Proceedings Paper:
Arrieta, M. and Esnaola, I. orcid.org/0000-0001-5597-1718 (2017) Smart meter privacy via the trapdoor channel. In: Proceedings of 2017 IEEE International Conference on Smart Grid Communications (SmartGridComm). 2017 IEEE International Conference on Smart Grid Communications (SmartGridComm), 23-27 Oct 2017, Dresden, Germany. IEEE , pp. 277-282. ISBN 9781538624012

https://doi.org/10.1109/SmartGridComm.2017.8340722

Smart Meter Privacy via the Trapdoor Channel

Miguel Arrieta* and Iñaki Esnaola*§
*Dept. of Automatic Control and Systems Engineering, University of Sheffield, Sheffield S1 3JD, UK
§Dept. of Electrical Engineering, Princeton University, Princeton, NJ 08544, USA
{marrieta2, esnaola}@sheffield.ac.uk

***Abstract*—A battery charging policy that provides privacy guarantees for smart meter systems with finite capacity battery is proposed. For this policy an upper bound on the information leakage rate is provided. The upper bound applies for general random processes modelling the energy consumption of the user. It is shown that the average energy consumption of the user determines the information leakage rate to the utility provider. The upper bound is shown to be tight by deriving the probability law of a random process achieving the bound.**

I. Introduction

The increasing appeal for economical and environmentally-friendly energy calls for more efficient energy generation, distribution, and consumption [1]. The introduction of a digital infrastructure into the traditional power system takes steps towards this vision by providing a cyber layer that elevates the existing power system to a cyberphysical system. This advanced sensing and communication infrastructure envisioned by the smart grid enables high resolution and real time management of the processes within the grid. This application layer also enables dynamic energy pricing, shifting user demand to match energy generation [2]. Moreover, the introduction of energy consumption indicators through Smart Meters (SMs) are reported to reduce the energy consumption of the user by up to 15% by raising awareness of the energy cost [3].

While the high-resolution information provided by the smart grid brings clear advantages it also raises privacy and security concerns [4], [5], [6]. By analysing the consumption profile of a user, techniques such as non-intrusive load monitoring (NILM) [7] track and recognise appliance usage patterns [8], [9]. Human presence, usage of individual appliances [10], [11], and even tuned TV channel [12] are among the list of recognizable elements [13]. This privacy breach hinders the implementation of some of the essential components of the smart grid [4], [5]. Within this paradigm, SMs are central components to the dilemma posed by the need for accurate monitoring while providing privacy. In 2009 two bills law aimed to enforce the usage of SMs were blocked by the Senate of the Netherlands motivated by the privacy concerns that emerge as a result of the increased penetration of SMs [14].

There is a growing body of literature addressing the conflict between efficient energy monitoring and privacy brought forth by the introduction of SMs. In [15], [16] obfuscation of the knowledge that the utility provider (UP) has about the energy consumption of the user is studied. Indeed, in the case in which the SM readings are the only source of information available to the UP, obfuscation yields some degree of privacy. Obfuscation is achieved by several mechanisms, such as aggregating the consumption of multiple users [15], compression of the energy consumption sequences [16] or homomorphic encryption [17] among others. A different approach to the problem arises in the setting in which users have access to alternative energy sources or energy storage devices [18]. In this case, the UP has perfect knowledge of the energy provided to the user, but the user employs the energy storing capability of the system to dissociate the energy consumed by the appliances from the energy provided by the UP. In [19], [20], [21] the case in which the user is assumed to have an alternative energy source with instantaneous power constraints is studied. In [20], [22], [23], [24] the user is assumed to have access to a rechargeable battery and the energy consumption is modelled as an independent and identically distributed (i.i.d.) random process.

In a practical setting, the energy consumption profiles of users exhibit non-stationary statistical structures and are not well described by memoryless random processes [23]. Moreover, information-theoretic privacy measures for random processes that are not i.i.d. are still not well understood [25]. The privacy utility tradeoff is characterized for stationary Gaussian energy consumption models in [26]. A first-order time-homogeneous Markov chain is considered in [27]. In this work we adopt a non-probabilistic framework by modelling the energy management system (EMS) with a finite capacity battery as a finite state channel without probabilistic structure. Inspired by the code construction in [28] we propose an energy charging policy and we characterize the privacy guarantees of the strategy for general random processes. We also particularize the analysis to the case in which the average energy consumption of the user is known. For this case we provide an upper bound to the amount of information that the user leaks to the UP and show that the average energy consumption governs the privacy that is achievable by the user. In this paper vectors are denoted by bold font, e.g. $\mathbf{x}$, random variables are denoted by upper-case, e.g. X, and vectors of n random variables are denoted by super-indexing the size, e.g. X^n.

II. Battery System Model

We consider the energy management system (EMS) operating in discrete-time illustrated in Fig. 1. The energy consumption of the user is modelled as a discrete-time random process $X_1, X_2, \ldots$ where X_i is a random variable taking values in $\mathcal{X} = \{0, 1, \ldots, \alpha\}$. In this setting X_i describes the energy consumed at time instant $i \in \mathbb{N}$. At each time instant, the energy consumption of the user is satisfied by requesting energy from the UP or by discharging the battery. This decision is taken by the energy management unit (EMU) at each time instant based on some power consumption performance and privacy criteria. The energy requested from the UP is also a discrete-time random process $Y_1, Y_2, ...$ where Y_i is a random variable taking values in $\mathcal{Y} = \{0, 1, ..., \gamma\}$ and describing the energy requested from the UP at time instant $i \in \mathbb{N}$.

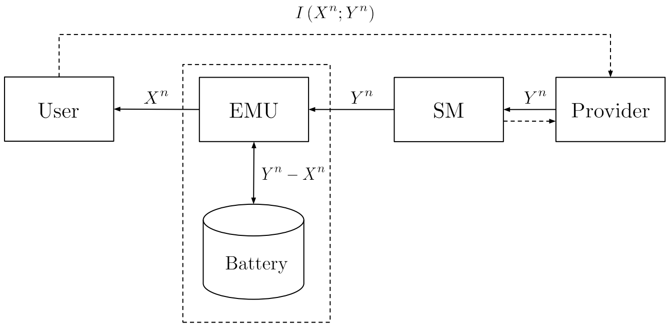

Fig. 1. Diagram of an Energy Management System with finite battery.

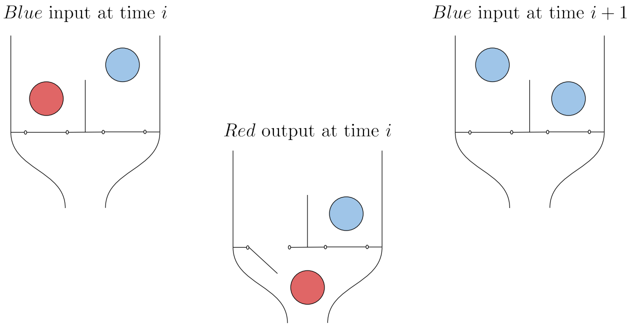

Fig. 2. [29] Diagram depicting the operation of a trapdoor channel with $\beta = 1$.

When $Y_i > X_i$, the excess energy is stored in the battery. Alternatively, when $Y_i \leq X_i$, the energy deficit is obtained from the battery. We assume the UP is able to satisfy the energy consumption of the user even in the case when there is no battery, i.e. $\gamma \geq \alpha$.

In the above model, the state S_i taking values in $\mathcal{S} = \{0, 1, ..., \beta\}$ describes the energy stored in the battery at time i. The energy stored in the battery S_i is a function of the previous energy consumption X^i, the energy request Y^i, and the initial state $s_0 \in \mathcal{S}$ given by

$$S_i = s_0 + \sum_{k=0}^{i-1} (Y_k - X_k). \tag{1}$$

Within this setting, a power outage occurs when $S_i + Y_i - X_i < 0$, and energy is wasted when $S_i + Y_i - X_i > \beta$. In the following we focus on EMUs that do not allow power outages nor energy wasting but provide a quantifiable privacy guarantee for the user. Given a particular realization $\mathbf{x} \in \mathcal{X}^n$ of the random process $X_1, X_2, \ldots, X_n$ modelling the energy consumption of the user up to time n, the set of energy requests that the EMU can implement is limited by the power outage and the energy waste constraints. The following definition describes the set of energy requests that the EMU can implement.

Definition 1: Given an energy consumption sequence $\mathbf{x} \in \mathcal{X}^n$ as the input of an EMU with a battery of capacity β, the *set of stable energy request sequences* that avoid power outages and energy waste is given by

$$\mathcal{Y}_\beta^n(s_0, \mathbf{x}) := \{\mathbf{y} \in \mathcal{Y}^n : s_i + y_i - x_i \in \mathcal{S} \text{ for all } i\}, \tag{2}$$

where $s_i \in \mathcal{S}_i$ is the state of the battery at time i, determined by $\mathbf{x}$ and $\mathbf{y}$ according to (1).

The task of the EMU is therefore to choose a particular sequence in the $\mathcal{Y}_\beta^n(s_0, \mathbf{x})$ for a given power consumption realization $\mathbf{x}$. The structure of the particular choice determines the policy implemented by the EMU and is captured by the following definition.

Definition 2: Given an EMU with a battery of capacity β the *set of stable battery policies* is the set of mappings between the energy consumption sequences and the set of stable energy request sequences given by

$$\mathcal{P}_\beta := \{P_\beta : \mathcal{S} \times \mathcal{X}^n \to \mathcal{Y}_\beta^n(s_0, \mathbf{x})\}. \tag{3}$$

Since Y^n is known by the UP, the information about the energy consumption of the user that the UP acquires via the energy request is given by the mutual information $I(X^n; Y^n)$. As in [22] this motivates the following definition of privacy.

Definition 3: Given an EMU operating with the stable battery policy P_β, the information about the consumption of the user, X^n, that is leaked to the UP is the *information leakage rate* given by

$$L_\beta(X^n, P_\beta) := \frac{1}{n} I(X^n; Y^n). \tag{4}$$

In a SM privacy context, the aim of the EMU is to choose a stable battery policy P_β that minimizes the information leakage rate, i.e. maximizes the privacy of the user. Note that the information leakage depends on the joint probability distribution of X^n and Y^n. In general, the evaluation of (4) yields involved expressions that are difficult to evaluate [25]. For that reason previous results [19], [20], [22] tend to consider simple probabilistic models, e.g. memoryless processes, to evaluate the mutual information. In the remaining of the paper we analyze the privacy guarantees for general discrete-time random processes modelling the user consumption. To that end, we model the EMS with a battery of capacity β as a non-probabilistic finite-state channel [28]. The rationale for this approach and the equivalence between the EMS and a non-probabilistic channel are discussed in the following section.

A. Equivalence with the trapdoor channel

The *trapdoor channel* [28], [29] is defined as a box containing b_0 *blue* balls and $\beta - b_0$ *red* balls. The operation of the channel is depicted in Fig. 2. At time i a new ball X_i coloured *blue* or *red* is thrown into the box. Immediately after, one of the $\beta + 1$ balls inside the box is selected and taken out of the box. Let Y_i denote the ball extracted at time i. Following this model, the number of *blue* balls inside the box at time i is given by

$$b_i = b_0 + \sum_{k=0}^{i-1} bl(X_k) - \sum_{k=0}^{i-1} bl(Y_k), \tag{5}$$

where the indicator function $bl(\cdot)$ equals 1 when its argument is coloured *blue* and 0 otherwise. Similarly, the number of *red* balls inside the box at time i is given by $r_i = \beta - b_i$. Replacing $b_i = \beta - r_i$ and $b_0 = \beta - r_0$ into (5) yields:

$$r_i = r_0 + \sum_{k=0}^{i-1} \big(bl(Y_k) - bl(X_k)\big). \tag{6}$$

The number of *red* balls inside the box is bounded between 0 and β. For a box of capacity β the set of *stable output balls* $\mathcal{Y}_{\beta}^{n}\left(r_0,\mathbf{x}\right)$ is defined as the set of outputs $\mathbf{y}^n \in \mathcal{Y}^n$ than can be pulled out of the box given an initial state r_0 and an input sequence $\mathbf{x} \in \mathcal{X}^n$, i.e.

$$\mathcal{Y}_{\beta}^{n}\left(r_0,\mathbf{x}\right) = \{\mathbf{y} : r_i + bl\left(y_i\right) - bl\left(x_i\right) \in \mathcal{R} \text{ for all } i\}, \quad (7)$$

where $\mathcal{R} = \{0, ..., \beta\}$.

It is easy to see that for the case in which $\gamma = \alpha = 1$ the EMS with a battery of capacity β described in Section II is equivalent to the *trapdoor channel* of capacity β. The set of balls inside the box determines the state of the trapdoor channel, and similarly, the amount of energy stored in the battery determines the state of the EMS channel. For the case in which $\alpha = \gamma = 1$ both systems are equivalent since requesting energy from the grid corresponds to extracting a ball from the trapdoor channel. Similarly, replacing a ball from the trapdoor channel in (6) corresponds to charging the battery of the EMS in (1).

III. Privacy With An Arbitrary Energy Consumption

In this section, we provide bounds on the information leakage rate when no restrictions are imposed on the probability law of X^n. We first propose the construction of a *stable battery policy* P_{β}^{*} and characterize an upper bound on the information leakage rate L_{β} induced by P_{β}^{*} and any arbitrary random process X^n. Furthermore, we show the tightness of the upper bound by constructing a random process $\hat{X}$ whose leakage is tight with respect to the upper bound. Moreover, the leakage rate induced by the random process $\hat{X}$ is shown to be independent of the employed battery policy $P_{\beta} \in \mathcal{P}_{\beta}$. This shows that the upper bound is tight with respect to the minimum *information leakage rate* a *stable battery policy* P_{β} can guarantee for general random processes X^n.

The approach to policy construction in this section is similar to the code construction in [28] where a trapdoor channel with a box of size $\hat{\beta}$ is considered. Therein, at every time instant a ball numbered $1, 2, ..., \hat{\alpha}$ is introduced into the box and one of the $\hat{\beta}+1$ inside the box is extracted. In [28, Section II] the case in which the box acts as a jammer trying to obstruct the communication process between a sender inserting the balls into the box and a receiver drawing the output balls is studied. Therein, the ball extracted from the box is selected in order to minimize the mutual information between the input and the output. Note that the extraction criteria is not probabilistic and is instead analyzed using combinatorial tools. Remarkably, in [28, Proposition 1] it is shown that the *Shannon capacity* $C_{\hat{\beta}}$ of such channels is lower bounded by

$$C_{\beta} \geq \frac{\log \hat{\alpha}}{\hat{\beta}+1}. \quad (8)$$

Moreover, when $\hat{\alpha} = 2$ the capacity is upper bounded by

$$C_{\beta} \leq \frac{1}{\hat{\beta}+1}. \quad (9)$$

In [28] the output is constrained to permutations of the input sequence. In our setting, the sum of the output sequence is bounded by the sum of the input sequence and the size of the battery. However, the output is not required to contain the same symbols as the input . For that reason, the approach in [28] requires some modification but the main idea remains. The derivation is presented in the next section.

A. Upper bound on the information leakage rate

We propose a battery policy based on the code construction in [28]. The codebook proposed in the trapdoor channel context is the counterpart of the battery policy in the smart meter case. The proposed policy structures the energy request sequences according to the output alphabet defined below.

Definition 4: Consider the set of codewords of length l constructed by repetitions of 0 or α symbols, i.e. $\mathcal{O}_l = \{(0, 0, \cdots, 0), (\alpha, \alpha, \cdots, \alpha)\}$. For $n = lm$, we define the *block repetition alphabet* as the set $\mathcal{O}_l^m$ of sequences obtained by the m-fold concatenation of codewords of length l. Specifically

$$\mathcal{O}_l^m = \mathcal{O}_l \times \mathcal{O}_l \times ... \times \mathcal{O}_l. \quad (10)$$

We now define a stable policy that maps the energy consumption of the user to the output sequences constructed with the *block repetition alphabet* $\mathcal{O}_l^m$.

Definition 5: A *block battery policy* P_{β}^{*} is a mapping of the form

$$P_{\beta}^{*} : \mathcal{S} \times \mathcal{X}^n \to \mathcal{O}_l^m \cap \mathcal{Y}_{\beta}^{l}\left(s_0,\mathbf{x}\right). \quad (11)$$

Note that a block battery policy is nothing more than a strategy to assign to each input sequence a stable energy request sequence constructed with a block repetition alphabet. With these definitions at hand we now provide the following privacy guarantee.

Theorem 1: Consider an EMS with a battery of capacity β and initial state $s_0 \in \mathcal{S}$. Let X^n be a random process with X_i taking values in $\mathcal{X} = \{0, 1, ..., \alpha\}$ for $i = 1, 2, \ldots n$ and P_{β}^{*} a *block battery policy* as described in Definition 5. Then for $l \leq \lfloor(\beta+1)/\alpha\rfloor$ at least one policy P_{β}^{*} exists such that

$$L_{\beta}\left(X^n, P_{\beta}^{*}\right) \leq \frac{1}{\lfloor(\beta+1)/\alpha\rfloor}. \quad (12)$$

Proof: Notice that the information leakage rate is upper bounded by

$$L_{\beta}\left(X^n, P_{\beta}^{*}\right) = \frac{1}{n} I\left(X^n; Y^n\right) \leq \frac{1}{n} H\left(Y^n\right). \quad (13)$$

Since Y^n takes values in $\mathcal{O}_l^m$ and $|\mathcal{O}_l^m| = 2^m$ the following holds:

$$\frac{1}{n} H\left(Y^n\right) \leq \frac{1}{n} \log|\mathcal{O}_l^m| = \frac{1}{n} \log\left(2^m\right) = \frac{m}{n} = \frac{1}{l}. \quad (14)$$

We now show that when $l \leq (\beta+1)/\alpha$ there exists at least one *block battery policy* P_{β}^{*} for every initial state s_0 and consumption $\mathbf{x}$. To prove this we establish that for every realization $\mathbf{x}$ and initial state s_0 there exist an energy request sequence determined by $\mathbf{y} \in \mathcal{O}_l^m$ such that $\mathbf{y}$ belongs to the set of stable energy requests $\mathcal{Y}_{\beta}^{n}\left(s_0,\mathbf{x}\right)$. The strategy is to notice that $\mathcal{O}_l^m \cap \mathcal{Y}_{\beta}^{n}\left(s_0,\mathbf{x}\right) \neq \emptyset$ for $m = 1$ and to then prove by induction that the non-emptiness of the intersection holds for $m \geq 1$.

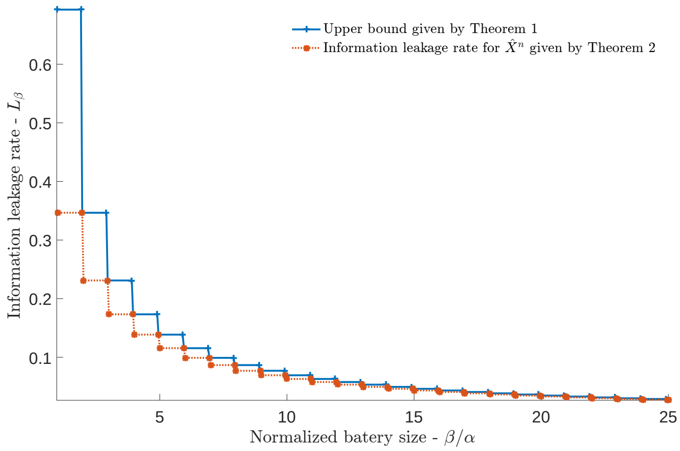

Fig. 3. Upper bound on the information leakage rate of an EMS as a function of the ratio between the battery size and the peak power consumption.

The intersection $\{(0,0,\cdots,0),(\alpha,\alpha,\cdots,\alpha)\} \cap \mathcal{Y}_\beta^l\left(s_0, X^l\right)$ is non-empty if and only if either the sequence $(0,0,\cdots,0)$ or $(\alpha,\alpha,\cdots,\alpha)$ belong to $\mathcal{Y}_\beta^n\left(s_0,\mathbf{x}\right)$. Jointly with (2) this implies that either

$$s_i + 0 - x_i \in \mathcal{S} \tag{15}$$

or

$$s_i + \alpha - x_i \in \mathcal{S} \tag{16}$$

holds for $i \leq l$. In the first case, described in (15), we have that $0 - x_i \leq 0$ for $i = 0, \cdots, l-1$. Hence, the energy stored in the battery, s_i, decreases monotonically. Therefore, all s_i belong to $\mathcal{S}$ when $s_i \geq 0$ on the last time step, i.e.,

$$0 \leq s_0 - \sum_{i=0}^{l-1} x_i. \tag{17}$$

Similarly, in the case described by (16), we have that $\alpha - x_i \geq 0$ and the energy stored increases monotonically. It is then sufficient to show that

$$s_0 - \sum_{i=0}^{l-1} x_i \leq \beta - \alpha l. \tag{18}$$

When $\beta - \alpha l \geq -1$ every integer s_i satisfies at least one of the inequalities given by (17) and (18). This ensures that either (17) or (18) hold for every $s_0 \in \mathcal{S}$ and $\mathbf{x} \in \mathcal{X}^l$, and therefore, the intersection $\mathcal{O}_l^m \cap \mathcal{Y}_\beta^n\left(s_0,\mathbf{x}\right)$ is non-empty. This completes the proof for $m = 1$. The induction for $m \geq 1$ is straightforward as the proof for $m = 1$ holds for every initial state s_0. ■

The upper bound derived in Theorem 1 is depicted for different battery sizes in Fig. 3. It is interesting to note that the privacy guarantees increase significantly for small values of β/α but the benefit vanishes as the size of the battery increases.

B. Tightness of the upper bound

We now study the tightness of the upper bound presented in Theorem 1. To this end, we construct a random process modelling the energy consumption of the user that is tight with respect to the result in Theorem 1 for every battery policy $P_\beta \in \mathcal{P}_\beta$.

Theorem 2: Consider an EMS with a battery of capacity β and initial state s_0. Let $\hat{X}^n$ be a random process taking uniformly distributed values in $\mathcal{O}_l^m$ with $l = \lceil(\beta+1)/\alpha\rceil$. Let P_β be a *stable battery policy*. Then

$$L_\beta\left(\hat{X}^n, P_\beta\right) = \frac{1}{\lceil(\beta+1)/\alpha\rceil}. \tag{19}$$

Proof: We expand L_β as

$$\frac{1}{n} I\left(X^n; Y^n\right) = \frac{1}{n} H\left(X^n\right) - \frac{1}{n} H\left(X^n | Y^n\right). \tag{20}$$

When X^n is uniformly distributed over the alphabet $\mathcal{O}_l^m$ it yields

$$\frac{1}{n} H\left(X^n\right) = \frac{1}{n} m = \frac{1}{l}. \tag{21}$$

We now show that the equivocation rate $\frac{1}{n} H\left(X^n|Y^n\right)$ is 0 when X^n takes values in $\mathcal{O}_l^m$ with $l > \beta/\alpha$. We prove by induction that when the input realization $\mathbf{x}$ belongs to $\mathcal{O}_l^m$ with $l > \beta/\alpha$, the sets $\mathcal{Y}_\beta^n\left(s_0,\mathbf{x}\right)$ of stable output words generated by different consumption sequences are disjoint, i.e.

$$\mathcal{Y}_\beta^n\left(s_0,\hat{\mathbf{x}}\right) \cap \mathcal{Y}_\beta^n\left(s_0,\mathbf{x}\right) = \emptyset \text{ for } \hat{\mathbf{x}} \neq \mathbf{x}. \tag{22}$$

As a result, any request sequence $\mathbf{y} \in \mathcal{Y}_\beta^n\left(s_0,\mathbf{x}\right)$ unequivocally determines the generating input $\mathbf{x}$. In other words, given an output sequence $\mathbf{y}$ there is no uncertainty about the input $\mathbf{x}$, and therefore, the equivocation rate $\frac{1}{n} H\left(X^n|Y^n\right)$ is 0.

For $m = 1$ there are two possible inputs $(0,0,\cdots,0)$ and $(\alpha,\alpha,\cdots,\alpha)$. When $\mathbf{x} = (0,0,\cdots,0) \in \mathcal{O}_l^1$ the energy stored in the battery at time l is given by

$$s_l = s_0 + \sum_{i=0}^{l-1} \left(y_i - 0\right). \tag{23}$$

Similarly, when $\mathbf{x} = (\alpha,\alpha,\cdots,\alpha) \in \mathcal{X}^l$ the energy stored in the battery at time l is given by

$$z_l = s_0 + \sum_{i=0}^{l-1} \left(y_i - \alpha\right). \tag{24}$$

Taking the difference between (23) and (24) yields:

$$s_l - z_l = \sum_{i=0}^{l-1} \alpha = l\alpha. \tag{25}$$

When $z_l \in \mathcal{S}$ we have that $s_l = z_l + l\alpha \geq l\alpha$, showing that for $l\alpha > \beta$ the events $z_l \in \mathcal{S}$ and $s_l \in \mathcal{S}$ do not occur simultaneously. This implies that the set of output words belonging to $\mathcal{Y}_\beta^n\left(s_0,(0,0,\cdots,0)\right)$ and $\mathcal{Y}_\beta^n\left(s_0,(\alpha,\alpha,\cdots,\alpha)\right)$ is empty for every initial state s_0. Therefore the sets are disjoint and $\frac{1}{n} H\left(X^n|Y^n\right) = 0$. The proof for $m > 1$ follows by induction and noticing that the proof above is valid for every initial state s_0. ■

IV. Privacy With An Average Energy Constraint

The information leakage rate bounds provided in Section III do not impose any moment restriction on the random process modeling the energy consumption of the user. Indeed, they depend only on the range of the energy consumption and on the size of the battery. However, one of the most widely used energy consumption metrics is the average energy

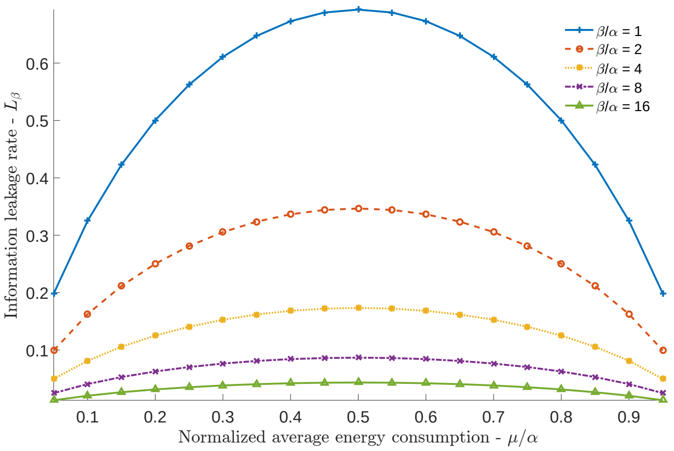

Fig. 4. Upper bound on the information leakage rate of an EMS when $n \to \infty$ as a function of the average energy consumption of the user for different values of the ratio between the battery size and the peak power consumption.

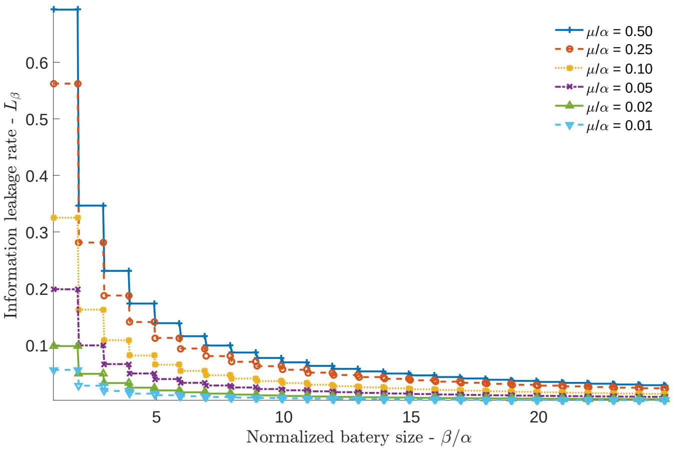

Fig. 5. Upper bound on the information leakage rate of an EMS when $n \to \infty$ as a function of the ratio between the battery size and the peak power consumption for different values of the average energy consumption of the user

consumption over an arbitrary time interval. In fact, it is common for SMs to display this information to the user. In the following, we particularize the results in Theorem 1 and Theorem 2 to the case in which the average energy consumption of the user is specified. Specifically, we analyze the impact of the average energy consumption on the privacy performance. We define the average energy consumption of the random process X^n as

$$\mu_n = \mathbb{E}\left[\frac{1}{n}\sum_{i=0}^{n-1} X_i\right]. \tag{26}$$

Note that since we do not impose any stationarity condition on the random process X^n, the average energy consumption is a function of the time index n. This agrees with the non-stationary nature observed in energy consumption profiles of users [23].

A. Upper bound on the information leakage rate

The following result provides an upper bound on the information leakage rate for random processes X^n with average energy consumption μ_n.

Theorem 3: Consider a battery system with capacity β and initial state s_0. Let X^n be a random process with *average energy consumption* μ_n. Let P_β^* be a *block battery policy*, then for $l \geq \lfloor(\beta+1)/\alpha\rfloor$ at least one policy P_β^* exists such that

$$L_\beta\left(X^n, P_\beta^*\right) \leq \frac{\max\left(H_2\left(\frac{\mu_n - \frac{\beta}{n}}{\alpha}\right), H_2\left(\frac{\mu_n + \frac{\beta}{n}}{\alpha}\right)\right)}{\lfloor(\beta+1)/\alpha\rfloor}, \tag{27}$$

where $H_2(p) = -p\log_2 p - (1-p)\log_2(1-p)$ denotes the binary entropy.

Proof: The entropy of a random process Y^n taking values in $\mathcal{O}_l^m$ is upper bounded by

$$\frac{1}{n}H(Y^n) = \frac{1}{n}\sum_{i=0}^{m-1} H\left(Y_{il}, \ldots, Y_{(i+1)l-1} | Y_0, \ldots, Y_{il-1}\right) \tag{28}$$

$$\leq \frac{1}{n}\sum_{i=0}^{m-1} H\left(Y_{il}, \ldots, Y_{(i+1)l-1}\right), \tag{29}$$

where (28) follows by applying the chain rule and (29) follows from the fact that conditioning reduces entropy. Notice that (29) is the entropy of m sequences Y^l taking values in $\mathcal{O}_l$, and therefore, the entropy of Y^n is upper bounded by

$$\frac{1}{n}H\left(Y^n\right) \leq \frac{1}{l}H_2\left(\frac{\mathbb{E}\left[\frac{1}{n}\sum_{i=0}^{n-1} Y_i\right]}{\alpha}\right), \tag{30}$$

for the case in which each sequence Y^l is independent and identically distributed, i.e. with distribution

$$\mathbb{P}\left[Y^l = (\alpha, \alpha, \ldots, \alpha)\right] = \frac{\mathbb{E}\left[\frac{1}{n}\sum_{i=0}^{n-1} Y_i\right]}{\alpha}, \tag{31}$$

and

$$\mathbb{P}\left[Y^l = (0, 0, \ldots, 0)\right] = 1 - \mathbb{P}\left[Y^l = (\alpha, \alpha, \ldots, \alpha)\right]. \tag{32}$$

We now bound the average energy requested from the grid as a function of the average energy consumption of the user and the battery size. Dividing (1) by n and taking the expected value yields

$$\mathbb{E}\left[\frac{1}{n}\sum_{i=0}^{n-1} Y_i\right] = \mathbb{E}\left[\frac{1}{n}\sum_{i=0}^{n-1} X_i\right] + \mathbb{E}\left[\frac{S_n - s_0}{n}\right], \tag{33}$$

or equivalently

$$\mu_n - \frac{\beta}{n} \leq \mathbb{E}\left[\frac{1}{n}\sum_{i=0}^{n-1} Y_i\right] \leq \mu_n + \frac{\beta}{n}. \tag{34}$$

Notice now that for $l \leq (\beta+1)/\alpha$, and for every initial state $s_0 \in \mathcal{S}$ and input realization $\mathbf{x} \in \mathcal{X}^n$ there exists a sequence $\mathbf{y} \in \mathcal{O}_l^m$ such that $\mathbf{y}$ belongs to the set of *stable energy requests* $\mathcal{Y}_\beta^n(s_0, \mathbf{x})$. This completes the proof. ■

The upper bound on the information leakage rate when the average energy consumption of the user is known and $n \to \infty$ is illustrated in Fig. 4 and Fig. 5. As expected, the binary entropy term in Theorem 3 introduces concavity in the upper bound as shown in Fig. 4. Interestingly, Fig. 5 shows that the information leakage rate reduction as the size of the battery increases is less significant for extreme values of the average energy consumption.

B. Tightness of the upper bound

Proceeding in a similar fashion as in Section III we now prove that the upper bound in Theorem 3 is tight for a certain class of random processes modelling the energy consumption.

Theorem 4: Consider a battery system with capacity β and initial state s_0. Let $\hat{X}^n$ be a random process with *average energy consumption* μ_n and taking values in $\mathcal{O}_l^m$ with $l = \lceil(\beta+1)/\alpha\rceil$. Let P_β be a stable battery policy, then

$$L_\beta\left(\hat{X}^n, P_\beta\right) = \frac{1}{\lceil(\beta+1)/\alpha\rceil} H_2\left(\frac{\mu_n}{\alpha}\right). \tag{35}$$

Proof: Borrowing from (30) the entropy rate of the random process X^n taking values in $\mathcal{O}_l^m$ is upper bounded by

$$\frac{1}{n} H\left(X^n\right) \leq \frac{1}{l} H_2\left(\frac{\mathbb{E}\left[\frac{1}{n}\sum_{i=0}^{n-1} X_i\right]}{\alpha}\right), \tag{36}$$

with equality when the X^l symbols forming X^n are i.i.d. We now recall that when X^n takes values in $\mathcal{O}_l^m$ with $l > \beta/\alpha$ the input $\mathbf{x}^n$ can be uniquely determined from the output sequence $\mathbf{y}^n$ and $H\left(X^n|Y^n\right) = 0$. We conclude the proof by selecting $l = \lceil(\beta+1)/\alpha\rceil$. ■

V. Conclusion

We have studied the information leakage rate of EMSs with finite battery capacity for general random processes modelling the energy consumption of the user. Inspired by the results on permuting channels we have proposed a battery charging policy with bounded information leakage rate for arbitrary random processes. We have particularized the analysis to the case in which the average energy consumption of the user is known and we have concluded that extreme values of the average energy consumption provide lower values of information leakage to the utility provider.

References

[1] A. Ipakchi and F. Albuyeh, "Grid of the future," *IEEE Power Energy Mag.*, vol. 7, no. 2, pp. 52–62, Mar. 2009.

[2] R. Hartway, S. Price, and C.K. Woo, "Smart meter, customer choice and profitable time-of-use rate option," *Energy*, vol. 24, no. 10, pp. 895–903, Oct. 1999.

[3] G. Wood and M. Newborough, "Dynamic energy-consumption indicators for domestic appliances: environment, behaviour and design," *Energy and Buildings*, vol. 35, no. 8, pp. 821–841, 2003.

[4] E. L. Quinn, "Privacy and the new energy infrastructure," *Available at SSRN 1370731*, 2008.

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

[6] I. Rouf, H. Mustafa, M. Xu, W. Xu, R. Miller, and M. Gruteser, "Neighborhood watch: Security and privacy analysis of automatic meter reading systems," in *Proc. Conf. Computer and Commun. Security*, New York, NY, USA, 2012.

[7] G. W. Hart, "Nonintrusive appliance load monitoring," *Proc. IEEE*, vol. 80, no. 12, Dec. 1992.

[8] E. McKenna, I. Richardson, and M. Thomson, "Smart meter data: Balancing consumer privacy concerns with legitimate applications," *Energy Policy*, vol. 41, no. C, pp. 807–814, 2012.

[9] A. Molina-Markham, P. Shenoy, K. Fu, E. Cecchet, and D. Irwin, "Private memoirs of a smart meter," in *Proc. Workshop Embedded Sensing Systems Energy-Efficiency Building*, New York, NY, USA, 2010.

[10] M. Enev, S. Gupta, T. Kohno, and S. N. Patel, "Televisions, video privacy, and powerline electromagnetic interference," in *Proc. Conf. Computer Commun. Security*, New York, NY, USA, 2011.

[11] A. Prudenzi, "A neuron nets based procedure for identifying domestic appliances pattern-of-use from energy recordings at meter panel," in *IEEE Power Engineering Society Winter Meeting*, 2002.

[12] U. Greveler, B. Justus, and D. Loehr, "Forensic content detection through power consumption," in *Proc. IEEE Int. Conf. Commun.*, Ottawa, ON, Canada, Jun. 2012.

[13] H. Y. Lam, G. S. K. Fung, and W. K. Lee, "A novel method to construct taxonomy electrical appliances based on load signaturesof," *IEEE Trans. Consum. Electron.*, vol. 53, no. 2, pp. 653–660, May 2007.

[14] C. Cuijpers and B. Koops, *Smart Metering and Privacy in Europe: Lessons from the Dutch Case*, pp. 269–293, Springer Netherlands, 2012.

[15] A. Bartoli, J. Hernandez-Serrano, M. Soriano, M. Dohler, A. Kountouris, and D. Barthel, "Secure lossless aggregation over fading and shadowing channels for smart grid m2m networks," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 844–864, Dec. 2011.

[16] 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.*, Brussels, Belgium, Oct. 2011.

[17] F. Li, B. Luo, and P. Liu, "Secure information aggregation for smart grids using homomorphic encryption," in *Proc. IEEE Int. Conf. Smart Grid Commun.*, Oct. 2010.

[18] L. Zhu, Z. Zhang, Z. Qin, J. Weng, and K. Ren, "Privacy protection using a rechargeable battery for energy consumption in smart grids," *IEEE Trans. Network*, vol. 31, no. 1, pp. 59–63, Jan. 2017.

[19] J. Gomez-Vilardebo and D. Gündüz, "Privacy of smart meter systems with an alternative energy source," in *Proc. IEEE Int. Symp. Inf. Theory*, Istanbul, Turkey, Jul. 2013.

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

[21] Z. Li and T. J. Oechtering, "Privacy on hypothesis testing in smart grids," in *Proc. IEEE Inf. Theory Workshop*, Jeju, South Korea, Oct. 2015.

[22] 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.

[23] G. Kalogridis, C. Efthymiou, S. Z. Denic, T. A. Lewis, and R. Cepeda, "Privacy for smart meters: Towards undetectable appliance load signatures," in *Proc. IEEE Int. Conf. Smart Grid Commun.*, Gaithersburg, MD, USA, Oct. 2010.

[24] G. Kalogridis, R. Cepeda, S. Z. Denic, T. Lewis, and C. Efthymiou, "Elecprivacy: Evaluating the privacy protection of electricity management algorithms," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 750–758, Dec. 2011.

[25] I. Esnaola, A. M. Tulino, and H. V. Poor, "A statistical physics approach to the wiretap channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Istanbul, Turkey, Jul. 2013.

[26] L. Sankar, S. Raj 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.

[27] G. Giaconi and D. Gündüz, "Smart meter privacy with renewable energy and a finite capacity battery," in *Proc. IEEE Int. Workshop Sig. Process. Advances Wireless Commun.*, Edinburgh, UK, Jul. 2016.

[28] R. Ahlswede and A. Kaspi, "Optimal coding strategies for certain permuting channels," *IEEE Trans. Inf. Theory*, vol. 33, no. 3, pp. 310–314, May 1987.

[29] R. B. Ash, *Information Theory*, Dover Publications Inc., New York, 1990.