Privacy against a Hypothesis Testing Adversary

Zuxing Li, *Member, IEEE*, Tobias J. Oechtering, *Senior Member, IEEE*,
and Deniz Gündüz, *Senior Member, IEEE*

Abstract—Privacy against an adversary (AD) that tries to detect the underlying privacy-sensitive data distribution is studied. The original data sequence is assumed to come from one of the two known distributions, and the privacy leakage is measured by the probability of error of the binary hypothesis test carried out by the AD. A management unit (MU) is allowed to manipulate the original data sequence in an online fashion, while satisfying an average distortion constraint. The goal of the MU is to maximize the minimal type II probability of error subject to a constraint on the type I probability of error assuming an adversarial Neyman-Pearson test, or to maximize the minimal error probability assuming an adversarial Bayesian test. The asymptotic exponents of the maximum minimal type II probability of error and the maximum minimal error probability are shown to be characterized by a Kullback-Leibler divergence rate and a Chernoff information rate, respectively. Privacy performances of particular management policies, the memoryless hypothesis-aware policy and the hypothesis-unaware policy with memory, are compared. The proposed formulation can also model adversarial example generation with minimal data manipulation to fool classifiers. Lastly, the results are applied to a smart meter privacy problem, where the user's energy consumption is manipulated by adaptively using a renewable energy source in order to hide user's activity from the energy provider.

Index Terms—Neyman-Pearson test, Bayesian test, information theory, large deviations, privacy-enhancing technology.

I. INTRODUCTION

Developments in information technology have drastically changed people's lives, e.g., smart homes, smart grids, and e-health. Most of these technologies are based on intelligent algorithms that provide better use of limited resources thanks to their ability to gather and process more information about users' behaviors and preferences. This population-scale collection of personal user data in return of various utilities is a growing concern, as the collected data can be misused beyond their intended application. Therefore, guaranteeing user privacy while continuing to deliver the benefits of such technologies is a fundamental research challenge that must be tackled in order to enable the adoption of these technologies without being concerned about privacy risks.

Privacy has been studied extensively in recent years for various information systems considering different privacy

measures [2]–[8]. Differential privacy [2], [9], [10] and mutual information [7], [11], [12] are probably the most commonly studied privacy measures. In [13], Fisher information is proposed as a privacy measure, and is shown to be a lower bound on the mean squared error of an unbiased estimation on the privacy. Kullback-Leibler divergence is used as a privacy measure in [14], [15], while total variation distance is considered in [16], which provides a bound on the privacy leakage measured by mutual information or maximal leakage [8].

Privacy is particularly difficult to guarantee against ADs employing modern data mining techniques, which are capable of identifying user behavior with large datasets. With such attacks in mind, we propose a novel formulation of the privacy leakage problem that can be applied to many information sharing applications. We assume that the original data is manipulated by the MU and then shared with a remote entity in an online manner. Our goals are to guarantee that the shared data satisfies a certain utility constraint, measured by the average distortion between the original data sequence and the shared sequence, and meanwhile to limit the inference capability of the AD, who might be the legitimate receiver of the shared data. We assume that the AD performs an optimal hypothesis test based on the shared data sequence (or the adversarial observation sequence) and tries to determine a privacy-sensitive underlying hypothesis. Hypothesis test has previously been considered in the privacy context [17]–[21]. However, different from these previous works, in which the goal is to increase the reliability of the hypothesis test while guaranteeing privacy, we measure the privacy risk by the error probability of the corresponding adversarial hypothesis testing problem.

Our problem formulation can also be considered as a model for generating adversarial examples to attack classifiers, which is a very popular research area due to the increasing adoption of machine learning techniques in all domains [22]–[25]. These attacks consist of generating examples that can fool a classifier, despite being very similar to the true distribution of the data that has been used to train the classifier. Deep neural networks are particularly vulnerable to such attacks [26]. In our setting, the AD can be considered as the classifier, and we are aiming at generating adversarial observation sequences that are similar to the original data sequence under a given distortion constraint, yet will make the detection of the AD as unreliable as possible.

A similar adversarial signal processing problem is studied in [27]–[29], which also considers independent identically distributed (i.i.d.) original data sequences, distortion-constrained data manipulation, and hypothesis test based on manipulated sequences. However, the results in this paper and in [27]–[29]

do not imply each other due to the following fundamental differences: i) We consider an MU to degrade the hypothesis testing accuracy of a passive but informed AD, and formulate the problems as worst-case analyses; while [27]–[29] consider an active attacker to degrade the hypothesis testing accuracy of a defender and formulate the problem as a zero-sum game; ii) we assume the AD to always perform an optimal hypothesis test without the restriction that the hypothesis test has to be based on the type of the manipulated data sequence as imposed in [27]–[29]; iii) we also study the scenario in which the MU does not know the hypothesis *a priori*.

We study the privacy performance by focusing on the asymptotic error exponent in two different settings. We consider i) the asymptotic exponent of the maximum minimal type II probability of error assuming an adversarial Neyman-Pearson test setting, and ii) the asymptotic exponent of the maximum minimal error probability assuming an adversarial Bayesian test setting. We show that the asymptotic error exponent can be characterized by a Kullback-Leibler divergence rate in the first setting, and by a Chernoff information rate in the second. In particular, we prove that the asymptotic error exponent achieved by the optimal memoryless hypothesis-aware policy reduces to a single-letter Kullback-Leibler divergence, or a single-letter Chernoff information. We also consider the hypothesis-unaware policies with memory, and show for both settings that, the asymptotic error exponent achieved by the optimal memoryless hypothesis-aware policy is an upper bound on the exponent achieved by the optimal hypothesis-unaware policy with memory. Here, we generalize the privacy problem model we introduced in [1], present the omitted proofs of Lemma 1, Lemma 2, Theorem 3, Theorem 4, and modify the proof of Theorem 1. Additionally, we analyze here the adversarial Bayesian hypothesis test setting.

As an application of the presented theoretical framework, we will consider a smart meter that reports the energy supply data to an energy provider (EP) at regular time intervals. The EP, which is the legitimate receiver of the smart meter data, can be considered as the AD in this setting, that tries to mine the smart meter data beyond the intended utility application, e.g., to infer the privacy-sensitive user behavior information: presence at home, or electrical appliance usage patterns. Various privacy-preserving techniques have been developed for the smart meter privacy problem in recent years, that can be classified into two groups [30]: The methods in the first group modify the smart meter readings in order to confuse the EP; while the second group of methods directly modify the energy supply pattern to achieve the same goal [7], [9], [10], [14], [31]–[34]. The advantage of the second approach is that the smart meters report truthful readings; therefore, the EP can reliably use their readings for energy provisioning. Both approaches can be formulated in the general framework proposed in this paper.

In the following, unless otherwise specified, we will denote a random variable by a capital letter, e.g., X, its realization by the lower-case letter, e.g., x, and its alphabet by the calligraphic letter, e.g., $\mathcal{X}$. Let X_t^k, x_t^k, and $\mathcal{X}_t^k$ denote a random sequence $(X_t, \ldots, X_k)$, its realization $(x_t, \ldots, x_k)$, and its alphabet $\mathcal{X}_t \times \cdots \times \mathcal{X}_k$, respectively. For simplification,
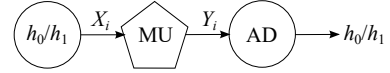

Fig. 1. The privacy problem model. The original data sequence $\{X_i\}$ is i.i.d. generated under the privacy hypothesis h_0 or h_1. The management unit (MU) follows a policy to determine the random adversarial observation sequence $\{Y_i\}$ that has to satisfy a long-term average distortion constraint. The informed adversary (AD) makes an optimal hypothesis test on the privacy based on the observations.

X^k, x^k, and $\mathcal{X}^k$ are used when $t = 1$. We use $\mathrm{D}(\cdot||\cdot)$ to denote Kullback-Leibler divergence, $\mathrm{D}_\tau(\cdot||\cdot)$ to denote τ-th order Rényi divergence, $\mathrm{C}(\cdot, \cdot)$ to denote Chernoff information, and $|\cdot|$ to denote set cardinality.

II. SYSTEM MODEL

The considered privacy problem is illustrated in Fig. 1. The original data sequence $\{X_i\}$ is assumed to come from one of two possible distributions depending on the binary privacy hypothesis H, which can be h_0 or h_1. Let p_0 and p_1 denote the prior probabilities of hypotheses h_0 and h_1, respectively, and without loss of generality, we assume $0 < p_0 \leq p_1 < 1$. In the following, we use the notation $\cdot|h_j$ for $\cdot|H = h_j$, $j \in \{0,1\}$, to denote a random variable under hypothesis h_j. Under hypothesis h_0 (resp. h_1), the data X_i at time slot i is i.i.d. generated according to $p_{X|h_0}$ (resp. $p_{X|h_1}$) where $p_{X|h_0}$ and $p_{X|h_1}$ are defined on the same finite alphabet $\mathcal{X}$ and satisfy $\mathcal{X} \subset \mathbb{R}$, $-\infty < \min \mathcal{X} < \max \mathcal{X} < \infty$, and $0 < \mathrm{D}(p_{X|h_0}||p_{X|h_1}) < \infty$.

At any time slot i, the MU follows a management policy γ_i to determine the (random) adversarial observation Y_i based on the data sequence x^i, the previous adversarial observations y^{i-1}, and the true hypothesis h as $Y_i = \gamma_i(x^i, y^{i-1}, h)$, which can also be represented by the corresponding conditional pmf $p_{Y_i|X^i, Y^{i-1}, H}$. Let $\mathcal{Y}$ denote the finite adversarial observation alphabet at any time slot, which satisfies $\mathcal{Y} \subset \mathbb{R}$ and $-\infty < \min \mathcal{Y} \leq \max \mathcal{Y} < \infty$. Let $\gamma^n \triangleq \{\gamma_i\}_{i=1}^n : \mathcal{X}^n \times \mathcal{H} \to \mathcal{Y}^n$ denote a management policy over an n-slot time horizon. The following average distortion constraint is further imposed:

$$\mathrm{E}\left[\frac{1}{n}\sum_{i=1}^n d(X_i, Y_i)\bigg| h_j\right] \leq s, \; j = 0, 1, \tag{1}$$

where $d(\cdot, \cdot) : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}_+$ is an additive distortion measure. A management policy that satisfies (1) over an n-slot time horizon is denoted by $\gamma^n(s)$. The distortion constraint may be imposed either to guarantee a utility requirement from the data, or due to the availability of limited resources to manipulate the data sequence $\{X_i\}$, e.g., smart meter privacy exploiting physical energy sources [33].

We consider that an AD has access to the observations y^n, and is fully informed about the prior probabilities of the hypotheses, the original data statistics, as well as the adopted management policy, i.e., the AD knows $p_0, p_1, p_{X^n|h_0}, p_{X^n|h_1}, \gamma^n(s)$, and hence the resulting adversarial observation statistics $p_{Y^n|h_0}, p_{Y^n|h_1}$. In this work, the informed AD is assumed to make an optimal hypothesis test on the binary privacy hypothesis, and the privacy risk is measured by the probability of error of this adversarial hypothesis test. In the following,

we will study the privacy leakage as a Neyman-Pearson hypothesis testing problem and a Bayesian hypothesis testing problem. The corresponding optimal privacy performances will be characterized in the asymptotic regime.

III. PRIVACY-PRESERVING DATA MANAGEMENT AGAINST A NEYMAN-PEARSON HYPOTHESIS TESTING ADVERSARY

In this section, the privacy leakage is modeled as a Neyman-Pearson hypothesis test performed by the informed AD. We assume that the AD has a maximum tolerance towards false positives, imposed by a maximum type I probability of error, and its goal is to minimize the type II probability of error under this constraint. Given a management policy $\gamma^n(s)$ and the resulting $p_{Y^n|h_0}$, $p_{Y^n|h_1}$, we define the minimal type II probability of error of the AD under an upper bound on the type I probability of error as

$$\beta(n, \varepsilon, \gamma^n(s)) \triangleq \min_{\mathcal{A}_n \subseteq \mathcal{Y}^n} \{p_{Y^n|h_1}(\mathcal{A}_n) | p_{Y^n|h_0}(\mathcal{A}_n^c) \le \varepsilon\},$$

where $\mathcal{A}_n$ and $\mathcal{A}_n^c$ denote the decision regions for h_0 and h_1 of the AD, respectively. On the other hand, the privacy-preserving design objective of the MU is to maximize the probability of error of the AD. More specifically, for a given constraint s on the average distortion that can be introduced, the MU uses the optimal management policy to achieve the maximum minimal type II probability of error subject to a type I probability of error constraint ε, which, with slight abuse of notation, is denoted by $\beta(n, \varepsilon, s)$ as

$$\beta(n, \varepsilon, s) \triangleq \max_{\gamma^n(s)} \{\beta(n, \varepsilon, \gamma^n(s))\}. \tag{2}$$

Remark 1. *Different from the game theoretic formulation in [27]–[29], we take a worst-case analysis approach by assuming an informed AD who always performs the optimal hypothesis test. This is more appropriate for the privacy problem studied here and will lead to a privacy guarantee independent of the knowledge of the AD about the system.*

In the following, the optimal privacy-preserving policy is characterized in the asymptotic regime as $n \to \infty$, by focusing on the asymptotic exponent of the maximum minimal type II probability of error subject to a type I probability of error constraint.

We define the Kullback-Leibler divergence rate $\theta(s)$ as

$$\theta(s) \triangleq \inf_{k, \gamma^k(s)} \left\{ \frac{1}{k} D(p_{Y^k|h_0} || p_{Y^k|h_1}) \right\}, \tag{3}$$

where the infimum is taken over all $k \in \mathbb{Z}_+$, and for each k, over all management policies that satisfy the average distortion constraint over a k-slot time horizon.

Lemma 1.

$$\theta(s) = \lim_{k \to \infty} \inf_{\gamma^k(s)} \left\{ \frac{1}{k} D(p_{Y^k|h_0} || p_{Y^k|h_1}) \right\}.$$

Proof: We first show that the following sequence $\left\{ \inf_{\gamma^k(s)} \left\{ D(p_{Y^k|h_0} || p_{Y^k|h_1}) \right\} \right\}_{k \in \mathbb{Z}_+}$ is *subadditive*. Given any n, $l \in \mathbb{Z}_+$, let $(\gamma^n(s), \gamma^l(s))$ denote a management policy, which uses $\gamma^n(s)$ over the first n slots, and $\gamma^l(s)$ over

the remaining l slots. Therefore, $(\gamma^n(s), \gamma^l(s))$ satisfies the average distortion constraint over an $(n+l)$-slot time horizon. We have

$$\inf_{\gamma^{n+l}(s)} \left\{ D(p_{Y^{n+l}|h_0} || p_{Y^{n+l}|h_1}) \right\}$$
$$\le \inf_{(\gamma^n(s), \gamma^l(s))} \left\{ D(p_{Y^{n+l}|h_0} || p_{Y^{n+l}|h_1}) \right\}$$
$$= \inf_{\gamma^n(s)} \left\{ D(p_{Y^n|h_0} || p_{Y^n|h_1}) \right\} + \inf_{\gamma^l(s)} \left\{ D(p_{Y^l|h_0} || p_{Y^l|h_1}) \right\},$$

where the equality follows from the chain rule for the Kullback-Leibler divergence. Then, it follows from Fekete's Lemma [35, Lemma 11.2] that

$$\theta(s) = \inf_{k, \gamma^k(s)} \left\{ \frac{1}{k} D(p_{Y^k|h_0} || p_{Y^k|h_1}) \right\}$$
$$= \lim_{k \to \infty} \inf_{\gamma^k(s)} \left\{ \frac{1}{k} D(p_{Y^k|h_0} || p_{Y^k|h_1}) \right\}.$$

∎

The following theorem characterizes the operational meaning of the Kullback-Leibler divergence rate $\theta(s)$ in the asymptotic assessment of the optimal privacy performance assuming an adversarial Neyman-Pearson hypothesis test.

Theorem 1. *Given $s > 0$,*

$$\limsup_{n \to \infty} \frac{1}{n} \log \frac{1}{\beta(n, \varepsilon, s)} \le \theta(s), \ \forall \varepsilon \in (0, 1), \tag{4}$$

and

$$\lim_{\varepsilon \to 1} \liminf_{n \to \infty} \frac{1}{n} \log \frac{1}{\beta(n, \varepsilon, s)} \ge \theta(s). \tag{5}$$

Proof: Given any $k \in \mathbb{Z}_+$, $\gamma^k(s)$, and the resulting $p_{Y^k|h_0}, p_{Y^k|h_1}$, let $\gamma^{kl}(s)$ denote a management policy which repeatedly uses $\gamma^k(s)$ for l times. From the definition in (2) and Stein's Lemma [36, Theorem 11.8.3], it follows that

$$\limsup_{l \to \infty} \frac{1}{kl} \log \frac{1}{\beta(kl, \varepsilon, s)} \le \lim_{l \to \infty} \frac{1}{kl} \log \frac{1}{\beta(kl, \varepsilon, \gamma^{kl}(s))}$$
$$= \frac{1}{k} D(p_{Y^k|h_0} || p_{Y^k|h_1}),$$

for all $\varepsilon \in (0, 1)$. For $k(l-1) < n \le kl$, we have

$$\beta(kl, \varepsilon, s) \le \beta(n, \varepsilon, s) \le \beta(k(l-1), \varepsilon, s).$$

It follows that

$$\limsup_{n \to \infty} \frac{1}{n} \log \frac{1}{\beta(n, \varepsilon, s)} \le \limsup_{l \to \infty} \frac{kl}{k(l-1)} \frac{1}{kl} \log \frac{1}{\beta(kl, \varepsilon, s)}$$
$$= \limsup_{l \to \infty} \frac{1}{kl} \log \frac{1}{\beta(kl, \varepsilon, s)}$$
$$\le \frac{1}{k} D(p_{Y^k|h_0} || p_{Y^k|h_1}),$$

for all $\varepsilon \in (0, 1)$, $k \in \mathbb{Z}_+$, and $\gamma^k(s)$. Therefore, we have the upper bound

$$\limsup_{n \to \infty} \frac{1}{n} \log \frac{1}{\beta(n, \varepsilon, s)} \le \theta(s), \ \forall \varepsilon \in (0, 1).$$

Given $0 < \delta' < \infty$ and for all $n \in \mathbb{Z}_+$, let $\varepsilon'(n) \triangleq$

$$\sup_{\gamma^n(s)} p_{Y^n|h_0} \left\{ y^n \left| \log \frac{p_{Y^n|h_0}(y^n)}{p_{Y^n|h_1}(y^n)} < D(p_{Y^n|h_0} || p_{Y^n|h_1}) - \delta' \right. \right\},$$

and let $\epsilon(n) \triangleq$

$$\sup_{\gamma^n(s)} p_{Y^n|h_0} \left\{ y^n \left| \log \frac{p_{Y^n|h_0}(y^n)}{p_{Y^n|h_1}(y^n)} \leq D(p_{Y^n|h_0}||p_{Y^n|h_1}) - \delta' \right. \right\}.$$

Note that $\varepsilon'(n) \leq \epsilon(n)$. We further have $\epsilon(n) < 1$ since $\epsilon(n) = 1$ will lead to the contradiction $D(p_{Y^n|h_0}||p_{Y^n|h_1}) \leq D(p_{Y^n|h_0}||p_{Y^n|h_1}) - \delta'$. It follows that $\varepsilon'(n) < 1$. Suppose that $\gamma^{n*}(s)$ leads to $p^*_{Y^n|h_0}$, $p^*_{Y^n|h_1}$, and achieves $\beta(n, \varepsilon'(n), s)$. If the AD uses the following hypothesis testing strategy

$$\mathcal{A}_n = \left\{ y^n \left| \frac{1}{n} \log \frac{p^*_{Y^n|h_0}(y^n)}{p^*_{Y^n|h_1}(y^n)} \geq t(n) \right. \right\}, \quad (6)$$

with the test threshold

$$t(n) = \frac{1}{n} D(p^*_{Y^n|h_0}||p^*_{Y^n|h_1}) - \frac{\delta'}{n}, \quad (7)$$

from the definition of $\varepsilon'(n)$, the corresponding type I probability of error satisfies the upper bound constraint

$$p^*_{Y^n|h_0}(\mathcal{A}^c_n) \leq \varepsilon'(n).$$

Since the hypothesis testing strategy in (6) is not necessarily optimal for the AD, the definition of the maximum minimal type II probability of error implies that

$$\beta(n, \varepsilon'(n), s) \leq p^*_{Y^n|h_1}(\mathcal{A}_n). \quad (8)$$

In [37, Lemma 4.1.1], it has been shown that

$$p^*_{Y^n|h_1}(\mathcal{A}_n) \leq \exp(-nt(n)). \quad (9)$$

The inequalities (8) and (9) jointly lead to

$$\beta(n, \varepsilon'(n), s) \leq \exp(-nt(n))$$
$$\leq \exp\left(-n\left(\inf_{\gamma^n(s)} \left\{ \frac{1}{n} D(p_{Y^n|h_0}||p_{Y^n|h_1}) \right\} - \frac{\delta'}{n}\right)\right),$$

i.e., for all $n \in \mathbb{Z}_+$,

$$\frac{1}{n} \log \frac{1}{\beta(n, \varepsilon'(n), s)} \geq \inf_{\gamma^n(s)} \left\{ \frac{1}{n} D(p_{Y^n|h_0}||p_{Y^n|h_1}) \right\} - \frac{\delta'}{n}.$$

In the asymptotic regime as $n \to \infty$, it follows that

$$\liminf_{n \to \infty} \frac{1}{n} \log \frac{1}{\beta(n, \varepsilon'(n), s)}$$
$$\geq \lim_{n \to \infty} \inf_{\gamma^n(s)} \left\{ \frac{1}{n} D(p_{Y^n|h_0}||p_{Y^n|h_1}) \right\} - \lim_{n \to \infty} \frac{\delta'}{n} = \theta(s),$$

where the equality follows from Lemma 1. Note that $\liminf_{n \to \infty} \frac{1}{n} \log \frac{1}{\beta(n, \varepsilon, s)}$ is a monotone non-decreasing function of ε. Then the lower bound holds:

$$\lim_{\varepsilon \to 1} \liminf_{n \to \infty} \frac{1}{n} \log \frac{1}{\beta(n, \varepsilon, s)} \geq \liminf_{n \to \infty} \frac{1}{n} \log \frac{1}{\beta(n, \varepsilon'(n), s)}$$
$$\geq \theta(s). \quad (10)$$

∎

When ε is close to one, the bounds on the asymptotic exponent of the maximum minimal type II probability of error are tight, which is made more concrete in the following corollary.

Corollary 1. *Given $s > 0$,*

$$\lim_{\varepsilon \to 1} \lim_{n \to \infty} \frac{1}{n} \log \frac{1}{\beta(n, \varepsilon, s)} = \theta(s).$$

Remark 2. *Given $s > 0$, the case $\varepsilon \to 1$ represents the worst privacy leakage scenario assuming the adversarial Neyman-Pearson hypothesis test, i.e., $\theta(s)$ is a privacy guarantee.*

In the following, we characterize the asymptotic privacy performances of two particular management policies in the worst case scenario, i.e., $\varepsilon \to 1$.

A. Memoryless Hypothesis-Aware Policy

A simple MU might have a limited processing capability, and at time slot i, applies a memoryless hypothesis-aware management policy π_i, which can be time-variant, to determine the random adversarial observation Y_i based on the current data x_i and the true hypothesis h as $Y_i = \pi_i(x_i, h)$. Let $\pi^n \triangleq \{\pi_i\}_{i=1}^n : \mathcal{X}^n \times \mathcal{H} \to \mathcal{Y}^n$ denote a memoryless hypothesis-aware management policy over an n-slot time horizon. If π^n satisfies the average distortion constraint in (1), it is denoted by $\pi^n(s)$. When the MU adopts the optimal memoryless hypothesis-aware policy, the maximum minimal type II probability of error subject to a type I probability of error upper bound ε is denoted by

$$\beta_L(n, \varepsilon, s) \triangleq \max_{\pi^n(s)} \{\beta(n, \varepsilon, \pi^n(s))\}. \quad (11)$$

We similarly define $\theta_L(s)$ as

$$\theta_L(s) \triangleq \inf_{k, \pi^k(s)} \left\{ \frac{1}{k} D(p_{Y^k|h_0}||p_{Y^k|h_1}) \right\}. \quad (12)$$

Following similar proof steps as in Theorem 1, we can specify the asymptotic exponent of the maximum minimal type II probability of error by the Kullback-Leibler divergence rate $\theta_L(s)$ when the MU adopts the optimal memoryless hypothesis-aware policy.

Corollary 2. *Given $s > 0$,*

$$\lim_{\varepsilon \to 1} \lim_{n \to \infty} \frac{1}{n} \log \frac{1}{\beta_L(n, \varepsilon, s)} = \theta_L(s). \quad (13)$$

We next show that the asymptotic exponent of the maximum minimal type II probability of error can also be characterized by a single-letter Kullback-Leibler divergence. To this end, we first define the following single-letter Kullback-Leibler divergence function of single-slot average distortion constraints under both hypotheses.

Given $\bar{s}, \tilde{s} > 0$, we define $\phi(\bar{s}, \tilde{s})$ as

$$\phi(\bar{s}, \tilde{s}) \triangleq \min_{(p_{Y|X,h_0}, p_{Y|X,h_1}) \in \mathcal{P}(\bar{s}, \tilde{s})} \{D(p_{Y|h_0}||p_{Y|h_1})\}, \quad (14)$$

where the minimization is over the convex domain

$$\mathcal{P}(\bar{s}, \tilde{s}) \triangleq \left\{ (p_{Y|X,h_0}, p_{Y|X,h_1}) \left| \begin{array}{l} E[d(X, Y)|h_0] \leq \bar{s} \\ E[d(X, Y)|h_1] \leq \tilde{s} \end{array} \right. \right\}.$$

In the definition of $\mathcal{P}(\bar{s}, \tilde{s})$, $E[d(X, Y)|h_0] \leq \bar{s}$ denotes the single-slot average distortion constraint under hypothesis

h_0; and $\mathrm{E}[d(X,Y)|h_1] \leq \tilde{s}$ denotes the single-slot average distortion constraint under hypothesis h_1.

Lemma 2. $\phi(\bar{s}, \tilde{s})$ *is a non-increasing, continuous, and jointly convex function for $\bar{s} > 0$ and $\tilde{s} > 0$.*

Proof: The non-increasing property of $\phi(\bar{s}, \tilde{s})$ is self-evident. On the two-dimensional convex open set of $\bar{s} > 0$, $\tilde{s} > 0$, its continuity will follow from the convexity [38]. Therefore, we only prove its convexity here. Assume that $(p_{Y|X,h_0}^{(1)}, p_{Y|X,h_1}^{(1)})$ leads to $\phi(\bar{s}_1, \tilde{s}_1) = \mathrm{D}(p_{Y|h_0}^{(1)}||p_{Y|h_1}^{(1)})$ and $(p_{Y|X,h_0}^{(2)}, p_{Y|X,h_1}^{(2)})$ leads to $\phi(\bar{s}_2, \tilde{s}_2) = \mathrm{D}(p_{Y|h_0}^{(2)}||p_{Y|h_1}^{(2)})$. For all $0 \leq \lambda \leq 1$,

$$
\begin{aligned}
&\lambda\phi(\bar{s}_1, \tilde{s}_1) + (1-\lambda)\phi(\bar{s}_2, \tilde{s}_2) \\
=&\lambda\mathrm{D}(p_{Y|h_0}^{(1)}||p_{Y|h_1}^{(1)}) + (1-\lambda)\mathrm{D}(p_{Y|h_0}^{(2)}||p_{Y|h_1}^{(2)}) \\
\geq&\mathrm{D}(\lambda p_{Y|h_0}^{(1)} + (1-\lambda)p_{Y|h_0}^{(2)}||\lambda p_{Y|h_1}^{(1)} + (1-\lambda)p_{Y|h_1}^{(2)}) \\
\geq&\phi(\lambda\bar{s}_1 + (1-\lambda)\bar{s}_2, \lambda\tilde{s}_1 + (1-\lambda)\tilde{s}_2),
\end{aligned}
$$

where the first inequality follows from the convexity of $\mathrm{D}(\cdot||\cdot)$; and the second follows from the definition of $\phi(\bar{s}, \tilde{s})$ in (14). ∎

Theorem 2. *Given $s > 0$,*

$$\theta_{\mathrm{L}}(s) = \phi(s, s). \tag{15}$$

Proof: For any $k \in \mathbb{Z}_+$, $\pi^k(s)$, and the resulting $p_{Y^k|h_0}$, $p_{Y^k|h_1}$, we have

$$
\begin{aligned}
&\frac{1}{k}\mathrm{D}\left(p_{Y^k|h_0}||p_{Y^k|h_1}\right) \\
\overset{(a)}{=}&\frac{1}{k}\sum_{i=1}^{k}\mathrm{D}\left(p_{Y_i|h_0}||p_{Y_i|h_1}\right) \\
\overset{(b)}{\geq}&\frac{1}{k}\sum_{i=1}^{k}\phi\left(\mathrm{E}[d(X_i,Y_i)|h_0], \mathrm{E}[d(X_i,Y_i)|h_1]\right) \\
\overset{(c)}{\geq}&\phi\left(\mathrm{E}\left[\frac{1}{k}\sum_{i=1}^{k}d(X_i,Y_i)\Big|h_0\right], \mathrm{E}\left[\frac{1}{k}\sum_{i=1}^{k}d(X_i,Y_i)\Big|h_1\right]\right) \\
\overset{(d)}{\geq}&\phi(s,s),
\end{aligned}
$$

where (a) follows since the policy $\pi^k(s)$ leads to $p_{Y^k|h_j} = \prod_{i=1}^{k}p_{Y_i|h_j}$ for $j = 0, 1$; (b) follows from the definition of $\phi(\bar{s}, \tilde{s})$; (c) and (d) follow from the convexity and the non-increasing property of $\phi(\bar{s}, \tilde{s})$, respectively.

Therefore, we have

$$\theta_{\mathrm{L}}(s) = \inf_{k,\pi^k(s)}\left\{\frac{1}{k}\mathrm{D}(p_{Y^k|h_0}||p_{Y^k|h_1})\right\} \geq \phi(s,s). \tag{16}$$

The proof of the opposite direction is straightforward. Let $(p_{Y|X,h_0}^*, p_{Y|X,h_1}^*)$ be the optimizer which achieves $\phi(s,s)$. It can be seen as a single-slot memoryless hypothesis-aware policy $\pi^1(s)$. From the definition of $\theta_{\mathrm{L}}(s)$ in (12), it follows that

$$\theta_{\mathrm{L}}(s) \leq \phi(s,s). \tag{17}$$

Alternatively, the inequality (17) follows since $\phi(s,s)$ is the asymptotic exponent of the minimal type II probability of error

achieved by a memoryless hypothesis-aware policy by using the single-slot policy $(p_{Y|X,h_0}^*, p_{Y|X,h_1}^*)$ at all slots.

The inequalities (16) and (17) jointly lead to Theorem 2. ∎

Remark 3. *Given $s > 0$, the asymptotic exponent of the maximum minimal type II probability of error, $\lim_{\varepsilon \to 1}\lim_{n \to \infty}\frac{1}{n}\log\frac{1}{\beta_{\mathrm{L}}(n,\varepsilon,s)}$, can be achieved by a memoryless hypothesis-aware policy which uses the single-slot policy $(p_{Y|X,h_0}^*, p_{Y|X,h_1}^*)$ corresponding to the optimizer of $\phi(s,s)$ at all time slots.*

B. Hypothesis-Unaware Policy with Memory

We now consider an MU which does not have access to the true hypothesis but has a large memory storage and a powerful processing capability. At time slot i, the MU follows a hypothesis-unaware management policy with memory, denoted by ρ_i, to determine the random adversarial observation Y_i based on the data sequence x^i and the past adversarial observations y^{i-1} as $Y_i = \rho_i(x^i, y^{i-1})$. Let $\rho^n \triangleq \{\rho_i\}_{i=1}^n : \mathcal{X}^n \to \mathcal{Y}^n$ denote a hypothesis-unaware management policy with memory over an n-slot time horizon. If ρ^n satisfies the average distortion constraint in (1), it is denoted by $\rho^n(s)$. When the MU uses the optimal privacy-preserving hypothesis-unaware policy with memory, the achieved maximum minimal type II probability of error subject to a type I probability of error upper bound ε is denoted by

$$\beta_{\mathrm{M}}(n,\varepsilon,s) \triangleq \max_{\rho^n(s)}\{\beta(n,\varepsilon,\rho^n(s))\}. \tag{18}$$

Similarly, we define a Kullback-Leibler divergence rate $\theta_{\mathrm{M}}(s)$ as

$$\theta_{\mathrm{M}}(s) \triangleq \inf_{k,\rho^k(s)}\left\{\frac{1}{k}\mathrm{D}(p_{Y^k|h_0}||p_{Y^k|h_1})\right\}. \tag{19}$$

As specified in the following corollary, the asymptotic exponent of the maximum minimal type II probability of error can be characterized by the Kullback-Leibler divergence rate $\theta_{\mathrm{M}}(s)$ when the MU uses the optimal privacy-preserving hypothesis-unaware policy with memory.

Corollary 3. *Given $s > 0$,*

$$\lim_{\varepsilon \to 1}\lim_{n \to \infty}\frac{1}{n}\log\frac{1}{\beta_{\mathrm{M}}(n,\varepsilon,s)} = \theta_{\mathrm{M}}(s). \tag{20}$$

Compared with the memoryless hypothesis-aware policy, the hypothesis-unaware policy with memory has all the past data and adversarial observations while it does not know the true hypothesis. We next compare the asymptotic privacy performances of the two policies.

Theorem 3. *Given $s > 0$,*

$$\theta_{\mathrm{M}}(s) \leq \theta_{\mathrm{L}}(s) = \phi(s,s). \tag{21}$$

To prove Theorem 3, we will construct a two-phase hypothesis-unaware policy with memory, which first learns the hypothesis, and then we bound its privacy performance

by $\theta_M(s)$ and $\phi(s,s)$. The complete proof is given in the appendix.

Remark 4. *The optimal privacy-preserving memoryless hypothesis-aware policy cannot outperform the optimal privacy-preserving hypothesis-unaware policy with memory. This is because the MU can learn the hypothesis with an arbitrarily small probability of error after observing a sufficiently long sequence of the original data.*

IV. PRIVACY-PRESERVING DATA MANAGEMENT AGAINST A BAYESIAN HYPOTHESIS TESTING ADVERSARY

In this section, we consider an adversarial Bayesian hypothesis testing formulation for the privacy problem. A particular Bayesian risk used here is the error probability of the AD. Thus, the minimal error probability of the AD measures the privacy leakage. The asymptotic privacy performance is studied under the adversarial Bayesian hypothesis testing framework.

The informed AD is assumed to use the optimal hypothesis testing strategy to achieve the minimal error probability as

$$\alpha(n,\gamma^n(s)) \triangleq \min_{\mathcal{A}_n \subseteq \mathcal{Y}^n} \left\{ p_0 \cdot p_{Y^n|h_0}(\mathcal{A}_n^c) + p_1 \cdot p_{Y^n|h_1}(\mathcal{A}_n) \right\},$$

where p_0, p_1 denote the prior probabilities of hypotheses h_0, h_1; and $\mathcal{A}_n$, $\mathcal{A}_n^c$ denote the decision regions for h_0, h_1 of the AD. Correspondingly, the MU uses the optimal management policy that maximizes the minimal error probability of the AD as

$$\alpha(n,s) \triangleq \max_{\gamma^n(s)} \left\{ \alpha(n,\gamma^n(s)) \right\}. \tag{22}$$

We will characterize the optimal privacy performance in the asymptotic regime as $n \to \infty$ by focusing on the asymptotic exponent of the maximum minimal error probability. To this end, we first introduce the Chernoff information. The Chernoff information of a probability distribution $P(Z)$ from another distribution $Q(Z)$ is defined as

$$C(P(Z),Q(Z)) \triangleq \max_{0 \le \tau \le 1} \left\{ C_\tau(P(Z),Q(Z)) \right\},$$

where

$$C_\tau(P(Z),Q(Z)) \triangleq -\log\left(\sum_{z \in \mathcal{Z}} P^\tau(z)Q^{1-\tau}(z) \right).$$

The convexity of Chernoff information is shown in the following propositions.

Proposition 1. *Given* $0 \le \tau \le 1$, *the function* $C_\tau(P(Z),Q(Z))$ *is jointly convex in* $P(Z)$ *and* $Q(Z)$.

Proof: Given $0 \le \tau < 1$, the function $C_\tau(P(Z),Q(Z))$ is related to the Rényi divergence $D_\tau(P(Z)||Q(Z))$ as

$$C_\tau(P(Z),Q(Z)) = (1-\tau) \cdot D_\tau(P(Z)||Q(Z)).$$

The convexity of $C_\tau(P(Z),Q(Z))$ follows from the convexity of Rényi divergence [39, Theorem 11] since $1-\tau$ is a positive scalar.

If $\tau = 1$, we have

$$C_1(P(Z),Q(Z)) = -\log\left(P\left(\{z \in \mathcal{Z}|Q(z) > 0\} \right) \right)$$
$$= D_0(Q(Z)||P(Z)).$$

In this case, the convexity of $C_1(P(Z),Q(Z))$ follows from the convexity of $D_0(Q(Z)||P(Z))$. ∎

Proposition 2. *The Chernoff information* $C(P(Z),Q(Z))$ *is jointly convex in* $P(Z)$ *and* $Q(Z)$.

The convexity of $C(P(Z),Q(Z))$ follows from the convexity of $C_\tau(P(Z),Q(Z))$ for all $0 \le \tau \le 1$ and the fact that pointwise maximum preserves convexity [40, Section 3.2.3].

We define the Chernoff information rate $\mu(s)$ as follows:

$$\mu(s) \triangleq \inf_{k,\gamma^k(s)} \left\{ \frac{1}{k} C(p_{Y^k|h_0}, p_{Y^k|h_1}) \right\}. \tag{23}$$

The following lemma shows that the infimum over $k \in \mathbb{Z}_+$ in the definition of the Chernoff information rate $\mu(s)$ is taken at the limit.

Lemma 3.

$$\mu(s) = \lim_{k \to \infty} \inf_{\gamma^k(s)} \left\{ \frac{1}{k} C(p_{Y^k|h_0}, p_{Y^k|h_1}) \right\}.$$

Proof: Given any $n, l \in \mathbb{Z}_+$, as defined before, $(\gamma^n(s), \gamma^l(s))$ is a management policy satisfying the average distortion constraint over an $(n+l)$-slot time horizon. We have

$$\inf_{\gamma^{n+l}(s)} \left\{ C(p_{Y^{n+l}|h_0}, p_{Y^{n+l}|h_1}) \right\}$$
$$\le \inf_{(\gamma^n(s),\gamma^l(s))} \left\{ C(p_{Y^{n+l}|h_0}, p_{Y^{n+l}|h_1}) \right\}$$
$$\overset{(a)}{=} \inf_{(\gamma^n(s),\gamma^l(s))} \left\{ \max_{0 \le \tau \le 1} \left\{ C_\tau(p_{Y^n|h_0}, p_{Y^n|h_1}) + C_\tau(p_{Y^l|h_0}, p_{Y^l|h_1}) \right\} \right\}$$
$$\le \inf_{(\gamma^n(s),\gamma^l(s))} \left\{ \max_{0 \le \kappa \le 1} \left\{ C_\kappa(p_{Y^n|h_0}, p_{Y^n|h_1}) \right\} + \max_{0 \le \sigma \le 1} \left\{ C_\sigma(p_{Y^l|h_0}, p_{Y^l|h_1}) \right\} \right\}$$
$$= \inf_{\gamma^n(s)} \left\{ C(p_{Y^n|h_0}, p_{Y^n|h_1}) \right\} + \inf_{\gamma^l(s)} \left\{ C(p_{Y^l|h_0}, p_{Y^l|h_1}) \right\},$$

where (a) follows from the independence property $p_{Y^{n+l}|h_j} = p_{Y^n|h_j} \cdot p_{Y^l|h_j}$, $j = 0,1$, satisfied by the management policy $(\gamma^n(s), \gamma^l(s))$. Therefore, the sequence $\left\{ \inf_{\gamma^k(s)} \left\{ C(p_{Y^k|h_0}, p_{Y^k|h_1}) \right\} \right\}_{k \in \mathbb{Z}_+}$ is *subadditive*. Then, Lemma 3 follows from Fekete's Lemma. ∎

Next, we show that the asymptotic exponent of the maximum minimal error probability is characterized by the Chernoff information rate $\mu(s)$.

Theorem 4. *Given* $s > 0$,

$$\lim_{n \to \infty} \frac{1}{n} \log \frac{1}{\alpha(n,s)} = \mu(s). \tag{24}$$

Proof: Given any $k \in \mathbb{Z}_+$, $\gamma^k(s)$, and the resulting $p_{Y^k|h_0}$, $p_{Y^k|h_1}$, as defined before, $\gamma^{kl}(s)$ is a management policy which repeatedly uses $\gamma^k(s)$ for l times and satisfies the average distortion constraint over a kl-slot time horizon. From the optimality in the definition (22) and the theorem [36, Theorem 11.9.1], it follows that

$$\limsup_{l \to \infty} \frac{1}{kl} \log \frac{1}{\alpha(kl,s)} \le \lim_{l \to \infty} \frac{1}{kl} \log \frac{1}{\alpha(kl,\gamma^{kl}(s))}$$
$$= \frac{1}{k} C(p_{Y^k|h_0}, p_{Y^k|h_1}).$$

For $k(l-1) < n \leq kl$, we have

$$\alpha(kl, s) \leq \alpha(n, s) \leq \alpha(k(l-1), s).$$

It follows that

$$\begin{aligned}
\limsup_{n \to \infty} \frac{1}{n} \log \frac{1}{\alpha(n, s)} &\leq \limsup_{l \to \infty} \frac{kl}{k(l-1)} \frac{1}{kl} \log \frac{1}{\alpha(kl, s)} \\
&= \limsup_{l \to \infty} \frac{1}{kl} \log \frac{1}{\alpha(kl, s)} \\
&\leq \frac{1}{k} \mathrm{C}(p_{Y^k|h_0}, p_{Y^k|h_1}),
\end{aligned}$$

for all $k \in \mathbb{Z}_+$ and $\gamma^k(s)$. Therefore, we have the upper bound

$$\limsup_{n \to \infty} \frac{1}{n} \log \frac{1}{\alpha(n, s)} \leq \mu(s). \tag{25}$$

Given any $n \in \mathbb{Z}_+$, suppose that $\gamma^{n*}(s)$ leads to $p^*_{Y^n|h_0}$, $p^*_{Y^n|h_1}$, and achieves $\alpha(n, s)$. An optimal hypothesis testing strategy of the AD is a deterministic likelihood-ratio test [41] given by

$$\mathcal{A}^*_n = \left\{ y^n \, \middle| \, \frac{p^*_{Y^n|h_0}(y^n)}{p^*_{Y^n|h_1}(y^n)} \geq \frac{p_1}{p_0} \right\}.$$

Based on the optimal strategy of the AD, we can then rewrite the maximum minimal error probability $\alpha(n, s)$, and derive upper bounds on it as follows. For all $0 \leq \tau \leq 1$,

$$\begin{aligned}
\alpha(n, s) &= \sum_{\mathcal{Y}^n} \min \left\{ p_0 \cdot p^*_{Y^n|h_0}(y^n), p_1 \cdot p^*_{Y^n|h_1}(y^n) \right\} \\
&\overset{(a)}{\leq} \sum_{\mathcal{Y}^n} p_0^\tau \cdot p^{*\tau}_{Y^n|h_0}(y^n) \cdot p_1^{1-\tau} \cdot p^{*1-\tau}_{Y^n|h_1}(y^n) \\
&\leq \sum_{\mathcal{Y}^n} p^{*\tau}_{Y^n|h_0}(y^n) p^{*1-\tau}_{Y^n|h_1}(y^n),
\end{aligned}$$

where the inequality (a) follows from [36, (11.244)]. Therefore, we have

$$\begin{aligned}
&\frac{1}{n} \log \frac{1}{\alpha(n, s)} \\
&\geq \frac{1}{n} \max_{0 \leq \tau \leq 1} \left\{ -\log \left(\sum_{\mathcal{Y}^n} p^{*\tau}_{Y^n|h_0}(y^n) p^{*1-\tau}_{Y^n|h_1}(y^n) \right) \right\} \\
&= \frac{1}{n} \mathrm{C}(p^*_{Y^n|h_0}, p^*_{Y^n|h_1}) \\
&\geq \inf_{\gamma^n(s)} \left\{ \frac{1}{n} \mathrm{C}(p_{Y^n|h_0}, p_{Y^n|h_1}) \right\}.
\end{aligned}$$

In the asymptotic regime, as $n \to \infty$, we have

$$\begin{aligned}
\liminf_{n \to \infty} \frac{1}{n} \log \frac{1}{\alpha(n, s)} &\geq \lim_{n \to \infty} \inf_{\gamma^n(s)} \left\{ \frac{1}{n} \mathrm{C}(p_{Y^n|h_0}, p_{Y^n|h_1}) \right\} \\
&= \mu(s),
\end{aligned} \tag{26}$$

where the final equality follows from Lemma 3.

The inequalities (25) and (26) jointly lead to Theorem 4. ∎

Remark 5. *Over a finite time horizon, the prior distribution of the hypothesis determines the test threshold of the optimal likelihood-ratio test of the AD, and further determines the exponent of the maximum minimal error probability. However, as shown in Theorem 4, the asymptotic exponent of the*

maximum minimal error probability does not depend on the prior distribution.

In the following, we characterize the privacy-preserving managements for the two particular cases, the memoryless hypothesis-aware policy and the hypothesis-unaware policy with memory, under the adversarial Bayesian hypothesis testing framework.

A. Memoryless Hypothesis-Aware Policy

When the MU uses the optimal memoryless hypothesis-aware policy $\pi^{n*}(s)$ against the adversarial Bayesian hypothesis test, the achieved maximum minimal error probability is denoted by

$$\alpha_{\mathrm{L}}(n, s) \triangleq \max_{\pi^n(s)} \{\alpha(n, \pi^n(s))\}. \tag{27}$$

We similarly define $\mu_{\mathrm{L}}(s)$ as

$$\mu_{\mathrm{L}}(s) \triangleq \inf_{k, \pi^k(s)} \left\{ \frac{1}{k} \mathrm{C}(p_{Y^k|h_0}, p_{Y^k|h_1}) \right\}. \tag{28}$$

Following similar proof steps as in Theorem 4, we can show that the asymptotic exponent of the maximum minimal error probability is specified by the Chernoff information rate $\mu_{\mathrm{L}}(s)$ when the MU uses the optimal privacy-preserving memoryless hypothesis-aware policy.

Corollary 4. *Given $s > 0$,*

$$\lim_{n \to \infty} \frac{1}{n} \log \frac{1}{\alpha_{\mathrm{L}}(n, s)} = \mu_{\mathrm{L}}(s). \tag{29}$$

We next define two single-letter expressions, which characterize upper and lower bounds on the asymptotic exponent of the maximum minimal error probability. Given $\bar{s}, \tilde{s} > 0$ and $0 \leq \tau \leq 1$, we define

$$\nu_\tau(\bar{s}, \tilde{s}) \triangleq \min_{(p_{Y|X, h_0}, p_{Y|X, h_1}) \in \mathcal{P}(\bar{s}, \tilde{s})} \left\{ \mathrm{C}_\tau(p_{Y|h_0}, p_{Y|h_1}) \right\}, \tag{30}$$

and define the single-letter Chernoff information as

$$\nu(\bar{s}, \tilde{s}) \triangleq \min_{(p_{Y|X, h_0}, p_{Y|X, h_1}) \in \mathcal{P}(\bar{s}, \tilde{s})} \left\{ \mathrm{C}(p_{Y|h_0}, p_{Y|h_1}) \right\}. \tag{31}$$

Lemma 4. *Given $0 \leq \tau \leq 1$, $\nu_\tau(\bar{s}, \tilde{s})$ is a non-increasing and jointly convex function for $\bar{s} > 0$ and $\tilde{s} > 0$.*

Lemma 5. *$\nu(\bar{s}, \tilde{s})$ is a non-increasing, continuous, and jointly convex function for $\bar{s} > 0$ and $\tilde{s} > 0$.*

The proofs of Lemmas 4 and 5 follow from the same arguments as in the proof of Lemma 2, and are therefore omitted.

Lemma 6. *Given $s > 0$, we have*

$$\max_{0 \leq \tau \leq 1} \{\nu_\tau(s, s)\} \leq \mu_{\mathrm{L}}(s) \leq \nu(s, s). \tag{32}$$

Proof: Given any $0 \leq \tau \leq 1$, $k \in \mathbb{Z}_+$, memoryless hypothesis-aware policy $\pi^k(s)$, and the resulting $p_{Y^k|h_0}$, $p_{Y^k|h_1}$, we have

$$
\begin{aligned}
&\frac{1}{k}\mathrm{C}(p_{Y^k|h_0}, p_{Y^k|h_1}) \\
=&\frac{1}{k} \max_{0 \leq \kappa \leq 1} \left\{ \mathrm{C}_\kappa \left(p_{Y^k|h_0}, p_{Y^k|h_1} \right) \right\} \\
\geq&\frac{1}{k} \mathrm{C}_\tau \left(p_{Y^k|h_0}, p_{Y^k|h_1} \right) \\
\overset{(a)}{=}&\frac{1}{k} \sum_{i=1}^{k} \mathrm{C}_\tau \left(p_{Y_i|h_0}, p_{Y_i|h_1} \right) \\
\overset{(b)}{\geq}&\frac{1}{k} \sum_{i=1}^{k} \nu_\tau \left(\mathrm{E}[d(X_i, Y_i)|h_0], \mathrm{E}[d(X_i, Y_i)|h_1] \right) \\
\overset{(c)}{\geq}&\nu_\tau \left(\mathrm{E}\left[\frac{1}{k} \sum_{i=1}^{k} d(X_i, Y_i) \middle| h_0 \right], \mathrm{E}\left[\frac{1}{k} \sum_{i=1}^{k} d(X_i, Y_i) \middle| h_1 \right] \right) \\
\overset{(d)}{\geq}&\nu_\tau(s, s),
\end{aligned}
\tag{33}
$$

where (a) follows since the policy $\pi^k(s)$ leads to $p_{Y^k|h_j} = \prod_{i=1}^{k} p_{Y_i|h_j}$ for $j = 0, 1$; (b) follows from the definition of $\nu_\tau(\bar{s}, \tilde{s})$; (c) follows from the convexity of $\nu_\tau(\bar{s}, \tilde{s})$; and finally (d) follows from the non-increasing property of $\nu_\tau(\bar{s}, \tilde{s})$.

For any $k \in \mathbb{Z}_+$, $\pi^k(s)$, and the resulting $p_{Y^k|h_0}$, $p_{Y^k|h_1}$, we have

$$
\frac{1}{k}\mathrm{C}(p_{Y^k|h_0}, p_{Y^k|h_1}) \geq \max_{0 \leq \tau \leq 1} \left\{ \nu_\tau(s, s) \right\},
$$

since the inequality (33) holds for all $0 \leq \tau \leq 1$. It further follows that

$$
\mu_{\mathrm{L}}(s) = \inf_{k, \pi^k(s)} \left\{ \frac{1}{k}\mathrm{C}(p_{Y^k|h_0}, p_{Y^k|h_1}) \right\} \geq \max_{0 \leq \tau \leq 1} \left\{ \nu_\tau(s, s) \right\}.
$$

The other inequality $\mu_{\mathrm{L}}(s) \leq \nu(s, s)$ in (32) follows from the definitions of $\mu_{\mathrm{L}}(s)$ and $\nu(s, s)$. ∎

Under the adversarial Bayesian hypothesis test setting, we have obtained a $\max\min$ single-letter lower bound and a $\min\max$ single-letter upper bound on the asymptotic exponent of the maximum minimal error probability when we focus on memoryless hypothesis-aware policy. In the following theorem, we show that the two bounds match each other, and the asymptotic exponent of the maximum minimal error probability can be specified by the single-letter Chernoff information $\nu(s, s)$.

Theorem 5. *Given $s > 0$,*

$$
\mu_{\mathrm{L}}(s) = \nu(s, s).
\tag{34}
$$

Proof: Given $s > 0$, the lower and upper bounds derived in Lemma 6 can be specified by a $\max\min$ expression and a $\min\max$ expression as follows:

$$
\begin{aligned}
&\max_{0 \leq \tau \leq 1} \left\{ \nu_\tau(s, s) \right\} \\
=&\max_{0 \leq \tau \leq 1} \left\{ \min_{(p_{Y|X,h_0}, p_{Y|X,h_1}) \in \mathcal{P}(s,s)} \left\{ \mathrm{C}_\tau(p_{Y|h_0}, p_{Y|h_1}) \right\} \right\},
\end{aligned}
$$

$$
\begin{aligned}
&\nu(s, s) \\
=&\min_{(p_{Y|X,h_0}, p_{Y|X,h_1}) \in \mathcal{P}(s,s)} \left\{ \max_{0 \leq \tau \leq 1} \left\{ \mathrm{C}_\tau(p_{Y|h_0}, p_{Y|h_1}) \right\} \right\}.
\end{aligned}
$$

If τ is fixed, $\mathrm{C}_\tau(p_{Y|h_0}, p_{Y|h_1})$ is a jointly convex function in $p_{Y|X,h_0}$ and $p_{Y|X,h_1}$, which follows from the convexity of $\mathrm{C}_\tau(\cdot, \cdot)$ shown in Proposition 1 and the convexity-preserving composition rule in [40, Section 3.2.4]. If $p_{Y|X,h_0}$ and $p_{Y|X,h_1}$ are fixed, $p_{Y|h_0}$ and $p_{Y|h_1}$ are fixed, and $\mathrm{C}_\tau(p_{Y|h_0}, p_{Y|h_1})$ is a concave function in τ, which follows from the result [39, Corollary 2].

From von Neumann's Minimax Theorem [42], it follows that

$$
\max_{0 \leq \tau \leq 1} \left\{ \nu_\tau(s, s) \right\} = \nu(s, s).
\tag{35}
$$

Lemma 6 and (35) jointly lead to Theorem 5. ∎

Remark 6. *Given $s > 0$, the asymptotic exponent of the maximum minimal error probability, $\lim_{n \to \infty} \frac{1}{n} \log \frac{1}{\alpha_{\mathrm{L}}(n,s)}$, can be achieved by a memoryless hypothesis-aware policy which uses the single-slot policy $(p^*_{Y|X,h_0}, p^*_{Y|X,h_1})$ corresponding to the optimizer of $\nu(s, s)$ at all time slots.*

B. Hypothesis-Unaware Policy with Memory

When the MU uses the optimal hypothesis-unaware policy with memory, the achieved maximum minimal error probability is denoted by

$$
\alpha_{\mathrm{M}}(n, s) \triangleq \max_{\rho^n(s)} \left\{ \alpha(n, \rho^n(s)) \right\}.
\tag{36}
$$

We define the Chernoff information rate $\mu_{\mathrm{M}}(s)$ as

$$
\mu_{\mathrm{M}}(s) \triangleq \inf_{k, \rho^k(s)} \left\{ \frac{1}{k}\mathrm{C} \left(p_{Y^k|h_0}, p_{Y^k|h_1} \right) \right\},
\tag{37}
$$

which characterizes the asymptotic exponent of the maximum minimal error probability as shown in the following corollary.

Corollary 5. *Given $s > 0$,*

$$
\lim_{n \to \infty} \frac{1}{n} \log \frac{1}{\alpha_{\mathrm{M}}(n, s)} = \mu_{\mathrm{M}}(s).
\tag{38}
$$

Similar to the adversarial Neyman-Pearson hypothesis testing case, the following theorem shows that the optimal memoryless hypothesis-aware policy cannot outperform the optimal hypothesis-unaware policy with memory against the adversarial Bayesian hypothesis test.

Theorem 6. *Given $s > 0$,*

$$
\mu_{\mathrm{M}}(s) \leq \mu_{\mathrm{L}}(s) = \nu(s, s).
\tag{39}
$$

The proof of Theorem 6 follows similarly to Theorem 3. We can construct a two-phase hypothesis-unaware policy with memory, which first learns the hypothesis. Then, we bound its privacy performance with $\mu_{\mathrm{M}}(s)$ and $\nu(s, s)$. The complete proof is given in the appendix.
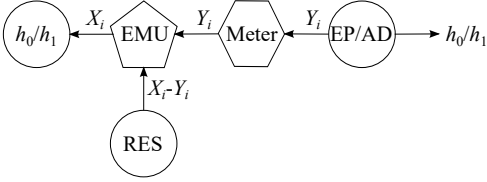
Fig. 2. The smart meter privacy problem in the presence of a renewable energy source (RES).

V. APPLICATION TO SMART METER PRIVACY PROBLEM

In this section, we consider an application of the theoretical framework we have introduced above to the smart meter privacy problem.

A. Smart Meter Privacy Model

The considered smart meter privacy problem is shown in Fig. 2. The i.i.d. data X_i denotes the *non-negative* energy demand at time slot i, from a finite energy demand alphabet $\mathcal{X}$. The binary hypothesis represents the private user behavior, e.g., the user is at home or not, or a particular appliance is being used or not. Assuming that the EP is an AD, the adversarial observation Y_i corresponds in this setting to the *non-negative* meter reading, or equivalently the energy supply from the EP, at time slot i. The finite energy supply alphabet satisfies $\mathcal{Y} \supseteq \mathcal{X}$.

The energy management unit (EMU) is a processor run by the user that manages the energy supplies to satisfy the energy demand. At every time slot, the EMU follows an instantaneous energy management policy to determine Y_i. It is assumed that the rest of the energy demand, $X_i - Y_i$, is satisfied from the renewable energy source (RES). We impose the following instantaneous constraint:

$$p_{Y_i|X_i}(y_i|x_i) = 0, \text{ if } y_i > x_i, \quad (40)$$

i.e., the RES cannot be charged with energy from the EP. We assume that the RES has a positive average energy generation rate s and is equipped with a sufficiently large energy storage. Accordingly, the following average energy constraint is imposed on the energy management policy over an n-slot time horizon for the long-term availability of renewable energy supply:

$$\mathrm{E}\left[\left.\frac{1}{n}\sum_{i=1}^{n}(X_i - Y_i)\right| h_j\right] \leq s, \ j = 0, 1. \quad (41)$$

The assumptions on the informed AD and the adversarial hypothesis test are the same as the general problem. Here the EP is considered as an AD, who has the authority to access the meter readings of energy supplies from the EP, and has all the statistical knowledge, but does not have access to the amount of energy supply from the RES, which is available locally to the user.

The privacy-preserving energy management is to modify the energy demand sequence $\{X_i\}$ into an energy supply sequence $\{Y_i\}$ by exploiting the RES in an online manner to prevent the EP from correctly inferring the user behavior.

B. Privacy-Preserving Energy Management

Note that an energy management policy satisfying (40) and (41) will satisfy the constraint (1) in the general problem. Therefore, the obtained theoretic results of asymptotic privacy performance in the general problem can be directly applied to the smart meter privacy problem. Here, we will focus on the memoryless hypothesis-aware energy management policy, which can be easily designed and implemented.

Under the adversarial Neyman-Pearson hypothesis test, and given a renewable energy generation rate $s > 0$, it follows from Corollary 2 and Theorem 2 that the exponent of the maximum minimal type II probability of error can be characterized by

$$\phi(s,s) = \min_{(p_{Y|X,h_0}, p_{Y|X,h_1}) \in \mathcal{P}_{\mathrm{E}}(s,s)} \left\{ \mathrm{D}(p_{Y|h_0}||p_{Y|h_1}) \right\}, \quad (42)$$

with the convex set

$$\mathcal{P}_{\mathrm{E}}(s,s)$$
$$\triangleq \left\{ (p_{Y|X,h_0}, p_{Y|X,h_1}) \left|
\begin{array}{l}
\mathrm{E}[X - Y|h_0] \leq s \\
\mathrm{E}[X - Y|h_1] \leq s \\
p_{Y|X,h_0}(y|x) = 0, \text{ if } y > x \\
p_{Y|X,h_1}(y|x) = 0, \text{ if } y > x
\end{array} \right. \right\}.$$

While solving the optimization problem in (42) leads to the optimal privacy performance, the energy supply alphabet $\mathcal{Y}$ can be very large, which means a highly complex optimization problem. On the other hand, the energy demand alphabet $\mathcal{X}$ is determined by a number of operation modes of the appliances and is typically finite. We show in the next theorem that the alphabet $\mathcal{Y}$ can be limited to the alphabet $\mathcal{X}$. This result can greatly simplify the numerical evaluation of $\phi(s,s)$ in the considered smart meter privacy context.

Theorem 7. *The energy supply alphabet can be limited to the energy demand alphabet under both hypotheses without loss of optimality for the evaluation of $\phi(s,s)$.*

Proof: Suppose that $\phi(s,s) = \mathrm{D}(p^*_{Y|h_0}||p^*_{Y|h_1})$ is achieved by $p^*_{Y|X,h_0}$ and $p^*_{Y|X,h_1}$. Let $\mathcal{X} = \{x_{(1)}, \ldots, x_{(|\mathcal{X}|)}\}$ with $x_{(i)} < x_{(k)}$ if $i < k$. Consider the following quantization operation which maps y to $\hat{y}$:

$$\hat{y} = \begin{cases} x_{(i)}, \text{ if } y \in (x_{(i-1)}, x_{(i)}], \ i \geq 2 \\ x_{(1)}, \text{ if } y \in [0, x_{(1)}] \end{cases}.$$

It can be verified that $p_{\hat{Y}|X,h_0}$, $p_{\hat{Y}|X,h_1}$ satisfy the constraints in the definition of $\mathcal{P}_{\mathrm{E}}(s,s)$. From the optimality in the definition of $\phi(s,s)$, we have

$$\phi(s,s) = \mathrm{D}(p^*_{Y|h_0}||p^*_{Y|h_1}) \leq \mathrm{D}(p_{\hat{Y}|h_0}||p_{\hat{Y}|h_1}).$$

In addition, due to the data processing inequality of Kullback-Leibler divergence [39, Theorem 9], we have

$$\phi(s,s) = \mathrm{D}(p^*_{Y|h_0}||p^*_{Y|h_1}) \geq \mathrm{D}(p_{\hat{Y}|h_0}||p_{\hat{Y}|h_1}).$$

Therefore,

$$\phi(s,s) = \mathrm{D}(p_{\hat{Y}|h_0}||p_{\hat{Y}|h_1}),$$

and the energy supply alphabet under both hypotheses can be constrained to $\mathcal{X}$ without loss of optimality. ∎

Under the adversarial Bayesian hypothesis test, and given a renewable energy generation rate $s > 0$, it follows from Corollary 4 and Theorem 5 that the exponent of the maximum minimal error probability can be characterized by

$$\nu(s,s) = \min_{(p_{Y|X,h_0}, p_{Y|X,h_1}) \in \mathcal{P}_E(s,s)} \left\{ C(p_{Y|h_0}, p_{Y|h_1}) \right\}. \quad (43)$$

Similarly, the following theorem shows that the supply alphabet $\mathcal{Y}$ can be limited to the demand alphabet $\mathcal{X}$ without loss of optimality for the numerical evaluation of $\nu(s,s)$ in the considered smart meter privacy context.

Theorem 8. *The energy supply alphabet can be limited to the energy demand alphabet under both hypotheses without loss of optimality for the evaluation of $\nu(s,s)$.*

Proof: Suppose that $\nu(s,s) = C(p^*_{Y|h_0}, p^*_{Y|h_1})$ is achieved by $p^*_{Y|X,h_0}$ and $p^*_{Y|X,h_1}$. We use the same quantization operation in the proof of Theorem 7 to map y to $\hat{y}$. Therefore, $p_{\hat{Y}|X,h_0}$ and $p_{\hat{Y}|X,h_1}$ satisfy the constraints in the definition of $\mathcal{P}_E(s,s)$. From the optimality in the definition of $\nu(s,s)$, it follows that

$$\nu(s,s) = C(p^*_{Y|h_0}, p^*_{Y|h_1}) \leq C(p_{\hat{Y}|h_0}, p_{\hat{Y}|h_1}).$$

In addition, from the data processing inequality of Rényi divergence [39, Theorem 9], we have

$C(p^*_{Y|h_0}, p^*_{Y|h_1})$
$= \max_{0 \leq \tau \leq 1} \left\{ C_\tau(p^*_{Y|h_0}, p^*_{Y|h_1}) \right\}$
$= \max\{ \max_{0 \leq \tau < 1} \{(1-\tau)D_\tau(p^*_{Y|h_0}||p^*_{Y|h_1})\}, D_0(p^*_{Y|h_1}||p^*_{Y|h_0}) \}$
$\geq \max\{ \max_{0 \leq \tau < 1} \{(1-\tau)D_\tau(p_{\hat{Y}|h_0}||p_{\hat{Y}|h_1})\}, D_0(p_{\hat{Y}|h_1}||p_{\hat{Y}|h_0}) \}$
$= \max_{0 \leq \tau \leq 1} \left\{ C_\tau(p_{\hat{Y}|h_0}, p_{\hat{Y}|h_1}) \right\}$
$= C(p_{\hat{Y}|h_0}, p_{\hat{Y}|h_1}).$

Therefore,

$$\nu(s,s) = C(p_{\hat{Y}|h_0}, p_{\hat{Y}|h_1}),$$

and the energy supply alphabet under both hypotheses can be constrained to $\mathcal{X}$ without loss of optimality. ∎

We also highlight here the connections and differences of our result with other smart meter privacy literature. While Kullback-Leibler divergence can be considered as yet another statistical similarity measure, our formulation here provides an operational meaning to its use as the privacy measure. This is in contrast to some of the other privacy measures considered in the literature, such as the mutual information used in [7], [31], [32], which are mainly proposed as distances between the energy demand data and the smart meter reading. A related work is [13] in the context of smart meter privacy, which uses the Fisher information as a privacy measure. The Fisher information is an approximation of the Kullback-Leibler divergence between two similar probability distributions [43]; therefore, the two privacy measures are closely related. But the authors in [13] focus on a parameter estimation problem, rather than a classification problem in our work, and exploit an energy storage device to provide privacy. It is difficult to compare our result with [7], [13], [31], [32] because of their different
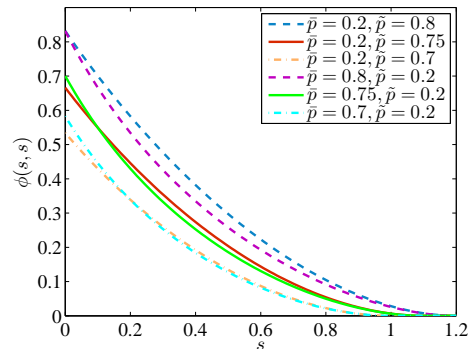

Fig. 3. Asymptotic privacy performance $\phi(s,s)$ for a binary demand model with different values of $\bar{p}$ and $\tilde{p}$.
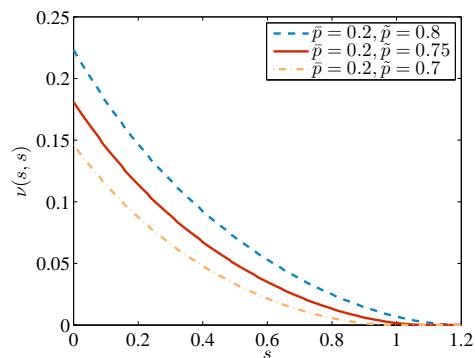

Fig. 4. Asymptotic privacy performance $\nu(s,s)$ for a binary demand model with different values of $\bar{p}$ and $\tilde{p}$.

problem settings, e.g., there is only one energy demand profile in [7], [13], [31], [32] while there are two energy demand profiles in our work. The connection between differential privacy and the minimum mutual information leakage problem has been revealed in [44]. The differential privacy model in [9], [10] considers the aggregation of meter readings rather than a sequence of meter readings in our work; and furthermore, the constraint of non-negative renewable energy supply in our work cannot provide the Laplace noise often used in the differential privacy model.

C. Binary Demand Example

We first present a simple example with binary energy demand alphabet[1] $\mathcal{X} = \{0,2\}$ for the purpose to numerically illustrate the main results in this paper. Based on Theorem 7 and Theorem 8, it is sufficient to consider a binary supply alphabet $\mathcal{Y} = \{0,2\}$, as well. Denote $p_{X|h_0}(0)$ by $\bar{p}$ and $p_{X|h_1}(0)$ by $\tilde{p}$.

Under the adversarial Neyman-Pearson test setting and using the optimal memoryless hypothesis-aware policy, the asymptotic privacy performance $\phi(s,s)$ is shown in Fig. 3 for different values of $\bar{p}$ and $\tilde{p}$. Confirming the claim in Lemma 2, it can be observed that $\phi(s,s)$ is convex and non-increasing.

[1]Certain appliances have binary energy demands, e.g., a fridge with "on" and "sleep" modes.

TABLE I
OPERATION MODES UNDER EACH HYPOTHESIS.

$p_{X\|h}$ \\ h \qquad x [W]	0	200	500	1200
h_0 (type A)	0.2528	0.3676	0	0.3796
h_1 (type B)	0.1599	0.0579	0.2318	0.5504

When $s = 0$, $X_i = Y_i$ for all $i \in \mathbb{Z}_+$ under both hypotheses, and $\phi(0,0) = \mathrm{D}(p_{X|h_0}\|p_{X|h_1})$. Intuitively, it is more difficult for the AD to identify the hypotheses when they lead to more similar energy demand profiles. It can be observed in Fig. 3 that $\phi(s,s)$ decreases as $\tilde{p}$ (resp. $\bar{p}$) gets closer to the fixed $\bar{p}$ (resp. $\tilde{p}$). Another interesting observation is that $\phi(s,s)$ curves for different settings of energy demand statistics $(\bar{p}, \tilde{p})$ might intersect. For instance, to achieve the privacy performance of 0.6, a lower renewable energy generation rate is required for $(\bar{p} = 0.2, \tilde{p} = 0.75)$ than that for $(\bar{p} = 0.75, \tilde{p} = 0.2)$; while to achieve the privacy performance of 0.3, a higher renewable energy generation rate is required for $(\bar{p} = 0.2, \tilde{p} = 0.75)$ than that for $(\bar{p} = 0.75, \tilde{p} = 0.2)$.

Under the adversarial Bayesian hypothesis test setting and using the optimal memoryless hypothesis-aware policy, the asymptotic privacy performance $\nu(s,s)$ is shown in Fig. 4 for different values of $\bar{p}$ and $\tilde{p}$. Confirming the claim in Lemma 5, the asymptotic privacy performance $\nu(s,s)$ is a convex and non-increasing function of s. From the same argument that more similar energy demand profiles make the AD more difficult to identify the hypotheses, it follows that $\nu(s,s)$ decreases as $\tilde{p}$ gets closer to the fixed $\bar{p}$. Note that the "opposite" settings, $(\bar{p} = 0.8, \tilde{p} = 0.2)$, $(\bar{p} = 0.75, \tilde{p} = 0.2)$, and $(\bar{p} = 0.7, \tilde{p} = 0.2)$, are not presented here since they lead to the same privacy performances as presented in the figure.

From the numerical results, the renewable energy generation rate that can guarantee a certain privacy performance can be determined.

D. Numerical Experiment

Here we present a numerical experiment with energy data from the REDD dataset [45] to illustrate the practical value of our theoretic results. We consider a kitchen with a dishwasher, which can be type A (h_0) or type B (h_1). From the energy data, we can identify four operation modes of a dishwasher. Table I shows the operation modes and the corresponding statistics obtained through training under the assumption of i.i.d. energy demands. We consider three renewable energy generation rates: $s = 0, 4000, 5000$ [W]. Given a value of s and under the adversarial Neyman-Pearson test setting, the optimal memoryless hypothesis-aware policy is implemented on the real energy demand data under each hypothesis. The resulting energy supplies under each hypothesis are shown in Fig. 5. When $s = 0$ [W], the energy supplies exactly follow the energy demands and it is easy for the AD to identify the hypothesis from the energy supply data. When $s = 4000$ [W], the optimal policy enforces that every operation mode under hypothesis h_0 and the same operation mode under hypothesis h_1 are statistically similar. When $s = 5000$ [W], most energy
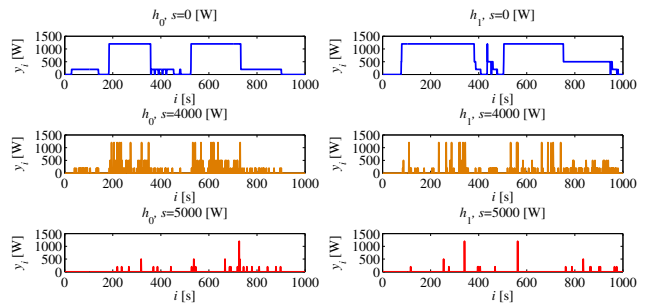

Fig. 5. Energy supplies under each hypothesis for different renewable energy generation rates when the optimal memoryless hypothesis-aware policy is used.

demands are satisfied from the RES, and it becomes very difficult for the AD to identify the hypothesis from the energy supply data.

VI. CONCLUSIONS

We have proposed novel formulations for privacy problem as adversarial hypothesis tests, where the fully-informed adversary makes an optimal hypothesis test on the privacy based on the adversarial observations; and the management unit manipulates the original data sequence into adversarial observations under an average distortion constraint to degrade the adversarial hypothesis testing accuracy. The privacy performance is evaluated by an asymptotic error exponent of the fully-informed adversary.

In the adversarial Neyman-Pearson hypothesis test setting, an asymptotically optimal privacy performance is shown to be characterized by a Kullback-Leibler divergence rate. Focusing on the worst case scenario, in which the type I probability of error upper bound approaches one, the asymptotic exponent of the maximum minimal type II probability of error achieved by the optimal memoryless hypothesis-aware policy is characterized by a single-letter Kullback-Leibler divergence; and it is also shown that the optimal memoryless hypothesis-aware policy cannot asymptotically outperform the optimal hypothesis-unaware policy with memory, since the management unit can learn the true hypothesis with an arbitrarily small probability of error from a sufficiently long sequence of original data. In the adversarial Bayesian hypothesis test setting, the informed adversary is assumed to have access to the prior knowledge on the hypotheses, and the privacy leakage is measured by the minimal error probability of the adversary. Asymptotic results are derived on the optimal privacy performance similar to the adversarial Neyman-Pearson hypothesis test setting by substituting Kullback-Leibler divergence with Chernoff information. These theoretic results can be directly applied to the smart meter privacy problem, where the energy supply alphabet can be constrained to the energy demand alphabet without loss of optimality for the evaluation of the single-letter privacy performances in both adversarial hypothesis test settings.

The asymptotic privacy performances derived in this work provide fundamental limits on any practical management policy and they therefore serve as references/benchmarks useful

for the performance assessment of any practical policy. The optimal memoryless hypothesis-aware policy is an i.i.d. policy, which can be easily designed and implemented in practice.

APPENDIX

A. Proof of Theorem 3

Proof: Let $o(n) \triangleq \lfloor \log n \rfloor$, $c(n) \triangleq o(n) + 1$, $q(n) \triangleq n - o(n)$, and $d_{\max} \triangleq \max_{x \in \mathcal{X}, y \in \mathcal{Y}} d(x, y)$. We choose any $\delta \in (0, s)$, $\omega \in (0, s)$, and type I probability of error upper bound $\varepsilon' \in (\max\{0, 1 - \min\{D(p_{X|h_0}||p_{X|h_1}), \frac{\delta}{d_{\max}}\}\}, 1)$. Then set $\xi = 1 - \varepsilon'$ and $\psi = \delta - d_{\max} \cdot \xi$. It can be verified that $0 < \xi < D(p_{X|h_0}||p_{X|h_1})$ and $0 < \psi \leq \delta$. We use these parameters to construct a hypothesis-unaware policy with memory, ρ_{p}^n, over an n-slot time horizon, which consists of two successive phases.

$o(n)$-slot learning phase. The goal of the MU is to learn the true hypothesis at the end of the first phase. To prevent privacy leakage during the learning phase, identical instantaneous policies are used at all time slots as:

$$y_i = \rho_i(x_i) = \min \mathcal{Y}, \ \forall i \leq o(n), \ \forall x_i \in \mathcal{X}.$$

Based on the observations of original data $x^{o(n)}$, the MU makes a decision $\hat{H}$:

$$\hat{H} = \begin{cases} h_0, & \text{if } x^{o(n)} \in \mathcal{A}_\xi^{o(n)}(p_{X|h_0}||p_{X|h_1}) \\ h_1, & \text{otherwise} \end{cases},$$

where $\mathcal{A}_\xi^{o(n)}(p_{X|h_0}||p_{X|h_1})$ denotes a relative entropy typical set as defined in [36], and any sequence $x^{o(n)} \in \mathcal{A}_\xi^{o(n)}(p_{X|h_0}||p_{X|h_1})$ satisfies

$$\left| \frac{1}{o(n)} \log \frac{p_{X^{o(n)}|h_0}\left(x^{o(n)}\right)}{p_{X^{o(n)}|h_1}\left(x^{o(n)}\right)} - D(p_{X|h_0}||p_{X|h_1}) \right| \leq \xi.$$

The MU can make a wrong decision. Let $p_{\bar{e}} = 1 - p_{X^{o(n)}|h_0}(\mathcal{A}_\xi^{o(n)}(p_{X|h_0}||p_{X|h_1}))$ denote the type I probability of error and $p_{\bar{e}} = p_{X^{o(n)}|h_1}(\mathcal{A}_\xi^{o(n)}(p_{X|h_0}||p_{X|h_1}))$ denote the type II probability of error for the MU.

$q(n)$-slot privacy-preserving phase. Depending on the decision $\hat{H}$ in the learning phase, identical instantaneous policies are used at all slots of the second phase as follows.
If $\hat{H} = h_0$ (resp. $\hat{H} = h_1$) and for all $i \in \{c(n), \dots, n\}$, $\rho_i : p_{Y_i|X_i}$ corresponds to the optimizer $p_{Y|X,h_0}^*$ (resp. $p_{Y|X,h_1}^*$) of $\phi(s - \delta, s - \omega)$.

Next, we check the average distortion constraint. Under hypothesis h_0 (resp. h_1), if a correct decision is made at the end of the learning phase, the instantaneous management policies used in the privacy-preserving phase guarantee that the single-slot average distortion upper bound $s - \delta$ (resp. $s - \omega$) is satisfied at all time slots of this phase; otherwise, bounded violation of the single-slot average distortion constraint might happen at all time slots of the privacy-preserving phase. When n is sufficiently large, we have

$$\mathrm{E}\left[\frac{1}{n}\sum_{i=1}^n d(X_i, Y_i)\middle| h_0\right]$$
$$\leq \frac{o(n)}{n} d_{\max} + \mathrm{E}\left[\frac{1}{n}\sum_{i=c(n)}^n d(X_i, Y_i)\middle| h_0\right]$$
$$\stackrel{(a)}{\leq} \frac{o(n)}{n} d_{\max} + \frac{q(n)}{n}(s - \delta + d_{\max} \cdot \xi)$$
$$\leq s - \delta + d_{\max} \cdot \xi + \frac{o(n)}{n} d_{\max}$$
$$= s - \left(\psi - \frac{o(n)}{n} d_{\max}\right)$$
$$\stackrel{(b)}{\leq} s,$$

and

$$\mathrm{E}\left[\frac{1}{n}\sum_{i=1}^n d(X_i, Y_i)\middle| h_1\right]$$
$$\leq \frac{o(n)}{n} d_{\max} + \mathrm{E}\left[\frac{1}{n}\sum_{i=c(n)}^n d(X_i, Y_i)\middle| h_1\right]$$
$$\leq \frac{o(n)}{n} d_{\max} + \frac{q(n)}{n}(s - \omega + d_{\max} \cdot p_{\bar{e}})$$
$$\stackrel{(c)}{\leq} s + \underbrace{d_{\max} \cdot e^{-o(n) \cdot (D(p_{X|h_0}||p_{X|h_1}) - \xi)} + \frac{o(n)}{n} d_{\max} - \omega}_{\Delta(n)}$$
$$\stackrel{(d)}{\leq} s,$$

where (a) follows since $p_{\bar{e}} < \xi$ when $o(n)$ is sufficiently large [36, Theorem 11.8.2]; (b) follows since $\psi - \frac{o(n)}{n} d_{\max} \geq 0$ when n is sufficiently large; (c) follows since $p_{\bar{e}} < e^{-o(n) \cdot (D(p_{X|h_0}||p_{X|h_1}) - \xi)}$ [36, Theorem 11.8.2]; and (d) follows since $\Delta(n) \leq 0$ when n is sufficiently large.

Therefore, ρ_{p}^n is a management policy which satisfies the average distortion constraint in (1) when n is sufficiently large. Then, we have

$$\liminf_{n \to \infty} \frac{1}{q(n)} \log \frac{1}{\beta(n, \varepsilon', \rho_{\mathrm{p}}^n)} \geq \liminf_{n \to \infty} \frac{1}{n} \log \frac{1}{\beta(n, \varepsilon', \rho_{\mathrm{p}}^n)}$$
$$\geq \liminf_{n \to \infty} \frac{1}{n} \log \frac{1}{\beta_{\mathrm{M}}(n, \varepsilon', s)}, \tag{44}$$

where the first inequality follows from $q(n) \leq n$; and the second inequality follows since the constructed management policy ρ_{p}^n is not necessarily optimal for all sufficiently large n.

From the point of view of the informed AD, the observations $y^{o(n)}$ in the learning phase do not reveal any information about the hypothesis. Therefore, the strategy of the AD only depends on the observation statistics in the privacy-preserving phase $p_{Y_{c(n)}^n|h_0}$ and $p_{Y_{c(n)}^n|h_1}$. Then, the term $\frac{1}{q(n)} \log \frac{1}{\beta(n, \varepsilon', \rho_{\mathrm{p}}^n)}$ in (44) represents the exponent of the minimal type II probability of error in the privacy-preserving phase. With the management policy ρ_{p}^n specified above, the sequence of adversarial observations in the privacy-preserving phase is a mixture of the i.i.d. sequences $Y_{c(n)}^n|\hat{H} = h_0, H = h_0$ and $Y_{c(n)}^n|\hat{H} = h_1, H = h_0$

under hypothesis h_0, or a mixture of the i.i.d. sequences $Y_{c(n)}^n|\hat{H} = h_1, H = h_1$ and $Y_{c(n)}^n|\hat{H} = h_0, H = h_1$ under hypothesis h_1. The probability distributions are

$$p_{Y_{c(n)}^n|h_0}\left(y_{c(n)}^n\right) = (1-p_{\bar{e}}) \cdot p_{Y_{c(n)}^n|\hat{H}=h_0, H=h_0}\left(y_{c(n)}^n\right)$$
$$+ p_{\bar{e}} \cdot p_{Y_{c(n)}^n|\hat{H}=h_1, H=h_0}\left(y_{c(n)}^n\right),$$

$$p_{Y_{c(n)}^n|h_1}\left(y_{c(n)}^n\right) = (1-p_{\tilde{e}}) \cdot p_{Y_{c(n)}^n|\hat{H}=h_1, H=h_1}\left(y_{c(n)}^n\right)$$
$$+ p_{\tilde{e}} \cdot p_{Y_{c(n)}^n|\hat{H}=h_0, H=h_1}\left(y_{c(n)}^n\right).$$

We define

$$\mathcal{B}(R,n) \triangleq \left\{ y_{c(n)}^n \left| \frac{1}{q(n)} \log \frac{p_{Y_{c(n)}^n|h_0}\left(y_{c(n)}^n\right)}{p_{Y_{c(n)}^n|h_1}\left(y_{c(n)}^n\right)} \leq R \right. \right\},$$

and

$$K(R) \triangleq \limsup_{n\to\infty} p_{Y_{c(n)}^n|h_0}\left(\mathcal{B}(R,n)\right).$$

Based on the information-spectrum results [46, Theorem 1], [37, Theorem 4.2.1], we have

$$\sup\{R|K(R) \leq \varepsilon'\} \geq \liminf_{n\to\infty} \frac{1}{q(n)} \log \frac{1}{\beta(n,\varepsilon',\rho_p^n)}. \quad (45)$$

In the asymptotic regime as $n \to \infty$, $Y_{c(n)}^n|h_1$ reduces to the i.i.d. sequence $Y_{c(n)}^n|\hat{H} = h_1, H = h_1$ since $\lim_{n\to\infty} p_{\tilde{e}} \leq \lim_{n\to\infty} e^{-o(n)\cdot(\mathrm{D}(p_{X|h_0}||p_{X|h_1})-\xi)} = 0$. Based on the case study [37, Example 4.2.1], the upper bound $\sup\{R|K(R) \leq \varepsilon'\}$ is characterized by the Kullback-Leibler divergences $\mathrm{D}_1 = \mathrm{D}(p_{Y|\hat{H}=h_0, H=h_0}||p_{Y|\hat{H}=h_1, H=h_1})$ and $\mathrm{D}_2 = \mathrm{D}(p_{Y|\hat{H}=h_1, H=h_0}||p_{Y|\hat{H}=h_1, H=h_1})$ as follows. If $\mathrm{D}_1 \geq \mathrm{D}_2$, we have

$$\mathrm{D}(p_{Y|\hat{H}=h_0, H=h_0}||p_{Y|\hat{H}=h_1, H=h_1}) \geq \sup\{R|K(R) \leq \varepsilon'\};$$

otherwise, since $1 - p_{\bar{e}} > 1 - \xi = \varepsilon'$ as $n \to \infty$, we have

$$\mathrm{D}(p_{Y|\hat{H}=h_0, H=h_0}||p_{Y|\hat{H}=h_1, H=h_1}) = \sup\{R|K(R) \leq \varepsilon'\}.$$

Then, it follows that

$$\phi(s-\delta, s-\omega) = \mathrm{D}(p_{Y|\hat{H}=h_0, H=h_0}||p_{Y|\hat{H}=h_1, H=h_1})$$
$$\geq \sup\{R|K(R) \leq \varepsilon'\}. \quad (46)$$

The inequalities (44), (45), and (46) jointly lead to

$$\phi(s-\delta, s-\omega) \geq \liminf_{n\to\infty} \frac{1}{n} \log \frac{1}{\beta_{\mathrm{M}}(n,\varepsilon',s)}. \quad (47)$$

Given $\delta, \omega \in (0,s)$, the inequality (47) holds for all type I probability of error upper bounds $\varepsilon' \in (\max\{0, 1 - \min\{\mathrm{D}(p_{X|h_0}||p_{X|h_1}), \frac{\delta}{d_{\max}}\}\}, 1)$. Therefore,

$$\phi(s-\delta, s-\omega) \geq \lim_{\varepsilon\to 1} \liminf_{n\to\infty} \frac{1}{n} \log \frac{1}{\beta_{\mathrm{M}}(n,\varepsilon,s)} = \theta_{\mathrm{M}}(s). \quad (48)$$

Since the inequality (48) holds for all $\delta, \omega \in (0,s)$, we further have

$$\phi(s,s) = \inf_{\delta,\omega\in(0,s)}\{\phi(s-\delta, s-\omega)\} \geq \theta_{\mathrm{M}}(s),$$

where the equality follows from the non-increasing and continuous properties of $\phi(\bar{s}, \tilde{s})$. ∎

B. Proof of Theorem 6

Proof: Let $o(n) \triangleq \lfloor \log n \rfloor$, $c(n) \triangleq o(n) + 1$, $q(n) \triangleq n - o(n)$, $d_{\max} \triangleq \max_{x\in\mathcal{X}, y\in\mathcal{Y}} d(x,y)$. We choose any $\delta \in (0,s)$, $\omega \in (0,s)$, and $\xi \in \left(0, \min\left\{\mathrm{D}(p_{X|h_0}||p_{X|h_1}), \frac{\delta}{d_{\max}}\right\}\right)$. Then set $\psi = \delta - d_{\max} \cdot \xi$. It can be verified that $0 < \psi < \delta$. We use these parameters to construct a hypothesis-unaware management policy with memory, ρ_q^n, over an n-slot time horizon, which similarly has two phases as ρ_p^n.

$o(n)$-slot learning phase. This phase is the same as the learning phase of ρ_p^n: Identical instantaneous privacy-unaware policies are independently used at all slots and always output an adversarial observation sequence of $\min \mathcal{Y}$ regardless of the sequence of original data; and the learning decision at the end of this phase is $\hat{H} = 0$ if the sequence $x^{o(n)}$ is in the relative entropy typical set $\mathcal{A}_\xi^{o(n)}\left(p_{X|h_0}||p_{X|h_1}\right)$ or is $\hat{H} = 1$ otherwise. Again, we denote the type I probability of learning error by $p_{\bar{e}}$ and the type II probability of learning error by $p_{\tilde{e}}$.

$q(n)$-slot privacy-preserving phase. Depending on the decision $\hat{H}$ in the learning phase, identical instantaneous management policies are used at all slots of the second phase as follows. If $\hat{H} = h_0$ (resp. $\hat{H} = h_1$) and for all $i \in \{c(n), \ldots, n\}$, $\rho_i : p_{Y_i|X_i}$ corresponds to the optimizer $p_{Y|X,h_0}^*$ (resp. $p_{Y|X,h_1}^*$) of $\nu(s-\delta, s-\omega)$.

It follows from the same analysis on ρ_p^n that the policy ρ_q^n satisfies the average distortion constraint in (1) when n is sufficiently large. Similar to (44), we have

$$\liminf_{n\to\infty} \frac{1}{q(n)} \log \frac{1}{\alpha(n,\rho_q^n)} \geq \liminf_{n\to\infty} \frac{1}{n} \log \frac{1}{\alpha(n,\rho_q^n)}$$
$$\geq \lim_{n\to\infty} \frac{1}{n} \log \frac{1}{\alpha_{\mathrm{M}}(n,s)} = \mu_{\mathrm{M}}(s). \quad (49)$$

Since $y^{o(n)}$ in the learning phase is a deterministic sequence of $\min \mathcal{Y}$, the strategy of the AD only depends on the observation statistics in the privacy-preserving phase $p_{Y_{c(n)}^n|h_0}$ and $p_{Y_{c(n)}^n|h_1}$. Then, the term $\frac{1}{q(n)} \log \frac{1}{\alpha(n,\rho_q^n)}$ in (49) represents the exponent of the minimal error probability in the privacy-preserving phase. With the management policy ρ_q^n specified above, the sequence of adversarial observations in the privacy-preserving phase is a mixture of the i.i.d. sequences $Y_{c(n)}^n|\hat{H} = h_0, H = h_0$ with a probability $1 - p_{\bar{e}}$ and $Y_{c(n)}^n|\hat{H} = h_1, H = h_0$ with a probability $p_{\bar{e}}$ under hypothesis h_0, or a mixture of the i.i.d. sequences $Y_{c(n)}^n|\hat{H} = h_1, H = h_1$ with a probability $1 - p_{\tilde{e}}$ and $Y_{c(n)}^n|\hat{H} = h_0, H = h_1$ with a probability $p_{\tilde{e}}$ under hypothesis h_1. In the asymptotic regime as $n \to \infty$, $Y_{c(n)}^n|h_1$ can be seen as the i.i.d. sequence $Y_{c(n)}^n|\hat{H} = h_1, H = h_1$ since $\lim_{n\to\infty} p_{\tilde{e}} \leq \lim_{n\to\infty} e^{-o(n)\cdot(\mathrm{D}(p_{X|h_0}||p_{X|h_1})-\xi)} = 0$. Then, we have

$$\liminf_{n\to\infty} \frac{1}{q(n)} \log \frac{1}{\alpha(n,\rho_q^n)}$$
$$= \mathrm{C}(p_{Y|\hat{H}=h_0, H=h_0}, p_{Y|\hat{H}=h_1, H=h_1}) + \Delta(\xi)$$
$$= \nu(s-\delta, s-\omega) + \Delta(\xi), \quad (50)$$

where $\nu(s-\delta, s-\omega)$ corresponds to the asymptotic exponent of the minimal error probability in the privacy-preserving phase

if $p_{\bar{e}} = 0$; and the term $\Delta(\xi)$ denotes the impact of the i.i.d. sequence $Y_{c(n)}^n|\hat{H} = h_1, H = h_0$ with a probability $p_{\bar{e}}$ under hypothesis h_0. The inequalities (49) and (50) jointly lead to

$$\mu_{\mathrm{M}}(s) \leq \nu(s-\delta, s-\omega) + \Delta(\xi). \tag{51}$$

Since $\limsup_{n\to\infty} p_{\bar{e}} \leq \xi$, the term $\Delta(\xi)$ satisfies

$$\lim_{\xi\to 0} \Delta(\xi) = 0. \tag{52}$$

Given $\delta, \omega \in (0, s)$, the inequality (51) holds for all $\xi \in \left(0, \min\left\{ \mathrm{D}(p_{X|h_0}||p_{X|h_1}), \frac{\delta}{d_{\max}} \right\}\right)$. Therefore,

$$\mu_{\mathrm{M}}(s) \leq \nu(s-\delta, s-\omega) + \lim_{\xi\to 0} \Delta(\xi) = \nu(s-\delta, s-\omega). \tag{53}$$

Since (53) holds for all $\delta, \omega \in (0, s)$, we further have

$$\mu_{\mathrm{M}}(s) \leq \inf_{\delta,\omega\in(0,s)} \{\nu(s-\delta, s-\omega)\} = \nu(s, s), \tag{54}$$

where the equality follows from the non-increasing and continuous properties of $\nu(\bar{s}, \tilde{s})$. ∎

References

[1] Z. Li, T. J. Oechtering, and D. Gündüz, "Smart meter privacy based on adversarial hypothesis testing," in *Proceedings of ISIT 2017*, pp. 774–778.

[2] C. Dwork, "Differential privacy," in *Proceedings of ICALP 2006*, pp. 1–12.

[3] D. Rebollo-Monedero, J. Forné, and J. Domingo-Ferrer, "From t-closeness-like privacy to postrandomization via information theory," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 11, pp. 1623–1636, 2010.

[4] F. P. Calmon and N. Fawaz, "Privacy against statistical inference," in *Proceedings of Allerton 2012*, pp. 1401–1408.

[5] B. Rassouli and D. Gündüz, "On perfect privacy," in *Proceedings of ISIT 2018*, pp. 2551–2555.

[6] L. Sweeney, "K-anonymity: A model for protecting privacy," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 10, no. 5, pp. 557–570, 2002.

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

[8] I. Issa, S. Kamath, and A. B. Wagner, "An operational measure of information leakage," in *Proceedings of CISS 2016*, pp. 234–239.

[9] G. Ács and C. Castelluccia, "I have a DREAM! (DiffeRentially privatE smArt Metering)," in *Proceedings of International Workshop on Information Hiding*, 2011, pp. 118–132.

[10] J. Zhao, T. Jung, Y. Wang, and X. Li, "Achieving differential privacy of data disclosure in the smart grid," in *Proceedings of INFOCOM 2014*, pp. 504–512.

[11] A. D. Wyner, "The wire-tap channel," *The Bell System Technical Journal*, vol. 54, no. 8, pp. 1355–1387, 1975.

[12] M. Mhanna and P. Piantanida, "On secure distributed hypothesis testing," in *Proceedings of ISIT 2015*, pp. 1605–1609.

[13] F. Farokhi and H. Sandberg, "Fisher information as a measure of privacy: Preserving privacy of households with smart meters using batteries," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 4726–4734, 2018.

[14] Z. Li and T. J. Oechtering, "Privacy on hypothesis testing in smart grids," in *Proceedings of ITW 2015 Fall*, pp. 337–341.

[15] V. S. S. Nadendla and P. K. Varshney, "Design of binary quantizers for distributed detection under secrecy constraints," *IEEE Transactions on Signal Processing*, vol. 64, no. 10, pp. 2636–2648, 2016.

[16] B. Rassouli and D. Gündüz, "Optimal utility-privacy trade-off with total variation distance as a privacy measure," in *Proceedings of ITW 2018*, pp. 460–464.

[17] L. Willenborg and T. d. Waal, *Elements of Statistical Disclosure Control*. Springer-Verlag New York, 2001.

[18] Z. Li and T. J. Oechtering, "Privacy-aware distributed Bayesian detection," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 7, pp. 1345–1357, 2015.

[19] J. Liao, L. Sankar, V. Y. F. Tan, and F. P. Calmon, "Hypothesis testing in the high privacy limit," in *Proceedings of Allerton 2016*, pp. 649–656.

[20] J. Liao, L. Sankar, V. Y. F. Tan, and F. du Pin Calmon, "Hypothesis testing under mutual information privacy constraints in the high privacy regime," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 4, pp. 1058–1071, 2018.

[21] S. Sreekumar, D. Gündüz, and A. Cohen, "Distributed hypothesis testing under privacy constraints," in *Proceedings of ITW 2018*, pp. 470–474.

[22] L. Huang, A. D. Joseph, B. Nelson, B. I. P. Rubinstein, and J. D. Tygar, "Adversarial machine learning," in *Proceedings of AISec 2011*, pp. 43–57.

[23] A. Kurakin, I. Goodfellow, and S. Bengio, "Adversarial machine learning at scale," eprint arXiv:1611.01236.

[24] A. Athalye, L. Engstrom, A. Ilyas, and K. Kwok, "Synthesizing robust adversarial examples," eprint arXiv:1707.07397.

[25] F. Tramèr, A. Kurakin, N. Papernot, I. Goodfellow, D. Boneh, and P. McDaniel, "Ensemble adversarial training: Attacks and defenses," eprint arXiv:1705.07204.

[26] I. J. Goodfellow, J. Shlens, and C. Szegedy, "Explaining and harnessing adversarial examples," eprint arXiv:1412.6572.

[27] M. Barni and B. Tondi, "The source identification game: An information-theoretic perspective," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 3, pp. 450–463, 2013.

[28] B. Tondi, M. Barni, and N. Merhav, "Detection games with a fully active attacker," in *Proceedings of WIFS 2015*, pp. 1–6.

[29] M. Barni and B. Tondi, "Source distinguishability under distortion-limited attack: An optimal transport perspective," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 10, pp. 2145–2159, 2016.

[30] G. Giaconi, D. Gündüz, and H. V. Poor, "Privacy-aware smart metering: Progress and challenges," eprint arXiv:1802.01166.

[31] O. Tan, D. Gündüz, and H. V. Poor, "Increasing smart meter privacy through energy harvesting and storage devices," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 7, pp. 1331–1341, 2013.

[32] D. Varodayan and A. Khisti, "Smart meter privacy using a rechargeable battery: Minimizing the rate of information leakage," in *Proceedings of ICASSP 2011*, pp. 1932–1935.

[33] G. Giaconi, D. Gündüz, and H. V. Poor, "Smart meter privacy with renewable energy and an energy storage device," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 1, pp. 129–142, 2018.

[34] J. Fan, Q. Li, and G. Cao, "Privacy disclosure through smart meters: Reactive power based attack and defense," in *Proceedings of DSN 2017*, pp. 13–24.

[35] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Cambridge University Press, 2011.

[36] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley-Interscience, 2006.

[37] T. S. Han, *Information-Spectrum Methods in Information Theory*. Springer Berlin Heidelberg, 2003.

[38] "Every convex function is locally Lipschitz," *The American Mathematical Monthly*, vol. 79, no. 10, pp. 1121–1124, 1972.

[39] T. van Erven and P. Harremoës, "Rényi divergence and Kullback-Leibler divergence," *IEEE Transactions on Information Theory*, vol. 60, no. 7, pp. 3797–3820, 2014.

[40] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[41] H. L. van Trees, K. L. Bell, and Z. Tian, *Detection, Estimation, and Modulation Theory, Part I: Detection, Estimation, and Filtering Theory*. Wiley, 2013.

[42] H. Nikaidô, "On von Neumann's minimax theorem," *Pacific Journal of Mathematics*, vol. 4, no. 1, pp. 65–72, 1954.

[43] H. Jeffreys, "An invariant form for the prior probability in estimation problems," *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 186, no. 1007, pp. 453–461, 1946.

[44] W. Wang, L. Ying, and J. Zhang, "On the relation between identifiability, differential privacy, and mutual-information privacy," *IEEE Transactions on Information Theory*, vol. 62, no. 9, pp. 5018–5029, 2016.

[45] J. Z. Kolter and M. J. Johnson, "REDD: A public data set for energy disaggregation research," in *Proceedings of the SustKDD Workshop on Data Mining Applications in Sustainability*, 2011, pp. 1–6.

[46] P.-N. Chen, "General formulas for the Neyman-Pearson Type-II error exponent subject to fixed and exponential Type-I error bounds," *IEEE Transactions on Information Theory*, vol. 42, no. 1, pp. 316–323, 1996.

Zuxing Li (S'14-M'18) received his B.Eng. degree in Information Security in 2009 from Shanghai Jiao Tong University, P. R. China, his M.Sc. degrees in Electrical Engineering in 2013 from Technical University of Catalonia, Spain and KTH Royal Institute of Technology, Sweden, and his Ph.D. degree in Electrical Engineering in 2017 from KTH Royal Institute of Technology. His research interests include statistical inference, information theory, privacy, and machine learning.

Tobias J. Oechtering (S'01-M'08-SM'12) received his Dipl-Ing degree in Electrical Engineering and Information Technology in 2002 from RWTH Aachen University, Germany, his Dr-Ing degree in Electrical Engineering in 2007 from the Technische Universität Berlin, Germany, and his Docent degree in Communication Theory in 2012 from KTH Royal Institute of Technology, Sweden. In 2008, he joined KTH Royal Institute of Technology and has been a Professor in the Department of Information Science and Engineering since November 2018. He is currently Associate Editor of IEEE Transactions on Information Forensics and Security since June 2016 and was Editor of IEEE Communications Letters during 2012-2015. Dr. Oechtering received the "Förderpreis 2009" from the Vodafone Foundation. His research interests include communication and information theory, physical layer privacy and security, statistical signal processing, as well as communication for networked control.

Deniz Gündüz (S'03-M'08-SM'13) received his B.S. degree in Electrical and Electronics Engineering from METU, Turkey in 2002, and his M.S. and Ph.D. degrees in Electrical Engineering from NYU Polytechnic School of Engineering (formerly Polytechnic University) in 2004 and 2007, respectively. After his PhD, he served as a postdoctoral research associate at Princeton University, and as a consulting assistant professor at Stanford University. He was a research associate at CTTC in Barcelona, Spain until September 2012. Then he joined the Electrical and Electronic Engineering Department of Imperial College London, UK, where he is currently a Reader (Associate Professor) in Information Theory and Communications, and leads the Information Processing and Communications Laboratory (IPC-Lab).

His research interests lie in the areas of communications and information theory, machine learning, and privacy. Dr. Gündüz is an Editor of the IEEE Transactions on Green Communications and Networking, and a Guest Editor of the IEEE Journal on Selected Areas in Communications, Special Issue on Machine Learning in Wireless Communication. He served as an Editor of the IEEE Transactions on Communications from 2013 until 2018. He is the recipient of the IEEE Communications Society - Communication Theory Technical Committee (CTTC) Early Achievement Award in 2017, a Starting Grant of the European Research Council (ERC) in 2016, IEEE Communications Society Best Young Researcher Award for the Europe, Middle East, and Africa Region in 2014, Best Paper Award at the 2016 IEEE Wireless Communications and Networking Conference (WCNC), and the Best Student Paper Awards at the 2018 IEEE Wireless Communications and Networking Conference (WCNC) and the 2007 IEEE International Symposium on Information Theory (ISIT). He is the General Co-chair of the 2019 London Symposium on Information Theory. In the past, he served as the General Co-chair of the 2018 International ITG Workshop on Smart Antennas and the 2016 IEEE Information Theory Workshop.