

# A Unified Approach to Differentially Private Bayes Point Estimation \*

Braghadeesh Lakshminarayanan \* Cristian R. Rojas \*

\* *Decision and Control Systems, KTH Royal Institute of Technology,  
10044 Stockholm, Sweden (e-mails: blak@kth.se, crro@kth.se).*

**Abstract:** Parameter estimation in statistics and system identification relies on data that may contain sensitive information. To protect this sensitive information, the notion of *differential privacy* (DP) has been proposed, which enforces confidentiality by introducing randomization in the estimates. Standard algorithms for differentially private estimation are based on adding an appropriate amount of noise to the output of a traditional point estimation method. This leads to an accuracy-privacy trade off, as adding more noise reduces the accuracy while increasing privacy. In this paper, we propose a new Unified Bayes Private Point (UBaPP) approach to Bayes point estimation of the unknown parameters of a data generating mechanism under a DP constraint, that achieves a better accuracy-privacy trade off than traditional approaches. We verify the performance of our approach on a simple numerical example.

**Keywords:** Differential privacy; Parameter estimation; Bayes point estimation.

## 1. INTRODUCTION

Parameter estimation deals with the problem of approximating the unknown parameters of a mathematical model that describes a given real phenomenon, using data collected from that phenomenon. This problem has been intensely studied in areas such as statistics (Casella and Berger, 2002), system identification (Söderström and Stoica, 1989; Ljung, 1999), and machine learning (Shalev-Shwartz and Ben-David, 2012).

An important subfield of parameter estimation is point estimation, where the goal is to approximate the unknown quantity by a single value. Some of the most commonly used point estimators are Maximum Likelihood (Casella and Berger, 2002; Lehmann and Casella, 1998), the Method of Moments (Gouriéroux and Monfort, 1996), the Minimum Mean Square Error estimator (Van Trees, 2004), the Minimum Variance Unbiased estimator (Kay, 1997), and the Best Linear Unbiased estimator (McElroy, 1967).

Given a point estimate, or in general a function of the data, it is possible to retrieve some information about the individual samples used to compute it. For example, it was demonstrated in genomic studies (Homer et al., 2008) that, under some conditions, it is possible to identify whether or not the DNA sample of an individual was present in a dataset based on aggregate statistics. For this reason, the National Institute of Health (NIH) removed access to some aggregate statistics such as p-values and chi-squared statistics, which were once openly available (NIH, 2014). This and related concerns on confidentiality related to data

handling has led to the development of point estimators subject to privacy considerations.

An important notion of privacy that is considered in the literature is *differential privacy* (DP) (Dwork and Roth, 2014). DP is a privacy constraint that can be imposed on algorithms in order to protect the sensitive information contained in their output. DP ensures that by seeing the output of an algorithm, almost no probabilistic inference can be made about the observations used by such algorithm to produce this output; the level of desired privacy can be tuned via a parameter  $\epsilon > 0$ . A well known approach, known as the Laplace mechanism, enforces the DP constraint by adding a suitable amount of Laplace distributed random noise to the output of the algorithm.

Within automatic control, privacy has been well studied. For instance, (Sankar et al., 2013; Varodayan and Khisti, 2011; Nekouei et al., 2022) consider information theory approaches to satisfy privacy constraints. DP has also been considered, e.g., in Le Ny and Pappas (2013); Wang et al. (2018).

One of the main issues that arise while enforcing DP in point estimation is the so-called accuracy-privacy trade off (Wang et al., 2017; Cao and Başar, 2020): Enforcing a higher level of privacy (for example, by adding Laplace noise of larger variance) reduces the accuracy of the estimator. Most of the works that consider DP in parameter estimation rely on the Laplace mechanism to enforce DP. However, it is not clear if this mechanism achieves an optimal accuracy-privacy trade off. In this paper, we provide an alternative approach that achieves an optimal accuracy-privacy trade off for Bayes point estimation by posing the problem of maximizing accuracy subject to a DP constraint as a convex optimization program. In particular, our contributions are the following:

\* This work has been supported by the Swedish Research Council under contract number 2016-06079 (NewLEADS) and by the Digital Futures project EXTREMUM. The authors are with the Division of Decision and Control Systems, KTH Royal Institute of Technology, 100 44 Stockholm, Sweden (e-mails: blak@kth.se; crro@kth.se).

- We formulate the problem of Bayes point estimation subject to a DP constraint as a convex optimization program;
- we provide an approach (UBaPP) to solve the above optimization program for the case where the parameter space and observations are discrete;
- we demonstrate the advantage of our approach via a simple numerical example based on Bernoulli samples.

The paper is organized as follows: Section 2 defines the notion of DP and reviews Bayes point estimation. In Section 3, we state the problem formulation, while in Section 4 we propose our new approach (UBaPP). Then, we demonstrate our approach through a numerical example in Section 5, and in Section 6 we conclude the paper and discuss future work.

## 2. PRELIMINARIES

In this section, we formalize the problem of Bayes point estimation, and introduce DP. Then, we define the *Laplace mechanism*, a procedure that enforces DP.

### 2.1 Bayes Point Estimation

Consider observing a physical process that generates independent and identically distributed (*i.i.d.*) samples  $x_1, \dots, x_n \in \mathbb{R}^p$  at discrete time instants  $1, \dots, n$  respectively, according to a probability distribution  $\mathbb{P}(\cdot|\theta)$  that is parameterized by an unknown parameter  $\theta \in \Theta \subseteq \mathbb{R}^{d_\theta}$ . Let  $\mathbf{x} := (x_1, \dots, x_n)^T \in \mathbb{R}^{n \times p} =: \mathcal{X}$ . Let  $S$  be a *sufficient statistic* (Lehmann and Casella, 1998)  $S: \mathcal{X} \rightarrow \mathcal{Y} := \mathbb{R}^m$  where  $\mathbf{y} = (y_1, \dots, y_m)^T = S(\mathbf{x})$ , with  $m \leq n$ . We call  $\mathbf{y}$  *observations* and  $\mathcal{Y}$  the *observation space*, and  $\mathbf{x}$  *input* and  $\mathcal{X}$  the *input space*.

Given  $\mathbf{x} \in \mathcal{X}$ , the goal in point estimation is to construct an *estimator* or *decision rule*<sup>1</sup>, which is a mapping  $\delta: \mathcal{Y} \rightarrow \Theta$ , such that the *risk*

$$R(\theta, \delta) := \mathbb{E}[L(\theta, \delta(S(\mathbf{x})))] \quad (1)$$

is as small as possible, where  $L: \Theta \times \Theta \rightarrow \mathbb{R}_0^+$  is a *loss function*.

The expectation in (1) is taken with respect to the probability distribution of  $\mathbf{x}$ , *i.e.*,  $\mathbb{P}^n(\cdot|\theta)$ , since the samples  $x_1, \dots, x_n$  are *i.i.d.*

The quantity  $L(\theta, \delta(S(\mathbf{x})))$  measures the cost incurred in estimating the unknown parameter as  $\delta(S(\mathbf{x}))$ , whereas the true value of the parameter is  $\theta$ . Notice that  $R(\theta, \delta)$  depends on  $\theta$ , which is unknown. Thus, in order to evaluate the performance of the estimator  $\delta$ , it is required to reduce  $R(\theta, \delta)$  to a function that depends only on  $\delta$ . For this purpose, Bayes point estimation (Kay, 1997) assumes a *prior* probability distribution  $\pi(\cdot)$  over  $\Theta$  and considers the average risk  $\mathbb{E}_{\theta \sim \pi}[R(\theta, \delta)]$ . Then, the notion of an “optimal” Bayes decision rule is defined as

$$\delta_{\text{Bayes}}^*(\mathbf{x}) := \arg \min_{\delta \in \Delta_{\mathbf{x}}} \mathbb{E}_{\theta \sim \pi(\cdot)}[R(\theta, \delta)], \quad (2)$$

where  $\Delta_{\mathbf{x}} := \{\delta: \mathcal{Y} \rightarrow \Theta\}$ .

<sup>1</sup> In this section we focus on deterministic estimators, while in Section 4 we extend these definitions to *randomized* estimators.

If the loss function is the squared error

$$L(\theta, \delta(S(\mathbf{x}))) = (\theta - \delta(S(\mathbf{x})))^2,$$

then (2) is the *conditional mean estimate* (Kailath et al., 2000)

$$\delta_{\text{Bayes}}^*(\mathbf{x}) = \mathbb{E}[\theta|S(\mathbf{x})], \quad (3)$$

where the expectation is with respect to the posterior distribution of  $\theta$  after observing  $\mathbf{x}$ .

We call (2) a *non-private Bayes point estimate*, since it does not consider any privacy constraints. Similarly, we call (3) the *non-private conditional mean estimate*.

### 2.2 Differential Privacy

To define DP, we first need to define the notion of *neighbouring inputs*. For this, we need to introduce some notation.

Let  $\mathbf{x} = (x_1, \dots, x_n)^T$  and  $\mathbf{x}' = (x'_1, \dots, x'_n)^T$  be elements of  $\mathcal{X} := \mathbb{R}^{n \times p}$ , called the *input space*. Here,  $x_i, x'_i \stackrel{i.i.d.}{\sim} \mathbb{P}$ , for  $i = 1, \dots, n$ , where  $\mathbb{P}$  is some probability distribution.

*Definition 1.* (Neighbouring inputs; Dwork and Roth 2014).  $\mathbf{x}$  and  $\mathbf{x}'$  are called *neighbouring inputs* if  $d(\mathbf{x}, \mathbf{x}') = 1$ , where  $d$  is the Hamming distance (Hamming, 1950), *i.e.*, if  $x_i \neq x'_i$  for some unique  $i \in \{1, \dots, n\}$ , and  $x_j = x'_j$  for all  $j \in \{1, \dots, n\} \setminus \{i\}$ .

*Definition 2.* ( $\varepsilon$ -Differential Privacy; Dwork and Roth 2014). Let  $\Theta$  be either a subset of  $\mathbb{R}^{d_\theta}$  or  $\mathbb{C}^{d_\theta}$ ,  $d_\theta \in \mathbb{N}$ . For  $\varepsilon > 0$ , a randomized algorithm  $\mathcal{A}: \mathcal{X} \rightarrow \Theta$  is  $\varepsilon$ -differentially private ( $\varepsilon$ -DP) if for each pair of neighboring inputs  $\mathbf{x}, \mathbf{x}' \in \mathcal{X}$  and  $T \subseteq \Theta$  it holds that

$$\Pr[\mathcal{A}(\mathbf{x}) \in T] \leq e^\varepsilon \Pr[\mathcal{A}(\mathbf{x}') \in T]. \quad (4)$$

### Interpretation

Definition 2 implies that for small values of  $\varepsilon$ , the probability distribution of the output  $t = \mathcal{A}(\mathbf{x})$  of the algorithm is almost the same (up to a multiplicative constant  $e^\varepsilon$ ) for two neighbouring inputs  $\mathbf{x}$  and  $\mathbf{x}'$ . Therefore, by looking at the output  $t$ , it is difficult to infer whether  $\mathbf{x}$  or  $\mathbf{x}'$  is its corresponding input, since the distribution of  $t$  is almost indistinguishable for  $\mathbf{x}$  and  $\mathbf{x}'$ , thereby guaranteeing privacy.

### 2.3 Laplace Mechanism

A standard approach to achieve  $\varepsilon$ -DP is the *Laplace mechanism* (Dwork and Roth, 2014), which we describe below. To this end, we need the notion of  $l_1$ -sensitivity.

*Definition 3.* ( $l_1$ -sensitivity; Dwork and Roth 2014). The  $l_1$ -sensitivity of a function  $g: \mathcal{X} \rightarrow \Theta$  is

$$\sigma_g := \sup_{\mathbf{x}, \mathbf{x}': d(\mathbf{x}, \mathbf{x}')=1} \|g(\mathbf{x}) - g(\mathbf{x}')\|_1,$$

where  $\|\cdot\|_1$  denotes the  $l_1$  norm (Horn and Johnson, 2013).

*Definition 4.* (Laplace Mechanism; Dwork and Roth 2014). Given a function  $g: \mathcal{X} \rightarrow \Theta$ , the *Laplace mechanism* is a randomized algorithm that outputs the vector  $J(\mathbf{x})$  whose  $i^{th}$  component is distributed according to  $[J(\mathbf{x})]_i \stackrel{i.i.d.}{\sim} \text{Lap}([g(\mathbf{x})]_i, \frac{\sigma_g}{\varepsilon})$ , for  $i = 1, \dots, d_\theta$ . Here,  $[g(\mathbf{x})]_i$  denotes the  $i^{th}$  component of  $g(\mathbf{x})$ , and  $\text{Lap}(a, b)$  is the Laplace distribution with probability density function (pdf)

$$f_{\text{Lap}}(z; a, b) = \frac{1}{2b} \exp\left(-\frac{|z - a|}{b}\right),$$

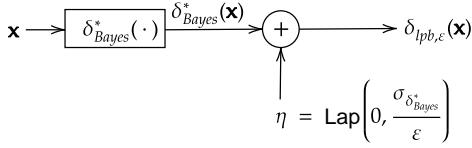


Fig. 1. Cartoon illustrating the standard approach to obtain Bayes point estimate under DP.

where  $b > 0$ .

**Remark 5.** Notice from Definition 4 that  $[J(\mathbf{x})]_i = [g(\mathbf{x})]_i + \eta_i$ , where  $\eta_i \stackrel{i.i.d.}{\sim} \text{Lap}(0, \frac{\sigma_g}{\epsilon})$ , i.e.,  $J(\mathbf{x}) = g(\mathbf{x}) + \boldsymbol{\eta}$ , where  $\boldsymbol{\eta} = (\eta_1, \dots, \eta_{d_\theta})^T$ . Hence, DP is enforced by explicitly randomizing the deterministic quantity  $g(\mathbf{x})$  via the addition of a Laplace noise vector  $\boldsymbol{\eta}$ .

### 3. PROBLEM FORMULATION

In this paper we consider the problem of Bayes point estimation (2) under a DP constraint (4). In the standard Laplace mechanism approach, this is done in two steps. First, given input  $\mathbf{x}$ , a non-private Bayes point estimate is obtained using (2). Second, Laplace noise is added to this non-private Bayes point estimate to obtain

$$\delta_{lpb,\epsilon}(\mathbf{x}) = \delta_{\text{Bayes}}^*(\mathbf{x}) + \text{Lap}\left(0, \frac{\sigma_{\delta_{\text{Bayes}}}^*}{\epsilon}\right). \quad (5)$$

This quantity, called the *Laplace Bayes Private Point* (LBAPP) estimate, satisfies the privacy constraint (4) (see Fig. 1.). Here,

$$\sigma_{\delta_{\text{Bayes}}}^* = \sup_{\mathbf{x}, \mathbf{x}' \in \mathcal{X}: d(\mathbf{x}, \mathbf{x}')=1} \|\delta_{\text{Bayes}}^*(\mathbf{x}) - \delta_{\text{Bayes}}^*(\mathbf{x}')\|_1. \quad (6)$$

Although  $\delta_{lpb,\epsilon}$  satisfies the privacy constraint, two important questions need to be addressed:

- Is the Laplace noise addition “optimal” in the sense of satisfying (2)?
- What if we do not have a closed form expression for  $\delta_{\text{Bayes}}^*(\mathbf{x})$  that is required to compute  $\sigma_{\delta_{\text{Bayes}}}^*$ ?

These two questions motivate the development, in the next section, of an alternative approach to find the optimal Bayes point estimator under a DP constraint.

### 4. PROPOSED APPROACH

Due to the definition of DP, we notice that in order to impose privacy, we require the estimate to be randomized. We will allow the randomization of the estimator to be “implicit” by replacing the *deterministic* estimate  $\delta(S(\mathbf{x}))$  with a *randomized private Bayes estimate*  $\delta_{p,\epsilon}(\cdot | S(\mathbf{x}))$ , which is a probability density function over the parameter space  $\Theta$  for each input  $\mathbf{x} \in \mathcal{X}$ . This is formalized below.

Let  $(\Theta, \mathcal{B}_\Theta)$  be a measurable space of parameters,  $(\mathcal{X}, \mathcal{B}_\mathcal{X})$  a measurable space of inputs endowed with a metric (e.g., the Hamming distance)  $d$ ,  $(\mathcal{Y}, \mathcal{B}_\mathcal{Y})$  a measurable space of observations,  $(q_\theta)_{\theta \in \Theta}$  a probability kernel on  $\mathcal{Y}$ , and  $\pi$  a prior distribution on  $\Theta$ . Let  $S$  be a sufficient statistic  $S: \mathcal{X} \rightarrow \mathcal{Y}$  that is *locally injective*, in the sense that  $S(\mathbf{x}) \neq S(\mathbf{x}')$  for each pair of neighbouring inputs  $\mathbf{x}, \mathbf{x}' \in \mathcal{X}$  (i.e.,  $d(\mathbf{x}, \mathbf{x}') = 1$ ). Then, given a *loss function*  $L: \Theta \times \Theta \rightarrow \mathbb{R}_0^+$



Fig. 2. UBAPP estimator. Unlike the Laplace mechanism that imposes an explicit randomization, UBAPP uses an implicit randomization.

and a randomized private Bayes estimator  $\delta_{p,\epsilon}$ , we define the *Bayes risk* of  $\delta_{p,\epsilon}$

$$R(\delta_{p,\epsilon}, \pi) = \int_{\theta \in \Theta} \int_{\mathbf{y} \in \mathcal{Y}} \int_{\tilde{\theta} \in \Theta} L(\theta, \tilde{\theta}) \delta_{p,\epsilon}(\tilde{\theta} | \mathbf{y}) q_\theta(\mathbf{y}) \pi(\theta) d\theta d\mathbf{y} d\tilde{\theta}. \quad (7)$$

Note that  $R(\delta_{p,\epsilon}, \pi)$  is a linear function of  $\delta_{p,\epsilon}$ . In the standard Bayes point estimation, the estimate is considered as a deterministic function of the samples, but here, due to the DP constraint, the estimate will have to rely on some type of randomization mechanism. This is contrary to the standard Laplace mechanism, where the randomization is explicit. We will construct an estimator defined in terms of  $\delta_{p,\epsilon}$  that is “optimal” in the sense of satisfying (2) subject to a DP constraint.

From the definition of DP (4), and that  $S$  is locally injective, it follows that for each  $\mathbf{x}, \mathbf{x}' \in \mathcal{X}$  with  $d(\mathbf{x}, \mathbf{x}') = 1$  ( $d$  is the Hamming distance), there exist  $\mathbf{y} = S(\mathbf{x})$ , and  $\mathbf{y}' = S(\mathbf{x}')$  such that,

$$\delta_{p,\epsilon}(\tilde{\theta} | \mathbf{y}) \leq e^\epsilon \delta_{p,\epsilon}(\tilde{\theta} | \mathbf{y}'), \text{ for each } \tilde{\theta} \in \Theta. \quad (8)$$

Note that (8) is a linear constraint on  $\delta_{p,\epsilon}$ .

By combining (7) and (8), we can define the optimal  $\epsilon$ -DP Bayes estimator  $\delta_{p,\epsilon}^*$  as the minimizer of the following optimization program:

$$\begin{aligned} \min_{\delta_{p,\epsilon} \in \mathcal{P}(\mathcal{Y}, \Theta)} & \int_{\Theta} \int_{\mathcal{Y}} \int_{\Theta} L(\theta, \tilde{\theta}) \delta_{p,\epsilon}(\tilde{\theta} | \mathbf{y}) q_\theta(\mathbf{y}) \pi(\theta) d\theta d\mathbf{y} d\tilde{\theta} \\ \text{s.t.} & \delta_{p,\epsilon}(\tilde{\theta} | S(\mathbf{x})) \leq e^\epsilon \delta_{p,\epsilon}(\tilde{\theta} | S(\mathbf{x}')), \text{ for each } \tilde{\theta} \in \Theta \\ & \text{and } \mathbf{x}, \mathbf{x}' \in \mathcal{X} \text{ s.t. } d(\mathbf{x}, \mathbf{x}') = 1, \\ & \int_{\Theta} \delta_{p,\epsilon}(\tilde{\theta} | \mathbf{y}) d\tilde{\theta} = 1, \text{ for each } \mathbf{y} \in \mathcal{Y}, \\ & \delta_{p,\epsilon}(\tilde{\theta} | \mathbf{y}) \geq 0, \text{ for each } \mathbf{y} \in \mathcal{Y}, \tilde{\theta} \in \Theta, \end{aligned} \quad (9)$$

where  $\mathcal{P}(\mathcal{Y}, \Theta)$  is the set of probability densities on  $\Theta$  conditioned on  $\mathcal{Y}$ .

From (9), it is evident that  $\delta_{p,\epsilon}^*$  corresponds to the optimal Bayes estimator (in the sense of minimizing the Bayes risk) satisfying the DP constraint (4). We denote the optimization program (9) as *Unified Bayes Private Point* (UBAPP) estimator (Fig. 2) and call its solution  $\delta_{p,\epsilon}^*$  the UBAPP estimate.

In the following subsection, we specialize the UBAPP estimator for the case where the parameter space  $\Theta$  and observation space  $\mathcal{Y}$  are both finite, thus arriving at a tractable convex optimization program.

#### 4.1 Finite Case

Consider the finite case when both  $\Theta = \{\theta_1, \dots, \theta_{|\Theta|}\}$ ,  $\mathcal{Y} = \{\mathbf{y}_1, \dots, \mathbf{y}_{|\mathcal{Y}|}\}$  are finite. Here,  $|\Theta|$  and  $|\mathcal{Y}|$  denote the

cardinality of  $\Theta$  and  $\mathcal{Y}$  respectively. Then, the randomized private Bayes estimate  $\delta_{p,\varepsilon}(\cdot|\mathbf{y})$  ( $\mathbf{y} = S(\mathbf{x})$  for some  $\mathbf{x} \in \mathcal{X}$ ) can be described in terms of a matrix  $\mathbf{P} \in \mathbb{R}^{|\Theta| \times |\mathcal{Y}|}$  such that

$$\mathbf{P}_{i,j} := \mathbb{P}[\hat{\theta} = \theta_i | \mathbf{y}_j] = \delta_{p,\varepsilon}(\theta_i | \mathbf{y}_j), \quad (10)$$

for  $i \in \{1, \dots, |\Theta|\}$ ,  $j \in \{1, \dots, |\mathcal{Y}|\}$ , and  $\mathbf{y}_j = S(\mathbf{x}_j)$  for some  $\mathbf{x}_j \in \mathcal{X}$ .

Combining (10) with (9), we obtain the optimization program for the finite case as

$$\begin{aligned} & \min_{\mathbf{P} \in \mathbb{R}^{|\Theta| \times |\mathcal{Y}|}} \sum_{i=1}^{|\mathcal{Y}|} \sum_{j=1}^{|\Theta|} \sum_{k=1}^{|\Theta|} L(\theta_j, \theta_k) \mathbf{P}_{k,i} \mathbb{P}[\mathbf{y}_i | \theta = \theta_j] \pi(\theta_j) \\ \text{s.t. } & \mathbf{P}_{k,i} \leq e^\varepsilon \mathbf{P}_{k,i'}, \text{ for all } k \in \{1, \dots, |\Theta|\} \\ & \text{and } i, i' \in \{1, \dots, |\mathcal{Y}|\} \text{ s.t. } d(\mathbf{x}_i, \mathbf{x}_{i'}) = 1, \\ & \mathbf{1}^T \mathbf{P} = \mathbf{1}^T, \\ & \mathbf{P} \geq 0, \end{aligned}$$

where  $\pi: \Theta \rightarrow [0, 1]$  is a prior probability mass function (pmf) on  $\Theta$ , and  $\mathbf{1} = (1, \dots, 1)^T \in \mathbb{R}^{|\Theta|}$ . Here,  $\mathbf{P} \geq 0$  means that all the entries of  $\mathbf{P}$  are non-negative. Note that  $\mathbb{P}[\mathbf{y}_i | \theta = \theta_j] = q_{\theta_j}(\mathbf{y}_i)$ .

By defining  $\mathbf{L} \in \mathbb{R}^{|\Theta| \times |\Theta|}$  as  $\mathbf{L}_{j,k} := L(\theta_j, \theta_k)$  for  $j, k \in \{1, \dots, |\Theta|\}$ ,  $\mathbf{Q} \in \mathbb{R}^{|\mathcal{Y}| \times |\Theta|}$  as  $\mathbf{Q}_{i,j} := \mathbb{P}[\mathbf{y}_i | \theta = \theta_j] = q_{\theta_j}(\mathbf{y}_i)$ , for  $i \in \{1, \dots, |\mathcal{Y}|\}$ ,  $j \in \{1, \dots, |\Theta|\}$ , and  $\boldsymbol{\pi} \in \mathbb{R}^{|\Theta|}$  as  $\boldsymbol{\pi}_j := \pi(\theta_j)$  ( $j \in \{1, \dots, |\Theta|\}$ ), we can re-write the optimization problem as

$$\begin{aligned} & \min_{\mathbf{P} \in \mathbb{R}^{|\Theta| \times |\mathcal{Y}|}} \text{tr}(\mathbf{Q} \text{diag}(\boldsymbol{\pi}) \mathbf{L} \mathbf{P}) \\ \text{s.t. } & \mathbf{P}_{k,i} \leq e^\varepsilon \mathbf{P}_{k,i'}, \text{ for all } k \in \{1, \dots, |\Theta|\}, \\ & \text{and } i, i' \in \{1, \dots, |\mathcal{Y}|\} \text{ s.t. } d(\mathbf{x}_i, \mathbf{x}_{i'}) = 1, \\ & \mathbf{1}^T \mathbf{P} = \mathbf{1}^T, \\ & \mathbf{P} \geq 0. \end{aligned} \quad (11)$$

Here,  $\text{tr}(\cdot)$  denotes the trace of a matrix and  $\text{diag}(\boldsymbol{\pi})$  denotes a diagonal matrix whose  $j^{\text{th}}$  entry is  $\boldsymbol{\pi}_j = \pi(\theta_j)$  ( $j \in \{1, \dots, |\Theta|\}$ ). The minimizer of optimization program (11) gives us the UBAPP estimate  $\delta_{p,\varepsilon}^*$  for the finite case, and it can be obtained using CVXPY (Diamond and Boyd, 2016). To evaluate the performance of the UBAPP estimator, we compute the theoretical mean-square error (MSE) of  $\delta_{p,\varepsilon}^*$  for different values of  $\varepsilon$ . We summarize the computation of  $\delta_{p,\varepsilon}^*$  and its MSE in Algorithm 1.

#### Algorithm 1 UBAPP estimator

- 1: **Input:**  $\mathcal{X}$ ,  $\varepsilon > 0$ ,  $\boldsymbol{\pi}$ , and  $\Theta$
- 2: Compute  $\mathbf{L}$  by evaluating  $L(\theta_j, \theta_k)$  for  $j, k \in \{1, \dots, |\Theta|\}$
- 3: Compute  $\mathbf{Q}$  by evaluating  $\mathbb{P}[\mathbf{y}_i | \theta = \theta_j]$  for each  $i \in \{1, \dots, |\mathcal{Y}|\}$ ,  $j \in \{1, \dots, |\Theta|\}$
- 4: Solve (11) by CVXPY solver and store the minimizer as  $\delta_{p,\varepsilon}^*$
- 5: **Output:**  $\delta_{p,\varepsilon}^*$ ,  $\text{MSE} = \text{tr}(\mathbf{Q} \text{diag}(\boldsymbol{\pi}) \mathbf{L} \delta_{p,\varepsilon}^*)$ .

Finally, we compare the MSE of the UBAPP estimator with that of the Laplace Private Bayes estimator  $\delta_{lpb,\varepsilon}$  (5), based on the Laplace mechanism, whose MSE is computed as in Algorithm 2.

---

#### Algorithm 2 MSE of LBAPP estimator

---

- 1: **Input:**  $\mathcal{X}$ ,  $\varepsilon > 0$ ,  $\boldsymbol{\pi}$ ,  $N$  (number of Monte-Carlo runs), and  $\Theta$
  - 2: **Initialize:**  $\text{cum\_mse} = 0$
  - 3: **for**  $i = 1, \dots, N$  **do**
  - 4:     Sample  $\theta_i \sim \boldsymbol{\pi}$
  - 5:     Sample  $\mathbf{x}_i \sim \mathbb{P}^n(\cdot | \theta_i)$
  - 6:     Compute  $\delta_{\text{Bayes}}^*(\mathbf{x}_i)$  from (2)
  - 7:     Compute  $\sigma_{\delta_{\text{Bayes}}^*}$  from (6)
  - 8:      $\hat{\theta}_i = \delta_{lpb,\varepsilon}(\mathbf{x}_i) = \delta_{\text{Bayes}}^*(\mathbf{x}_i) + \text{Lap}\left(0, \frac{\sigma_{\delta_{\text{Bayes}}^*}}{\varepsilon}\right)$
  - 9:      $\text{cum\_mse} \leftarrow \text{cum\_mse} + (\hat{\theta}_i - \theta_i)^2$
  - 10: **Output:**  $\text{MSE} = \frac{\text{cum\_mse}}{N}$ .
- 

## 5. SIMULATIONS

In this section, we compare our approach with the standard Laplace mechanism. In particular, we demonstrate that the accuracy-privacy trade off for our approach is better than for the Laplace mechanism, especially in the high privacy regime (*i.e.*, for small values of  $\varepsilon$ ). For this purpose, we consider the simple numerical example outlined next.

### 5.1 Setup

Suppose we are interested in estimating the unknown parameter  $\theta \in \Theta = [0, 1]$  of a Bernoulli distribution  $\text{Ber}(\theta)$  from which  $K$  independent samples are generated in  $\{0, 1\}$ . The simulation setup is as follows:

- Let  $\pi$  be the uniform distribution over  $\Theta$ .
- Let  $x_1, \dots, x_K \stackrel{i.i.d.}{\sim} \text{Ber}(\theta_i)$  be outcomes of Bernoulli trials, where  $\theta_i \in \Theta$ , and  $\mathbf{x} = (x_1, \dots, x_K)^T$ . Then, define the sufficient statistic  $S$  as

$$\mathbf{y} = S(\mathbf{x}) = \sum_{i=1}^K x_i.$$

This means that  $\mathbf{y} \in \mathcal{Y} = \{0, \dots, K\}$ .

- Let  $L$  be the square loss, *i.e.*,  $L(\theta_i, \theta_j) = (\theta_i - \theta_j)^2$ .

For this setup, we first derive the LBAPP estimator  $\delta_{lpb,\varepsilon}$  in Subsection 5.2. Then, in Subsection 5.3, we describe the procedure to obtain the UBAPP estimator  $\delta_{p,\varepsilon}^*$ .

### 5.2 LBAPP estimator

Let  $\mathbb{P}(\theta | \mathbf{y})$  denote the posterior distribution of  $\theta$ . Then, it can be shown that

$$\mathbb{P}(\theta | \mathbf{y}) \propto \theta^{\sum_{i=1}^K x_i + 1} (1 - \theta)^{(K+1 - \sum_{i=1}^K x_i) - 1}.$$

This implies that  $\mathbb{P}(\theta | \mathbf{y})$  is a Beta distribution. Therefore, from (3), the non-private Bayes estimate is given by

$$\delta_{\text{Bayes}}^*(\mathbf{x}) = \frac{\sum_{i=1}^K x_i + 1}{K + 2}. \quad (12)$$

We now add Laplace noise to obtain  $\delta_{lpb,\varepsilon}(\mathbf{x})$ . To this end, we need to compute  $\sigma_{\delta_{\text{Bayes}}^*}$  using (6). Let  $\mathbf{x} = (x_1, \dots, x_K)^T$ , and  $\mathbf{x}' = (x'_1, \dots, x'_K)^T$ . Then,

$$\begin{aligned}
\sigma_{\delta_{\text{Bayes}}^*} &= \sup_{\mathbf{x}, \mathbf{x}' : d(\mathbf{x}, \mathbf{x}') = 1} \|\delta_{\text{Bayes}}^*(\mathbf{x}) - \delta_{\text{Bayes}}^*(\mathbf{x}')\|_1 \\
&= \frac{1}{K+2} \sup_{\mathbf{x}, \mathbf{x}' : d(\mathbf{x}, \mathbf{x}') = 1} \left| \sum_{i=1}^K x_i - \sum_{i=1}^K x'_i \right| \\
&= \frac{1}{K+2}.
\end{aligned} \tag{13}$$

Finally, using (13) and (12) in (5), we obtain

$$\delta_{lpb, \varepsilon}(\mathbf{x}) = \frac{\sum_{i=1}^K x_i + 1}{K+2} + \text{Lap}\left(0, \frac{1}{(K+2)\varepsilon}\right).$$

We compute the MSE of LBaPP estimator using Algorithm 2, where we take  $N = 5000$ .

### 5.3 UBaPP estimator

To obtain the UBaPP estimator  $\delta_{p, \varepsilon}^*$  we will use formulation (11), for which we only need to discretize  $\Theta$ , as  $\mathcal{Y}$  is already finite. For this purpose, we consider a grid of  $M_\theta$  equally spaced points and denote the set of such equally spaced points by  $\tilde{\Theta}$ . Due to the discretization, we note that there should be an additional factor  $1/M_\theta$  in the objective function of (11), but this factor can be ignored since it does not depend on  $\delta_{p, \varepsilon}$ . Also,  $\pi$  is now the uniform distribution over  $\tilde{\Theta}$ , and hence, the  $j^{\text{th}}$  entry of  $\text{diag}(\pi)$  is  $\pi_j = 1/M_\theta$  ( $j \in \{1, \dots, |\tilde{\Theta}|\}$ ). For this simulation setup, it is easy to see that  $\mathbf{Q}$  is Binomial, *i.e.*,

$$\mathbf{Q}_{i,j} = \mathbb{P}[\mathbf{y}_i | \theta_j] = \binom{K}{\mathbf{y}_i} \theta_j^{\mathbf{y}_i} (1 - \theta_j)^{K - \mathbf{y}_i},$$

where  $\theta_j \in \tilde{\Theta}$ , and  $\mathbf{y}_i \in \mathcal{Y} = \{0, \dots, K\}$ , for  $j \in \{1, \dots, |\tilde{\Theta}|\}$ , and  $i \in \{0, \dots, |\mathcal{Y}|\}$  respectively. Also, for the simulation, we take  $M_\theta = 5000$ .

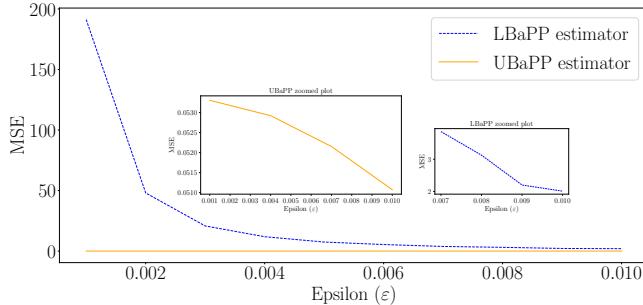


Fig. 3. Plot of MSE vs.  $\varepsilon$  for the LBaPP and UBaPP estimators in the high privacy regime. A zoomed plot for the UBaPP and LBaPP estimators is included.

### 5.4 Plots

We now study the performance of using our approach to differentially private Bayes point estimation, UBaPP. In particular, we focus on

- the effect of the privacy parameter  $\varepsilon$  on the MSE, for a fixed number of Bernoulli trials;
- the effect of the number of Bernoulli trials ( $K$ ) on the MSE, for a fixed  $\varepsilon$ .

To analyze the MSE of these two estimators for different values of  $\varepsilon$ , we consider two ranges of  $\varepsilon$ , which are classified

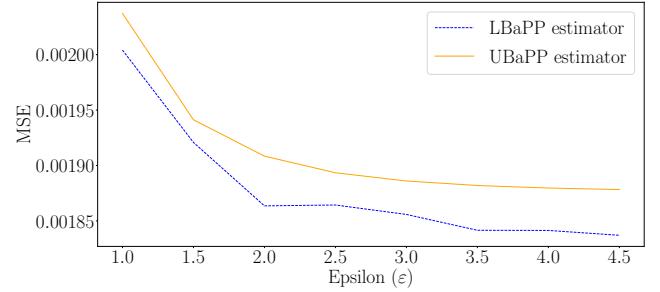


Fig. 4. Plot of MSE vs.  $\varepsilon$  for the LBaPP and UBaPP estimators in the moderate-low privacy regime.

as ‘‘High privacy regime’’ and ‘‘Moderate-low privacy regime’’. The high privacy regimes corresponds to low values of  $\varepsilon$ , *i.e.*, to  $\varepsilon \in \{0.001, 0.002, \dots, 0.010\}$ , while the moderate-low privacy regimes corresponds to high values of  $\varepsilon$ , *i.e.*,  $\varepsilon \in \{1, 1.5, \dots, 5\}$ . The resulting plots of the MSE vs.  $\varepsilon$  of the LBaPP and UBaPP estimators for the high and moderate-low privacy regimes are shown in Figs. 3 and 4, respectively.

Next, to study the effect of the number of Bernoulli trials on the MSE of both estimators, we plot MSE of LBaPP and UBaPP for different values of  $K$ . We consider two different values of  $\varepsilon$ , namely,  $\varepsilon = 0.001$  and  $\varepsilon = 5$ , corresponding to the high privacy and moderate-low privacy regimes, respectively. The corresponding plots are shown in Figs. 5 and 7, respectively.

Finally, to understand the randomization induced by the UBaPP estimator, we plot in Fig. 6 heat maps of UBaPP estimate  $\delta_{p, \varepsilon}^*$  for different values of  $\varepsilon$ .

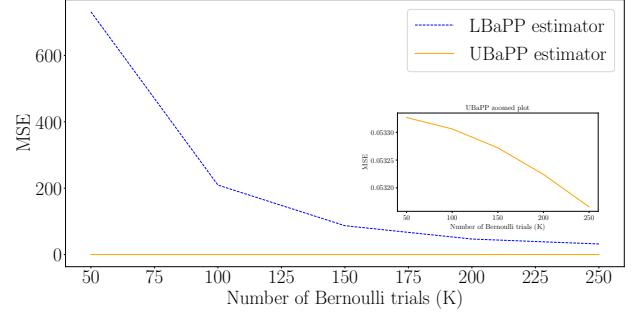


Fig. 5. Plot of MSE vs. number of Bernoulli trials ( $K$ ) for the LBaPP and UBaPP estimators in a fixed high privacy regime. A zoomed plot for the UBaPP estimator is included.

### 5.5 Discussion

We observe from Fig. 3 that UBaPP provides high accuracy (*i.e.*, low MSE) in the high privacy regime, when compared to that of LBaPP. Also, Fig. 4 shows that in the moderate-low privacy regime, the accuracy of UBaPP is very similar to that of LBaPP estimator: even though the MSE of UBaPP is higher than that of LBaPP, the difference is only of order approx.  $10^{-3}$ . This difference is due to the discretization of the parameter space  $\Theta$  used in the implementation of UBaPP, whereas the computation

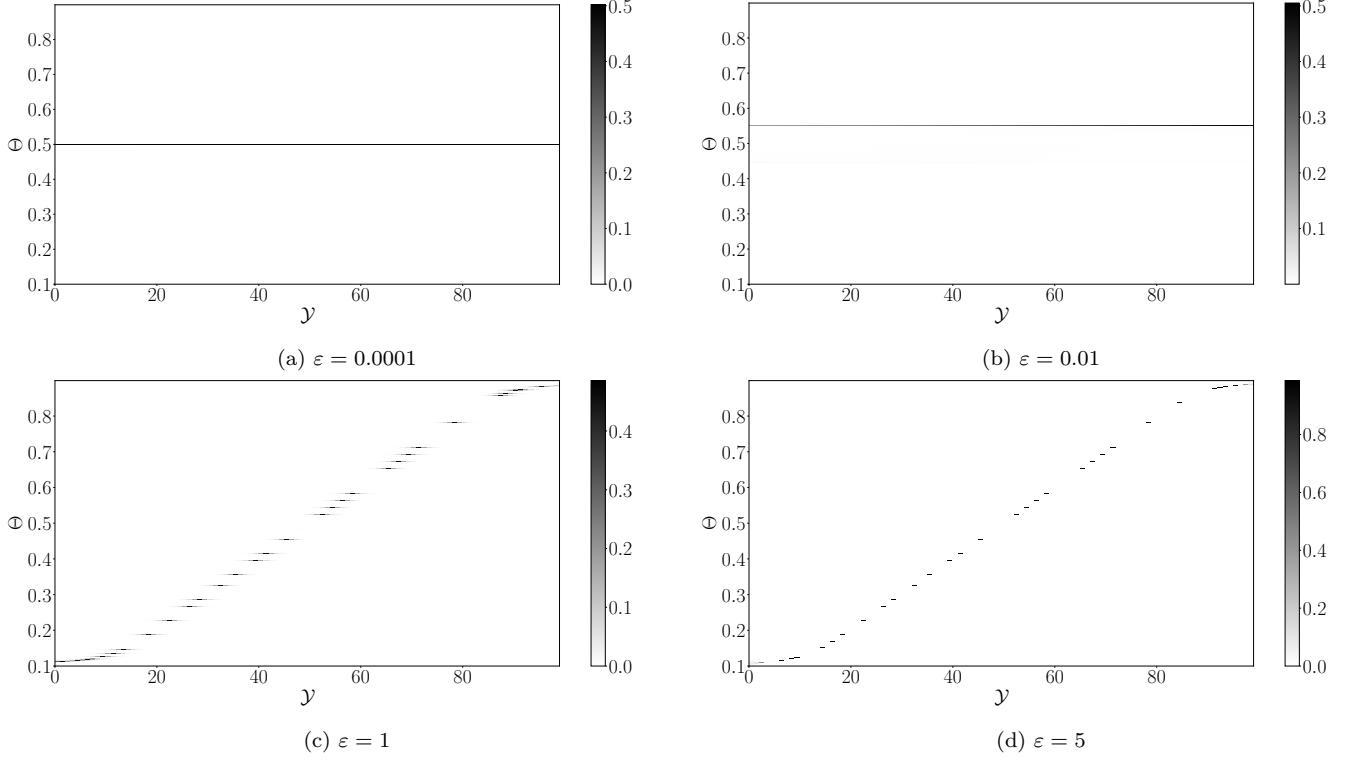


Fig. 6. Heat maps of UBaPP estimate  $\delta_{p,\varepsilon}^*$  for different values of  $\varepsilon$ .

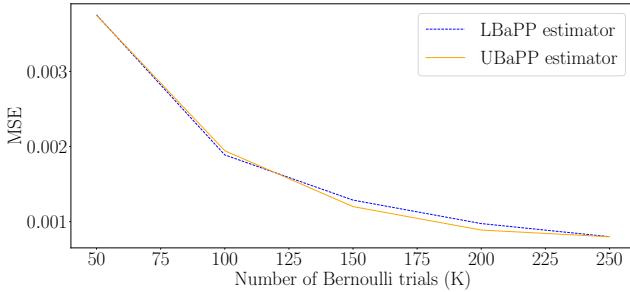


Fig. 7. Plot of MSE vs. number of Bernoulli trials ( $K$ ) for the LBaPP and UBaPP estimators in a fixed moderate-low privacy regime.

of the LBaPP estimates is exact (*i.e.*, no discretization is employed).

Regarding the dependence on the number of trials  $K$ , from Fig. 5 it is clear that in the high privacy regime, the MSE of UBaPP is less dependent on  $K$  than the MSE of LBaPP. This means that, for low values of  $K$ , UBaPP has significantly higher accuracy than LBaPP. On the other hand, according to Fig 7, the accuracy of LBaPP and UBaPP is similar in the moderate-low privacy regime.

To understand the difference between our approach, UBaPP, and the Laplace mechanism LBaPP, notice from Fig. 6a that, for low values of  $\varepsilon$ , UBaPP outputs a deterministic estimate around  $\theta = 0.5$  (which corresponds to the mean of the prior distribution on  $\theta$ ) for every  $y \in \mathcal{Y}$ ; since low values of  $\varepsilon$  imply high privacy, no valid inference about the data (input)  $\mathbf{x}$  can be made based on the estimate of  $\theta$ , so the estimator becomes independent

of the data by outputting the same deterministic estimate for all  $y \in \mathcal{Y}$ . Also, as seen in Figs. 6b-6c, when  $\varepsilon$  is increased (which implies a shift from high to moderate levels of privacy), UBaPP introduces some level of randomization, because for a given value of  $y \in \mathcal{Y}$  the probability distribution  $\delta_{p,\varepsilon}^*$  is not concentrated at a single value of  $\theta \in \Theta$ . Finally, as  $\varepsilon$  is further increased to very low privacy levels, Fig. 6d shows that UBaPP becomes deterministic again, as it tends to the standard (non-private) Bayes point estimator of  $\theta$ , which is known to be deterministic for convex loss functions (Berger, 1985). Also, note from Fig. 6d that the deterministic estimate varies with  $y \in \mathcal{Y}$ , which suggests that UBaPP estimator becomes strongly dependent of the data, and hence it becomes non-private.

In contrast, the Laplace mechanism adds randomization to the standard non-private Bayes point estimator of  $\theta$ , with a variance that increases as  $\varepsilon \rightarrow 0$ . Thus, for large values of  $\varepsilon$  it coincides, like UBaPP, with the standard (non-private) Bayes point estimator of  $\theta$  (shown in Fig. 6d). However, for very small values of  $\varepsilon$  (*i.e.*, very high privacy), the variance of the estimator grows unbounded, which implies that its MSE tends to infinity, as shown in Fig. 3, whereas the MSE of UBaPP tends to a constant as  $\varepsilon \rightarrow 0$ .

In conclusion, we see from the numerical study that UBaPP is more accurate than LBaPP in the high privacy regime, while for moderate-low privacy constraints both estimators yield similar performance (save for the discretization needed to implement UBaPP).

## 6. CONCLUSION

In this paper, we have studied the problem of Bayes point estimation under differential privacy. We have argued that the standard approach based on the Laplace mechanism

may not give accurate estimates under high privacy constraints. We then proposed an optimal approach that combines risk minimization (minimum MSE) and differential privacy into a single convex optimization program, and specialized this approach to the case of finite parameter and observation space. Via a simple numerical study, we have shown that our approach yields more accurate estimates in the high privacy regime than the Laplace mechanism, and that both approaches have similar performance under low privacy constraints.

As future work, we plan the extension of our approach to continuous (and even potentially high dimensional) parameter and observation spaces.

## REFERENCES

- Berger, J.O. (1985). *Statistical Decision Theory and Bayesian Analysis*, 2nd Ed. Springer-Verlag.
- Cao, X. and Başar, T. (2020). Differentially private parameter estimation: Optimal noise insertion and data owner selection. In *59th IEEE Conference on Decision and Control (CDC)*, 2887–2893.
- Casella, G. and Berger, R. (2002). *Statistical Inference*, 2nd Ed. Duxbury.
- Diamond, S. and Boyd, S. (2016). CVXPY: A Python-embedded modeling language for convex optimization. *Journal of Machine Learning Research*, 17(83), 1–5.
- Dwork, C. and Roth, A. (2014). The algorithmic foundations of differential privacy. *Foundations and Trends in Theoretical Computer Science*, 9(3–4), 211–407.
- Gouriéroux, C. and Monfort, A. (1996). *Statistics and Econometric Models*. Cambridge University Press.
- Hamming, R. (1950). Error detecting and error correcting codes. *The Bell System Technical Journal*, 29(2), 147–160.
- Homer, N., Szelinger, S., Redman, M., Duggan, D., Tembe, W., Muehling, J., Pearson, J., Stephan, D., Nelson, S., and Craig, D. (2008). Resolving individuals contributing trace amounts of DNA to highly complex mixtures using high-density SNP genotyping microarrays. *PLOS Genetics*, 4(8), e1000167.
- Horn, R. and Johnson, C. (2013). *Matrix Analysis*, 2nd Ed. Cambridge University Press.
- Kailath, T., Sayed, A., and Hassibi, B. (2000). *Linear Estimation*. Prentice Hall.
- Kay, S. (1997). *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall.
- Le Ny, J. and Pappas, G. (2013). Differentially private filtering. *IEEE Transactions on Automatic Control*, 59(2), 341–354.
- Lehmann, E. and Casella, G. (1998). *Theory of Point Estimation*, 2nd Ed. Springer-Verlag.
- Ljung, L. (1999). *System Identification: Theory for the User*, 2nd Ed. Prentice Hall.
- McElroy, F. (1967). A necessary and sufficient condition that ordinary least-squares estimators be best linear unbiased. *Journal of the American Statistical Association*, 62(320), 1302–1304.
- Nekouei, E., Sandberg, H., Skoglund, M., and Johansson, K. (2022). A model randomization approach to statistical parameter privacy. *IEEE Transactions on Automatic Control*.
- NIH (2014). NIH genomic data sharing policy. <https://tinyurl.com/bdd6se45>. Accessed: 2014-08-27.
- Sankar, L., Rajagopalan, S., and Poor, H. (2013). Utility-privacy tradeoffs in databases: An information-theoretic approach. *IEEE Transactions on Information Forensics and Security*, 8(6), 838–852.
- Shalev-Shwartz, S. and Ben-David, S. (2012). *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press.
- Söderström, T. and Stoica, P. (1989). *System Identification*. Prentice Hall.
- Van Trees, H.L. (2004). *Detection, Estimation, and Modulation Theory. Part I: Detection, Estimation, and Linear Modulation Theory*. John Wiley & Sons.
- Varodayan, D. and Khisti, A. (2011). Smart meter privacy using a rechargeable battery: Minimizing the rate of information leakage. In *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*.
- Wang, J., Zhu, R., and Liu, S. (2018). A differentially private unscented Kalman filter for streaming data in IoT. *IEEE Access*, 6, 6487–6495.
- Wang, Y., Mitra, S., and Dullerud, G.E. (2017). Differential privacy and minimum-variance unbiased estimation in multi-agent control systems. *IFAC-PapersOnLine*, 50(1), 9521–9526.