Privacy of Smart Meter Systems with an Alternative Energy Source

Jesus Gomez-Vilardebo
Centre Tecnològic de Telecomunicacions de Catalunya (CTTC)
Castelldefels, Spain
jesus.gomez@cttc.es

Deniz Gündüz
Imperial College London
London, UK
d.gunduz@imperial.ac.uk

Abstract— **Smart meter (SM) measurements provide near real-time information on the electricity consumption of a user to the utility provider (UP). This data can be used to extract private information on the energy consumption patterns of the user. Assuming that the user has access to an alternative energy source (AES) in addition to the power grid, SM privacy problem is studied from an information theoretic perspective. The energy requirement of the user (input load) at each time instant can be satisfied either from the power grid (output load) or from the AES. It is assumed that the output load can be perfectly tracked by the UP, and the privacy is measured through the information leakage rate. For given average and peak power constraints on the AES, privacy-power function is defined, and its equivalence to the rate-distortion function with a difference distortion measure is shown. Focusing on continuous input loads, the privacy-power function is characterized when there is only peak power limitation on the AES, while the Shannon lower bound is provided for the general case. The bound is shown to be achievable for the exponential input distribution.**

I. Introduction

Smart grids improve upon the current (legacy) power grids by providing improved real-time monitoring and control. To facilitate power monitoring and control, smart grids need an advanced metering infrastructure, which establishes a two-way communication network with the SMs (SMs). Compared to conventional electricity meters, SMs measure and report the energy consumption of the user to the utility providers (UP) much more frequently. This high resolution information on the consumption can increase the efficiency of energy networks significantly. This prompted hasty adoption of SMs worlwide [1]. However, SMs also triggired a growing concern on consumer privacy [2]. Through the SM reading, the UP can have direct access to user's daily life habits, such as times when individual lights are turned on and off, types of equipments used, or the time when nobody is at home [3].

Currently, there is a growing literature on advanced mechanisms to protect the privacy of the users' energy profiles. We identify two main lines of research: One research line assumes that the user has access to the SM readings and can manipulate them before forwarding to the UP. Bohli et al. [4] propose sending the aggregated energy consumption of a group of users, [5] proposes noise addition and [6] proposes compression of smart-meter data. The main limitation of these studies is the assumption that the UP depends solely on the SM reading to measure the user's energy consumption profile. However, the UP can have other means to keep track of a user's energy consumption directly. In the second line of research users are assumed to be equipped with a certain technology that allows them to store or produce energy, and hence, alter the energy consumption profile observed by the UP. In this case, the SM readings are not tempered, i.e., the UP can perfectly track the energy it provides to the user over time. While privacy protection using energy storage devices to filter out the real energy consumption is studied in [7]–[9], use of an alternative energy source (AES) is proposed in [9].

Here we explore the privacy protection solely based on an AES, which can be electric car batteries, energy harvesting devices such as solar panels, or even connection to another independent power grid. Obviously, if the AES is sufficient to provide all the required energy by the appliances, the privacy problem can be resolved in a straightforward manner. However, in general, the AES will be limited, in terms of the average or the peak power it can support, and as we show in this paper, how the user utilizes the energy provided by the AES is critical from the privacy perspective.

We measure privacy with the amount of leaked information about users' energy consumption to the UP, which is quantified by the mutual information between the users' real energy consumption and the energy provided by the UP. Mutual information has previously been proposed as a measure of privacy in SMs in [6] and [8]. We characterize the optimal privacy depending on the average and peak power supported by the AES, called the *privacy - power function*. We provide a single-letter characterization of the privacy - power function when the input load is an independent and identically distributed (i.i.d.) random variable. We show the equivalence of this function to the rate - distortion function with a difference distortion measure. For discrete input distributions, we show that the privacy - power function can be written as a convex optimization problem with linear constraints. For continuous input distributions, we derive the Shannon lower bound (SLB) on the privacy - power function, and show that the SLB is achieved for exponential input distributions, and for certain average and peak power values for other input load distributions.

This work was partially supported by the Catalan Government under SGR2009SGR1046 and by the Spanish Government under projects TEC2010-17816 and TSI-020400-2011-18.

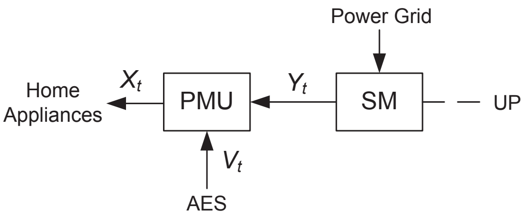

Fig. 1: Smart metering privacy model via an AES.

The rest of the paper is organized as follows. In Section II, we introduce the system model. In Section III, we characterize the privacy - power function when the appliances have an i.i.d. energy demand over time. The derivation of the privacy - power function is considered in Section IV and Section V for discrete and continuous input loads, respectively, and particularized to the exponential distribution in Section VI. Finally, Section VII concludes the paper.

II. System Model

We consider the discrete time SM model depicted in Fig. 1.

The input load $X_t \in \mathcal{X}$ denotes the energy required by the appliances at time instant t. The energy management unit (EMU) satisfies this energy demand at each t either from the power grid or from an AES. SM measures the output load, which is the power requested from the grid, denoted by $Y_t \in \mathcal{Y}$, and reports it to the UP. The remaining part of the demand, denoted by $V_t \in \mathcal{V}$, comes from the AES. We require, $X_t \geq Y_t \geq 0$ at each time instant t, i.e., the EMU is not allowed to ask more energy than the applications require even though doing so would increase its privacy at the expense of wasting energy. For the AES, there is a peak power constraint: $X_t - Y_t \leq \hat{P} \ \ \forall t$, and an average power constraint: $\lim_{n\to\infty} P_n \leq \bar{P}$, where

$$P_n = \mathbb{E}\left[\frac{1}{n}\sum_{t=1}^{n}(X_t - Y_t)\right]. \tag{1}$$

The expectation in (1) is taken over the probability distributions of the input and output loads. Recall that, we consider average and peak power constraints on the AES rather than modeling the instantaneous energy arrival process at the AEP. This is based on the assumption that the AES, which can be an energy harvester or an independent power grid, has its own storage unit. Hence, we are only concerned about the long-term stability of this storage unit.

III. The Privacy - Power Function

Let us denote the input and output load sequences after n units of time as $X^n = [X_1, ..., X_n]$ and $Y^n = [Y_1, ..., Y_n]$, respectively. We measure the privacy by the information rate leaked to the UP about the input load. Assuming that the statistical behavior of the energy demand is known by the UP, its initial uncertainty about the real energy consumption can be measured by the entropy rate $\frac{1}{n}H(X^n)$. This uncertainty is reduced to $\frac{1}{n}H(X^n|Y^n)$ once the UP observes the output load. Hence, the information leaked to the UP can be measured by the reduction in the uncertainty, or equivalently, by the mutual information rate between the input and output loads, defined as

$$I_n \triangleq \frac{1}{n}I(X^n;Y^n). \tag{2}$$

We consider energy management policies that decide on the amount of power that will be received from the AES at each time instant t based on the input load up to time t, X^t, and the output load up to the previous time instant, Y^{t-1}. We allow stochastic policies that satisfy the peak power constraint.

Definition 1: A length-n energy management policy is composed of, possibly random, power allocation functions

$$f_t : \mathcal{X}^t \times \mathcal{Y}^{t-1} \to \mathcal{Y},$$

for $t = 1, \dots, n$, such that $0 \leq X_t - Y_t \leq \hat{P}$ for all $1 \leq t \leq n$. The privacy achieved by this policy is given by the *information leakage rate*, I_n in (2), while the required *average power* from the AES is given by P_n in (1).

Definition 2: An information leakage rate - average power - peak power triplet $(I, \bar{P}, \hat{P})$ is said to be *achievable* if there exists a sequence of energy management policies such that $\lim_{n\to\infty} I_n \leq I$, and $\lim_{n\to\infty} P_n \leq P$.

Definition 3: The information leakage rate - average power - peak power *region* is the closure of the set of all achievable triplets $(I, \bar{P}, \hat{P})$.

Definition 4: The *privacy - power function* $\mathcal{I}(\mathcal{P})$, with $\mathcal{P} \triangleq (\bar{P}, \hat{P})$, is the infimum of the information leakage rates such that $(I, \bar{P}, \hat{P})$ is in the information leakage rate - average power - peak power region.

Our goal is to give a mathematically tractable expression for the privacy-power function, and identify the optimal energy management policy that achieves it.

In the rest of the paper, we consider, for simplicity, i.i.d. input load sequences. In the next theorem, we characterize the privacy-power function in a single-letter format.

Theorem 1: The privacy - power function $\mathcal{I}(\mathcal{P})$ for an i.i.d. input load X with distribution $f_X(x)$ is given by

$$\mathcal{I}(\mathcal{P}) \triangleq \inf_{\substack{f_{Y|X}(y|x): \mathbb{E}[X-Y] \leq \bar{P}, \\ 0 \leq X-Y \leq \hat{P}}} I(X;Y). \tag{3}$$

Some basic properties of the privacy-power function $\mathcal{I}(\mathcal{P})$ are characterized in the following lemma. The proof follows from standard techniques based on time-sharing arguments.

Lemma 1: The privacy - power function $\mathcal{I}(\mathcal{P})$ is a non-increasing convex function of $\bar{P}$ for a given $\hat{P}$.

Next, we give an outline for a proof of Theorem 1.

Proof: The achievability is trivial. Given a conditional probability distribution $f_{Y|X}(y|x)$ that satisfies the conditions in (3), we generate each Y_t independently according to $f_{Y|X}(y_t|x_t)$. The mutual information leakage rate is then given by $I(X;Y)$, whereas the average and peak power constraints are satisfied. For the converse, we assume that there is a series of power allocation functions that satisfy the average and peak power constraints. The information leakage rate of the resulting output load series will satisfy the following chain of inequalities:

$$\frac{1}{n}I(X^n;Y^n) \geq \frac{1}{n}\sum_{t=1}^{n} I(X_t;Y_t), \quad (4)$$

$$\geq \frac{1}{n}\sum_{t=1}^{n} \mathcal{I}(\mathbb{E}[X_t - Y_t], \hat{P}), \quad (5)$$

$$\geq \mathcal{I}\left(\frac{1}{n}\sum_{t=1}^{n}\mathbb{E}[X_t - Y_t], \hat{P}\right), \quad (6)$$

$$\geq \mathcal{I}(\mathcal{P}), \quad (7)$$

where (4) follows as conditioning reduces entropy; (5) follows from the definition of the privacy-power function $\mathcal{I}(\cdot)$ in (3); and (6) follows from Lemma 1 and Jensen's inequality. ■

Remark 1.1: We see from Theorem 1 that the optimal energy management policy is memoryless; that is, it can be achieved by simply looking at the instantaneous input load, and generating the output load randomly using the optimal conditional probability. This results in a stochastic energy management policy rather than a deterministic one. On the other hand, if the user knew all the future energy demand over a block of n time instants, the same privacy performance could be achieved by a deterministic block-based energy management policy.

We note here the correspondance between the privacy-power function in (3) and the rate-distortion function [10]. The privacy-power function in (3) is indeed a rate-distortion function with the following difference distortion measure:

$$d(x,y) = \begin{cases} x-y & \text{if } 0 \leq x-y \leq \hat{P}, \\ \infty & \text{otherwise.} \end{cases}$$

This correspondence allows us to use various tools from rate-distortion theory to study privacy in a SM system.

IV. Discrete Input Distributions

In the previous section we have characterized the privacy-power function for i.i.d. input loads as an optimization problem in a single-letter format. Now we will show that this problem can always be efficiently solved for discrete input distributions. If the input and output alphabets are both discrete, the characterization of the privacy - power function $\mathcal{I}(\mathcal{P})$ in (1) is a convex optimization problem since the mutual information is a convex function of the conditional probability values, $f_{Y|X}(y_m|x_k)$, for $y_m \in \mathcal{Y}$, $x_k \in \mathcal{X}$, and the constraints are linear. Then, (3) can be solved numerically e.g. using the efficient Blahut-Arimoto (BA) algorithm [10].

In the following, we show that the output alphabet can be constrained to the input alphabet, i.e., $\mathcal{Y} = \mathcal{X}$, without loss of optimally. This also implies that for any given discrete input alphabet the optimal output alphabet is also discrete.

Lemma 2: Consider a conditional probability distribution $f_{Y|X}(y|x)$ obtaining mutual information $I(X;Y)$ and satisfying the instantaneous output power constraints $x - \hat{P} \leq y \leq x$. For any $\Omega \subseteq \mathcal{Y}$ and $\bar{y} \in \mathcal{Y}$ we define a new conditional distribution $f_{\hat{Y}|X}(\hat{y}|x)$ as follows:

$$f_{\hat{Y}|X}(\hat{y}|x) = \begin{cases} 0 & \text{if } \hat{y} \in \Omega \setminus \bar{y} \\ \int_{\Omega} f_{Y|X}(y|x)dy & \text{if } \hat{y} = \bar{y}, \\ f_{Y|X}(y|x) & \text{if } \hat{y} \notin \Omega. \end{cases}$$

For any x satisfying $\bar{y} - \hat{P} \leq x \leq \bar{y}$, if the original distribution $f_{Y|X}(y|x)$ assigns an output in the set Ω, the new distribution assigns the output to $\bar{y}$. $f_{\hat{Y}|X}(\hat{y}|x)$ satisfies the instantaneous output power constraint, and we have $I(X;\hat{Y}) \leq I(X;Y)$.

Proof: The proof, omitted here due to space limitations, will be included in the longer version of the paper [11]. ■

Theorem 2: Without loss of optimality the output load alphabet $\mathcal{Y}$ can be constrained to the input load alphabet, i.e., $\mathcal{Y} = \mathcal{X}$.

Proof: Assume that the optimal privacy-power function is achieved by the conditional probability distribution $f_{Y|X}(y|x)$. Define a new conditional distribution $f_{\hat{Y}|X}(\cdot|\cdot)$, as indicated in Lemma 2. Select any interval Ω with $\Pr(X \in \Omega) = 0$ and any $\bar{y} \in \mathcal{X}$ satisfying $\bar{y} > y$ for $\forall y \in \Omega$. From Lemma 2, we have that $f_{\hat{Y}|X}(\hat{y}|x)$ leaks at most the same amount of information to the UP. In addition, since the output load is not reduced for any input load value, the power load demanded from the AES can only have a smaller average value. We can apply this operation for all subsets Ω, until we reduce the output alphabet to the input alphabet. ■

V. Continuous Input Distributions

For a continuous input distribution, the optimal output alphabet is potentially continuous. Consequently, efficient algorithms, such as the BA algorithm, do not yield the optimal solution. In this case, we are able to characterize the privacy - power function only for certain subsets of the power region $\mathcal{P}$. In particular, we characterize the optimal privacy - power function if the AES is limited only by the peak power. We also provide the Shannon lower bound (SLB) on the privacy - power function $\mathcal{I}_{SLB}(\mathcal{P})$, and identify a power region $\mathcal{P}_{SLB}$ when it is achievable.

A. Privacy with a Peak Power Limited AES

We again make use of Lemma 2. Given a peak power constraint $\hat{P}$ and an input load alphabet $\mathcal{X}$, we set the output alphabet to $\hat{\mathcal{Y}} = \{\hat{y}_l\}$, with $\hat{y}_1 = \min(\mathcal{X})$, and for $l > 1$

$$\hat{y}_l = \begin{cases} \hat{y}_{l-1} + \hat{P} & \text{if } \hat{y}_{l-1} + \hat{P} \in \mathcal{X}, \\ \min\left(\mathcal{X} \cap \{x : x > y_{l-1} + \hat{P}\}\right) & \text{otherwise,} \end{cases}$$

until $\hat{y}_l + \hat{P} \leq \max(\mathcal{X})$. Next, given any conditional distribution $f_{Y|X}(y|x)$ satisfying the output load constraints, we define a discrete conditional distribution $p_{\hat{Y}|X}(\hat{y}|x)$ by choosing $\Omega_l = [0, \infty)$ and $\bar{y} = \hat{y}_l$ in Lemma 2. That is, we assign the output $\hat{y}_l$ when the input satisfies $\hat{y}_l \leq x < \hat{y}_l + \hat{P}$. We repeat this operation for all l. Recall that by choosing $\Omega_l = (\hat{y}_{l-1}, \hat{y}_{l+1})$ we obtain exactly the same distribution $p_{\hat{Y}|X}(\hat{y}|x)$. Lemma 2 ensures that this distribution obtains a lower mutual information, i.e., $I(X, \hat{Y}) \leq I(X, Y)$ and satisfies the instantaneous power constraints. In addition, since the output load levels are separated at least by $\hat{P}$, and $\hat{y}_1 = \min(\mathcal{X})$, the mutual information $I(X, \hat{Y})$ can not be further reduced. This means that we can limit the output alphabet to the discrete set composed of $\hat{y}_l$'s.

Given an input load x, there is a unique output load $\hat{y} \in \hat{\mathcal{Y}}$ satisfying the power constraints $x - \hat{P} \leq \hat{y} \leq x$. Consequently,

$$p_{\hat{Y}|X}(\hat{y}_l|x) = \begin{cases} 1 & \hat{y}_l \leq x < \hat{y}_{l+1}, \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

The output probability is given by

$$p_Y(\hat{y}_l) = \int_{\mathcal{X}} p_{Y|X}(\hat{y}_l|x) f(x) dx = \int_{\hat{y}_l}^{\hat{y}_{l+1}} f(x) dx, \tag{9}$$

and $H(Y|X) = 0$. Finally, the privacy - power function is given by $\mathcal{I}_0(\mathcal{P}_0) = H(\hat{Y})$ where $\mathcal{P}_0 = (\bar{P}_0, \hat{P})$ with

$$\bar{P}_0 = \mathbb{E}[X] - \sum_{\hat{y}_l \in \hat{\mathcal{Y}}} \hat{y}_l p_Y(\hat{y}_l). \tag{10}$$

B. The Shannon Lower Bound

In the following, we derive the Shannon lower bound (SLB) [10] on the privacy - power function $\mathcal{I}_{SLB}(\mathcal{P})$, and identify some special cases in which it is achievable. We begin by presenting the distribution that maximizes the entropy among those random variables V with mean $\bar{P}$ and satisfying $0 \leq V \leq \hat{P}$. From [10, Ch. 11], we know that this distribution is the truncated exponential distribution $V \sim \mathsf{ExpT}(\bar{P}, \hat{P})$ with

$$f_V(v) = \begin{cases} \frac{1}{\lambda_0} e^{-\frac{v}{\lambda_1}}, & 0 \leq v \leq \hat{P}, \\ 0 & \text{otherwise.} \end{cases} \tag{11}$$

The variable $\lambda_0 \geq 0$ and $\lambda_1 \geq 0$ are chosen to satisfy

$$\begin{aligned} \int_0^\infty f_V(v) dv &= \frac{\lambda_1}{\lambda_0} p = 1, \\ \mathbb{E}[V] &= \lambda_1 - \hat{P}\frac{q}{p} = \bar{P}, \end{aligned} \tag{12}$$

where $q = e^{-\frac{\hat{P}}{\lambda_1}}$ and $p = 1 - q$. This distribution has differential entropy

$$\mathsf{h}\left(\mathsf{ExpT}(\hat{P}, \bar{P})\right) = \ln(\lambda_0) + \frac{\bar{P}}{\lambda_1},$$

and its Laplace transform $\mathcal{L}f_V(s) = \mathcal{L}\left(f_V(v)\right)(s)$ reads

$$\mathcal{L}f_V(s) = \frac{1}{p} \frac{1 - qe^{-\hat{P}s}}{1 + \lambda_1 s}. \tag{13}$$

Theorem 3: Consider an AES with an average power constraint $\bar{P}$ and a peak power constraint $\hat{P}$. The privacy - power function $\mathcal{I}(\mathcal{P})$ for an i.i.d. input load X with differential entropy $\mathsf{h}(X)$ is lower bounded by

$$\mathcal{I}_{SLB}(\mathcal{P}) = \mathsf{h}(X) - \ln(\lambda_0) - \frac{\bar{P}}{\lambda_1}, \tag{14}$$

where λ_0 and λ_1 are obtained from (12).

Proof: The proof follows from the SLB [10]. ■

Next, we present the necessary and sufficient conditions for any piecewise continuous input distribution $f_X(x)$ to achieve the SLB, together with the conditional probability distribution $f_{Y|X}(y|x)$ achieving it. We denote by $u(x)$, the unit step function which assigns 0 for $x < 0$, and 1 for $x \geq 0$. The Dirac delta function is denoted by $\delta(x)$. We use $f'(x)$ to denote the first order derivative of $f(x)$ and $f(x_i^+) = \lim_{x \to x_i^+} f(x)$ and $f(x_i^-) = \lim_{x \to x_i^-} f(x)$ and $x \to x_i^+$ and $x \to x_i^-$ mean that $x \to x_i$ from left and right, respectively. Finally, we define $\Delta_f(x_i) = f(x_i^+) - f(x_i^-)$.

Theorem 4: Suppose that the input distribution $f_X(x)$ is continuous on $\mathcal{R}_+$ except for a countable number of jump discontinuities or non-differentiable points $\mathcal{X}_D = \{x_1, ..., x_D\}$. Then, the SLB in (14) is achieved for all $\hat{P}$ and $\bar{P}$ satisfying $f_Y(y) \geq 0$ for all $y \in \mathcal{R}_+$ by using the conditional output distribution $f_{Y|X}(y|x) = f_V(x-y)\frac{f_Y(y)}{f_X(x)}$ where the output distribution is given by

$$f_Y(y) = \sum_{l=0}^{\infty} pq^l g_Y(y - l\hat{P}), \tag{15}$$

and $g_Y(y) = g_{Y_C}(y) + g_{Y_D}(y)$ is a mixture of a continuous and a discrete distribution specified as follows:

$$\begin{aligned} g_{Y_C}(y) &= f_X(y) + \lambda_1 f'_X(y), \; y \in \mathcal{R}_+/\mathcal{X}_D, \\ g_{Y_D}(y) &= \lambda_1 \sum_{i=0}^{D} \Delta_X(x_i)\delta(y - x_i), \; y \in \mathcal{X}_D. \end{aligned}$$

Proof: To prove this result, we need to find the conditional distribution $f_{Y|X}(y|x)$ that satisfies the SLB with equality [10]. We require the random variables $V = X - Y$ and Y to be independent, and V to be distributed according to a truncated exponential distribution $V \sim \mathsf{ExpT}(\bar{P}, \hat{P})$ with mean $\bar{P}$ and peak value $\hat{P}$. We first obtain the output distribution $f_Y(y)$ from its Laplace transform $\mathcal{L}f_Y(s) = \mathcal{L}(f_Y(y))(s)$. First, recall $\mathcal{L}f_V(s)$ in (13) and that $\mathcal{L}g_Y(s) = \mathcal{L}\left(g_Y(y)\right)(s)$ is given by $\mathcal{L}g_Y(s) = \mathcal{L}f_X(s)\left(1 + \lambda_1 s\right)$. Then, observe that

$$\begin{aligned} \mathcal{L}f_Y(s) &= \frac{\mathcal{L}f_X(s)}{\mathcal{L}f_V(s)}, & (16) \\ &= p\frac{\mathcal{L}g_Y(s)}{1 - qe^{-\hat{P}s}}, & (17) \\ &= \sum_{l=0}^{\infty} pq^l \mathcal{L}g_Y(s) e^{-l\hat{P}s}, & (18) \\ &= \sum_{l=0}^{\infty} pq^l \mathcal{L}\left(g_Y(y - l\hat{P})\right)(s). & (19) \end{aligned}$$

It follows that $f_Y(y)$ is given by (15). The conditional distribution $f_{Y|X}(y|x)$ is obtained using the fact that $f_{X|Y}(x|y) = f_V(x - y)$. Finally, it can be shown that $\int_0^\infty f_Y(y)dy = 1$; and thus, achievability only requires $f_Y(y) \geq 0$, $\forall y \in \mathcal{R}^+$. ■

Remark 4.1: If the achievability condition in Theorem 4 is satisfied for a given $\lambda_1^{\max}$, it is satisfied for any $\lambda_1 \leq \lambda_1^{\max}$. Then from the dependence of $\bar{P}$ and $\hat{P}$ on λ_1, it follows that, given $\hat{P}$, there is a unique critical average power level, $\bar{P}_0$, such that $\mathcal{I}(\bar{P}, \hat{P}) = \mathcal{I}_{SLB}(\bar{P}, \hat{P})$ for all $\bar{P} \leq \bar{P}_0$ and

$\mathcal{I}(\bar{P},\hat{P}) > \mathcal{I}_{SLB}(\bar{P},\hat{P})$ for all $\bar{P} > \bar{P}_0$. Similarly, given $\bar{P}$, there is a unique critical peak power $\hat{P}_0$, such that $\mathcal{I}(\bar{P},\hat{P}) = \mathcal{I}_{SLB}(\bar{P},\hat{P})$ for all $\hat{P} \geq \hat{P}_0$, and $\mathcal{I}(\bar{P},\hat{P}) > \mathcal{I}_{SLB}(\bar{P},\hat{P})$ for all $\hat{P} < \hat{P}_0$.

Remark 4.2: For an AES with an unlimited peak power constraint, i.e., $\hat{P} = \infty$, we have $\lambda_1 \to \bar{P}$, $\lambda_0 \to \bar{P}$, and V follows an exponential distribution $\mathsf{Exp}(\bar{P})$. Then, the SLB reduces to $\mathcal{I}_{SLB}(\bar{P}) = \mathsf{h}(X) - \ln(e\bar{P})$, and the output distribution simplifies to $f_Y(y) = g_Y(y)$.

VI. Exponential Distribution

In this section, we use the results obtained in the previous section to characterize the privacy - power function for an exponential input distribution. Let $X \sim \mathsf{Exp}(m)$, i.e., $f_X(x) = \frac{1}{m}e^{-\frac{x}{m}}u(x)$. From (15) we obtain the output probability distribution as

$$g_{Y_C}(y) = \left(1 - \frac{\lambda_1}{m}\right)\sum_{l=0}^{\infty} pq^l f_X(y - l\hat{P}), \tag{20}$$

$$g_{Y_D}(y) = \frac{\lambda_1}{m}\sum_{l=0}^{\infty} pq^l \delta(y - l\hat{P}). \tag{21}$$

For $Y \in \left\{l\hat{P} : l = 0, 1, ..., \infty\right\}$, Y follows a discrete geometric distribution, $\mathsf{Geom}(p) = pq^l$. Otherwise, Y follows a mixture of weighted and shifted continuous exponential distributions each with mean m. The SLB achievability condition $f_Y(y) \geq 0$ for all $y \in \mathcal{R}_+$ requires $m \geq \lambda_1$, or equivalently, $\bar{P} \leq \bar{P}_0$ with $q_0 = e^{-\frac{\hat{P}}{m}}$, and

$$\bar{P}_0 = m - \hat{P}\frac{q_0}{1 - q_0}. \tag{22}$$

This average power $\bar{P}_0$ coincides with the average power (10) needed to obtain the minimum information leaked for peak power only limited AES; thus,

$$\mathcal{I}_0\left(\mathcal{P}_0\right) = \mathsf{h}(Y) = \mathsf{h}(\mathsf{Geom}(p_0)). \tag{23}$$

Observe that for the exponential input distribution, the SLB is achievable for all possible peak and average power constraints. For the particular case of an unlimited peak power constraint, this rate distortion function was first derived in [12] and reduces to $\mathcal{I}(\bar{P}, \infty) = \ln\left(\frac{m}{\bar{P}}\right)$ if $\bar{P} \leq m$ and $\mathcal{I}(\bar{P}, \infty) = 0$ otherwise. For an exponential distribution with mean 1, Fig. 2 depicts the privacy - power function for different $\hat{P}$ values.

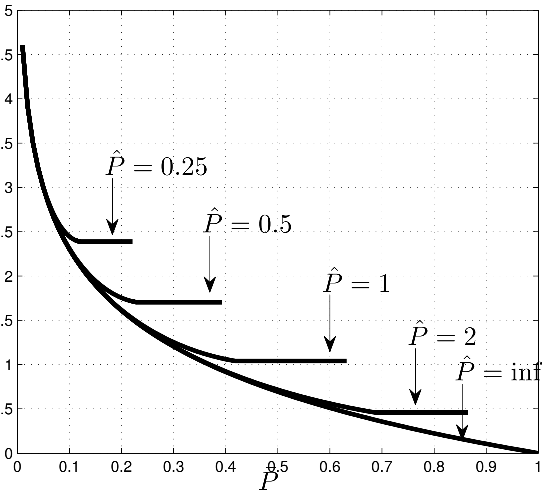

Fig. 2: Privacy-power function for $X \sim \mathsf{Exp}(1)$

VII. Conclusions

We have proposed a mathematical model for studying privacy in a SM system in the presence of an AES. We have shown that the user can hide its energy consumption profile from the UP by utilizing an AES in a stochastic manner. Using an information theoretic privacy measure, we have characterized the optimal privacy that can be achieved for given average and peak power constraints on the AES. We have shown that, for i.i.d. input loads, the privacy-power function has a single-letter expression. In addition, for discrete input alphabets we have shown that the privacy-power function can be evaluated numerically as the solution to a convex optimization problem. For continuous input distributions, we have obtained the privacy - power function for a peak power only limited AES, as well as for any AES when the input is exponentially distributed. We have also characterized the Shannon lower bound on the privacy - power function, and identified the conditions under which it is tight.

References

[1] P. Wunderlich, D. Veit, and S. Sarker, "Adoption of information systems in the electricity sector: The issue of smart metering," in *Proc. Americas Conference on Information Systems*, Seattle, WA, Aug 2012.

[2] P. McDaniel and S. McLaughlin, "Security and privacy challenges in the smart grid," *IEEE Security Privacy*, vol. 7, no. 3, pp. 75–77, May-Jun. 2009.

[3] U. Greveler, B. Justus, and D. Loehr, "Multimedia content identification through smart meter power usage profiles," in *Computers, Privacy and Data Protection (CPDP)*, Brussels, Belgium, Jan. 2012.

[4] J.-M. Bohli, C. Sorge, and O. Ugus, "A privacy model for smart metering," in *Proc. IEEE Int'l Conf. on Comm.*, Capetown, South Africa, May 2010.

[5] V. Rastogi and S. Nath, "Differentially private aggregation of distributed time-series with transformation and encryption," in *Proc. Intl. Conf. Data Management*, Indianapolis, Indiana, Oct. 2010.

[6] S. R. Rajagopalan, L. Sankar, S. Mohajer, and H. V. Poor, "Smart meter privacy: A utility-privacy tradeoff framework," in *Proc. IEEE Int'l Conf. Smart Grid Comm.*, Brussels, Belgium, Oct 2011.

[7] G. Kalogridis, C. Efthymiou, S. Denic, T. A. Lewis, and R. Cepeda, "Privacy for smart meters: Towards undetectable appliance load signatures," in *Proc. IEEE Int'l Conf. Smart Grid Comm.*, Gaithersburg, MD, Oct. 2010.

[8] D. Varodayan and A. Khisti, "Smart meter privacy using a rechargeable battery: Minimizing the rate of information leakage," in *Proc. IEEE Int. Conf. Acoust. Speech Sig. Proc.*, Prague, Czech Republic, May 2011.

[9] O. Tan, D. Gündüz, and H. V. Poor, "Smart meter privacy in the presence of energy harvesting and storage devices," in *Proc. IEEE Int'l Conf. Smart Grid Comm.*, Tainan City, Taiwan, Nov 2012.

[10] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley-Interscience, 1991.

[11] J. Gomez-Vilardebo and D. Gunduz, "Privacy of smart meter systems with an alternative energy source (in preparation)."

[12] S. Verdú, "The exponential distribution in information theory," *Probl. Inform. Transm.*, vol. 32, pp. 86 – 95, Jan. - Mar. 1996.