
Proceedings Paper:
Arrieta, M., Esnaola, I. orcid.org/0000-0001-5597-1718 and Effros, M. (2019) Universal privacy guarantees for smart meters. In: 2019 IEEE International Symposium on Information Theory (ISIT). 2019 IEEE International Symposium on Information Theory (ISIT), 07-12 Jul 2019, Paris, France. IEEE , pp. 2154-2158. ISBN 9781538692929

https://doi.org/10.1109/isit.2019.8849679

Universal Privacy Guarantees for Smart Meters

Miguel Arrieta*, Iñaki Esnaola*†, and Michelle Effros§
*Dept. of Automatic Control and Systems Engineering, University of Sheffield, Sheffield S1 3JD, UK
†Dept. of Electrical Engineering, Princeton University, Princeton, NJ 08544, USA
§Dept. of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125, USA

***Abstract*—Smart meters enable improvements in electricity distribution system efficiency at some cost in customer privacy. Users with home batteries can mitigate this privacy loss by applying charging policies that mask their underlying energy use. A battery charging policy is proposed and shown to provide universal privacy guarantees subject to a constraint on energy cost. The guarantee bounds our strategy's maximal information leakage from the user to the utility provider under general stochastic models of user energy consumption. The policy construction adapts coding strategies for non-probabilistic permuting channels to this privacy problem.**

I. Introduction

Smart meters (SMs) provide advanced monitoring of consumer energy usage, thereby enabling optimized management and control of electricity distribution systems [1]. Unfortunately, the data collected by SMs can reveal information about consumers' activities. For instance, an individual's energy usage pattern may leak information about the times at which they run individual appliances [2]. Two approaches have been proposed to tackle the privacy threat posed by such information leakage. One strategy involves manipulating user data before sending it to the utility provider (UP) [3]; this approach improves privacy at the cost of reduced operational insight for the UP. The other strategy employs rechargeable batteries at each consumer site to try to decouple energy usage from energy requests [4]; allowing devices to run off of either the battery or the UP and allowing the battery to charge at times of both activity and inactivity improves privacy at the cost of introducing individual batteries and, potentially, increasing consumer costs (e.g., if energy is requested when it best conceals the consumers' usage without regard to the energy bill). This paper investigates the latter approach.

Understanding the privacy implications of any strategy requires an appropriate privacy metric. A variety of metrics are used to study privacy in energy distribution systems. These include statistical distance metrics [4], differential privacy [5], distortion metrics [6], and information metrics like mutual information, which can be applied under a variety of assumptions on users' energy, including i.i.d. [7], [4], [8], [9], [10], stationary [11], [12], and first-order time-homogeneous Markov random processes [13]; see [14] for a comprehensive review. Alternative privacy metrics such as maximal leakage [15] have operational descriptions and relate to information measures like Sibson mutual information; its generalization, maximal α-leakage [16], establishes additional relationships to Arimoto mutual information, mutual information, and Renyi entropy [15], [16]. Many of these measures can be understood as measures of an adversary's ability to gain insight into an unknown random variable X by observing Y, with measures differing only in the loss functions they use to quantify that insight [16].

We here use mutual information to measure privacy both because its interpretation in terms of an adversary that minimizes log-loss with respect to an evolving soft-decision model [16] is well-matched to the evolving nature of energy distribution over time and because mutual information provides a useful bridge to adjacent fields such as hypothesis testing [17], estimation [18], and learning [19].

Since user energy consumption may be non-stationary, we seek privacy guarantees that apply across general random process models of energy consumption. Moreover, given that no battery can store unlimited energy, we impose finite capacity bounds on batteries. We therefore model the energy management unit (EMU) as a deterministic finite-state channel. We then adapt the Ahlswede-Kaspi coding strategy proposed for permuting channels [20] to the SM privacy setting. This work generalizes the battery policy proposed in [21] by including the price of the energy requested from the grid and minimizing information leakage subject to a bound on the resulting energy bill.

We denote vectors by bold letters, e.g. $\mathbf{x}$, and random variables by uppercase letters, e.g. X. The operator $\sigma(\cdot)$ denotes the sum over vector elements, e.g. $\sigma(\mathbf{x}) = \sum_i x_i$. Intervals on the integers are denoted by double brackets, e.g. $[\![a, b]\!] = \{a, a+1, \ldots, b-1, b\}$. The n-fold cartersian product of the interval is denoted by $[\![a, b]\!]^n = [\![a, b]\!] \times \ldots \times [\![a, b]\!]$. Given a vector $\mathbf{x}$ of size n and a set of indices $\mathcal{A} \subseteq [\![1, n]\!]$, we denote by $\mathbf{x}_\mathcal{A}$ the vector $\mathbf{x}_\mathcal{A} = \{x_i : i \in \mathcal{A}\}$. The support of the probability distribution P_X is denoted by $\text{supp}(P_X)$, and the positive part operator is $(a)^+ = \max(0, a)$.

II. Energy Management System with a Finite Battery Model

Figure 1 depicts an energy management system and the random processes therein. The privacy guarantee is defined in terms of the information leakage from the user to the provider, and the task of the EMU is to choose a battery policy that minimizes the leakage while satisfying the operation and cost constraints. Formal definitions follow.

We model user energy consumption as a discrete-time random process X^n on alphabet $\mathcal{X}^n = [\![0, \alpha]\!]^n$. The random

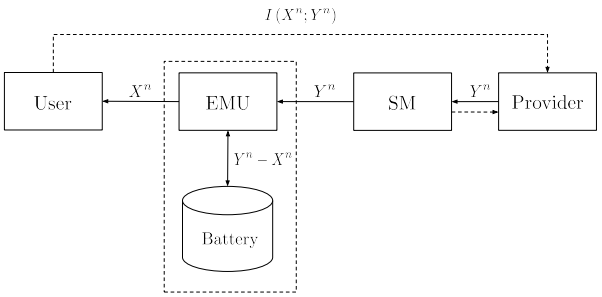

Fig. 1. Energy Management System with Finite Battery Model

variable X_i describes the energy consumed by the user at time step i with $i = 0, 1, ..., n-1$. For exposition simplicity we assume $\mathcal{X} \subseteq \mathbb{Z}$; the results generalize to arbitrary discrete alphabets. We use $P_{X^n} \in \mathcal{P}_{X^n}$ to denote the energy consumption pattern distribution, where $\mathcal{P}_{X^n}$ is a fixed family of such distributions. Since user energy consumption profiles tend to exhibit non-stationarities [4], $\mathcal{P}_{X^n}$ may contain non-stationary random processes.

The EMU maps consumption sequence $X^n \in \mathcal{X}^n$ to a request sequence $Y^n \in \mathcal{Y}^n$ using a battery policy $P_{Y^n|X^n}$ that is not allowed to vary with X^n; random variable Y_i describes the energy requested from the UP at time step $i = 0, 1, ..., n-1$. We again focus on integer random variables ($\mathcal{Y} \subseteq \mathbb{Z}$) for simplicity. We require $\mathcal{Y} \supseteq \mathcal{X}$ so that the UP can satisfy the user's energy consumption even when no battery is available. We allow $\mathcal{Y}$ to contain negative values to model scenarios where users can sell energy back to the grid.

To be considered feasible, battery policy $P_{Y^n|X^n}$ must create a request sequence that meets the energy demands of the user and does not request energy it cannot use or store. Let β denote the finite capacity of a given battery (in energy units) and S_i denote the amount of energy stored in that battery, the "energy state," at time i. Then S_i takes values in $\mathcal{S} = [\![0, \beta]\!]$ and is governed by the charging dynamics

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

where $s_0 \in \mathcal{S}$ is the initial battery state. A power outage occurs when $S_i + Y_i - X_i < 0$; energy is wasted when $S_i + Y_i - X_i > \beta$. Under this model, the battery resembles a box, energy units resemble balls that can be inserted (stored) and removed (consumed), and the set $\mathcal{Y}^n(s_0, \mathbf{x})$ of feasible requests, defined formally below, contains all sequences of insertions and removals allowed by the box. This feasibility constraint resembles [20][Eq. 2.4] from the work of Ahlswede and Kaspi; this link is studied in [21].

Definition 1: Given a battery with initial state $s_0 \in \mathcal{S}$ and capacity β, the *set of feasible energy requests* for energy consumption sequence $\mathbf{x} \in \mathcal{X}^n$ is

$$\mathcal{Y}^n(s_0, \mathbf{x}) \stackrel{\Delta}{=} \{\mathbf{y} \in \mathcal{Y}^n : s_i \in [\![0, \beta]\!] \ \ \forall\, i \in [\![0, n]\!]\}. \tag{2}$$

The *set of feasible battery policies* is

$$\Omega(s_0) \stackrel{\Delta}{=} \{P_{Y^n|X^n} : \text{supp}(P_{Y^n|X^n=\mathbf{x}}) \subseteq \mathcal{Y}^n(s_0, \mathbf{x}) \ \ \forall\, \mathbf{x} \in \mathcal{X}^n\}. \tag{3}$$

Our aim in feasible policy design is to minimize privacy subject to a constraint on policy cost. Towards this end, we next define our measures of information leakage (where privacy is high when information leakage is low) and cost.

We measure a battery policy's information leakage by its worst-case performance.

Definition 2: The *information leakage* of policy $P_{Y^n|X^n}$ is

$$\bar{\mathcal{I}}(P_{Y^n|X^n}) = \max_{P_{X^n} \in \mathcal{P}_{X^n}} \frac{1}{n} I(X^n; Y^n). \tag{4}$$

We measure the cost of a policy $P_{Y^n|X^n}$ as the difference between the user's energy bill under that policy and the user's energy bill under the feasible battery policy that minimizes the energy bill. (Under this definition, cost can be negative only for infeasible policies.) To calculate energy bills, we model the energy market price as a deterministic sequence, $\mathbf{m} \in \mathbb{R}^n$. Under this definition, the cost of an energy request sequence $\mathbf{y}$ is $\mathbf{m}^T\mathbf{y}$. We assume that the market price is constant over each of K blocks of time. The price and duration of the k-th block, $k = 0, 1, \ldots, K-1$, are m_k and l_k, respectively, giving

$$\mathbf{m} = (\underbrace{m_0, \ldots, m_0}_{l_0}, \underbrace{m_1, \ldots, m_1}_{l_1}, \ldots, \underbrace{m_{K-1}, \ldots, m_{K-1}}_{l_{K-1}}). \tag{5}$$

Definition 3: Consider an EMU with battery capacity β, initial state $s_0 \in \mathcal{S}$, and market price $\mathbf{m}$. The *system cost of energy consumption sequence* $\mathbf{x} \in \mathcal{X}^n$ *under battery policy* $P_{Y^n|X^n}$ is

$$g(Y^n, \mathbf{x}) = \mathbb{E}_{P_{Y^n|X^n=\mathbf{x}}}[\mathbf{m}^T Y^n - \mathbf{m}^T \mathbf{y}^*(\mathbf{x})], \tag{6}$$

where $\mathbf{y}^*(\mathbf{x}) = \text{argmin}_{\mathbf{y} \in \mathcal{Y}^n(s_0, \mathbf{x})} \mathbf{m}^T\mathbf{y}$. For any $\Delta \geq 0$, the set of *feasible Δ-affordable battery policies* is

$$\Gamma(\Delta) \stackrel{\Delta}{=} \left\{P_{Y^n|X^n} \in \Omega(s_0) : g(Y^n, \mathbf{x}) \leq \Delta \ \ \forall\, \mathbf{x} \in \mathcal{X}^n\right\}. \tag{7}$$

Finally, the privacy-cost function defines the optimal trade-off between privacy and cost achievable by feasible battery policies.

Definition 4: Given an EMU with battery capacity β, initial state s_0 and market price $\mathbf{m}$, the *privacy cost function* is defined, for each $\Delta \geq 0$, as

$$\mathcal{I}(\Delta) \stackrel{\Delta}{=} \min_{P_{Y^n|X^n} \in \Gamma(\Delta)} \bar{\mathcal{I}}(P_{Y^n|X^n}). \tag{8}$$

To bound $\mathcal{I}(\Delta)$, we adapt techniques developed by Ahlswede and Kaspi [20] from channel capacity to privacy-cost. While the resulting solution employs a non-causal battery policy, detailed analysis of [20] shows that knowing just $\beta+1$ time steps ahead suffices to achieve optimality, where β is the battery capacity. Thus, we envision practical implementations that rely on consumption predictions. This approach also provides insight on what prediction capabilities are needed.

III. Geometry of the Feasible Sets

A. Shared Output Sequences

Lemma 1 characterizes a necessary and sufficient condition under which a set $\mathcal{A}$ of input pairs $(s_0, \mathbf{x})$ share a common feasible output sequence $\mathbf{y}_\mathcal{A}$. Such shared output sequences are good for privacy since a UP that sees $\mathbf{y}_\mathcal{A}$ cannot distinguish which input pair $(s_0, \mathbf{x}) \in \mathcal{A}$ caused it. Conversly, when two inputs $(s_0, \mathbf{x}), (\hat{s}_0, \hat{\mathbf{x}})$ share no feasible output $\mathbf{y}_\mathcal{A}$, the EMU cannot hide from the UP which pair caused the request. The following measure of distance is useful for that analysis.

Definition 5: The distance between two input pairs $(s_0, \mathbf{x}), (\hat{s}_0, \hat{\mathbf{x}}) \in \mathcal{S} \times \mathcal{X}^n$ is defined as

$$d_n\Big((s_0,\mathbf{x}),(\hat{s}_0,\hat{\mathbf{x}})\Big) = \max_{i\in[\![0,n-1]\!]} \Big|\big(s_0-\sigma(\mathbf{x}^i)\big)-\big(\hat{s}_0-\sigma(\hat{\mathbf{x}}^i)\big)\Big|. \quad (9)$$

Lemma 1 shows that the distance between input pairs determines the existence of a shared feasible output $\mathbf{y}$. The result emphasizes the central role that battery capacity β plays in privacy.

Lemma 1: Let $\mathcal{A}$ denote a subset of the input pair alphabet $\mathcal{S} \times \mathcal{X}^n$. The following two statements are equivalent.

a) The distance between every two pairs $(s_0, \mathbf{x}), (\hat{s}_0, \hat{\mathbf{x}}) \in \mathcal{A}$ is less than or equal to the capacity of the battery, i.e.

$$d_n\Big((s_0, \mathbf{x}), (\hat{s}_0, \hat{\mathbf{x}})\Big) \leq \beta \text{ for all } (s_0, \mathbf{x}), (\hat{s}_0, \hat{\mathbf{x}}) \in \mathcal{A}. \quad (10)$$

b) All sequences in $\mathcal{A}$ share a feasible request $\mathbf{y}_\mathcal{A}$, i.e.

$$\mathbf{y}_\mathcal{A} \in \bigcap_{(s_0,\mathbf{x})\in\mathcal{A}} \mathcal{Y}^n(s_0, \mathbf{x}). \quad (11)$$

Proof: Let the sequence $\mathbf{y}_\mathcal{A}$ be such that for all i:

$$\sigma(\mathbf{y}^i_\mathcal{A}) = -\min_{(s_0,\mathbf{x})\in\mathcal{A}} (s_0 - \sigma(\mathbf{x}^i)). \quad (12)$$

Thus, for any $(\hat{s}_0, \hat{\mathbf{x}}) \in \mathcal{A}$, the battery state at time $i+1$ is

$$s_{i+1} = (\hat{s}_0 - \sigma(\hat{\mathbf{x}}^i)) - \min_{(s_0,\mathbf{x})\in\mathcal{A}} (s_0 - \sigma(\mathbf{x}^i)). \quad (13)$$

Now $d_n\big((s_0, \mathbf{x}), (\hat{s}_0, \hat{\mathbf{x}})\big) \leq \beta$ implies that $s_{i+1} \in [\![0, \beta]\!]$ for all i, so $\mathbf{y}_\mathcal{A}$ is a feasible sequence. The converse follows since for any sequence $\mathbf{y}$ and any two input pairs $(s_0, \mathbf{x}), (\hat{s}_0, \hat{\mathbf{x}}) \in \mathcal{A}$ such that $d_n\big((s_0, \mathbf{x}), (\hat{s}_0, \hat{\mathbf{x}})\big) > \beta$, the absolute difference between the corresponding battery states at some time step i satisfies

$$\big|s_{i+1} - \hat{s}_{i+1}\big| = \big|(s_0 - \sigma(\mathbf{x}^i)) - (\hat{s}_0 - \sigma(\hat{\mathbf{x}}^i))\big| > \beta. \quad (14)$$

Thus s_{i+1} and $\hat{s}_{i+1}$ cannot both belong to $\mathcal{S} = [\![0, \beta]\!]$. ■

B. Cardinality bounds

Building on Lemma 1, Theorem 1 gives an upper bound on the number of distinguishable input pairs $(s_0, \mathbf{x}^n) \in \mathcal{S}_0 \times \mathcal{X}^n$, where $\mathcal{S}_0 \subseteq \mathcal{S}$ is the set of possible initial battery states. The result is derived by building a covering $\{\mathcal{A}_i\}$ of $\mathcal{S}_0 \times \mathcal{X}$ such that all input pairs in each $\mathcal{A}_i$ share a common feasible request $\mathbf{y}_i$. The result shows that the minimal time $\lambda \triangleq \lfloor(\beta+1)/\alpha\rfloor$ needed to fully discharge a battery of capacity β under maximal consumption $\alpha \triangleq \max \mathcal{X}$ is a central parameter in the construction of privacy preserving battery policies. The proof is inspired by the code construction presented by Ahlswede and Kaspi [20, Proposition 1].

Theorem 1: Let the input alphabet be $\mathcal{S}_0 \times \mathcal{X}^n$, with $\overline{\mathcal{S}_0}$ and $\underline{\mathcal{S}_0}$ denoting the maximum and minimum values of $\mathcal{S}_0$, respectively. There exists a set of request sequences $\mathcal{V}^n(\mathcal{S}_0) \subseteq \mathcal{Y}^n$ such that

$$\log\big|\mathcal{V}^n(\mathcal{S}_0)\big| \leq \left\lceil \frac{n - \big\lfloor(\beta + \underline{\mathcal{S}_0} - \overline{\mathcal{S}_0})/\alpha\big\rfloor}{\lambda} \right\rceil. \quad (15)$$

Moreover, for every input pair $(s_0, \mathbf{x}) \in \mathcal{S}_0 \times \mathcal{X}^n$, at least one sequence $\mathbf{v} \in \mathcal{V}^n(\mathcal{S}_0)$ is feasible, that is

$$\mathcal{Y}^n(s_0, \mathbf{x}) \cap \mathcal{V}^n(\mathcal{S}_0) \neq \emptyset. \quad (16)$$

Proof: At time step i, the value of $s_0 - \sigma(\mathbf{x}^i)$ for any input pair $(s_0, \mathbf{x}) \in \mathcal{S}_0 \times \mathcal{X}^i$ with $\mathcal{X} = [\![0, \alpha]\!]$ is bounded by

$$\underline{\mathcal{S}_0} - i\alpha \leq s_0 - \sigma(\mathbf{x}^i) \leq \overline{\mathcal{S}_0}. \quad (17)$$

At time step $l = \big\lfloor(\beta + \underline{\mathcal{S}_0} - \overline{\mathcal{S}_0})/\alpha\big\rfloor$, the distance between any two input pairs $(s_0, \mathbf{x}), (\hat{s}_0, \hat{\mathbf{x}}) \in \mathcal{S}_0 \times \mathcal{X}^l$ is bounded by

$$d_l\Big((s_0, \mathbf{x}), (\hat{s}_0, \hat{\mathbf{x}})\Big) \leq \overline{\mathcal{S}_0} - (\underline{\mathcal{S}_0} - l\alpha) \leq \beta. \quad (18)$$

Therefore, Lemma 1 guarantees the existence of a request $\mathbf{y}_0$ that is feasible for every input pair in $\mathcal{S}_0 \times \mathcal{X}^l$. Following a similar reasoning, consider the set of possible input pairs during the subsequent λ times steps, i.e. $\mathcal{S} \times \mathcal{X}^\lambda$ with $\mathcal{S} = [\![0, \beta]\!]$. Define a cover of the input alphabet, $\mathcal{S} \times \mathcal{X}^\lambda \subseteq (\mathcal{A}_1 \bigcup \mathcal{A}_2)$, with subsets given by

$$\mathcal{A}_1 = \left\{(s_0, \mathbf{x}) \in \mathcal{S} \times \mathcal{X}^\lambda : s_0 - \sigma(\mathbf{x}) \in [\![0, \beta]\!]\right\}, \quad (19)$$

and

$$\mathcal{A}_2 = \left\{(s_0, \mathbf{x}) \in \mathcal{S} \times \mathcal{X}^\lambda : s_0 - \sigma(\mathbf{x}) \in [\![-\lambda\alpha, -1]\!]\right\}. \quad (20)$$

Note $\mathcal{A}_1 \bigcup \mathcal{A}_2$ contains all sequences in $\mathcal{S}\times\mathcal{X}^\lambda$ as (17) implies that $s_0 - \sigma(\mathbf{x}) \in [\![-\lambda\alpha, \beta]\!]$. The distance between any two input pairs in $\mathcal{A}_i$ with $i = 1, 2$ is bounded by β. Therefore, by Lemma 1, there exists a shared feasible sequence $\mathbf{y}_i$ for all pairs in $\mathcal{A}_i$. Setting $\kappa = \lceil(n-l)/\lambda\rceil$ and

$$\mathcal{V}^n(\mathcal{S}_0) = \{\mathbf{y}_0\} \times \underbrace{\{\mathbf{y}_1, \mathbf{y}_2\} \times \ldots \times \{\mathbf{y}_1, \mathbf{y}_2\}}_{\kappa} \quad (21)$$

completes the proof. ■

To map input pairs $(s_0, \mathbf{x})$ to energy request in $\mathcal{V}^n(\mathcal{S}_0)$ it suffices to forecast, at the start of each block of length λ, whether the battery will deplete during the current block, i.e. $s_0 - \sigma(\mathbf{x}^\lambda) \lesseqgtr 0$. In [22], it is shown that the upper bound in Theorem 1 is tight. The construction of the set of request sequences given by (21) describes the forecasting capabilities required to implement optimal battery policies.

C. Impact of the Output Alphabet on Information Leakage

Lemma 2 shows that the privacy cost function $\mathcal{I}(\Delta)$ does not vary when the EMU operates with a constrained output alphabet $\mathcal{Y}_c$. This result is consistent with prior results reported for privacy based on hypothesis testing [23, Theorem 1] and multi-user scenarios [24, Theorem 2].

Lemma 2: Define output alphabet $\mathcal{Y}_c^n = [\![-\beta/\underline{l}, \beta/\underline{l} + \alpha]\!]^n$ where $\underline{l} = \min_k l_k$ and l_k is the length of the k-th market price period as defined in (5). Let $\mathcal{I}(\Delta)$ and $\mathcal{I}_c(\Delta)$ represent the privacy-cost functions under output alphabets $\mathcal{Y}^n$ and $\mathcal{Y}_c^n$ for any output alphabet $\mathcal{Y}^n \supset \mathcal{Y}_c^n$. Then

$$\mathcal{I}_c(\Delta) = \mathcal{I}(\Delta). \tag{22}$$

Proof: Let $\Gamma(\Delta)$ and $\Gamma_c(\Delta)$ denote the set of feasible Δ-affordable battery policies under output alphabets $\mathcal{Y}^n$ and $\mathcal{Y}_c^n$. It follows from [22] that a function $F : \mathcal{Y}^n \to \mathcal{Y}_c^n$ exists such that if $P_{Y^n|X^n} \in \Gamma(\Delta)$ then $F \circ P_{Y^n|X^n} \in \Gamma_c(\Delta)$. Noting that the function F induces the Markov chain

$$X^n \to Y^n \to F(Y^n) \tag{23}$$

yields $I(X^n; F_n(Y^n)) \leq I(X^n; Y^n)$ by the data processing inequality. The converse follows by noting that $\Gamma_c(\Delta) \subseteq \Gamma(\Delta)$. ■

We note that the proof for the existence of the function F presented in [22] requires forecasting of $\underline{l}$ time steps ahead.

IV. Universal Privacy bounds

In the following, we bound the information leakage given in Definition 4. We first study the case for which only the feasibility constraint is imposed.

Theorem 2: The privacy cost function $\mathcal{I}(\infty)$ is bounded by

$$\mathcal{I}(\infty) \leq \frac{1}{n} \left\lceil \frac{n - \lfloor \beta/\alpha \rfloor}{\lambda} \right\rceil. \tag{24}$$

Proof: Theorem 1 shows the existence of a set $\mathcal{V}^n(\{s_0\})$ with cardinality bounded by

$$\log|\mathcal{V}^n(\{s_0\})| \leq \left\lceil \frac{n - \lfloor (\beta + s_0 - s_0)/\alpha \rfloor}{\lambda} \right\rceil = \left\lceil \frac{n - \lfloor \beta/\alpha \rfloor}{\lambda} \right\rceil, \tag{25}$$

such that the intersection $\mathcal{V}^n(\{s_0\}) \cap \mathcal{Y}(s_0, \mathbf{x})$ is not empty for every input pair $(s_0, \mathbf{x})$. Letting the output Y^n take values in $\mathcal{V}^n(\{s_0\}) \cap \mathcal{Y}(s_0, \mathbf{x})$ completes the proof. ■

Theorem 3 presents our main result, where we bound the information leakage for arbitrary cost constraints Δ. The proof proceeds by constructing a battery policy that combines two components for every request sequence. One of the components guarantees the feasibility constraint, while the other guarantees the cost constraint.

Theorem 3: Consider an EMU with battery capacity β, initial state s_0, market price $\mathbf{m}$, and output alphabet $\mathcal{Y}^n$ satisfying $\mathcal{Y}_c^n \subseteq \mathcal{Y}^n$ with $\mathcal{Y}_c^n$ defined in Lemma 2, then

$$\mathcal{I}(\Delta) \leq \mathcal{I}(\infty) + \mathcal{I}_\Gamma(\Delta), \tag{26}$$

where

$$\mathcal{I}_\Gamma(\Delta) = \min_{P_{\hat{S}_\gamma|\hat{S}_\omega} \in \Gamma_\omega(\Delta)} \max_{P_{\hat{S}_\omega} \in \mathcal{P}_{\hat{S}_\omega}} \frac{1}{n} I(\hat{S}_\gamma - \hat{S}_\omega; \hat{S}_\omega). \tag{27}$$

Here $\hat{S}_\omega$ and $\hat{S}_\gamma$ are random processes in $[\![0, \beta]\!]^K$ with joint distribution determined by

$$\Gamma_\omega(\Delta) = \left\{ P_{\hat{S}_\gamma|\hat{S}_\omega} : \mathbb{E}(\hat{S}_\gamma \boldsymbol{\delta}) \leq \Delta - \beta\sigma((\boldsymbol{\delta})^+) \right\}, \tag{28}$$

where $\boldsymbol{\delta} \in \mathbb{Z}^K$ denotes the vector of market price differences, with entries given by $\boldsymbol{\delta}_0 = -m_0$, $\boldsymbol{\delta}_k = m_{k-1} - m_k$ for $k = 1, 2, \ldots, K-1$ and $\boldsymbol{\delta}_K = m_{K-1}$.

Proof: We prove the result for $\mathcal{Y}^n = \mathbb{Z}^n$; Lemma 2 generalizes the proof for every $\mathcal{Y}^n$ satisfying $\mathcal{Y}_c^n \subseteq \mathcal{Y}^n$. The proof follows by dividing the optimization process into two steps. In the first step, we present a battery policy ω such that the resulting request sequence V_ω^n satisfies the power outage and energy waste constraints, i.e., $\omega \in \Omega(s_0)$ as defined in (3). These policies are discussed on Theorem 2. In the second step, we define a random vector V_γ^n such that $Y^n = V_\omega^n + V_\gamma^n$ also satisfies the cost constraints. Specifically, we set

$$V_\gamma^n = \sum_{t \in \mathcal{T}} \Big((\mathbf{e}_t - \mathbf{e}_{t+1})(S_\gamma - S_\omega)_t \Big), \tag{29}$$

where $\mathcal{T}$ denotes the ordered set of time steps at which a market transition takes place, i.e., $\mathcal{T} = \{0, l_0, l_0+l_1, \ldots, n-1\}$. This implies that

$$\begin{aligned} g(Y^n, \mathbf{x}) &= \mathbb{E}[(S_\gamma)_\mathcal{T} \boldsymbol{\delta} + \mathbf{m}^T \mathbf{x} - \mathbf{m}^T \mathbf{y}^*(\mathbf{x})] && (30) \\ &= \mathbb{E}[(S_\gamma)_\mathcal{T} \boldsymbol{\delta}] + \beta\sigma((\boldsymbol{\delta})^+), && (31) \end{aligned}$$

where (30) follows by (29) and the battery charging dynamics (1) and (31) follow by noting that $\mathcal{Y}^n = \mathbb{Z}^n$. Selecting the transformation γ determining $(S_\gamma)_\mathcal{T}$ from the set described in (28) yields

$$\begin{aligned} I(X^n; Y^n) &\leq I(X^n; V_\omega^n) + I(X^n; V_\gamma^n | V_\omega^n) && (32) \\ &= I(X^n; V_\omega^n) + H(V_\gamma^n | V_\omega^n) - H(V_\gamma^n | V_\omega^n, X^n, S_\omega) && (33) \\ &= I(X^n; V_\omega^n) + H(S_\gamma - S_\omega | V_\omega^n) - H(S_\gamma - S_\omega | S_\omega) && (34) \\ &\leq I(X^n; V_\omega^n) + I(S_\gamma - S_\omega; S_\omega), && (35) \end{aligned}$$

where (33) follows as X^n and V_ω^n determine S_ω by the battery charging dynamics (1); (34) follows by (29) and noting that $S_\gamma - S_\omega$ is independent of V_ω^n and X^n given S_ω. Thus

$$\begin{aligned} n\mathcal{I}(\Delta) &= \min_{P_{Y^n|X^n} \in \Gamma(\Delta)} \max_{\mathcal{P}_{X^n}} I(X^n; Y^n) && (36) \\ &\leq \min_{\gamma \in \Gamma_\omega(\Delta)} \min_{\omega \in \Omega(s_0)} \max_{\mathcal{P}_{X^n}} \Big(I(X^n; V_\omega^n) + I(S_\gamma - S_\omega; S_\omega) \Big) && (37) \\ &\leq \min_{\omega \in \Omega(s_0)} \max_{\mathcal{P}_{X^n}} I(X^n; V_\omega^n) + \min_{\gamma \in \Gamma_\omega(\Delta)} \max_{\mathcal{P}_{S_\omega}} I(S_\gamma - S_\omega; S_\omega). && (38) \end{aligned}$$

This completes the proof. ■

While direct computation of the information leakage in (8) relies on finding an n-dimensional joint distribution satisfiying $\Gamma(\Delta)$, the bound presented in (26) relies on a K-dimensional distribution and the simplified version of $\Gamma(\Delta)$ defined in (28). This significantly eases the computation of the information leakage as described in Section V. Note also that (27) implies that $\mathcal{I}_\Gamma(0) \leq |\mathcal{S}_\omega| = K/n \log_2(\beta+1)$ and $\mathcal{I}_\Gamma(\Delta) = 0$ for any

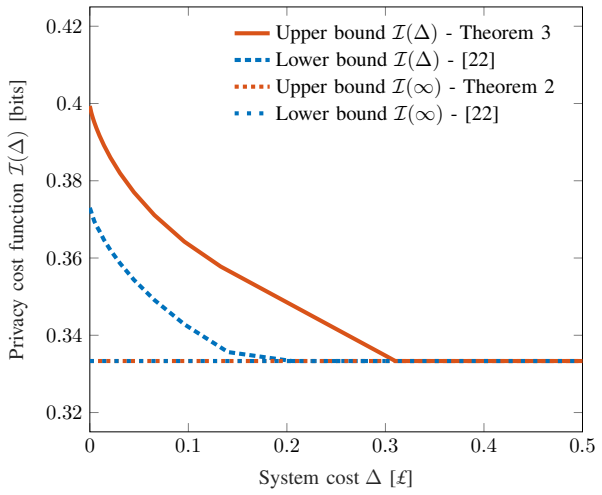

Fig. 2. Upper and lower bounds on the privacy cost function as a function of the privacy budget.

$\Delta \geq \Delta_{\max}$ with $\Delta_{\max} = \beta\|\boldsymbol{\delta}\|_1 - \beta m_0$. Interestingly, a time-sharing argument presented in [22] yields

$$\mathcal{I}(\Delta) \leq \frac{1}{n}\left\lceil \frac{n - \lfloor \beta/\alpha \rfloor}{\lambda} \right\rceil + \left(1 - \frac{\Delta}{\Delta_{\max}}\right)^+ \frac{K}{n} \log_2(\beta+1). \quad (39)$$

V. Numerical results

In this section, we numerically assess the upper bounds on the privacy cost described in Theorem 2 and Theorem 3. For comparison purposes, we also include the lower bounds on the privacy cost given in [22]. We model the market price after the UK Economy 7 tariff, where users are charged an off-peak price of 0.071 £/kWh within a 7 hour block and a peak price of 0.152 £/kWh otherwise [25]. We assume the user has an LG Chem RESU 6.5 battery with a capacity of 4.2 kWh and a peak power of 4.2 kW. For simplicity we match the users' maximum power consumption to the peak power of the battery, i.e., 4.2 kW [14]. The SM sends the UP integrated energy readings every 30 min following UK specifications for SMs [14]. Thus, we set the time elapsed between time steps i and $i+1$ to 30 min. Defining 2.1 kWh as 1 unit of energy yields the following parameters in our system model: battery capacity $\beta = 4.2$ kWh$/2.1$ kWh $= 2$; maximum consumption between time steps $\alpha = 4.2$ kW $\times\, 0.5$ h$/2.1$ kWh $= 1$; market lengths $l_0 = 7$ h$/0.5$ h $= 14$ and $l_1 = 17$ h$/0.5$ h $= 34$; corresponding market prices of $m_0 = 0.152$ £$/$ kWh $\times\, 2.1$ kWh $= 0.3192$ £ and $m_1 = 0.071$£$/$ kWh $\times\, 2.1$ kWh $= 0.1791$ £ per unit of energy.

Figure 2 depicts the bounds on the privacy cost $\mathcal{I}(\Delta)$ for different values of the system cost Δ and initial battery state $s_0 = 0$ during a one day period, i.e. $n = 24$ h$/0.5$ h $= 48$. Following (39), when the user does not wish to increase the system cost for privacy, the privacy cost is bounded by $\mathcal{I}(0) = 0.4$ bits. For large values of the system cost Δ the cost constraint is always satisfied, i.e. $\mathcal{I}_\Gamma(\Delta) = 0$, and the privacy leakage is governed by the feasibility constraints.

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

[3] 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.

[4] 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.

[5] J. Zhao, T. Jung, Y. Wang, and X. Li, "Achieving differential privacy of data disclosure in the smart grid," in *Proc. IEEE Conf. on Computer Communications*, Apr. 2014, pp. 504–512.

[6] G. Giaconi, D. Gündüz, and H. V. Poor, "Joint privacy-cost optimization in smart electricity metering systems," *arXiv preprint arXiv:1806.09715*, 2018.

[7] 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, pp. 1932–1935.

[8] 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.

[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. Conf. Smart Grid Commun.*, Tainan, Taiwan, Nov. 2012, pp. 664–669.

[10] J. Gómez-Vilardebó and D. Gündüz, "Privacy of smart meter systems with an alternative energy source," in *Proc. IEEE Int. Symp. Inf. Theory*, Istanbul, Turkey, Jul. 2013, pp. 2572 –2576.

[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.*, Brussels, Belgium, Oct. 2011, pp. 190–195.

[12] 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.

[13] 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.

[14] G. Giaconi, D. Gündüz, and H. V. Poor, "Privacy-aware smart metering: Progress and challenges," *IEEE Signal Process. Mag.*, vol. 35, no. 6, pp. 59–78, Nov. 2018.

[15] I. Issa, S. Kamath, and A. B. Wagner, "An operational measure of information leakage," in *Proc. Annu. Conf. on Information Science and Systems*, Princeton, NJ, USA, Mar. 2016, pp. 234–239.

[16] J. Liao, O. Kosut, L. Sankar, and F. du Pin Calmon, "A tunable measure for information leakage," *arXiv preprint arXiv:1806.03332*, 2018.

[17] H. V. Poor and S. Verdú, "A lower bound on the probability of error in multihypothesis testing," *IEEE Trans. Inf. Theory*, vol. 41, no. 6, pp. 1992–1994, Nov. 1995.

[18] D. Guo, S. Shamai, and S. Verdú, "Mutual information and minimum mean-square error in gaussian channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1261–1282, Apr. 2005.

[19] A. Xu and M. Raginsky, "Information-theoretic analysis of generalization capability of learning algorithms," in *Proc. Advances in Neural Information Processing Systems*, 2017, pp. 2524–2533.

[20] 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.

[21] M. Arrieta and I. Esnaola, "Smart meter privacy via the trapdoor channel," in *Proc. IEEE Int. Conf. Smart Grid Commun.*, Dresden, Germany, Nov. 2017, pp. 277–282.

[22] M. Arrieta, I. Esnaola, and M. Effros, "Universal mutual information privacy guarantees for smart meters," *arXiv preprint*, 2019.

[23] Z. Li and T. J. Oechtering, "Privacy on hypothesis testing in smart grids," in *Proc. IEEE Inf. Theory Workshop*, Jeju, South Korea, Oct. 2015, pp. 337–341.

[24] J. Gómez-Vilardebó and D. Gündüz, "Smart meter privacy for multiple users in the presence of an alternative energy source," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 1, pp. 132–141, 2015.

[25] [Online]. Available: https://www.moneysavingexpert.com/utilities/economy-7/, [Accessed: 21-Jan-2019]