Spectral Differential Privacy: Application to Smart Meter Data
Kendall Parker®, Member, IEEE, Matthew Hale®, Member, IEEE, and Prabir Barooah®, Senior Member, IEEE

Abstract—We present spectral differential privacy (SpDP), a novel form of differential privacy (DP) designed to protect the frequency content of time-series data that come from wide sense stationary (WSS) stochastic processes. This notion is motivated by privacy needs in applications with time-series data over unbounded time, such as smart meters. First, a notion of DP on the space of (discretized) spectral densities is introduced. A Gaussian-like mechanism for SpDP is then presented that provides DP to the spectral density. Next, a novel streaming implementation is developed to enable real-time use of the proposed mechanism. The privacy guarantee provided by SpDP is independent of the time duration over which data are collected or shared. In contrast, time-domain trajectory-level DP (TrDP) will require noise with large variance to provide privacy over an extended time duration. The technique is numerically evaluated using smart meter data from a single home to compare the utility of SpDP to that of time-domain TrDP. The noise added by SpDP is substantially smaller than that added by time-domain TrDP, particularly when privacy over long time horizons is sought by TrDP.

Index Terms—Differential privacy (DP), smart meter, spectrum, trajectory.

I. INTRODUCTION

The Internet of Things (IoT) is a central hub of interconnected, data-driven technologies supporting a variety of critical infrastructure. In the power grid, smart meters connect to the IoT for smarter energy management. They can benefit utility companies in billing, consumption monitoring, and load forecasting. For consumers, they can be a tool for plan for conservation or monitor electricity use [1]. However, smart meter data are often collected at high temporal resolutions and can reveal sensitive information about its source. Such data have been well documented to reveal consumer presence, absence, or lifestyle patterns [2]–[4]. Hence, the nature of this data creates serious privacy concerns.

Differential privacy (DP) is a formal privacy framework that can strike a balance between the need for detailed data and addressing privacy concerns. Using a statistical notion of privacy, DP adds carefully calibrated noise to sensitive data (or functions thereof) to protect it [5]. The value of DP to smart metering and related grid applications has been recognized by many researchers. For instance, DP protects the location of smart homes from traffic analysis attacks in [6]. A data aggregation scheme that ensures DP in the presence of general measurement report failures was explored in [7].

While DP originated in the context of databases, it was extended to data in the form of trajectories or signals, termed trajectory-level differential privacy (TrDP) [8]. A challenge in trajectory-level DP (TrDP) is adequately protecting sensitive events as the data length grows. In many applications, such as analytics with smart meter data, there may be instances when an upper bound on the required time duration for an analytic is unknown, and thus, a need may arise to protect many sensitive events over time. Privacy noise grows with the duration and magnitude of sensitive events to privatize, and thus, the protection of many events over arbitrarily long time horizons can require arbitrarily large noise. Indeed, the noise scale in TrDP may grow to the point that the privatized trajectory is useless. This pitfall has been addressed for classical DP, where the author weakens privacy guarantees in the distant past [9]. The performance of the mechanism is dependent on the discount factor, which may be difficult to determine for varying applications.

In this work, we propose a new notion of DP targeting this weakness of TrDP, by going to the frequency domain. In our approach, which we term spectral DP (SpDP), a signal’s power spectral density (PSD) is treated as sensitive information, and a new definition and approach to privacy are defined to protect the spectral representation of sensitive events, rather than the time-domain representation. A privatized PSD in this setting is the result of a privacy mechanism applied directly to a PSD. By computing the PSD offline, this mechanism can be implemented offline. Our motivation is the privacy of smart meter data, and a meter must transmit time-domain data in streaming (as opposed to batch) fashion. We, therefore, also provide a streaming implementation to compute and share a signal in real-time so that the PSD of the transmitted signal is the same as the privatized PSD that the SpDP mechanism generates.

Motivation comes from the observation that the frequency content of smart grid signals is highly sensitive because of the capability to exploit usage patterns and routines in consumers [10]–[12]. Through analysis of frequency content, the energy distribution of smart meter data per unit time can be classified into frequency bins, thus identifying prominent times...
of activity. To the best of our knowledge, a frequency-domain representation of DP has only been seen in [13], where authors study local DP with singular spectrum analysis. Though results show that the utility of private data is better retained than time-domain DP, the computation time for SSA begs to question how difficult implementation will be without additional hardware at the meter.

Most related works utilize DP to protect aggregate statistics for a collection of $n$ users. Eibl and Engel [14] used infinite divisibility of the Laplace distribution to have each consumer add gamma-distributed noise, leading to DP for aggregated information but not for individuals. Likewise, the privacy-preserving aggregation system in [15], which uses the fog-computing architecture, provides privacy to individual users using additively homomorphic encryption. Only the aggregate statistics have $\epsilon - \delta$ DP guarantees. A similar privacy preservation scheme using distributed DP is used in [16], where additive homomorphic encryption is added to protect individual statistics and to ensure each participant’s privacy. In contrast, our work will provide DP guarantees to all users individually, at all time instances without concern for encryption key management. The mechanism will be developed and evaluated for a single consumer with extension to a neighborhood of users discussed in Section V.

The main advantage of our proposed privacy framework is that the noise in the resulting time-domain data that is shared, i.e., the output of the streaming implementation, is bounded irrespective of data length. In contrast, noise in the privatized data shared by TrDP grows without bound as the time duration increased. Data privatized with spectral DP (SpDP), therefore, has more utility for downstream analytics when long time intervals are involved.

Two additional contributions of this article are: 1) a novel nonindependent identically distributed additive Gaussian mechanism and 2) a data-based calibration method for choosing the adjacency parameters for SpDP and time-domain TrDP. A related work that does not provide DP guarantees but privatizes smart meter data with correlated noise is [17]. The adjacency parameter is a design choice, and few guidelines exist in the literature on how to choose its numerical value.

The remainder of this article is organized as follows. Section II summarizes TrDP and its weaknesses, defines the required mathematical preliminaries for SpDP, and explains the problem solved by SpDP. Following this, Section III defines all elements of SpDP along with streaming implementation. Section IV discusses numerical results, followed by conclusions and future work in Section V.

II. PRELIMINARIES

The symbols $\mathbb{R}, \mathbb{R}_+, \mathbb{Z},$ and $\mathbb{Z}_+$ denote the sets of real, nonnegative real, integers, and nonnegative integers, respectively. A discrete sequence $x$ is a function $x: \mathbb{Z} \rightarrow \mathbb{R}^n$, i.e., $x_k \in \mathbb{R}^n$ for every $k \in \mathbb{Z}_+$. The prototypical signal of interest is the power demand of a consumer. For that signal, the index $k$ in $x_k$ is a discrete time index, corresponding to the time ticks when data are sent.

The $\ell_2$ norm of a sequence $x$ is $\|x\|_{\ell_2} = (\sum_{k=0}^{\infty} \|x_k\|_2^2)^{1/2}$, and the symbol $\ell_2$ denotes the set of all sequences $x$ that have finite $\ell_2$ norm. Furthermore, the notation $\ell_2^n$ denotes the set of all sequences $x: \mathbb{N} \rightarrow \mathbb{R}^n$ such that every finite truncation of the sequence has a finite $\ell_2$ norm. In other words, $x \in \ell_2^n$ if $\|x_k\|_2 < \infty$ for all $k$. Often we will consider a finite truncation of a sequence, such as $x_{1:n} := (x_1, \ldots, x_n)$. For consistency, we will use $\|x_{1:n}\|_{\ell_2}$ to denote the 2-norm of the vector of truncated values, i.e., $\|x_{1:n}\|_{\ell_2} = (\sum_{i=1}^{n} \|x_i\|_2^2)^{1/2}$.

A. Summary of TrDP

DP was designed in part to prevent differential attacks. Even with a secure aggregation scheme, say at the power utility, an adversary can acquire the aggregation of $n$ users and that of $n-1$ users, and compromise the privacy of the differential user [7]. Hence, DP masks the differences between adjacent pieces of data by ensuring that they produce approximately indistinguishable outputs when a privacy mechanism is applied to them. That is, given some private output sequence, it should be unlikely for its recipient to make meaningful distinctions between input sequences that could have produced it. We refer to the input $x$ as the sensitive data and the output—a realization of the DP mechanism $M(x)$—as the privatized data. In this work, a system will directly add noise to its outputs before sharing them. This is the input perturbation approach to DP, and has the advantage of masking sensitive data before it is shared.

The essential notions from (TrDP) with the Gaussian mechanism are stated in the following proposition; see [8] for a thorough exposition. We first define adjacency using an adjacency parameter $B > 0$. This design parameter is selected based on the size of the difference in demand signals that needs to be hidden so that differential attacks of adjacent data sets are unlikely.

**Proposition 1:** Fix $B > 0$.

1) Two sequences $x, y \in \ell_2^n$ are said to be adjacent if $\|x - y\|_{\ell_2} \leq B$.

2) A mechanism $M$ is $(\epsilon, \delta)$-differentially private with respect to this adjacency relationship if

$$\mathbb{P}[M(x) \in A] \leq e^\epsilon \mathbb{P}[M(y) \in A] + \delta$$

(1)

for all measurable $A \subseteq \ell_2^n$ and all adjacent $x$ and $y$.

3) The Gaussian mechanism $M(x) = x + w$ is $(\epsilon, \delta)$ differentially private with $w(k) \sim N(0, \sigma^2 I)$, if

$$\sigma \geq \frac{B}{2\epsilon} \left( Q^{-1}(\delta) + \sqrt{Q^{-1}(\delta)^2 + 2\epsilon} \right)$$

(2)

and $Q(a) = (1/\sqrt{2\pi}) \int_{-\infty}^{a} \exp(-u^2/2) du$ is the Gaussian tail integral.

The interpretation is as follows: the parameter $\epsilon$ controls information leakage about sensitive data. Smaller values of $\epsilon$ imply less leakage and hence, stronger privacy. The parameter $\delta$ can be interpreted as the probability that $\epsilon$-DP fails. In the literature, typical values are $\epsilon \in (0, \log 3)$ and $\delta \in [0, 0.5]$ [5], [8].

To apply the notion of TrDP to time-domain demand data sequences, consider two power demand trajectories (in kW)
of length $K$ from the same consumer$^1$: $d_{1:K} := \{d_k\}_{k=1}^N$ and $d_{K+1:2K} := \{d_k\}_{k=K+1}^{2K}$. As $K$ increases, the distance $D(K) := \|d_{1:K} - d_{K+1:2K}\|_2$ increases without bound. Fig. 1 is evidence in support of this claim for electrical power demand data collected at a 5-min interval from a single home.

Since the time series used to find the distance $D(K)$ are from the same consumer over two consecutive time intervals, a reasonable notion of adjacency should qualify them as adjacent. Fig. 1(b) shows that the time-domain distance between these trajectories is large and grows without bound over time. Thus, in a TrDP framework, a large adjacency parameter $B$ is required to qualify the time series as adjacent. For given privacy parameters $\epsilon$ and $\delta$, a large $B$ necessitates the use of high-variance noise for privacy (cf. Proposition 1). As privacy is demanded for data of even larger duration $K$, the value of $B(K)$ must also be chosen larger. Since the variance of privacy noise is proportional to $B^2$, the accuracy of any analytics with the privatized data degrades as $K$ increases. Consequently, the utility of privatized data decreases commensurately as $K$ increases. To provide $(\epsilon, \delta)$-DP for a fixed $\epsilon$ and $\delta$ independent of the time duration $K$, the adjacency parameter—and thus, the noise added—must be infinitely large. As a result, the utility of privatized data for analytics becomes zero. If $B$ is held fixed while the length of trajectories increases without bound, then only a small number of events and/or events of short duration (relative to the length of the trajectory) can be protected from differential attacks through time-domain TrDP.

Ny and Pappas [8] addressed adjacency parameter selection for arbitrary finite-horizon (and finite-dimensional) settings. The subtlety there is that $B$ must be fixed based on the events that need protection; as the duration of interest increases, a new adjacency parameter must be fixed to protect events in the longer trajectory.

This discussion shows the difficulty of using standard TrDP to provide privacy to smart meter data, or to any smart grid applications requiring time-domain data: two data streams generated by the same behavior and same consumer will have small differences that do not vanish as time increases. Thus, the distance between these two time series will grow without bound as the time duration of the series grows. An exception is time series $x$ with asymptotically vanishing $x_k$’s so that $x \in \ell_2$, but such time series are not relevant to smart grid applications.

### B. Mathematical Preliminaries on PSDs

Here, we summarize mathematical preliminaries needed to formally define the problem that is the subject of this work. Throughout this work, we assume sensitive time-domain data are a wide sense stationary (WSS) stochastic process to make use of the well-established spectral theory for stationary time series. A pure WSS model may not be appropriate for many types of smart grid data due to seasonal and other quasiperiodic variations. However, such data can be modeled as a deterministic time-varying mean plus a WSS process.

The PSD of a zero mean WSS process $x$ is the Fourier transform of its autocorrelation, or, a limit of the average of the square of the Fourier transform of the truncated data

$$\Phi_{xx}(\omega) := \sum_{m=-\infty}^{\infty} R_{xx}[m]e^{-j m \omega}$$

$$= \lim_{T \to \infty} \frac{1}{T} \mathbb{E} \left[ \left| \sum_{k=1}^{T} x_k e^{-jk \omega} \right|^2 \right]$$

where $\omega$ is the (continuous) frequency variable. The equality between (3) and (4) is the Wiener–Khinchin theorem [19]. The process $x$ is assumed to be zero mean throughout this article for notational convenience. Otherwise, every definition that involves an expectation must subtract the mean.

We will work with a sampled version of the function $\Phi_{xx}(\omega)$. We choose an integer $N$ and sample the PSD at $2N$ frequency points, where the $n$th frequency $\omega_n$ is

$$\omega_n = \frac{2 \pi F_s}{2N} n \quad \text{(rad/sample)}, \quad n = 0, \ldots, 2N - 1$$

where $F_s$ is the sampling frequency of the time-domain data in samples per unit time. Denote the part that goes up to the Nyquist frequency as $\Phi_N$

$$\Phi_N = [\Phi_{xx}(e^{j0\omega}), \Phi_{xx}(e^{j\omega}), \ldots, \Phi_{xx}(e^{j(2N-1)\omega})]^T \in \mathbb{R}_{+}^{N+1}.$$  

The length of the sampled PSD is a design variable selected based on the frequencies one wants to resolve in the PSD. The superscript $N$ will be often omitted to reduce clutter, and the sampled PSD will be referred to as $\Phi$.

There are several reasons why it is meaningful to consider the sampled PSD as the sensitive data to be privatized. First, as long as $N$ is large enough to resolve the frequencies of interest, the sampled version will contain most of the information from

---

$^1$We use $x$ to denote an arbitrary time series or stochastic process. Here, we consider demand signals specifically, which is indicated by a change in notation to $d$ for the signal.
the function $\Phi_{xx}(\omega)$. The second reason considers computation needs. All numerical algorithms to estimate the PSD do not estimate the continuous function $\Phi_{xx}(\omega)$. Rather, they estimate the vector $\phi$ for some $N$. The expectation of this unbiased estimate approaches the true spectral density of the process $x$ for large $K$ [20]. In addition, in order to have good estimation accuracy $N$ must be far less than $K$. Moreover since the streaming implementation of SpDP developed in Section III and the analysis of privatized data require the sampled PSD, privacy guarantees provided directly on $\phi$ are useful.

C. Problem Definition

The privacy goal we consider in this work is preventing differential attacks of a single consumer’s trajectories to prevent adversaries from exploiting repetitive behaviors to uncover sensitive information. Since the frequency content of signals is sensitive information [10]–[12], we focus on privatizing the PSD of the power demand rather than the time-domain data itself. In light of the weakness of time-domain DP described previously, the key advantage of privacy in the Fourier domain is that the PSD is defined over a frequency interval that is independent of the length of time involved. The highest frequency is the Nyquist frequency, which is half of the sampling frequency of the data. Thus, the noise needed to privatize intuitively adjacent PSDs is not dependent on (and does not grow with) the time interval.

To illustrate this advantage, Fig. 2 shows three different PSD estimates of a consumer’s power demand. These estimates are computed from varying lengths of time-domain data, but the frequency range only goes up to the Nyquist frequency, which is $0.5 \times (1/300) = 1.67 \times 10^{-3}$ Hz (≈ 6 hour$^{-1}$) in this case since the data are sampled every 5 min. In addition, one can see from the figure that PSD estimates obtained from various data traces of the same consumer do not differ much. Later, it will be shown that this feature allows a uniform $B$ (irrespective of time interval involved) to quantify adjacency for DP when the PSD is considered the sensitive data to be privatized.

With this, the following problem is the subject of this work.

Problem 1: Given a time-domain signal $x = \{x_k\}_{k \in \mathbb{N}}$ and its sampled PSD estimate $\hat{\phi}$, do the following.

1) Design an $(\epsilon, \delta)$-DP mechanism $\mathcal{M}$ to prevent differential attacks of frequency-domain sensitive data $\phi$.
2) Develop a streaming implementation of $\mathcal{M}$ that generates samples $\tilde{x}_k$ in real-time so that the PSD of $\tilde{x}$ is $\phi$.

The time-domain data that the streaming implementation produces must still be useful in downstream analytics. Problem 1.1 is solved with correlated Gaussian and the SpDP mechanism in Section III-B. Problem 1.2 is solved in Section III-C.

D. Correlated Gaussian Mechanism for TrDP

The Gaussian mechanism mentioned in Proposition 1 uses independent identically distributed noise, which is standard in the literature on TrDP [5]. We now present an extension to the nonindependent identically distributed case, which will be useful in our sequel.

Proposition 2: Fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and let $d$ be a sensitive signal and let privacy parameters $\epsilon > 0$, $\delta \in (0, 0.5)$ be given. Consider the correlated Gaussian mechanism $\mathcal{M} : \mathbb{E}^n \times \Omega \to \mathbb{E}^n$ defined by $\mathcal{M}(d) = d + W$ with $W \sim \mathcal{N}(0, \Sigma)$. This mechanism is $(\epsilon, \delta)$-differentially private if $\lambda_{\min}(\Sigma)$, the minimum eigenvalue of the covariance matrix $\Sigma$, satisfies $\lambda_{\min}(\Sigma) \geq 2$, where

$$\lambda := \frac{B^2}{(\sqrt{A} + \sqrt{A^2 + 2\epsilon})^2}, \quad A = Q^{-1}(\delta).$$

Due to its technical nature, proof of the proposition is in the Appendix. With correlated noise, the probability of having a large difference between $\phi$ and its private counterpart is lessened compared to a more common independent identically distributed mechanism. With this, the output time-domain signal from streaming implementation will maintain better correlation with the time-domain sensitive data when correlated noise is used. This will be further discussed in Section III-C.

III. SPECTRAL DIFFERENTIAL PRIVACY

This section develops SpDP, including an adjacency relation between the sampled power spectra, a formal mechanism, and streaming implementation to generate a private time series from the differentially private PSD. We use the input perturbation approach to privacy, which provides privacy to a database with a single entry.

A. Definitions

Indistinguishability of signals in SpDP is made parallel to TrDP using an adjacency relation on the PSD estimates rather than the time series. Subscripts $i$ and $j$ indicate PSD estimates of two different trajectories from a single consumer.

Definition 1: Fix $B > 0$. The adjacency relation $\text{Adj}_B$ is defined $\forall \phi_i, \phi_j \in \mathbb{R}^N_+$ for $p \geq 1$ as

$$\text{Adj}_B(\phi_i, \phi_j) = \begin{cases} 1, & \|\phi_i - \phi_j\|_p \leq B \\ 0, & \text{otherwise.} \end{cases}$$

Following the privacy goal, each user applies this definition to mask sensitive events in their data individually. Therefore, the user specifies the boundedness defining $\text{Adj}_B$. Namely, a user selects an adjacency parameter $B > 0$, and their true PSD is made approximately indistinguishable from all other PSDs within distance $B$ by the privacy mechanism.
The benefits of an adjacency relation on the PSDs are best illustrated in comparison with TrDP. Fig. 3 shows the distance between PSDs of demand as a function of duration of data used to estimate the PSD. The same figure also shows the distance between demand data (in the time domain), which increases monotonically as the data duration increases. The 2-norm was used in both cases to compute this distance. For PSD estimates, the distance does not grow as one considers longer durations of data. In fact, the distance between PSDs appears to settle down to a constant. With a fixed length PSD, the adjacency parameter needed to do so will not increase with time-series length. This results in constant distances between PSDs and bounded noise added to the PSD for DP.

A mechanism $\mathcal{M}(\cdot)$ in SpDP is a randomized mapping with domain and co-domain $\mathbb{R}^N$. We define

$$\tilde{\phi} = \mathcal{M}(\phi)$$

so that $\mathcal{M}$ provides $(\epsilon, \delta)$-DP to frequency-domain sensitive data $\phi$ for given $\epsilon$ and $\delta$. We formally define SpDP below with probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and we use the Borel $\sigma$-algebra over $\mathbb{R}^N$, denoted $\mathcal{B}_N$.

**Definition 2:** A mechanism $\mathcal{M} : \mathbb{R}^N \times \Omega \to \mathbb{R}^N$ preserves $(\epsilon, \delta)$-DP if, for all adjacent PSD signals $\phi_i, \phi_j \in \mathbb{R}^N$

$$\mathbb{P}[\mathcal{M}(\phi_i) \in S] \leq \epsilon \mathbb{P}[\mathcal{M}(\phi_j) \in S] + \delta \quad \forall S \in \mathcal{B}_N.$$  

SpDP is made parallel to TrDP in that the privacy mechanism is applied to the PSD in a parallel manner so that: 1) each realization of the private PSD is itself a valid PSD and 2) the aforementioned definition is satisfied. This is consistent with the standard interpretation of DP applied to time-domain data; with SpDP, an adversary will be unlikely to make meaningful distinctions between signals’ frequency content. The claim is that SpDP provides SpDP to all PSDs (and likewise provides its attendant immunity to postprocessing and robustness to side information) with improved utility over TrDP because it requires less noise.

**B. SpDP Mechanism**

Algorithm 1 describes a mechanism for SpDP. It uses concepts from positive dynamical systems. A positive dynamical system is one that, if the initial condition and input are nonnegative, has nonnegative states and outputs [21].

![Fig. 3. Distances between pairs of time series and sampled PSD estimates. Data: Pecan Street Project [18].](image)

**Algorithm 1:** Mechanism $\mathcal{M}_{\text{SpDP}}$ to Provide $(\epsilon, \delta)$-Spectral Differential Privacy

**Input:** Sensitive PSD of demand $\phi$, noise covariance matrix $\Sigma$, positive filter $P(z)$

**Output:** A differentially private PSD $\tilde{\phi}$

/* Apply Gaussian mechanism */
1 Set $\phi' \leftarrow \phi + \eta$, with $\eta \sim \mathcal{N}(0, \Sigma)$.
/* Make values non-negative */
2 Set $\phi' \leftarrow (\phi')^+$ for all $\omega \in [0, \pi]$
/* Apply $P(z)$ non-causally */
3 Set $\tilde{\phi} \leftarrow P(z)[\phi']$

**Theorem 1:** For a given adjacency parameter $B > 0$ and privacy parameters $\epsilon > 0$ and $\delta \in (0, 0.5)$, Algorithm 1 provides $(\epsilon, \delta)$-DP to the PSD $\phi$ and produces a valid PSD $\phi$, if $\lambda_{\min}(\Sigma) \geq \frac{1}{B}$ where $\lambda_{\min}$ is defined in (7).

**Proof:** Consider the intermediate array $\phi' := \phi + \eta$, where $\eta$ is a vector of $N$ samples from a Gaussian distribution with mean 0 and covariance $\Sigma$. Because of the eigenvalue condition in the hypothesis, $(\epsilon, \delta)$-DP of $\phi'$ follows immediately from Proposition 2. However, the sum $\phi + \eta$ will be negative at some indices (frequencies) with nonzero probability due to the zero mean nature of $\eta$, and hence, the sum may not be a valid sampled PSD. Applying postprocessing by thresholding negative values to zero makes the signal nonnegative at all frequencies, which ensures the output is a valid PSD. Filtering with a positive dynamical system makes the result smoother while maintaining nonnegativity, which ensures that the output of the mechanism is a valid PSD. That $\tilde{\phi}$ is $(\epsilon, \delta)$-differentially private follows from immunity to postprocessing of DP: $\tilde{\phi}$ is $(\epsilon, \delta)$-differentially private, and subsequent operations are merely postprocessing, which means their outputs have the same level of privacy [5].

Fig. 4 demonstrates a numerical example of $\mathcal{M}_{\text{SpDP}}$ applied with Algorithm 1. The noise was generated using

![Fig. 4. Numerical example of an SpDP mechanism: a sampled PSD estimate of a consumer’s electric power demand is the sensitive data, while the privatized PSD is the output of the SpDP mechanism of Algorithm 1. The sensitive PSD is obtained from a consumer demand data in Pecan Street Project [18].](image)
the procedure described in Section II-D. The figure shows the frequency-domain sensitive data \( \phi \), which is the sampled PSD of a consumer’s demand and the privatized PSD data \( \tilde{\phi} \) obtained by applying the mechanism \( \mathcal{M}_{\text{SpDP}} \). These PSDs were estimated using one month of time-domain data. Values of the parameters used in this numerical example are provided in Table I. The positive dynamical system used in this example is a discrete time low-pass filter with cutoff frequency \( \omega_N \)

\[
P(z) = K_p \frac{\omega_N z^{-1}}{1 + (\omega_N - 1)z^{-1}}. \quad (10)
\]

C. Streaming Implementation

Estimates of a consumer’s PSD from power demand data can be determined from past usage, and hence, privatizing the PSD occurs offline. However, meters must transmit time-domain data, not PSDs, so a streaming implementation is necessary for a time-domain application of SpDP. Thus, a streaming implementation generates values in time whose PSD is the same as the private PSD that the SpDP mechanism generates. For streaming implementation, we assume the frequency-domain sensitive data \( \phi \), the privatized PSD \( \tilde{\phi} \), and the time-domain sensitive data \( x \) are available to the smart meter that will perform the streaming implementation. While \( \phi \) and \( \tilde{\phi} \) are known \textit{a priori}, \( x \) is available only in real-time. Also, the development of a streaming implementation requires smart meters that are tamper resistant and trusted with the ability to perform filtering. Fig. 5 illustrates the streaming implementation described in the following algorithm.

The next proposition shows the steps above are a valid streaming implementation of the mechanism described in Section III-B.

**Proposition 3:** If \( \gamma[n] > 0 \) for \( n = 0, \ldots, N \), then Algorithm 2 is feasible, and the sampled PSD of the released time-domain data \( \tilde{x} \) is \( \tilde{\phi} \), and the cross-correlation function between \( x \) and \( \tilde{x} \) is \( \gamma \) in \( \mathbb{R}^{N+1} \), where

\[
\gamma[n] := \tilde{\phi}[n] - |F(e^{j\omega_0})|^2 \phi[n], n = 0, \ldots, N. \quad (11)
\]

| Parameter | Description | Value | Units |
|-----------|-------------|-------|-------|
| \( N \) | Number of NPFT Points | 256 | none |
| \( \epsilon \) | Privacy Level | log(2) | none |
| \( \delta \) | Privacy Failure Parameter | 0.01 | none |
| \( \beta \) | Degree of Noise Correlation | 0.5 | none |
| \( \omega_N \) | Positive Filter Cutoff Frequency | 0.39 | Hz |
| \( \omega_F \) | Cutoff Frequency of \( F(z) \) | 2 \( \times \) 10\(^{-4} \) | Hz |
| \( K_F \) | Gain of \( F(z) \) | 0.8 | None |
| \( B_{\text{SpDP}} \) | SpDP Adjacency Parameter | 0.12 | kW\(^2\)-Hour |
| \( B_{\text{SpDP}}^{\text{4H}} \) | 4-Hour Adjacency Parameter | 0.2 | kW |
| \( B_{\text{SpDP}}^{\text{1W}} \) | 1-Week Adjacency Parameter | 39 | kW |

**Algorithm 2: Streaming Implementation of SpDP**

**Input:** Time-domain sensitive signal \( x = \{x_k\}_{k \in \mathbb{N}} \), its PSD \( \phi \), its privatized form \( \tilde{\phi} \), and the filter \( F(z) \)

**Output:** A time-domain signal \( \tilde{x} = \{\tilde{x}_k\}_{k \in \mathbb{N}} \) related to \( x \) whose PSD is \( \tilde{\phi} \)

\[
/* \text{Offline:} */
\]

1. Compute \( \gamma \in \mathbb{R}^{N+1} \), where

\[
\gamma[n] := \tilde{\phi}[n] - |F(e^{j\omega_0})|^2 \phi[n], n = 0, \ldots, N. \quad (11)
\]

2. Determine \( H(z) \) such that \( |H(e^{j\omega_0})|^2 = \gamma[n], n = 0, \ldots, N \) where \( \omega_0 \) is defined in (5).

\[
/* \text{Online:} */
\]

3. for \( k = 0, 1, \ldots \)

4. Generate white Gaussian noise \( w_k \) with 0 mean and unit variance.

5. for \( k = 0, 1, \ldots \)

6. Compute \( \tilde{H}(z) \)

7. \( \tilde{x}_k \leftarrow \tilde{x}_k + c_k \)

8. Release \( \tilde{x}_k \)

9. end for

then them can be used to determine \( H \), which will by design satisfy the requirement in the algorithm, that \( |H(e^{j\omega_0})|^2 = \gamma[n], n = 0, \ldots, N \). This proves feasibility. To show that the released time-domain data \( \tilde{x} \) have the desired PSD

\[
\text{PSD}(\tilde{x}) = \text{PSD}(x^F + c) = \text{PSD}(x^F) + \text{PSD}(c)
\]

where the second equality follows from the fact that \( c \) is independent of \( x^F \). Denoting by \( \Phi_{xc}(\omega) \) the PSD of a stochastic process \( x \), along with standard properties relating PSDs of inputs and outputs of filters, and the fact that the PSD of
a independent identically distributed unit variance Gaussian process is 1 at all frequencies [19], we get from the above that

\[
\Phi_{\tilde{x}x}(\omega) = |F(e^{j\omega})|^2 \Phi_{\tilde{x}x}(\omega) + |H(e^{j\omega})|^2
\]

\[
\implies \Phi_{\tilde{x}x}(\omega_n) = |F(e^{j\omega_n})|^2 \Phi_{\tilde{x}x}(\omega_n) + |H(e^{j\omega_n})|^2
\]

\[
n = 0, \ldots, N
\]

\[
\Phi_{\tilde{x}x}(\omega_n) + \gamma[n], \quad n = 0, \ldots, N
\]

\[
\Phi[n], \quad n = 0, \ldots, N
\]

where the last equality follows by substituting (11) in the previous equation. This shows that the PSD of the process \(\tilde{x}\) at the frequencies \(\omega_n\) has the desired value, and so the algorithm is a streaming implementation of the SpDP mechanism. The statement about the cross correlation follows from standard results on the cross-correlations between inputs and outputs of a linear filter (see [19, Ch. 9]) upon recognizing that \(c\) is independent of \(x\).

To satisfy the positivity requirement of \(\gamma[n]\), one has to choose the reduction filter \(F\) appropriately. A poorly designed filter can make \(\gamma < 0\) at some \(n\), in which case its spectral factorization into \(H\) is not theoretically possible. In that case, Algorithm 2 is not implementable. In practice, \(F\) is a low-pass filter, with adaptable cutoff frequency to ensure positivity of \(\gamma\). The filter design can be performed after the mechanism has been applied and the private PSD \(\Phi\) is available, and so the positivity condition can always be maintained. Since the filter \(F\) is used in streaming implementation, it does not affect the privacy guarantees on the frequency-domain data \(\Phi\). The positivity requirement is also maintained through the use of the correlated Gaussian mechanism. If there are large differences between the filtered version of \(\Phi\) and \(\Phi\), the spectral factorization of \(\gamma\) would result in time-domain noise with large standard deviation. This is due to the resulting large magnitude of \(\gamma\). Instead, with correlated noise privatizing \(\Phi\), there is less probability of large values of \(\gamma\) and thus, the resulting noise from streaming implementation is of smaller standard deviation.

Apart from feasibility of streaming implementation, the design of \(F\) along with the differentially private noise \(\eta\) determines the degree of correlation between released time-domain data \(\tilde{x}\) and time-domain sensitive data \(x\). Fig. 5(b) illustrates this: if the filter \(F\) has low gain at some frequency, the gap \(\gamma\) between the PSDs of sensitive demand (i.e., \(\Phi\)) and filtered demand (i.e., \(|F(e^{j\omega_n})|^2 \Phi\)) will be large at that frequency. Recall this gap is filled by the colored noise \(c\), since the PSD of \(c\) is \(|H(e^{j\omega_n})|^2 = \gamma[n]\). Thus, the released data will have a large noise variance \(c\) compared to time-domain sensitive data \(x\) at that frequency. Depending on the level of noise in the particular realization \(\Phi\) of the mechanism \(M\), the reduction filter may have to be designed with extremely low gain at certain frequencies. In that case, the time-domain privatized data \(\tilde{x}\) produced by streaming implementation will have low correlation with the time-domain sensitive data \(x\). Correspondingly, downstream analytics with the released time-domain data \(\tilde{x}\) will then be less accurate than those done with the sensitive data \(x\). The loss of accuracy will increase as the gain of \(F\) is reduced. In contrast, as the gain of \(F\) approaches 1, the loss of accuracy approaches 0 but streaming implementation may be infeasible.

### IV. Numerical Evaluation

We evaluate the proposed paradigm on consumer demand from a single home in Pecan Street [18]. The privacy goal considered in our numerical evaluation is to protect demand data from a single home with a 5-min sampling period. By first applying the SpDP mechanism \(\mathcal{M}_{\text{SpDP}}\), and then performing the streaming implementation, we conduct a full application of SpDP. The result is that the privatized data \(\tilde{d}\) is streamed by the smart meter in real-time based on time-domain sensitive demand data \(d\) immediately as the sensitive data are measured. The Lisa technology package data analysis (LTPDA) was used to generate correlated noise \(c\) noise from private PSD \(\Phi\) [23]. Numerical values of the parameters used in the study are in Table I.

#### A. Choice of \(B\)

The adjacency parameter \(B\) depends on the choice of norm used to define distances between trajectories. Beyond this, \(B\) is a design choice and few guidelines exist on how to choose it. Since time-domain TrDP and SpDP use vastly different norms to define distances, the choice of \(B\) must differ in these two distinct privacy paradigms.

Recall that the privacy goal is to protect demand data from a single home. For SpDP, we choose \(B\) as

\[
B_{\text{SpDP}} = \max_{i,j \in \mathcal{N}} \|\Phi_i - \Phi_j\|_2
\]

where \(\Phi_i\) is the estimate of the sampled PSD computed from the \(i\)th time-domain data set \(d_i := d_{k_i}, d_{k_i+1}, \ldots, d_{k_i+N_i}\), and \(\mathcal{N}\) is a set of such time-domain data, with each data set potentially of distinct duration \(N_i\). Each numerical estimate of the PSD from a particular time-domain data set can be thought of a distinct PSD itself, which is a frequency-domain characterization of the potential behavior of a consumer.

The true (unknown) PSD of the consumer does not vary with time duration. Thus, to determine \(B_{\text{SpDP}}\), we want the time-domain data sets that produce similar PSD estimates, with differences among estimates attributable to estimation errors. It turned out that PSDs from 7 to 12 week durations produced the most similar PSD estimates and the result from (12) was 0.12 kW²×h and this value was chosen as \(B_{\text{SpDP}}\). This adjacency parameter can be used in SpDP irrespective of time duration and protects events up to a frequency of 6 hr⁻¹, which is the Nyquist frequency corresponding to the sampling period of 5 min.

For time-domain TrDP with the same privacy goal, an appropriate choice of \(B\) will be

\[
B_{\text{TrDP}}(K) = \max_{\ell,m} \|d^{(\ell)}_{1:K} - d^{(m)}_{1:K}\|_2
\]

which depends on the time interval \(K\). Superscripts \(\ell\) and \(m\) indicate two nonoverlapping power demand trajectories of length \(K\) from a single consumer. As discussed in Section II-A, calibrating privacy in this way for TrDP requires an infinitely large adjacency parameter for data over an unbounded time interval, and so the choice of \(B\) necessarily depends on the time duration \(K\) over which privacy is to be provided.
Fig. 6. 24-h snapshot of time-domain sensitive data and privatized results from (a) SpDP, which are the same no matter the time duration considered, (b) TrDP, providing privacy for 4 h, and (c) TrDP, providing privacy for one week. Data: Pecan Street Project [18].

For illustration, we consider two scenarios with two distinct \( K \)'s, denoted \( K^{(1)} \) and \( K^{(2)} \). The first is \( K^{(1)} = 576 \) kW, corresponding to 4 h of data. This yields \( B^{(1)}_{\text{TrDP}} = 0.2 \) kW. The second is for \( K^{(2)} = 2016 \), corresponding to one week of data, which yields \( B^{(2)}_{\text{TrDP}} = 39 \) kW.

The correlation coefficient between the time-domain sensitive data and the released time-domain data is taken as a measure of utility. A higher correlation coefficient corresponds to more useful data for downstream analytics. When the coefficient is 1 that means the time-domain sensitive data itself are released and there is no loss in accuracy of analytics, though it would also mean no privacy is afforded to the consumer.

### B. Results

The results of the full implementation of SpDP and implementation of TrDP for the two scenarios, applied to the demand data from a house in Pecan Street, are in Fig. 6. Fig. 6(a) shows time-domain sensitive data \( d \) and released data \( \hat{d} \) after applying the SpDP mechanism (Algorithm 1) and streaming implementation (Algorithm 2). Fig. 6(b) and (c) shows the same sensitive data along with data privatized using TrDP and the Gaussian mechanism in Proposition 1. The additional parameters used in SpDP are in Table I.

TrDP shows superior performance over SpDP in the first scenario (TrDP designed for 4 h) in terms of utility, but SpDP outperforms TrDP in utility for the second scenario (TrDP designed for a week). In fact, for any time duration higher than 4 h SpDP outperformed TrDP. It should be emphasized that SpDP provides DP for all time durations. Table II further highlights the utility differences of SpDP and TrDP results. SpDP and TrDP designed for 4 h preserve the mean of the sensitive time series, and their standard deviations are within 15% of the sensitive time series. Contrarily, TrDP designed for one week has a standard deviation over 100 times larger and a mean almost two times larger than the sensitive time series.

Table III shows the standard deviation of added time-domain noise and correlation coefficients between the time-domain sensitive data and released data for each result in Fig. 6. From these results, TrDP only outperforms SpDP at extremely short time durations, illustrated by the higher correlation coefficient. The signal-to-noise ratio (SNR) is the ratio of signal level to the noise level in decibels. Ratios greater than 1 indicate there is more useful information in the signal than there is unwanted data, i.e., the noise. For both SpDP and TrDP designed for 4 h, the SNR indicates good utility, whereas for TrDP designed for one week the sensitive power demand is buried in the noise. Increasing the length of the signal for which privacy is sought results in poor performance of TrDP and little to no utility of the privatized signal.

Extending SpDP to a collection of users will reap a similar utility result. Since the standard deviation of added noise required by SpDP’s released time series is significantly smaller than TrDP, an aggregation of multiple consumers will magnify the poor performance of TrDP.

### C. Parameter Effects on Data Utility

Three parameters are critical to utility of the released time series in SpDP: 1) \( \epsilon \); 2) \( \delta \); and 3) adjacency parameter \( B \). In practice, these parameters should be selected on a per-user basis so that the streaming implementation
output is in accordance with their privacy needs. Though the utility of the SpDP will be impacted by varied parameter selection, the effectiveness of the mechanism is not diminished.

In both SpDP and TrDP, stronger privacy guarantees are provided with a smaller privacy level $\epsilon$ and smaller $\delta$; the expense is in the scale of added noise. Decreasing either privacy parameter without retuning of the filters in SpDP would result in mechanism output that is too noisy to satisfy the positivity condition needed for a streaming implementation. This can be combated by proper tuning of the positive filter without risk of losing information or privacy guarantees, which is a benefit of SpDP over TrDP. The amount of tuning required will vary based on the positive filter design and utility needs of the released time series.

The adjacency parameter $B_{\text{SpDP}}$ determines the events that are protected by SpDP and the scale of noise required for DP. SpDP protects events (in the frequency domain) up to the Nyquist frequency, which is specified by the sampling period, and this includes all possible events for this fixed length PSD. If $B_{\text{SpDP}}$ is reduced, the utility of the released time series will be improved due to less noise needed for privacy, but fewer events will be protected. An increased adjacency parameter will have the opposite effect.

V. CONCLUSION AND FUTURE WORK

A new notion of DP, called SpDP, was presented. SpDP treats the PSD of the underlying process that generates the data as the sensitive information that needs protection. The frequency content of time-series data can reveal patterns as potential sources of exploitation by adversaries. DP guarantees on the frequency content of data are designed in this work to circumvent this issue. A key advantage of SpDP is that the noise needed to provide a certain level of privacy to the downstream analytics. Furthermore, an assessment of analytics, such as billing and real-time monitoring, will be performed to quantify the error in the released data estimate. Additionally, the impact of the WSS assumption on mechanism design is of interest since this assumption may not hold for many smart grid processes due to seasonal or other cyclic phenomenon.

APPENDIX

PROOF OF PROPOSITION 2

Proof: Define $u = M(x)$ and random variables $W \sim \mathcal{N}(0, \Sigma)$ with observations $w$ and $U \sim \mathcal{N}(x, \Sigma)$ with observations $u$. For $S \in \ell_2$, denote $v := x - x'$ where $x$ and $x'$ are adjacent signals according to Proposition 1

$$P(U \in S) = \frac{1}{(2\pi)^{K/2} |\Sigma|^{1/2}} \int e^{-\frac{1}{2}(u-x)^T \Sigma^{-1}(u-x)} dv$$

$$= \frac{1}{(2\pi)^{K/2} |\Sigma|^{1/2}} \int \left[ e^{-\frac{1}{2}(u-x')^T \Sigma^{-1}(u-x')} \times e^{(x-x')^T \Sigma^{-1} v - \frac{1}{2} v^T \Sigma^{-1} v} \right] du$$

$$\leq e^\epsilon \mathbb{P}(M(x') \in S) + \frac{1}{(2\pi)^{K/2} |\Sigma|^{1/2}} \int_{S^c \cap S_2} \left[ e^{-\frac{1}{2}(u-x)^T \Sigma^{-1}(u-x)} \right. \times \left. 1[u((u-x)^T \Sigma^{-1} v + \frac{1}{2} v^T \Sigma^{-1} x - \epsilon - 0)] \right] du.$$

Here, we completed the square and partitioned the sample space such that $\Omega = S_1 \cup S_2$ where $S_1 = \{u((u-x')^T \Sigma^{-1} v - (1/2)v^T \Sigma^{-1} v - \epsilon \leq 0\}$ and $S_2 = \{u((u-x')^T \Sigma^{-1} v - (1/2)v^T \Sigma^{-1} v - \epsilon > 0\}$. For $S_2$ note that $((u-x')^T \Sigma^{-1} v = (u-x)^T \Sigma^{-1} v + v^T \Sigma^{-1} v$. This substitution is needed because the mean of $U$ is defined by $x$, not $x'$.

Since $\Sigma$ is a valid covariance matrix, there exists a positive-definite, symmetric square root of its inverse $\Sigma^{-1}$ that can be defined as $L := \Sigma^{-1/2}$. Let $y^T = (u-x)^T L$ and represent observations of random variable $Y \sim \mathcal{N}(0, I_k)$. From here, we seek to bound the right-hand side of (16) by $\delta$

$$P(U \in S_2) = \mathbb{P}(y^T L v > \epsilon - \frac{1}{2} v^T L^2 v)$$

Defining an intermediate random variable $Z = (Y^T L v)/\|L\|_2$ with distribution $\mathcal{N}(0, 1)$, the integral of (16) can be rewritten as $P(Z \geq \epsilon/\|L\|_2) - \|L\|_2/2) \leq \delta$, which is equivalent to $Q((\epsilon/\|L\|_2) - |\|L\|_2/2|) \leq \delta$ for $Q(t) = (1/\sqrt{2\pi}) \int_t^{\infty} e^{-(t^2/2)} dt$. For ease of notation, let $A = Q^{-1}(\delta)$

$$\epsilon \leq \frac{\|L\|_2^2}{2} \geq A\|L\|_2$$

$$\|L\|_2 \leq -A + \sqrt{A^2 + 2\epsilon}$$

Note that $\|L\|_2 \leq \|L\|_2/\|v\|_2 \leq \|L\|_2 B$ since $\|v\|_2 = \|x - x'\|_2 \leq B$. Let the sorted eigenvalues of $\Sigma$ be $\lambda_{min} \geq \cdots \geq \lambda_{max}$ so the eigenvalues of $L$ are $\left(1/\sqrt{\lambda_{max}}\right) \geq \cdots \geq \left(1/\sqrt{\lambda_{min}}\right)$. Then

$$\|L\|_2 B \leq -A + \sqrt{A^2 + 2\epsilon}$$

$$B \leq \frac{1}{\sqrt{\lambda_{min}}} \leq -A + \sqrt{A^2 + 2\epsilon}$$

$$\lambda_{max} \geq \frac{B^2}{2\epsilon} \left(\frac{-A + \sqrt{A^2 + 2\epsilon}}{2}\right)^2.$$

If $\lambda = \lambda_{min}$, then $(\epsilon, \delta)$-DP follows.
REFERENCES

[1] S. Thorve, L. Kotut, and M. Semaan, “Privacy preserving smart meter data,” in Proc. 7th Int. Workshop Urban Comput. (UrbComp), Aug. 2018, pp. 1–5.

[2] Office of the General Counsel, “Data access and privacy issues related to smart grid technologies,” United States Dept. Energy, Washington, DC, USA, Rep., 2010.

[3] The European Data Protection Supervisor, “Opinion of the european data protection supervisor on the commission recommendation on preparations for the roll-out of smart metering systems,” Eur. Union, Brussels, Belgium, Rep. C(2012)1342., Jun. 2012.

[4] V. Y. Pillitteri and T. L. Brewer, “Guidelines for smart grid cybersecurity,” Nat. Inst. Stand. Technol., Gaithersburg, MD, USA, Rep. NIST 7628 Rev 1, Sep. 2014.

[5] C. Dwork and A. Roth, “The algorithmic foundations of differential privacy,” Found. Trends Theor. Comput. Sci., vol. 9, nos. 3–4, pp. 211–407, Aug. 2014. [Online]. Available: http://dx.doi.org/10.1561/0400000042

[6] J. Liu, C. Zhang, and Y. Fang, “Epic: A differential privacy framework to defend smart homes against Internet traffic analysis,” IEEE Internet Things J., vol. 5, no. 2, pp. 1206–1217, Apr. 2018.

[7] H. Bao and R. Lu, “A new differentially private data aggregation with fault tolerance for smart grid communications,” IEEE Internet Things J., vol. 2, no. 3, pp. 248–258, Jun. 2015.

[8] J. Le. Ny and G. J. Pappas, “Differentially private filtering,” IEEE Trans. Autom. Control, vol. 59, no. 2, pp. 341–354, Feb. 2013.

[9] F. Farokhi, “Temporally discounted differential privacy for evolving datasets on an infinite horizon,” in Proc. ACM/IEEE 11th Int. Conf. Cyber Phys. Syst. (ICCPS), Sydney, NSW, Australia, 2020, pp. 1–8.

[10] F. Siddiqui, S. Zeadally, C. Alcaraz, and S. Galvao, “Smart grid privacy: Issues and solutions,” in Proc. 21st Int. Conf. Comput. Commun. Netw. (ICCCN), Munich, Germany, 2012, pp. 1–5.

[11] S. R. Rajagopalan, L. Sankar, S. Mohajer, and H. V. Poor, “Smart meter privacy: A utility-privacy framework,” in Proc. IEEE Int. Conf. Smart Grid Commun. (SmartGridComm), Brussels, Belgium, 2011, pp. 190–195.

[12] L. Sankar, S. R. Rajagopalan, S. Mohajer, and H. V. Poor, “Smart meter privacy: A theoretical framework,” IEEE Trans. Smart Grid, vol. 4, no. 2, pp. 837–846, Jun. 2013.

[13] L. Ou, Z. Qin, S. Liao, T. Li, and D. Zhang, “Singular spectrum analysis for local differential privacy of classifications in the smart grid,” IEEE Internet Things J., vol. 7, no. 6, pp. 5246–5255, Jun. 2020.

[14] G. Eibl and D. Engel, “Differential privacy for real smart metering data,” Comput. Sci. Res. Develop., vol. 32, no. 1, pp. 173–182, Mar. 2017. [Online]. Available: https://doi.org/10.1007/s00450-016-0310-y

[15] L. Lyu, K. Nandakumar, B. Rubinstein, J. Jin, J. Bedo, and M. Palaniswami, “PFPA: Privacy preserving fog-enabled aggregation in smart grid,” IEEE Trans. Ind. Informat., vol. 14, no. 8, pp. 3733–3744, Aug. 2018.

[16] L. Lyu, Y. W. Law, J. Jin, and M. Palaniswami, “Privacy-preserving aggregation of smart metering via transformation and encryption,” in Proc. IEEE Trustcom/BigDataSE/ICCESS, Sydney, NSW, Australia, 2017, pp. 472–479.

[17] M. Savi, C. Rottondi, and G. Verticale, “Evaluation of the precision-privacy tradeoff of data perturbation for smart metering,” IEEE Trans. Smart Grid, vol. 6, no. 5, pp. 2409–2416, Sep. 2015.

[18] J. D. Rhodes et al., “Experimental and data collection methods for a large-scale smart grid deployment: Methods and first results,” Energy, vol. 65, pp. 462–471, Feb. 2014.

[19] A. Papoulis and U. Pillai, Probability, Random Variables and Stochastic Processes, 4th ed. New York, NY, USA: McGraw-Hill, Nov. 2001.

[20] R. G. Brown and P. Y. Hwang, Introduction to Random Signals and Applied Kalman Filtering, vol. 2. New York, NY, USA: Wiley, 1992.

[21] L. Farina and S. Rinaldi, Positive Linear Systems: Theory and Applications, P. Hilton, H. Hochstadt, M. B. Allen III, D. A. Cox, and J. T. P. Lax, Eds. New York, NY, USA: Wiley, 2000.

[22] P. Stoica and R. L. Moses, Spectral Analysis of Signals. Upper Saddle River, NJ, USA: Pearson Educ., 2005.

[23] E. M. K. Danzmann, Lisa Technology Package Data Analysis (LTPDA), Max Planck Soc., Munich, Germany, 2014. [Online]. Available: https://www.lisamission.org/ltpda/