoring under reasonable privacy requirements?

1.1 Challenges

Studying the utility of differential privacy PETs for grid monitoring is challenging, due to the complexity of grid monitoring systems and privacy requirements customers may have. Specifically, we face three challenges:

Defining Candidates for Utility Metrics To answer (RQ1), candidates for utility metrics are required. This is challenging due to the plurality of results [47] a grid monitoring system usually outputs.

Assessing Utility To answer (RQ2), given a utility metric, a subsequent challenge is to identify utility thresholds stating that the grid monitoring results are accurate enough for DSOs. This definition should be DSO-independent to support the generality of our results.

Defining Reasonable Privacy To answer (RQ1) and (RQ2), it is essential to study the utility for reasonable privacy requirements. Intuitively, reasonable privacy is given if realistic privacy requirements are fulfilled. However, selecting them and studying the utility with respect to a *plurality* of realistic requirements is challenging without conducting an unlimited number of experiments.

1.2 Contributions

To tackle these challenges and answer our research questions, we provide the following contributions.

- We identify three steps of differentially private grid monitoring, namely, (1) measurement sanitization with the PET, (2) load-flow analysis as first step in each grid monitoring system, (3) subsequent analyses like voltage and line loading analysis, and candidates for utility metrics for all three. For each step, the metrics are in line with the metrics used by experts for the individual steps.

- To give an intuition on the utility of grid monitoring results, we consider the frequent case in which measurement devices are not fully accurate, since the DSOs already accept this resulting reduced utility.

- We identify realistic power patterns from literature and decode them into privacy parameters. To study the utility for a plurality of them with a limited number of experiments, we leverage the differentially private Uniform mechanism [30] that unpacks multiple requirements into one scaling parameter.

- To answer our research questions, we perform experiments on a real-world grid topology with realistic measurements. With respect to the first question, our study indicates that the utility of grid monitoring decreases faster than metrics measuring the sanitization error suggest, indicating that grid monitoring specific metrics are needed to assess utility meaningfully. With respect to the second one, our study suggests that it is hard to achieve reasonable utility and privacy in grid monitoring at the same time.

Ref.	$w = 1$	$w \in [2, \infty)$	$w \to \infty$
[31]	✓	✗	✗
[17]	✓	✗	✗
[15]	✗	✗	✓
[2]	✓	✓	✗
[50]	✓	✗	✗
[22]	✓[a]	✗	✗
[23]	✓	✓[b]	✗
[33]	✗	✗	✓
This paper	✓	✓	✓

Table 1: Comparing related work on differential privacy for measurement data with respect to the DP variant considered.

[a] Privacy potentially violated by using faithfulness value β computed on the measured data during post processing.

[b] Considers a weaker variant of w-event DP with disjunctive w-periods instead of rolling windows as considered in [30].

1.3 Outline

This paper is structured as follows: In Section 2, we sketch related work on differentially private grid monitoring. In Section 3, we identify the analyses conducted by grid monitoring systems and select the data used in our study. Next, in Section 4, we identify candidates for utility metrics and state how we generate measurement errors. Then, in Section 5, we select the PET used in our study and identity reasonable privacy requirements. Section 6 provides and discusses the results of our study with respect to our research questions. Last, in Section 7, we conclude our paper and discuss implications on future work.

2 Related Work

In this section, we sketch related work on differentially private PETs for measurements, utility metrics for PETs as well as studies on security and privacy with respect to additional data used in grid monitoring systems.

Differentially Private PETs For streams, w-event differential privacy (DP) is the current state of the art [30]. It is a probabilistic definition claiming the indistinguishably with a factor e^{ϵ} of power measurements that differ at most by a share Δ at each time stamp within a window of w consecutive time stamps. For $w = 1$ and $w \to \infty$, w-event DP is equivalent to event-level DP ($w = 1$) [14] or user-level DP ($w \to \infty$) [12]. Both has been investigated in related work for measurement data (see Table 1). However, event-level DP features only limited privacy for streams [8, 2], and user-level differential privacy limited utility, already for finite time series [15]. While w-event DP for $w > 1$ has proven its worth for, e.g., location streams [56, 34, 37], so far, it has been sparsely investigated for measurement data. [2] considers the concept even before the proposal of w-event DP. However, it features only a limited number of experiments with respect to w-event DP. Additionally, it considers a special definition called distributed DP that generally results in lower utility than w-event DP.

Measuring Utility of PETs All perturbation methods have the drawback that they introduce an error into the measurements, that influence the utility of grid monitoring systems. Related work focusing on w-event DP for measurement streams assesses this error *either* by the sanitization error [4, 15, 2], *or* by the error of a specific analysis. The latter includes local energy market analysis [31], specific forecasting algorithms [17], peak-load analysis [36] or state estimation [50]. To the best of our knowledge, these two types of utility metrics have not been systematically related to each other before, nor intuitions on "high enough" utility are given.

Security and Privacy with Respect to Additional Data Besides privacy with respect to measurements, customers may have additional privacy requirements, like secure data transmission [38, 49] or protecting the Smart Meter against attacks [3, 1]. Additionally, in case an untrusted service provider hosts the grid monitoring system, DSOs have to transmit the additional data, like grid topology data, needed as input for grid monitoring systems. In this context, [44, 21] focus on differential privacy for grid topology data, like line parameters. However, these approaches are orthogonal to our research questions.

3 Fundamentals on Grid Monitoring and Identification of Study Data

In this section, we first sketch fundamentals on grid monitoring systems, resulting in requirements on the data used in our study. Then, we introduce related work on differentially private grid monitoring.

3.1 Grid Monitoring Systems

In this section, we first sketch fundamentals on grid topologies and measurements serving as an input into grid monitoring systems. Second, we introduce grid monitoring systems as detailed as needed to identify requirements on study data and to define utility metrics in the remainder.

3.1.1 Grid Topologies and Measurements

Subsequently, we sketch fundamentals on grid topologies, relevant measurands and measurement scenarios together with notation.

Grid Topology A low-voltage grid topology is given by

1. a grid topology graph $G = (N, E)$ in which the edges E are the lines, and N the nodes,
2. a function *typeN* $: N \rightarrow$ {Trafo, customer connection box, junction box, sleeve} assigning nodes a type, and
3. a function *typeE* $: E \rightarrow \mathbb{R}^4$ assigning lines quintuples of resistance (R1), reactance (X1), capacitance (C1) and ampacity (I_{max}) in ampere.

Without loss of generality, we assume that G is a tree, as this simplifies explanations and applies for most grids. A *feeder* is a single branch in the tree. Each node and line has a specific type. For nodes, it holds that the unique root of the tree is of type Trafo. It serves as connection point to the parent medium-voltage grid. The leaf nodes of the grid are customer connection boxes (CCBs) connecting residential and industrial customers with the grid. The nodes between the substation and the customer connection boxes are either junction boxes or sleeves. The type of a line specifies its electrical parameter resistance, reactance, capacitance and ampacity. In this paper, we assume that the topology is fully known. If not, one can use approaches to complete the topology, e.g., [35].

Measurands In a grid, DSOs deploy measurement devices at nodes that measure specific measurands. Typically, they are measured per phase. However, grid monitoring systems usually consider a one-phase representation of the three-phase grid [47].

Let $n \in N$ be a node, and $e \in E$ be a line. We denote with $\mathsf{V}(n, t)$ the average voltage magnitude in volts at the secondary side of node n in a time interval ending at time stamp t. Similarly, with $\mathsf{P}(n, t)$ and $\mathsf{Q}(n, t)$, we denote the total active power in kilowatts (kW) and reactive power in kilovar (kVar) injected at time t and node n into the grid. Additionally, $\mathsf{I}(e, t)$ is the average current magnitude in ampere in a time interval ending at time stamp t at the secondary side of line e. In case they are clear from the context, we omit the parameters n and e.

Measurement Scenarios The measurement scenario of a grid specifies which measurands are measured at which grid node or line. Considering real-world scenarios, we identified two measurement scenarios, namely PQ and P only, imposing different privacy requirements of a PET. Subsequently, we describe them.

Measurement Scenario PQ This is the measurement scenario stated in Table 2. Here, the voltages at the Trafo, as well as active and reactive power at all customer connection boxes are measured.

Measurement Scenario P only In contrast to scenario PQ, the meters at the customer connection boxes measure only active power, but not reactive power. This is a frequent setting in real-world. For instance, Smart Meters in Germany are, by default, configured accordingly [7].

3.1.2 Grid Monitoring Systems

As illustrated in Figure 2, a grid monitoring system usually implements a two-step process: A load-flow analysis determining non-measured measurands followed by a plurality of subsequent grid analyses calculating system indicators [5, 28]. Below, we sketch both steps briefly based on [28], and state specifics relevant for ensuring the reproducibility of our results.

Step 1 – Load-flow Analysis The load-flow analysis calculates voltages at all nodes except the Trafo, and the currents at all lines. The algorithm is stated in Algorithm 1. There, based on grid topology and measurements of arbitrary granularity, the algorithm obtains a set of linearized, originally non-linear, power balance equations. The unknown variables correspond to voltages at non-Trafo nodes, and the known ones to active and reactive power at the customer connection boxes. By solving this system with the iterative

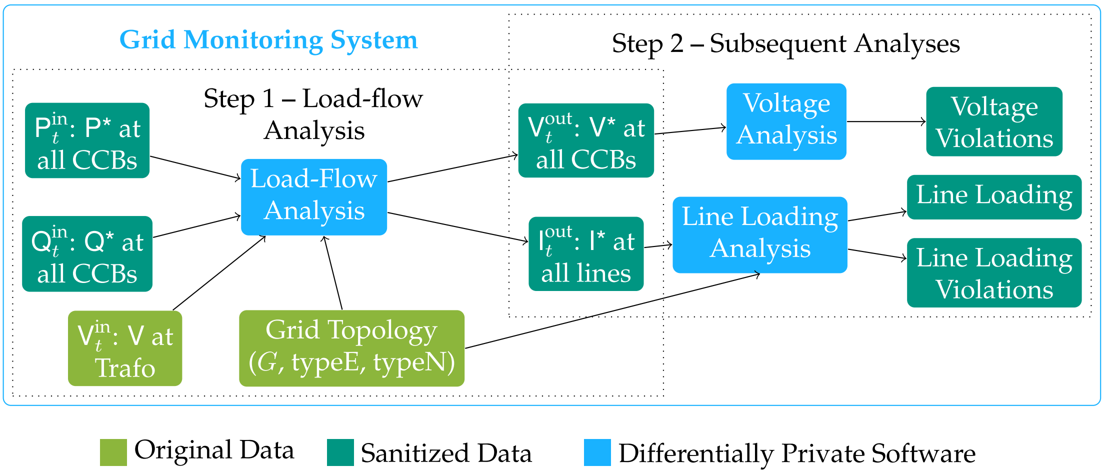

Grid Monitoring System
Step 1 – Load-flow Analysis
Step 2 – Subsequent Analyses
P_t^{in}: P* at all CCBs
Q_t^{in}: Q* at all CCBs
V_t^{in}: V at Trafo
Load-Flow Analysis
Grid Topology (G, typeE, typeN)
V_t^{out}: V* at all CCBs
I_t^{out}: I* at all lines
Voltage Analysis
Voltage Violations
Line Loading Analysis
Line Loading
Line Loading Violations
🟩 Original Data
🟩 Sanitized Data
🟦 Differentially Private Software

Figure 2: Illustration of the two-step process of grid monitoring.

Node Type	V	P	Q	I
Transformer	measured	-	-	-
Junction box	calculated	-	-	-
CCB	calculated	measured	measured (in scenario PQ)	-
Lines	-	-	-	calculated

Table 2: Measured measurands, and the ones calculated by load-flow analysis. Reactive power at CCBs are only measured in scenario PQ, and are replaced by pseudo-measurements in scenario P only.

Newton-Raphson method [6] that minimizes the mismatches in active and reactive power at customer connection boxes, the algorithm obtains the voltages of all nodes. Based on the obtained voltages, the algorithm calculates the currents. In a grid, except the power at the Trafo, the power, voltage and current measurements in one feeder are independent of the measurements in another feeder. Consequently, the load-flow analyses on different feeders are independent. In our study, we use the load-flow analysis that is implemented in a grid model developed in [47], and successfully validated in [43]. It is based on the MATLAB implementation in [46]. We stop as soon as the mismatches are smaller than 10^{-5}, or after 100 iterations otherwise. Reactive power measurements are a necessary input into the load-flow analysis. Consequently, in measurement scenario P only, so-called pseudo measurements are generated using background knowledge on the customer connection [40]. Since this background knowledge is private information as well, in our study, we use for reactive power the pseudo-measurement 0 kilovar, which is a common choice in case no background knowledge is available.

Step 2 – Subsequent Grid Analyses For a low-voltage feeder, two grid analyses, namely, voltage analysis and line loading analysis are relevant [28]. Subsequently, we introduce both analyses.

Voltage Analysis By European standards [19], the DSO must ensure that the voltage mag-

Algorithm 1 Load-flow Analysis at time stamp t [46]

```
function LOADFLOW$_t$($G$, typeN, typeE, $V_t^{in}$, $P_t^{in}$, $Q_t^{in}$)
    $V_t^{init} \leftarrow V_t^{in}$
    for $c \in N$ with typeN(c) == CCB do
        $V_t(c,t) \leftarrow 0$                                  ▷ Initialization
        $V_t^{init} \leftarrow V_t^{init} \cup \{V(c,t)\}$
    end for
    $V_t^{out} \leftarrow$ Newton-Raphson($G$, typeE, $V_t^{init}$, $P_t^{in}$, $Q_t^{in}$)
    $I_t^{out} \leftarrow \{\frac{|V(c,t)-V(c_j,t)|}{\text{typeE}(e_{i,j}).R1} \mid e_{i,j} \in E\}$
    return $V_t^{out}$, $I_t^{out}$
end function
```

nitudes fluctuates at maximum $+/-10\%$ of the nominal voltage of 400 V. To alert DSOs before they violate this hard limit, the voltage analysis verifies whether the voltages at all nodes are in range $+x\%$ (*over*voltage violation) and $-y\%$ (*under*voltage violation) of 400 V, and reports violations in case they are not. In line with industry standards, we use $x = y = 5$.

Line Loading Analysis Lines can manage a certain nominal current that is given by their ampacity. If the actual current is higher, they overheat. The line loading analysis therefore calculates for $e \in E$ the load in percent by

$$\text{Load}(e,t) = \frac{|(\mathsf{I}(e,t)|}{\text{typeE}(e).\mathsf{I}_{\text{max}}} \tag{1}$$

and alerts the DSO in case the load is above a certain limit. That way, DSOs can react before the lines overheat. A common load limit that we use is 90% [52].

3.2 Selection of Study Data

In this section, we select the grid topology and measurement data for our study. For each of them, we first state requirements we impose on the data to be used in our study based on the previous sections. Then, we state our selection. A natural overall requirement is that the grid topology and measurements match, meaning that a data set containing only measurements or only a grid topology is not appropriate.

Grid Topology We derive the following requirements on a low-voltage grid topology used in the study. First, we need the grid topology of at least one feeder of a low-voltage grid. A single feeder is also sufficient, because the results of the load-flow analysis are independent for each feeder. Second, to ensure the validity of our results, there should exist a successfully validated digital representation. Third, ideally, it should be a real-world grid to support the validity of our results. Considering these requirements, as grid topology, we use a feeder of a real-world grid topology from the Danish DSO Thy Mors Energi[1] (TME). Several studies before [27, 43] use it as a reference grid model as well. The grid contains 25 customer connection boxes. Four of them correspond to industrial customers (e.g., a farm), the others correspond to residential customers. Additionally, the feeder contains one Trafo and 10 junction boxes.

[1] www.thymors.dk

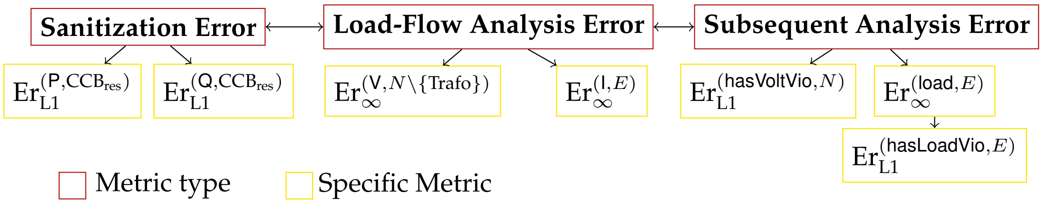

Sanitization Error
Load-Flow Analysis Error
Subsequent Analysis Error
$\mathrm{Er}_{\mathrm{L1}}^{(\mathsf{P},\mathrm{CCB}_{\mathrm{res}})}$
$\mathrm{Er}_{\mathrm{L1}}^{(\mathsf{Q},\mathrm{CCB}_{\mathrm{res}})}$
$\mathrm{Er}_{\infty}^{(\mathsf{V},N\setminus\{\mathrm{Trafo}\})}$
$\mathrm{Er}_{\infty}^{(\mathsf{I},E)}$
$\mathrm{Er}_{\mathrm{L1}}^{(\mathsf{hasVoltVio},N)}$
$\mathrm{Er}_{\infty}^{(\mathsf{load},E)}$
$\mathrm{Er}_{\mathrm{L1}}^{(\mathsf{hasLoadVio},E)}$
Metric type
Specific Metric

Figure 3: Identified utility metric types together with associated error metrics.

Measurements We derive the following requirements on measurements used in the study. First, to be able to calculate utility metrics, a fully-measured measurement scenario (i.e., all measurands are measured) is needed. Second, previous work on differential privacy for measurement data [15] indicates that measurement streams of one day are needed and also sufficient. Third, ideally, violations should be present in order to study the relationship between privacy requirements and violations in the voltage analysis. Considering these requirements, we use the measurements from the 25 hours undervoltage trace described in [43][2] and aggregate them to quarter-hourly values. The data set features a fully-measured measurement scenario (i.e., all electrical parameters are measured) and contains undervoltage violations. The measurements are simulated with the hardware-in-the-loop simulator OPAL-RT[3], and were also used in previous studies [43]. For details regarding the simulation process, we refer to [43, 27].

4 Defining Reasonable Utility for Grid Monitoring

In this section, we first define candidates for utility metrics for grid monitoring. In our experiments, we investigate which of them are appropriate. Second, we state how we generate measurement errors used to give an intuition on the utility achieved by a PET.

4.1 Measuring Utility – Candidate Utility Metrics

Considering Figures 1 and 2, we identified three types of utility metrics serving as candidates. As shown in Figure 3, these are: the sanitization error, load-flow analysis error and subsequent analysis error. In the remainder of this section, we first identify the specific metrics commonly used in related work for each of the types. No subsequent analysis error metrics have been proposed in literature before. Consequently, second, we propose novel subsequent analysis error metrics. We state them by using the following notation: $\mathcal{E}$ is a measurand or system indicator (like line loading), and $\mathcal{X}$ is a set of lines or nodes. Measurands or system indicators *with* superscript $*$ belong to sanitized measurements or system indicators calculate by using sanitized measurements. Consequently, e.g., P and Q are the measured powers, and P^* and Q^* the sanitized ones. Additionally, p is the number of time intervals in the measurement stream prefix.

[2] `www.bit.ly/vbn-data-expval.`
[3] `www.opal-rt.com`

Sanitization Error Metric For the sanitization error, the mean absolute error is frequently used [30, 56, 10, 51]. For measurand $\mathcal{E}$ and set of lines or nodes $\mathcal{X}$ it is defined by

$$\mathrm{Er}_{\mathrm{L1}}^{(\mathcal{E},\mathcal{X})} = \frac{1}{p \cdot |\mathcal{X}|} \sum_{t=1}^{p} \sum_{x \in \mathcal{X}} |\mathcal{E}(x,t) - \mathcal{E}^*(x,t)|, \tag{2}$$

i.e., the average over the L1-norm between $\mathcal{E}$ and $\mathcal{E}^*$ at each time stamp and node or line. Since we sanitize active and reactive power of residential customers, in our study, we focus on $\mathrm{Er}_{\mathrm{L1}}^{(\mathsf{P},\mathrm{CCB}_{\mathrm{res}})}$ and $\mathrm{Er}_{\mathrm{L1}}^{(\mathsf{Q},\mathrm{CCB}_{\mathrm{res}})}$, in which $\mathrm{CCB}_{\mathrm{res}}$ is the sub-set of customer connection boxes of N belonging to residential customers.

Load-flow Analysis Error Metric The load-flow analysis calculates the voltages and currents that are subsequently used in voltage and line loading analysis. This suggests to measure the utility of the load-flow analysis by the error in calculated voltages and current. To define these errors, power engineers usually rely on the maximum norm [42]. Consequently, to quantify the voltage and current error, we use

$$\mathrm{Er}_{\infty}^{(\mathcal{E},\mathcal{X})} = \frac{1}{p} \sum_{t=1}^{p} \max_{x \in \mathcal{X}} |\mathcal{E}(x,t) - \mathcal{E}^*(x,t)|. \tag{3}$$

Note that our study, measurement streams contain the measured voltages and currents at all nodes and lines, facilitating us to calculate this error. Since the load-flow analysis calculates the voltages at all nodes except the Trafo, we are interested in $\mathrm{Er}_{\infty}^{(\mathsf{V},N\setminus\{\mathrm{Trafo}\})}$. For currents, we are interested in the current error at all lines, i.e., $\mathrm{Er}_{\infty}^{(\mathsf{I},E)}$.

Subsequent Analyses Error Metrics To the best of our knowledge, there is no related work measuring the utility of voltage and line loading analysis. Consequently, we propose novel error metrics considering the benefit of the DSO of both analyses.

Voltage Analysis To calculate this error, we first compare the number of violations for each node on the original and sanitized measurements. Then, we sum these numbers yielding error in the total number of voltage violations. Formally, let the variable $\mathsf{hasVoltVio}(n,t)$ indicate whether there is a voltage violation at node n and time stamp t. With that notation, we measure the error in the total number of voltage violations that is given by $\mathrm{Er}_{\mathrm{L1}}^{(\mathsf{hasVoltVio},N)}$. This means that one obtains the calculation formula by using $\mathcal{E} = \mathsf{hasVoltVio}$ and $\mathcal{X} = N$ in Equation 2.

Line Loading Analysis The line loading analysis first calculates the load of each line, and then checks whether the load is below the defined limit. We define an error metric for each of the two steps. First, we calculate the loading error in percentage points based on comparing the loading for each line on the original and sanitized measurements. Formally, we define loading error in percentage points (% P) by $\mathrm{Er}_{\infty}^{(\mathsf{load},E)}$. This means that one obtains the calculation formula by using $\mathcal{E} = \mathsf{load}$ and $\mathcal{X} = E$ in Equation 3. Second, we calculate the error in the total number of loading violations based on comparing the loading violations for each line on the original and sanitized measurements. Formally, let $\mathsf{hasLoadVio}(e,t)$ be the indicator variable for a line loading violation of line e at time stamp t. With that notation, we measure the error in the total number of loading violations by $\mathrm{Er}_{\mathrm{L1}}^{(\mathsf{hasLoadVio},E)}$. This means that one obtains the calculation formula by using $\mathcal{E} = \mathsf{hasLoadVio}$ and $\mathcal{X} = E$ in Equation 2.

4.2 Assessing Utility – Measurement Errors

A PET is not the only influence factor that may reduce the utility of grid monitoring system. Specifically, measurement devices are typically not completely accurate, meaning that even the analysis results on measured data are erroneous. Consequently, to assess whether the utility provided by the PET is reasonable, we propose to compare the utility of the PET with the utility that can be achieved if measurement errors are present. This utility is already accepted by DSOs. As measurement error, in line with previous work [42], we use measurement-dependent Gaussian noise with $\sigma = 0.01$. Specifically, let $\mathcal{G}$ be the Gaussian distribution. Then, the resulting active and reactive power measurement containing measurement errors are given by $(1 + \mathcal{G}(0, \sigma)) \cdot \mathsf{P}(t)$ and $(1 + \mathcal{G}(0, \sigma)) \cdot \mathsf{Q}(t)$.

5 Defining and Ensuring Reasonable Privacy for Measurement Streams

In this section, we first introduce the differentially private PET for measurement streams we use in our study. The PET is designed such that it allows to evaluate various privacy requirements with a limited number of experiments, i.e., PET runs. Second, we define reasonable privacy by identifying meaningful privacy requirements from literature.

5.1 Differential Privacy for Measurement Streams

Below, we first briefly introduce the definition w-event differential privacy (DP) for streams as far as required for this paper. For details, we refer to [30]. Then, we state the PET that we use satisfying w-event DP.

5.1.1 Definition of w-Event Differential Privacy

Intuitively, w-event DP guarantees that anyone who inspects the measurements is not able to distinguish whether a certain power pattern, like an appliance usage cycle, of a maximum length of w time stamps is present or not. To this end, the definition of w-event DP is based on the notion of neighboring databases and stream prefixes. Both are itself based on power shares and window lengths. Formally, let $S = (D_1, D_2, ..)$ be a power measurement stream of a customer connection box collecting database D_t at time stamp t as illustrated in Figure 4. Two databases D_t, D'_t are neighbors if (a) the total active power P(t) differs by at most Δ^{P} and (b) the total reactive power Q(t) differs by at most Δ^{Q} [9, 26, 2]. Now, let $S_p = (D_1, .., D_p)$ be a prefix of the stream S of length p. According to Definition 1, two stream prefixes are w-neighbors, if (1) the databases collected at each time are pairwise the same or neighbors, and (2) all neighboring databases fit into a window of size w.

Definition 1 (w-Neighboring Stream Prefixes [30])**.** Let w be a positive integer. Further, let $t, t_1, t_2 \leq p$ be three time stamps. Then, two stream prefixes S_p, S'_p are *w-neighboring*, if

1. for each D_t, D'_t with $D_t \neq D'_t$, it holds that D_t, D'_t are neighboring
2. for each $D_{t_1}, D_{t_2}, D'_{t_1}, D'_{t_2}$ with $t_1 < t_2$, $D_{t_1} \neq D'_{t_1}$ and $D_{t_2} \neq D'_{t_2}$, it holds that $t_2 - t_1 + 1 \leq w$.

	Consumer/	D_1		D_2		D_3		$\cdots$
Stream S	Generator	P	Q	P	Q	P	Q	$\cdots$
	Consumer 1	0.00	0.00	1.71	0.20	1.84	0.21	$\cdots$
	Consumer 2	0.00	0.00	0.00	0.00	0.60	0.10	$\cdots$
$Q(D_t) = (\mathsf{P}(t)\vert\mathsf{Q}(t))$		0.00	0.00	1.71	0.20	2.44	0.31	$\cdots$

Figure 4: Example for a measurement stream of a consumer connection.

Based on the definition of neighboring stream prefixes, Definition 2 defines w-event differential privacy. w-event ϵ-differential privacy is given if the power measurements of all w-neighboring stream prefixes are hard to distinguish. The privacy budget ϵ quantifies the hardness. It usually lies between 0.1 and 1.0 [51].

Definition 2 (w-Event ϵ-Differential Privacy [30])**.** Let $\mathcal{M}$ be a randomized mechanism that takes as input a stream prefix of arbitrary size. We say that $\mathcal{M}$ satisfies *w-event ϵ-differential privacy* if for all $R \in \text{Range}(\mathcal{M})$, all w-neighboring stream prefixes S_p, S'_p, and all p, holds that

$$\Pr[\mathcal{M}(S_p) = R] \leq e^{\epsilon} \cdot \Pr[\mathcal{M}(S'_p) = R]. \tag{4}$$

5.1.2 A PET ensuring w-Event Differential Privacy

To implement a PET, we leverage the Uniform mechanism [30] as stated in Algorithm 2. At each time stamp t, it sanitizes the power measurements $\mathsf{P}(t)$ and reactive $\mathsf{Q}(t)$ of a customer according to given, time-invariant, privacy requirements. The privacy requirements consist of the privacy level ϵ, window size w, the power shares Δ^{P} and Δ^{Q}, and noise splitting parameters $\alpha^{\mathsf{P}}, \alpha^{\mathsf{Q}}$ satisfying $\alpha^{\mathsf{P}} + \alpha^{\mathsf{Q}} = 1$. The noise splitting parameters split the available budget ϵ between active and reactive power. This is required because active and reactive power are correlated [29]. Given these inputs, the PET first calculates the noise scales for the Laplace distribution in Lines 4 and 5. Then, it perturbs the measurements yielding the sanitized measurements. Last, the PET outputs the sanitized measurements. [30] proofs that this PET satisfies w-event differential privacy.

Considering the scales in Lines 4 and 5, we observe that multiple privacy requirements result in the same noise scale. For instance, $\lambda^{\mathsf{P}}(\Delta^{\mathsf{P}} = 1, w = 5, \alpha^{\mathsf{P}} = 1, \epsilon = 1) = 5 = \lambda^{\mathsf{P}}(\Delta^{\mathsf{P}} = 5, w = 1, \alpha^{\mathsf{P}} = 1, \epsilon = 1)$. This enables to investigate the utility with respect to an infinite number of privacy requirements by one PET run.

5.2 Assessing Reasonable Privacy – Privacy Requirements

To limit the number of experiments, we aim to perform experiments for different noise scales. To this end, we focus on noise scales representing appliance usages from residential customers. Consequently, we sanitize only the measurements from the residential customers. In the remainder, we select the requirements and then state the resulting noise scales used in our study.

Algorithm 2 PET: Uniform Mechanism at time stamp t

```
function UNIFORM_t($\mathsf{P}(t)$, $\mathsf{Q}(t), \epsilon, w, \Delta^{\mathsf{P}}, \Delta^{\mathsf{Q}}, \alpha^{\mathsf{P}}, \alpha^{\mathsf{Q}}$)
    $\epsilon^{\mathsf{P}} \leftarrow \alpha_{\mathsf{P}} \cdot \epsilon$
    $\epsilon^{\mathsf{Q}} \leftarrow \alpha_{\mathsf{Q}} \cdot \epsilon$
    $\lambda^{\mathsf{P}} \leftarrow \lambda^{\mathsf{P}}(\Delta^{\mathsf{P}}, w, \alpha^{\mathsf{P}}, \epsilon) = \frac{\Delta^{\mathsf{P}} \cdot w}{\epsilon \cdot \alpha^{\mathsf{P}}}$
    $\lambda^{\mathsf{Q}} \leftarrow \lambda^{\mathsf{Q}}(\Delta^{\mathsf{Q}}, w, \alpha^{\mathsf{Q}}, \epsilon) = \frac{\Delta^{\mathsf{Q}} \cdot w}{\epsilon \cdot \alpha^{\mathsf{Q}}}$
    return PERTURB($\mathsf{P}(t), \mathsf{Q}(t), \lambda^{\mathsf{P}}, \lambda^{\mathsf{Q}}$)
end function
function PERTURB($\mathsf{P}(t), \mathsf{Q}(t), \lambda^{\mathsf{P}}, \lambda^{\mathsf{Q}}$)
    $\mathsf{P}^*(t) \leftarrow \mathsf{P}^*(t) + \mathrm{Lap}(\lambda^{\mathsf{P}})$        ▷ Laplace mechanism
    $\mathsf{Q}^*(t) \leftarrow \mathsf{Q}^*(t) + \mathrm{Lap}(\lambda^{\mathsf{Q}})$
    return $\mathsf{P}^*(t), \mathsf{Q}^*(t)$
end function
```

5.2.1 Selection of Privacy Requirements

Below, we select the privacy requirements. Table 4 provides an overview of the requirements we discuss. With respect to appliances, we consider all 14 appliances from [24] simulating Smart Meter time series from residential customers. In compliance with related work, we keep $\epsilon \in (0, 1.0]$.

Selection of Window Size w We focus on window sizes that correspond to typical appliance usage cycles. As stated in Table 3, we consider individual appliance usages, as well as combinations. An example for an individual appliance usage is the usage of the washing machine. According to [24], it usually runs for 2 hours and 15 min. Considering that our measurement streams feature an interval length of 15 min., this corresponds to $w = \frac{2h15m}{15m} = 7$. We consider additionally combinations of appliances, because appliance usages are typically correlated. An example is that the dryer usually runs after the washing machine. To cover this, we have to select the sum of the cycle duration of both appliances. For instance, if the washing machine runs 7 time stamps, and the dryer 9 time stamps, we have to select $w = 16$. If two appliances run in parallel, we have to select the maximum of both cycle durations. For compliance with related work, we integrate the respective window size for ensuring event-level ($w = 1$) and user-level ($w = \infty$) DP in our study (see Table 3). For user-level DP, that is not implementable for infinite streams, we interpret the stream as a finite time series. Then, we choose $w = p = 100$, which is the number of time stamps of our measurement stream prefix.

Selection of Shares Δ^{P} and Δ^{Q} Related work uses shares according to the classical local setting that protects the existence of all producer and consumer. To calculate these shares, we rely on the highest difference in active and reactive power of two residential customers at one time stamp that can occur in our measurement trace. Specifically, for $\mathcal{E} \in \{\mathsf{P}, \mathsf{Q}\}$, the share [2] is given by

$$\Delta^{\mathcal{E}} = |\max_{\mathrm{c} \in \mathrm{CCB}_{\mathrm{res}}, t} \mathcal{E}(t, c) - \min_{\mathrm{c} \in \mathrm{CCB}_{\mathrm{res}}, t} \mathcal{E}(t, c)|. \tag{5}$$

For our study data introduced in Section 3.2, these shares are given by $\Delta^{\mathsf{P}} = 4.85$ kW and $\Delta^{\mathsf{Q}} = 3.37$ kVar. With respect to appliance usages, the shares must be selected in line

Requirement	Δ^{P}	Δ^{Q}	w
Measurement Scenario PQ			
user-level DP	Eq. 5 ($\mathcal{E} = \mathsf{P}$)	Eq. 5 ($\mathcal{E} = \mathsf{Q}$)	p
w-event DP	Eq. 5 ($\mathcal{E} = \mathsf{P}$)	Eq. 5 ($\mathcal{E} = \mathsf{Q}$)	[2,p)
event-level DP	Eq. 5 ($\mathcal{E} = \mathsf{P}$)	Eq. 5 ($\mathcal{E} = \mathsf{Q}$)	1
Measurement Scenario P only			
Individual appliance i	power_share(i)	-	$\frac{\text{cycle\_duration}(i)}{15 \text{ min.}}$
i_1 followed by i_2	$\max\limits_{i \in \{i_1, i_2\}} \{\text{power\_share}(i)\}$	-	$\sum\limits_{i \in \{i_1, i_2\}} \frac{\text{cycle\_duration}(i)}{15 \text{ min.}}$
i_1 parallel to i_2	$\sum\limits_{i \in \{i_1, i_2\}} \text{power\_share}(i)$	-	$\max\limits_{i \in \{i_1, i_2\}} \{\frac{\text{cycle\_duration}(i)}{15 \text{ min.}}\}$

Table 3: Overview of shares and window sizes assuming a 15 min. stream sampling rate.

with the power usually consumed by the appliances or appliance combination as stated in Table 3. Since we are not aware of any publication stating the reactive power consumed by appliances, we consider privacy requirements with respect to appliance usages only in combination with measurement scenario P only, in which reactive power is not given.

Selection of Splitting Parameters $\alpha^{\mathsf{P}}, \alpha^{\mathsf{Q}}$ In line with related work [30], in the measurement scenario PQ, we use $\alpha^{\mathsf{P}} = \alpha^{\mathsf{Q}} = 0.5$. In the measurement scenario P only, in which we do not need to hide Q, we use the full budget for sanitizing active power. This means that we use $\alpha^{\mathsf{P}} = 1$.

5.2.2 Resulting Noise Scales

Table 4 states the resulting minimum and maximum noise scales per measurement scenario. For the measurement scenario P only, we only consider hiding individual appliance usages explicitly, as the maximum noise scale in this case is already similar to the minimum one for achieving user-level DP measurement scenario PQ. In case we do not achieve reasonable utility for user-level nor event-level DP, we consider reduced noise scales, limiting the maximum noise scale in scenario P only to the minimum noise scale needed to achieve event-level DP in scenario PQ. Note that we cover combinations of appliances implicitly, since the noise scales lie in the range between user-level privacy and consideration of individual appliances.

6 Results

In this section, we present and discuss the results of our study regarding our two research questions. Before that, we present grid monitoring results in the ground truth scenario to give an intuition on grid monitoring results in general. Second, we answer our first research question (RQ1) by relating the utility metrics of different types identified in Section 4.2 to each other. Finally, we state the results with respect to our second research question (RQ2) using the appropriate utility metrics identified. In all experiments, we use different values of w, ϵ, Δ^{P} and Δ^{Q} broke down to the noise scales λ^{P} and λ^{Q} as described in Section 5.2.1.

Name	Measrmt.	Requirement	ϵ	α^{P}	α^{Q}	λ^{P}		λ^{Q}	
	Scenario			min	max	min	max		
user	PQ	user-level DP	$[0.1, 1.0]$	0.5	0.5	970[a]	9.700[b]	674	6,740
event	PQ	event-level DP	$[0.1, 1.0]$	0.5	0.5	9.70[c]	97[d]	6.74	67.40
app	P only	indiv. appliances	$[0.1, 1.0]$	1.0	0.0	0.14[e]	980[f]	-	-
app′	P only	app reduced	$[0.1, 1.0]$	1.0	0.0	0.14	9.7[g]	-	-

[a] $\frac{\Delta^{\mathsf{P}}\cdot w}{\epsilon\cdot\alpha^{\mathsf{P}}} = \frac{4.85\cdot 100}{1.0\cdot 0.5}$

[b] $\frac{\Delta^{\mathsf{P}}\cdot w}{\epsilon\cdot\alpha^{\mathsf{P}}} = \frac{4.85\cdot 100}{0.1\cdot 0.5}$

[c] $\frac{\Delta^{\mathsf{P}}\cdot w}{\epsilon\cdot\alpha^{\mathsf{P}}} = \frac{4.85\cdot 1}{1.0\cdot 0.5}$

[d] $\frac{\Delta^{\mathsf{P}}\cdot w}{\epsilon\cdot\alpha^{\mathsf{P}}} = \frac{4.85\cdot 1}{0.1\cdot 0.5}$

[e] Refrigerator cycle with cycle_duration= 15 min. ($w = 1$) [24], power_share=140 W ($\Delta^{\mathsf{P}} = 0.14$ kW) [24] and $\epsilon = 1.0$.

[f] Space heating with cycle_duration= 210 min. ($w = 14$) [24], power_share=7,000 W ($\Delta^{\mathsf{P}} = 7$ kW) [24] and $\epsilon = 0.1$.

[g] Maximum noise scale from event-level DP.

Table 4: Privacy requirements and resulting minimum and maximum noise scales.

Note, as our experiments depend on random numbers, in line with the benchmark requirements proposed in [25, 51], we execute each experiment multiple times, and report the average errors. As preliminary experiments revealed that the average error converges after 10 runs, we stick to this number of runs.

6.1 Illustration of Grid Monitoring Results

Below, we present grid monitoring results in the measurement scenario PQ to give an intuition on grid monitoring results in general. The grid monitoring results in scenario serve as ground truth for all subsequent experiments. Additionally, we state the errors in measurement scenario P only compared to scenario PQ, and give an intuition on the sensitivity of the metrics. They serve as a lower bound on the errors we can achieve with a PET in scenario P only.

Figure 5 shows the results of the load-flow analysis as well as the subsequent analyses. These are the voltages and currents calculated by the load-flow analysis, as well as the resulting line loadings, and voltage respectively line loading violations determined by subsequent analyses. We observe that the voltages at all nodes are between 360 and 400 V, which is expected. Additionally, we see the voltage drop that was provoked during the simulation of the undervoltage trace. In total, we have 1,555 undervoltage violations: All 25 customer connections and 10 junction boxes have either 44 or 45 time stamps voltages below 380 V. The line currents are between close to zero and 23 A. The highest currents occur for the lines near the Trafo when the voltages in the nodes drop. This is expected, as the measurements are simulated by an online modification of the *secondary* voltages at the transformer [28]. Consequently, the loads of the lines increase during the voltage drop as well. However, the line load violation limit (90%) is never exceeded, meaning that no line loading violations are present. Using pseudo-measurements for reactive power as in scenario P only reduces utility. Below, beside the pure error numbers, we give a first intuition on the high sensitivity of the subsequent analyses error metrics. Table 5, Line $\mathrm{SC}^{\mathsf{P\ only}}$, shows the errors with respect to all defined metrics in the measurement scenario P only. We

Scenario	$\mathrm{Er}_{\infty}^{(\mathsf{V}, N \setminus \{\mathrm{Trafo}\})}$	$\mathrm{Er}_{L1}^{(\mathsf{hasVoltVio}, N)}$	$\mathrm{Er}_{\infty}^{(\mathsf{I}, E)}$	$\mathrm{Er}_{\infty}^{(\mathsf{load}, E)}$	$\mathrm{Er}_{L1}^{(\mathsf{hasLoadVio}, E)}$
Measurement Scenario with P and Q measured					
$\mathrm{SC}^{\mathsf{PQ}}$	——— Ground Truth ———				
$\mathrm{SC}^{\mathsf{PQ}}_{\sigma=0.01}$	0.013 V	0.2	0.13 A	0.24 %P	0
$\mathrm{SC}^{\mathsf{PQ}}_{\mathrm{event}}$	$[25; 17{\cdot}10^3]$ V	[299; 1,756]	$[113; 9 \cdot 10^6]$ A	$[201; 7 \cdot 10^6]$ %P	$[1; 3\,]{\cdot}10^3$
Measurement Scenario with P measured only					
$\mathrm{SC}^{\mathsf{P\ only}}$	0.23 V	13	2.64 A	1.55 %P	0
$\mathrm{SC}^{\mathsf{P\ only}}_{\sigma=0.01}$	0.24 V	4.8	4.0 A	3.1 %P	0
$\mathrm{SC}^{\mathsf{P\ only}}_{\mathrm{app'}}$	[0.4; 24.4] V	[17; 236.9]	[3.3; 95.4] A	[1; 245.5] %P	[0; 751]

Table 5: Overview of utility in all experiments. Error ranges relate to parameter variations. Non-integer violation numbers relate to reported mean errors over 10 experiment repetitions.

observe a small voltage $\mathrm{Er}_{\infty}^{(\mathsf{V}, N \setminus \{\mathrm{Trafo}\})}$ and current $\mathrm{Er}_{\infty}^{(I, E)}$ error. Additionally, we observe an error in the number of voltage violations, but not in the line loading violations. The rationale for the former is that, as Figure 5 (a) reveals, many voltage values are close to the undervoltage violation threshold. Consequently, even a small error in the voltages results in a difference in the number of violations. Inversely, a high current error is required to achieve a difference in the number of loading violations.

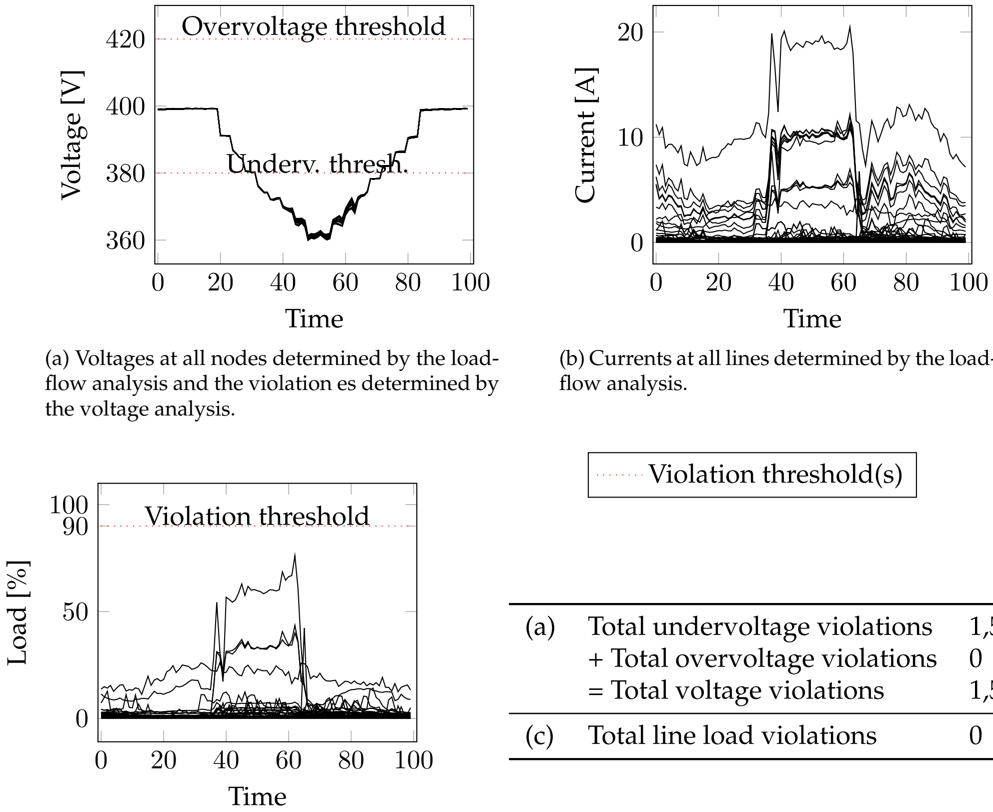
6.2 (RQ1) Utility Metric Relationships in Grid Monitoring

The DSO is interested in the results of the subsequent analyses. However, in literature, the utility of mechanisms with respect to sanitization error are known [51]. This raises Question (RQ1) asking how the sanitization, or the error of load-flow analysis as an intermediate step, relate to the subsequent analysis error. Considering the increasing number of subsequent analyses that exist, using the sanitization error or load-flow analysis error would be preferred if appropriate. Consequently, as illustrated in Figure 3, we subsequently relate these three types of utility metrics to each other. Ideally, they have a pairwise linear relation. In this case, the metrics are would be interchangeable, and all are appropriate to measure the utility of grid monitoring. To this end, we consider measurement scenario P only, because we have a higher variety of privacy requirements identified than for PQ.

The key results indicate the following. First, the relations between the metric types are highly non-linear. Second, the subsequent analysis errors are more measurement data dependent than the load-flow analysis errors. Consequently, we propose to use the load-flow analysis error in future work.

6.2.1 Sanitization Error vs. Load-flow Analysis Error

Below, we relate the sanitization error to both error metrics used to measure the load-flow analysis error, namely, the voltage and the current error. For the Uniform mechanism, it holds that the sanitization error converges towards λ^{P} or λ^{Q} [30], i.e., $\mathrm{Er}_{\mathrm{L1}}(\mathsf{P}, \mathrm{CCB}_{\mathrm{res}}) \rightarrow \lambda^{\mathsf{P}}, \mathrm{Er}_{\mathrm{L1}}(\mathsf{Q}, \mathrm{CCB}_{\mathrm{res}}) \rightarrow \lambda^{\mathsf{Q}}$. Due to the lack of reasonable privacy requirements w.r.t. Q in the



(a) Voltages at all nodes determined by the load-flow analysis and the violation es determined by the voltage analysis.

(b) Currents at all lines determined by the load-flow analysis.

(c) Line loads at all lines determined by the line-loading analysis.

(a)	Total undervoltage violations	1,555
	+ Total overvoltage violations	0
	= Total voltage violations	1,555
(c)	Total line load violations	0

(d) Sums of violations identified by voltage and line loading analysis.

Figure 5: Grid monitoring results in the ground truth measurement scenario PQ. In (a), each curve represents one node. In (b) and (c), each curve represents one line.

most scenarios (see Section 5.2), in our experiments, we have a focus on the measurement scenario PQ. But in this scenario, it holds that $\mathrm{Er}_{\mathrm{L1}}(\mathsf{Q}, \mathrm{CCB}_{\mathrm{res}}) = 0$, since Q is not measured and therefore also not sanitized. Consequently, we focus on λ^{P} as sanitization error.

Figure 6 (a) reveals that the load-flow analysis errors increase, as the sanitization error increases. However, not in a linear way. Both, the voltage and current error increase faster than the sanitization error. This means that the higher the sanitization error is, the less meaningful is it to assess the utility of grid monitoring. This applies especially to the current error increasing even faster than the voltage error. We consequently propose to use not only the sanitization error to assess the utility of newly proposed w-event DP PETs, but to use an analysis-specific metric in addition.

6.2.2 Load-flow Analysis Error vs. Subsequent Analysis Error

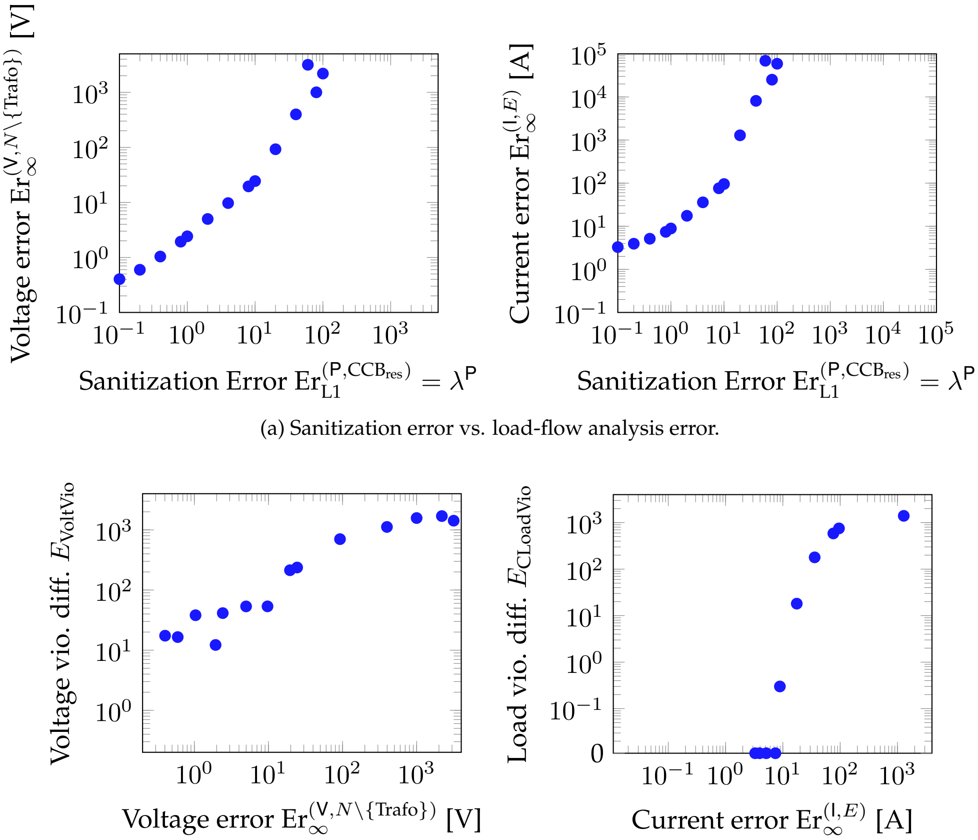
Knowing that the sanitization error is not appropriate, it remains the question whether the load-flow analysis error is the appropriate, or whether subsequent analysis metrics should be used. The reason is that subsequent analyses process the outputs of the load-flow analysis before the results are useful for the DSO. Consequently, we now discuss the relationship of the voltage and current error (output of load-flow analysis) on the difference in the voltage and line loading violations (output of subsequent analyses). They are illustrated in Figure 6 (b). Generally, we observe that the relationship is again non-linear. In contrast to Figure 6 (a), the relationship is even more complex. Specifically, regarding the voltage violations, for voltage errors < 10, the subsequent analyses errors change only slightly. The reason is that for many time stamps and nodes, the voltages are next to the violation limit (see Figure 5). Regarding the line loading violations, for readability, we focus on the error in the total number of violations. Note that for DSOs, the loading violations are more important that the 'raw' line loadings. Considering the results, we observe that even for a current error of 7 A, the violations are still correctly identified. The reason is that the loads in the undervoltage trace are far below the violation thresholds (see Figure 5). This indicates that errors of subsequent analyses are highly sensitive with respect to measurements and chosen thresholds. Measuring utility with subsequent analyses error metrics in research studies requires to compute the errors with respect to various measurements and thresholds, being a high effort. As a consequence, we propose to use load-flow analysis error metrics as a compromise. They do not reflect the error in grid monitoring as accurate as subsequent analyses error metrics, but significantly more accurate than the sanitization error.

6.3 (RQ2) Privacy-Utility Trade-Off

We first assess reasonable utility by determining the utility if measurement errors are present. Second, we compare this utility with the utility achieved by the PET. To this end, we focus, but not limit ourselves, to load-flow analysis utility metrics, as the previous section suggests. Table 5 gives an overview of the results discussed in this section. The scenarios are notated with $\text{SC}^{\text{Measurement Scenario}}_{\text{Noise}}$. The superscript states the measurement scenario $\in \{\textsf{PQ}, \textsf{P only}\}$. The subscript states which noise is introduced into the measurements, if any. In this context, $\sigma = 0.01$ stands for noise relating to measurement errors. Additionally, for noise resulting from a PET, the names are in line with Table 4. Note that both dimensions, i.e., measurement scenario and introduced noise, are orthogonal to each other. All errors are computed by comparing the analysis results with the ground truth $\text{SC}^{\textsf{PQ}}$. Key outcome is that it is hard to achieve reasonable utility while keeping reasonable privacy.

6.3.1 Utility in the Presence of Measurement Errors

In this experiment, for each measurement scenario, namely, PQ and P only, we determine the utility if we inject measurement errors. We only use this utility to give an intuition on the utility achieved by a PET in the remainder. The utility in scenario P only is stated in Table 5, Lines $\text{SC}^{\textsf{PQ}}_{\sigma=0.01}$ and $\text{SC}^{\textsf{P only}}_{\sigma=0.01}$. For the measurement scenario PQ, we observe that the measurement errors cause only a small error in voltages and currents. Additionally, they do not affect the number of line loading violation. Specifically, the voltage, current and loading errors are a magnitude smaller than the errors in $\text{SC}^{\textsf{P only}}$. For the measurement scenario P only, we compare the error in $\text{SC}^{\textsf{P only}}$ and $\text{SC}^{\textsf{P only}}_{\sigma=0.01}$, since the former is a lower bound of the latter. It reveals that the measurement errors have only a marginal impact on the voltage error, but nearly double the current and loading error. Interestingly, the difference in the number of voltage violation error *decreases*, which is unexpected. However, the reason is that the voltages calculated deviate more upwards to the ground truth, be-



(a) Sanitization error vs. load-flow analysis error.

(b) Load-flow analysis error vs. subsequent analysis error.

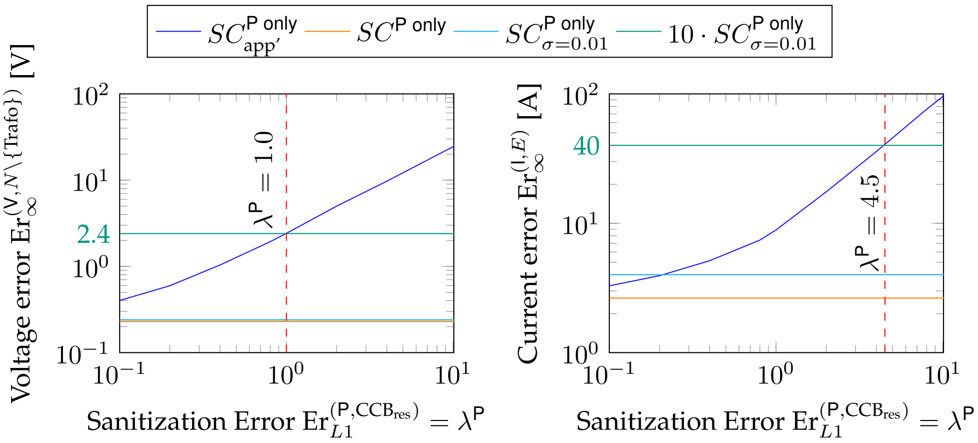
Figure 6: Relations between error metrics of different types.

cause there are more imbalances in the measurements than in $\mathrm{SC}^{\mathsf{P\ only}}$. Consequently, less undervoltage violations are present.

6.3.2 Intuition on the Utility of the PET

Subsequently, we compare the utility of the PET with the utility that can be achieved if measurement errors are present.

Measurement Scenario PQ For measurement scenario PQ, we first considered event-level differential privacy, since we expect higher utility than from a user-level differentially private PET. The noise scales used in our experiment are compliant with Table 4, Line 2. In Table 5, the Line $\mathrm{SC}_{\mathrm{event}}^{\mathsf{PQ}}$ shows the resulting errors. The lower numbers apply to $\epsilon = 1.0$, and the higher numbers to $\epsilon = 0.1$. We observe that already for $\epsilon = 1.0$ inducing the lowest privacy guarantee, the voltage and current errors are three orders of magnitudes higher than for the measurement error scenario $\mathrm{SC}_{\sigma=0.01}^{\mathsf{PQ}}$. Additionally, we observe *over*-voltage violations, that are not in line with the undervoltage trace (not visible in the table). As a result, we assess the errors are too high to achieve reasonable utility in this scenario.



(a) Voltage error for varying noise scales.

(b) Current error for varying noise scales.

Figure 7: Utility for varying noise scales.

Measurement Scenario P only The maximum noise scale Table 4 for hiding individual appliances is with $\lambda^{\text{P}} = 980$ two orders of magnitudes larger than the minimum noise scale for event-level DP, that already does not yield reasonable utility. Consequently, in our experiments, we consider the reduced setting and limit the upper bound to $\lambda^{\text{P}} = 9.7$, hiding the total power at one time stamp only. To investigate whether the PET achieves reasonable utility, we now compare the error in $\text{SC}^{\text{P only}}_{\text{app}'}$ with the ones in $\text{SC}^{\text{P only}}_{\sigma=0.01}$. Figure 7 shows both. We observe that *all* considered privacy requirements yield a higher voltage error than in $\text{SC}^{\text{P only}}_{\sigma=0.01}$ by far. This means that the usage of a PET yields – even for low privacy requirements – worse utility than the utility resulting from measurement errors. However, if a higher error is acceptable for the DSO, the figures are useful to derive the achievable privacy requirements for a predefined error value, and vice versa. For example, as illustrated in Figure 7 (a), if a voltage or current error that is one order of magnitude higher than the errors in $\text{SC}^{\text{P only}}_{\sigma=0.01}$ is still acceptable, $\lambda^{\text{P}} = \min\{1.0, 4.5\} = 1.0$ is the maximum possible noise scale, that in turn corresponds to, e.g., the privacy requirement "protecting one refrigerator cycle with $\epsilon = 0.14$". To sum up, the results suggest that it is hard to achieve reasonable privacy and utility, because the utility with respect to *all* considered utility metrics for weak privacy requirements is already low.

7 Conclusions

In this paper, we study the utility of differentially private grid monitoring. Specifically, we ask (1) which utility metrics are appropriate and (2) how utility of a PET relates to utility under measurement errors. To this end, we identify candidates for utility metrics for all three steps of grid monitoring. To define reasonable privacy, we use privacy requirements relating to appliance usages given in literature. Based on these definitions, we perform a case study on a real-world grid and realistic measurements. With respect to the first question, we observe that the utility of grid monitoring decreases faster than the sanitization error that is frequently used in related work on differential privacy as utility metric. With respect to the second question, the study indicates that already under weak privacy requirements, the utility is worse than under measurement errors.

Our work has the following implications on future work: First, our study indicates that in future work, it is recommended to not only consider the sanitization error as utility metric, but analysis-specific error as well. Second, our study suggests that it is hard to achieve reasonable utility and privacy in grid monitoring at the same time. Consequently, we investigate two possible approaches for achieving it: The first one is allowing a finer granular specification of privacy requirements by generalizing policy-based notions [26, 32]. The second one is to design load-flow analysis algorithms that take the additional error due to the PET into account, like previously investigated for measurement errors [53].

References

[1] Mohamed S Abdalzaher, Mostafa M Fouda, and Mohamed I Ibrahem. "Data privacy preservation and security in smart metering systems". In: *Energies* 15.19 (2022), p. 7419.

[2] Gergely Ács and Claude Castelluccia. "I have a dream! (differentially private smart metering)". In: *Proceedings of the 13th International Workshop on Information Hiding (IH)*. Springer. 2011, pp. 118–132.

[3] David Bačnar et al. "On security and privacy in smart metering systems". In: *2022 7th International Conference on Smart and Sustainable Technologies (SpliTech)*. IEEE. 2022, pp. 1–6.

[4] Pedro Barbosa et al. "Lightweight privacy for smart metering data by adding noise". In: *Proceedings of the 29th Annual ACM Symposium on Applied Computing (SAC)*. 2014, pp. 531–538.

[5] Daniel Bayer and Marco Pruckner. "A digital twin of a local energy system based on real smart meter data". In: *Energy Informatics* 6.1 (2023), pp. 1–26.

[6] Adi Ben-Israel. "A Newton-Raphson method for the solution of systems of equations". In: *Journal of Mathematical analysis and applications* 15.2 (1966), pp. 243–252.

[7] Bundesamt für Sicherheit in der Informationstechnik. *Technische Richtlinie BSI TR-03109-1 Anforderungen an die Interoperabilität der Kommunikationseinheit eines intelligenten Messsystems*. Tech. rep. Bonn: BSI, 2021.

[8] Yang Cao et al. "Quantifying differential privacy in continuous data release under temporal correlations". In: *IEEE Transactions on Knowledge and Data Engineering (TKDE)* 31.7 (2018), pp. 1281–1295.

[9] Konstantinos Chatzikokolakis et al. "Broadening the scope of differential privacy using metrics". In: *Proceedings of 13th International Symposium on Privacy Enhancing Technologies Symposium (PETS)*. Springer. 2013, pp. 82–102.

[10] Yan Chen et al. "Pegasus: Data-adaptive differentially private stream processing". In: *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security (CCS)*. ACM. 2017, pp. 1375–1388.

[11] Yuwen Chen et al. "A privacy-preserving noise addition data aggregation scheme for smart grid". In: *Energies* 11.11 (2018), p. 2972.

[12] Cynthia Dwork. "Differential privacy: A survey of results". In: *Proceedings of the 5th International Conference on Theory and Applications of Models of Computation (TAMC)*. Springer. 2008, pp. 1–19.

[13] Cynthia Dwork, Aaron Roth, et al. "The Algorithmic Foundations of Differential Privacy". In: *Foundation and Trends ® in Theoretical Computer Science* 9.3-4 (2014), pp. 211–407.

[14] Cynthia Dwork et al. "Differential privacy under continual observation". In: *Proceedings of the forty-second ACM symposium on Theory of Computing (STOC)*. 2010, pp. 715–724.

[15] Günther Eibl and Dominik Engel. "Differential privacy for real smart metering data". In: *Computer Science-Research and Development (CSRD)* 32.1 (2017), pp. 173–182.

[16] Günther Eibl and Dominik Engel. "Influence of data granularity on smart meter privacy". In: *IEEE Transactions on Smart Grid* 6.2 (2014), pp. 930–939.

[17] Günther Eibl et al. "The influence of differential privacy on short term electric load forecasting". In: *Energy Informatics* 1.1 (2018), pp. 93–113.

[18] Fatima Zahra Errounda and Yan Liu. "Continuous location statistics sharing algorithm with local differential privacy". In: *Proceedings of the 2018 IEEE International Conference on Big Data (Big Data)*. IEEE. 2018, pp. 5147–5152.

[19] European Committee for Electrotechnical Standardization. *DIN EN50160: Voltage characteristics of electricity supplied by public distribution systems*. Berlin: Beuth Verlag, 2019.

[20] Jingyao Fan, Qinghua Li, and Guohong Cao. "Privacy disclosure through smart meters: Reactive power based attack and defense". In: *Proceedings of the 47th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*. IEEE. 2017, pp. 13–24.

[21] Ferdinando Fioretto, Terrence W K Mak, and Pascal Van Hentenryck. "Differential Privacy for Power Grid Obfuscation". In: *IEEE Transactions on Smart Grid* 11.2 (2020), pp. 1356–1366.

[22] Ferdinando Fioretto and Pascal Van Hentenryck. "Constrained-based differential privacy: Releasing optimal power flow benchmarks privately". In: *Proceedings of the 15th International Conference on the Integration of Constraint Programming, Artificial Intelligence, and Operations Research (CPAIOR 2018)*. Springer. 2018, pp. 215–231.

[23] Ferdinando Fioretto and Pascal Van Hentenryck. "Differential private stream processing of energy consumption". In: *arXiv preprint arXiv:1808.01949* (2018).

[24] Sebastian Gottwalt et al. "Demand side management—A simulation of household behavior under variable prices". In: *Energy Policy* 39.12 (2011), pp. 8163–8174.

[25] Michael Hay et al. "Principled evaluation of differentially private algorithms using dpbench". In: *Proceedings of the 2016 ACM SIGMOD International Conference on Management of Data*. 2016, pp. 139–154.

[26] Xi He, Ashwin Machanavajjhala, and Bolin Ding. "Blowfish privacy: Tuning privacy-utility trade-offs using policies". In: *Proceedings of the 2014 ACM SIGMOD International Conference on Management of Data*. ACM. 2014, pp. 1447–1458.

[27] Florin Iov and Catalin Ciontea. *Net2DG Deliverable D5.1 – First integrated deployment at Lab and Testbed*. Tech. rep. Net2DG, 2019.

[28] Florin Iov et al. *Net2DG Deliverable D5.3 - Final Consolidated Results*. Tech. rep. Net2DG, 2021.

[29] Peter Kairouz, Sewoong Oh, and Pramod Viswanath. "The composition theorem for differential privacy". In: *Proceedings of the 32nd International conference on Machine Learning (ICML).* Proceedings of Machine Learning Research. 2015, pp. 1376–1385.

[30] Georgios Kellaris et al. "Differentially private event sequences over infinite streams". In: *Proceedings of the VLDB Endowment (PVLDB)* 7.12 (2014), pp. 1155–1166.

[31] Stephan Kessler, Erik Buchmann, and Klemens Böhm. "Deploying and evaluating pufferfish privacy for smart meter data". In: *Proceedings of the IEEE 12th Intl Conf on Ubiquitous Intelligence and Computing and IEEE 12th Intl Conf on Autonomic and Trusted Computing and IEEE 15th Intl Conf on Scalable Computing and Communications and Its Associated Workshops (UIC-ATC-ScalCom)*. IEEE. 2015, pp. 229–238.

[32] Daniel Kifer and Ashwin Machanavajjhala. "Pufferfish: A framework for mathematical privacy definitions". In: *ACM Transactions on Database Systems (TODS)* 39.1 (2014), pp. 1–36.

[33] Franklin Leukam Lako, Paul Lajoie-Mazenc, and Maryline Laurent. "Privacy-preserving publication of time-series data in smart grid". In: 2021 (2021), pp. 1–21.

[34] Haoran Li et al. "Differentially private histogram publication for dynamic datasets: an adaptive sampling approach". In: *Proceedings of the 24th ACM international Conference on Information and Knowledge Management (CIKM)*. ACM. 2015, pp. 1001–1010.

[35] Haoran Li et al. "Distribution Grid Topology and Parameter Estimation Using Deep-Shallow Neural Network with Physical Consistency". In: *IEEE Transactions on Smart Grid* (2023). Early access.

[36] Xiaojing Liao et al. "Towards secure metering data analysis via distributed differential privacy". In: *2014 44th Annual IEEE/IFIP International Conference on Dependable Systems and Networks*. IEEE. 2014, pp. 780–785.

[37] Xiang Liu et al. "Trajectory Privacy Protection on Spatial Streaming Data with Differential Privacy". In: *2018 IEEE Global Communications Conference (GLOBECOM)*. IEEE. 2018, pp. 1–7.

[38] Lingjuan Lyu et al. "Privacy-preserving aggregation of smart metering via transformation and encryption". In: *Proceedings of the 2017 IEEE Trustcom/BigDataSE/ICESS*. IEEE. 2017, pp. 472–479.

[39] Heiko Maaß et al. "Data processing of high-rate low-voltage distribution grid recordings for smart grid monitoring and analysis". In: *EURASIP Journal on Advances in Signal Processing* 2015.1 (2015), pp. 1–21.

[40] Efthymios Manitsas et al. "Distribution system state estimation using an artificial neural network approach for pseudo measurement modeling". In: *IEEE Transactions on Power Systems* 27.4 (2012), pp. 1888–1896.

[41] Andrés Molina-Markham et al. "Private memoirs of a smart meter". In: *Proceedings of the 2nd ACM workshop on embedded sensing systems for energy-efficiency in building (BuildSys)*. ACM, 2010, pp. 61–66.

[42] Karthikeyan Nainar and Florin Iov. "Smart Meter Measurement-Based State Estimation for Monitoring of Low-Voltage Distribution Grids". In: *Energies* 13.20 (2020), p. 5367.

[43] Karthikeyan Nainar et al. "Experimental validation and deployment of observability applications for monitoring of low-voltage distribution grids". In: *Sensors* 21.17 (2021), p. 5770.

[44] Lakshminarayanan Nandakumar et al. “Protecting the grid topology and user consumption patterns during state estimation in smart grids based on data obfuscation”. In: *Energy Informatics* 2.1 (2019), pp. 1–23.

[45] Jimmy J Nielsen et al. “Secure real-time monitoring and management of smart distribution grid using shared cellular networks”. In: *IEEE Wireless Communications* 24.2 (2017), pp. 10–17.

[46] Praviraj PG. *Newton-Raphson Loadflow*. MATLAB Central File Exchange. `https://www.mathworks.com/matlabcentral/fileexchange/21059-newton-raphson-loadflow`. May 2020.

[47] Daniel Vázquez Pombo et al. *Net2DG Deliverable D2.1 – Algorithms for grid estimation and observability applications*. Tech. rep. Net2DG, 2018.

[48] Andreas Reinhardt, Frank Englert, and Delphine Christin. “Averting the privacy risks of smart metering by local data preprocessing”. In: *Pervasive and Mobile Computing* 16 (2015), pp. 171–183.

[49] Sushmita Ruj and Amiya Nayak. “A decentralized security framework for data aggregation and access control in smart grids”. In: *IEEE Transactions on Smart Grid* 4.1 (2013), pp. 196–205.

[50] Henrik Sandberg, György Dán, and Ragnar Thobaben. “Differentially private state estimation in distribution networks with smart meters”. In: *Proceedings of the 54th IEEE Conference on Decision and Control (CDC)*. IEEE. 2015, pp. 4492–4498.

[51] Christine Schäler, Thomas Hütter, and Martin Schäler. “Benchmarking the Utility of w-event Differential Privacy Mechanisms – When Baselines Become Mighty Competitors”. In: *Proceedings of the VLDB Endowment (PVLDB)* 16.8 (2023), pp. 1830–1842.

[52] Christine Schäler et al. “Increased Renewable Hosting Capacity of a Real Low-Voltage Grid Based on Continuous Measurements – Results from an Actual PV Connection Request”. In: *Proceedings of the 17th European Dependable Computing Conference (EDCC) Workshops* (2021), pp. 90–98.

[53] Hans-Peter Schwefel et al. “Using smart meter measurements to manage accuracy of current calculations in lv feeders”. In: *Proceedings of the 2019 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*. IEEE. 2019, pp. 1–7.

[54] Latanya Sweeney. “k-anonymity: A model for protecting privacy”. In: *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 10.05 (2002), pp. 557–570.

[55] Christine Tex, Martin Schäler, and Klemens Böhm. “Swellfish privacy: Supporting time-dependent relevance for continuous differential privacy”. In: *Information Systems* 109 (2022), p. 102079.

[56] Qian Wang et al. “Real-time and spatio-temporal crowd-sourced social network data publishing with differential privacy”. In: *IEEE Transactions on Dependable and Secure Computing (TDSC)* 15.4 (2016), pp. 591–606.