Author/s:
Farokhi, F;Sandberg, H

Title:
Ensuring privacy with constrained additive noise by minimizing Fisher information

Date:
2019-01-01

Citation:
Farokhi, F. & Sandberg, H. (2019). Ensuring privacy with constrained additive noise by minimizing Fisher information. Automatica, 99, pp.275-288. https://doi.org/10.1016/j.automatica.2018.10.012.

Persistent Link:
https://hdl.handle.net/11343/285059

Ensuring Privacy with Constrained Additive Noise by Minimizing Fisher Information

Farhad Farokhi [a,b], Henrik Sandberg [c]

[a] *CSIRO's Data61, Australia*

[b] *Department of Electrical and Electronic Engineering at the University of Melbourne, Australia*

[c] *Department of Automatic Control, KTH Royal Institute of Technology, Sweden*

Abstract

The problem of preserving the privacy of individual entries of a database when responding to linear or nonlinear queries with constrained additive noise is considered. For privacy protection, the response to the query is systematically corrupted with an additive random noise whose support is a subset or equal to a pre-defined constraint set. A measure of privacy using the inverse of the trace of the Fisher information matrix is developed. The Cramér-Rao bound relates the variance of any estimator of the database entries to the introduced privacy measure. The probability density that minimizes the trace of the Fisher information (as a proxy for maximizing the measure of privacy) is computed. An extension to dynamic problems is also presented. Finally, the results are compared to the differential privacy methodology.

Key words: Privacy; Additive constrained noise; Fisher information.

1 Introduction

The constant state of connectedness has enabled the use of new technologies, such as participatory sensing and big-data analysis. These technologies can vastly improve the efficiency of existing infrastructures with little investment. This has come at the price of the erosion of privacy in the society. An example of this lack of privacy is the potential use of data from smart electrical meters by adversaries, such as criminals, advertising agencies, and governments, for monitoring the presence and the activities of occupants [2]. Other examples can include the use of detailed travel data for traffic estimation in intelligent transportation systems [3], privacy violations caused by sharing information in distributed control systems [4], and privacy concerns in cloud computing and control [5]. These concerns have therefore motivated an urgent need for creating appropriate mechanisms that can protect the privacy of the individuals whose information is stored in various databases. Specifically, it is of interest to provide responses to queries from policy makers on the aggregated data as accurately as possible while not leaking the private information of the individuals.

To combat these problems, the problem of preserving the privacy of individual entries of a database using a constrained additive noise is considered in this paper. It is assumed that anyone, including an adversary, can submit queries to a (trusted) server possessing the entire database. The server returns a response to the query that is systematically corrupted by an additive noise whose support is a subset, or equal, to a desired constraint set. The Cramér-Rao bound [6, p. 169] is then used to relate the variance of the estimation error of unbiased estimators of the database by adversaries from the provided responses to the trace of the inverse of the Fisher information matrix.

We start with finding a maximizer of the inverse of the trace of the Fisher information matrix. This optimiza-

⋆ An early version of this paper has appeared at the 56th IEEE Conference on Decision and Control, 2017 [1]. The work of F. Farokhi was supported by the McKenzie Fellowship from the University of Melbourne, a competitive grant (MyIP: ID6874) from Defence Science and Technology Group (DSTG), and the VESKI Victoria Fellowship from the Government of Victoria. The work of H. Sandberg was supported by the EU CHIST-ERA project COPES and the Swedish Civil Contingencies Agency through the CERCES project.

Email addresses: ffarokhi@unimelb.edu.au, farhad.farokhi@data61.csiro.au (Farhad Farokhi), hsan@kth.se (Henrik Sandberg).

tion problem is nonconvex. It is proved that finding the probability density function of the noise boils down to solving a nonlinear partial differential equation, which is complex in general, even for the simplest case (of linear queries and when the probability density of the noise is independent of the content of the database). Thus, we opt for maximizing a lower bound of the inverse of the trace of the Fisher information matrix, which is the inverse of the trace of the Fisher information matrix scaled by a constant. This is equivalent to minimizing the trace of the Fisher information matrix. This optimization problem is proved to be convex. It is shown that the optimal noise distribution can be calculated by solving a linear partial differential equation (that can be sometimes further simplified with the aid of separation of variables). These results are subsequently generalized to the case where the support set of the noise distribution is unbounded. Noting that, for unbounded sets, the solution to the problem is to add a noise with infinite variance (because the trace of the Fisher information matrix can be pushed to zero), the need for minimizing the trace of the Fisher information (for ensuring privacy) is balanced with the quality of the response (captured by the variance of the additive noise). It can be shown that the Gaussian noise is optimal if the noise is not constrained. These results are demonstrated on three illustrative examples involving smart meter privacy, and computing the average and variance of private databases.

The problem formulation and parts of the results are extended to dynamic estimation problems, where the initial condition of the system is assumed to be the variable that needs to be kept private. This is motivated by a traffic estimation problem in which the initial condition of the system (modelling the vehicle) corresponds to the location of driver's house, which is private.

Finally, the optimal privacy-preserving policies of this paper are compared with differentially-private policies (specifically, the Laplace mechanism) and the optimal privacy-preserving policies when using mutual information as a measure of privacy. It is observed that the optimal policies in the unconstrained-noise formulation are also (ϵ, δ)-differentially private. Further, the optimal policies in this paper coincide with the optimal privacy-preserving policies when using mutual information as a measure of privacy for the unconstrained case, thus inheriting strong information-theoretic guarantees.

A common approach for ensuring that the privacy of participants in large databases (or rather the content of the entries of the database owned by those participants) is the application of differential privacy [4, 7–11]. Those studies often advocate the addition of noises with slow-decaying probability density functions to the response of the queries submitted to the server. This is done so that an adversary cannot accurately infer the private information of the individuals stored in the database. The Laplace noise is frequently utilized in the differential privacy literature; see, e.g., [8,12]. However, other noise distributions are also common for achieving the differential privacy, or variants thereof. This has prompted various studies to seek the optimal noise distribution for differential privacy [13–15].

Although several information-theoretic interpretations of differential privacy have been presented [16–18], the available literature does not offer an operational meaning for the concept as well as a systematic approach for setting the differential privacy parameter (except a broad sweep). Further, in practice, it may not be possible to use an additive noise with infinite support as the noise might need to satisfy certain constraints, e.g., it must belong to a bounded set for smart metering [19].

Application of differential privacy in control systems has also gained attention recently. Differential private filtering is discussed in [9], where releasing filtered signals while respecting privacy of user data streams is considered. Distributed control of multi-agent systems in the presence of privacy constraints is studied in [4]. An attainable lower bound on the entropy of output is presented in the case where an additive noise is used to ensure differential privacy for discrete-time systems [20]. Differential privacy in the specific case of consensus-seeking algorithms is also considered in [21–26]. A thorough review of these results can be found in a recent tutorial paper [27].

Differential privacy has also been successfully utilized in numerical optimization. In [11], parameters of individual constraints in resource allocation problems is kept private. Preserving the privacy of decisions and cost functions in distributed optimizations is investigated in [28, 29]. Applications of differential privacy can also be found in other related problems, such as machine learning [30, 31], mechanism design [32, 33], and transportation systems [34, 35].

Recently, several studies have used mutual information (or entropy) and the least mean square estimation error as measures of privacy [36–42]. Similar to differential privacy literature, most results based on mutual information do not provide an intuitive or interpretable bound on the statistics of the estimation error by the adversary (with the exception of [43] which uses rate distortion theory to get an interpretable bound on the performance of the adversary). They also require *a priori* assumptions on the distribution of the database, which might not be available in practice due to complexity and scale of the database. The privacy results using the least mean square estimation error also restrict the behaviour of the adversary and assume the underlying random variables are Gaussian, which might not be the case in practice as well.

In [19], Fisher information is utilized as a measure of privacy to design privacy-preserving charging policies for

batteries in households with smart meters. In this paper, those results are extended to develop a general framework using Fisher information as a measure of privacy. This paper extends [19] in the following ways. The results of [19], in the language of this paper, are about releasing all the entries of the database, i.e., the query submitted to the server is an identity function. That very special and restrictive case does not even cover linear queries, let alone providing the optimal privacy-preserving policy for non-linear queries and dynamic estimation.

Finally, it is worth mentioning that the statistics community has previously used Fisher information as a measure of privacy [44, 45]. However, in those studies, minimizing the Fisher information to obtain privacy-preserving policies over the set of density functions whose support sets is appropriately constrained is not discussed.

The rest of the paper is organized as follows. The problem formulation and some preliminary results are presented in Section 2. The optimal privacy-preserving probability density functions for the additive constrained noise are developed in Section 3. These results are then generalized to dynamic estimation problems in Section 4. The relationships between the presented framework and the existing results in the literature are discussed in more depth in Section 5. Finally, Section 6 concludes the paper and presents viable avenues for future work.

2 Background and Problem Formulation

Let $x \in \mathcal{X} \subseteq \mathbb{R}^n$ be a variable that should be kept private, i.e., a database that is possessed by a (trusted) server. This data is only available to the server. In what follows, the vector x is assumed to be deterministic and fixed, i.e., no prior is required nor available. Anyone, including an adversary, can submit queries of the form $f(x)$ with $f : \mathbb{R}^n \to \mathbb{R}^m$ to the server. The server in return provides a noisy response to the query given by

$$y = f(x) + w, \tag{1}$$

where $w \in \mathcal{W}(x) \subseteq \mathbb{R}^m$ denotes the noise that the server adds to the data to protect the entries of the private variable x. The following standing assumption is made.

Assumption 1 *f is continuously differentiable.*

In this paper, two specific families of constraints on the support set of the additive noise $\mathcal{W}(x)$ are considered:

(i) the set $\mathcal{W}(x)$ is independent of x; or
(ii) the set $\mathcal{W}(x)$ takes the special form of $\{-f(x)\} \oplus \mathcal{Y}$ for some set $\mathcal{Y} \subseteq \mathbb{R}^m$.

Here, $\mathcal{C} \oplus \mathcal{B}$ denotes the set $\{a + b \,|\, a \in \mathcal{C}, b \in \mathcal{B}\}$ for any two sets $\mathcal{C}$ and $\mathcal{B}$. The family of constraints following the form of (i) models the case where the noise itself is constrained, e.g., the additive noise should be positive or bounded. However, the family of constraints in (ii) captures the case where the output y must be constrained inside the set $\mathcal{Y}$. In what follows, we use $\mathcal{W}$ to denote either of these families of constraints on the support set of the additive noise.

Note that the server has the intention to respond as accurately as possible to the query $f(x)$ as this typically corresponds to statistical properties (e.g., mean) of the database, which are valuable to, e.g., policy makers. However, it does not want the entries of x (the private data of the people) to be released online nor estimated. Examples of applications where the server wants to provide accurate answers to the submitted queries while keeping the entries of the database hidden can be found in [46–48].

The server's policy (which is the object of interest in this paper) is the probability density function $\gamma(\cdot|x) : \mathcal{W} \to \mathbb{R}_{\geq 0}$ of the noise w, where $\mathbb{R}_{\geq 0} := \{x \in \mathbb{R} \,|\, x \geq 0\}$. This implies that $\mathbb{P}\{w \in \mathcal{W}' \,|\, x\} = \int_{w' \in \mathcal{W}'} \gamma(w'|x) \mathrm{d}w'$ for any Lebesgue-measurable set $\mathcal{W}' \subseteq \mathcal{W}$. In this paper, it is desired to seek a policy γ that makes the problem of inferring the private variable x difficult (according to an appropriate measure described below). The set of all admissible policies Γ is restricted according to the following standing assumption.

Assumption 2 *(i) $\gamma(w|x)$ is such that $\mathbb{P}\{w \in \mathbb{R}^m \setminus \mathcal{W} \,|\, x\} = 0$, ($ii$) $\gamma(w|x)$ is twice continuously differentiable in (w, x) over $\mathcal{W} \times \mathcal{X}$, and ($iii$) $\gamma(w|x) = 0$ for all $w \in \partial\mathcal{W}$.*

Assumption 2 (i) ensures that, with probability one, the noise is restricted to the set $\mathcal{W}$, i.e., $\mathbb{P}\{w \in \mathcal{W} \,|\, x\} = 1$. This is to ensure that the constraints on the noise or the output are satisfied almost surely. Assumption 2 (ii) is required for the use of the Cramér-Rao bound as well as the use of results from calculus of variations for finding the optimal probability density function. Finally, as observed later in the paper, Assumption 2 (iii) is necessary for the Cramér-Rao bound (see Proposition 1). The latter part of this assumption is satisfied if the set $\mathcal{W}$ is unbounded (as, in the limit, a probability density function is always zero otherwise it does not integrate to one). Note that, for bounded constraint sets $\mathcal{W}$, the set of probability density functions that are zero over $\partial\mathcal{W}$ can approximate any probability density function arbitrarily closely. This is proved in [1].

Under the aforementioned policy of the server, the probability density of y for a given x is then equal to

$$p(y|x) = \gamma(y - f(x)|x), \ \forall y \in \mathcal{Y} = \{f(x)\} \oplus \mathcal{W}. \tag{2}$$

Further, for any continuously differentiable function $g : \mathbb{R}^n \to \mathbb{R}$, the notation $\partial g(x)/\partial x$ is used to denote a

column vector containing the partial derivatives of the function. For any multivariate function $g(x)$, $G(x)$ denotes its Jacobian, i.e., a matrix with the element in the i-th row and the j-th column being equal to $\partial g_i(x)/\partial x_j$. Before stating the next preliminary result, the Fisher information matrix $\mathcal{I}(x) \in \mathbb{R}^{n\times n}$ is defined as

$$\mathcal{I}(x)=\int_{y\in\mathcal{Y}} p(y|x)\left[\frac{\partial \log(p(y|x))}{\partial x}\right]\left[\frac{\partial \log(p(y|x))}{\partial x}\right]^\top \mathrm{d}y. \tag{3}$$

The following results immediately follows from the use of the Cramér-Rao bound providing a lower bound on the adversary's estimation error of the private variable independent of the policy.

Proposition 1 *Under Assumption 2, for any unbiased estimate of x denoted by $\hat{x}(y)$, it holds that*

$$\mathbb{E}\{\|x-\hat{x}(y)\|_2^2\} \geq \mathrm{Tr}(\mathcal{I}(x)^{-1}), \tag{4}$$

if $\mathcal{I}(x)$ is invertible. Furthermore, for any biased estimate of x denoted by $\hat{x}(y)$ such that $\mathbb{E}\{\hat{x}(y)\} = g(x)$, it holds that

$$\begin{aligned}\mathbb{E}\{\|x-\hat{x}(y)\|_2^2\} \geq& \mathrm{Tr}(G(x)^\top \mathcal{I}(x)^{-1}G(x)) \\ &+ \|x-g(x)\|_2^2,\end{aligned} \tag{5}$$

if $\mathcal{I}(x)$ is invertible.

Proof: The proof follows from the use of the Cramér-Rao bound [6, p. 169]. ∎

Here, it is desirable to find a policy γ for the server that makes estimation of x as difficult as possible. This can be pursued through multiple avenues. When an unbiased estimator of the database exists, following Proposition 1, in order to make the task of inferring about x from the measurement y difficult, the trace of the inverse of the Fisher information matrix should be maximized. However, if $m < n$, there may not exist any unbiased estimator of the database x because there are more unknowns n than measurements m. In this case, the goal becomes to maximize $\mathrm{Tr}(G(x)^\top \mathcal{I}(x)^{-1}G(x))$. It can be shown that

$$\begin{aligned}\mathrm{Tr}(G(x)^\top \mathcal{I}(x)^{-1}G(x)) &= \mathrm{Tr}(\mathcal{I}(x)^{-1}G(x)G(x)^\top) \\ &\geq \mathrm{Tr}(\mathcal{I}(x)^{-1})\lambda_{\min}(G(x)G(x)^\top),\end{aligned} \tag{6}$$

where, for any matrix, $\lambda_{\min}(\cdot)$ denotes its smallest eigenvalue. This inequality shows that, even if a biased estimator is utilized, the trace of the inverse of the Fisher information matrix can be maximized to preserve the privacy of the entries of the database, albeit if $G(x)$ assumes full row rank for all $x \in \mathcal{X}$; otherwise the lower bound in (6) is zero and maximizing $\mathrm{Tr}(\mathcal{I}(x)^{-1})$ does not result in any tangible privacy guarantee. Therefore, the following assumption is made.

Assumption 3 *$G(x)$ is full row rank for all $x \in \mathcal{X}$.*

The sensibility of this problem formulation relies on the validity of Assumption 3, which is in general difficult, if not impossible, to check (see Example 2 for a case in which this assumption can be easily checked). Assumption 3 is not strictly-speaking necessary, at least for the convexified problem discussed later (see Problem 2) as the problem formulation can be alternatively motivated by using a worst-case privacy guarantee. Consider the case where, for any $1 \leq i \leq n$, the server aims at protecting the content of x_i even if all other entries of the database $x_{-i} = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$ are leaked (i.e., this is a worst-case analysis for privacy protection). This is to ensure that even if the owners of all the other entries of the database are colluding with the adversary, they cannot extract the private data of an individual (to a reasonable extent). To do so, let $\hat{x}_i(y, x_{-i})$ denote an unbiased estimator of x_i based on y for a given x_{-i}. By fixing x_{-i} as knowns, Proposition 1 can be used to deduce that

$$\mathbb{E}\{\|x_i-\hat{x}_i(y,x_{-i})\|_2^2\} \geq 1/\mathcal{I}_i(x), \tag{7}$$

where

$$\mathcal{I}_i(x) = \int_{y\in\mathcal{Y}} p(y|x)\left[\frac{\partial \log(p(y|x))}{\partial x_i}\right]\left[\frac{\partial \log(p(y|x))}{\partial x_i}\right]^\top \mathrm{d}y$$

with $p(y|x) = \gamma(w - f(x_i, x_{-i})|x_i, x_{-i})$. Hence,

$$\begin{aligned}\min_i \mathbb{E}\{\|x_i-\hat{x}_i(y,x_{-i})\|_2^2\} &\geq \min_i(1/\mathcal{I}_i(x)) \\ &= 1/(\max_i \mathcal{I}_i(x)).\end{aligned}$$

Further, note that

$$\begin{aligned}\max_i \mathcal{I}_i(x) &\leq \sum_{i=1}^n \mathcal{I}_i(x) \\ &= \int_{y\in\mathcal{Y}} p(y|x)\sum_{i=1}^n\left[\frac{\partial \log(p(y|x))}{\partial x_i}\right] \\ &\qquad \times \left[\frac{\partial \log(p(y|x))}{\partial x_i}\right]^\top \mathrm{d}y \\ &= \mathrm{Tr}\bigg(\int_{y\in\mathcal{Y}} p(y|x)\left[\frac{\partial \log(p(y|x))}{\partial x}\right] \\ &\qquad \times \left[\frac{\partial \log(p(y|x))}{\partial x}\right]^\top \mathrm{d}y\bigg) \\ &= \mathrm{Tr}(\mathcal{I}(x)).\end{aligned}$$

Hence, it can be deduced that

$$\min_i \mathbb{E}\{\|x_i - \hat{x}_i(y, x_{-i})\|_2^2\} \geq 1/\mathrm{Tr}(\mathcal{I}(x)). \quad (8)$$

This shows that, even for this stronger notion of privacy (in the absence of Assumption 3), minimizing $\mathrm{Tr}(\mathcal{I}(x))$ (instead of maximizing $\mathrm{Tr}(\mathcal{I}(x)^{-1})$) provides a reasonable solution.

Before presenting the problem formulation, it should be noted that maximizing $\mathrm{Tr}(\mathcal{I}(x)^{-1})$ is not well-defined when x is not known *a priori*. Therefore, instead, we maximize the cost function

$$\overline{\mathcal{J}} := \int_{x\in\mathcal{X}} \mathrm{Tr}(\mathcal{I}(x)^{-1})p(x)\mathrm{d}x, \quad (9)$$

where $p : \mathcal{X} \to \mathbb{R}_{\geq 0}$ is a weight associated with x such that $p(x) \neq 0$ for some set with non-zero Lebesgue measure (since otherwise $\overline{\mathcal{J}}$ is identical to zero). Note that $p(x)$ is not a prior but a weight that captures how difficult the server wants the estimation of the database to become for a given x. Note that it can always be assumed that $\int_{x\in\mathcal{X}} p(x)\mathrm{d}x = 1$. This is without loss of generality as $p(x)$ can be always scaled by $\int_{x\in\mathcal{X}} p(x)\mathrm{d}x > 0$ to achieve the equality. This is of course a design parameter.

Problem 1 *Find* $\gamma^* \in argmax_{\gamma\in\Gamma}\, \overline{\mathcal{J}}$.

Remark 1 (Well-Defined Problem Formulation) *Problem 1 is well-defined if the support set of the additive noise $\mathcal{W}$ is bounded. If this is not the case, the optimal solution is to push the variance of the additive noise to infinity (as that pushes $\mathrm{Tr}(\mathcal{I}^{-1})$ to infinity). Note that if $\mathcal{W}$ is bounded, $\mathrm{Tr}(\mathcal{I}^{-1})$ is also bounded. Hence, for the case where $\mathcal{W}$ is unbounded, an additional term capturing the quality of the provided response by the server can be included in the utility function to ensure the existence of non-trivial and implementable solutions. This case is investigated later in the paper.*

Remark 2 (Existence of Solutions) *Investigating existence of solutions to Problem 1 is a daunting task due to the nature of the set Γ. In the remainder of this section, a convex approximation of this problem is presented. In this case, the necessary condition for optimality is also sufficient. This allows us to formulate the problem of finding (sub)optimal privacy-preserving policies as solving a linear partial differential equation with Dirichlet boundary conditions. For the relaxed problem, therefore, the existence of an optimal solution can be cast as the existence of solutions to a linear partial differential equation. For some cases, it is possible to find a solution to the partial differential equation satisfying all the boundary conditions (thus guaranteeing the existence of solutions constructively).*

Remark 3 (Data Processing Inequality) *Mutual information is often lauded as a measure of privacy due to data processing inequality, i.e., additional manipulation of the transmitted messages based on ones private information can only decrease the amount of the leaked information. This is also considered a beneficial property of the differential privacy, that is, additional manipulations of the outcomes of a differentially private policy cannot decreases the privacy guarantees of the process [49]. This property also holds for the Fisher information [50] pointing to that $\mathrm{Tr}(\mathcal{I}(x)^{-1})$ can only be increased upon further manipulations of the transmitted response y.*

Remark 4 (Side Channel Information) *Privacy studies, including the presented framework based on Fisher information as well as those based on mutual information and differential privacy, are often fragile to admitting side channel information, e.g., measurements of the private variable already available. In this paper's problem formulation, the primary goal is to maximize $\mathbb{E}\{\|x - \hat{x}(y)\|_2^2\}$ (or its lower bound given by the Fisher information) which clearly states that the adversary's estimator $\hat{x}(y)$ is only a function of y. The setup however can be generalized following the same line of reasoning to maximize $\mathbb{E}\{\|x - \hat{x}(y, z)\|_2^2\}$, where z models the side channel information. If the nature of the side-channel information is not fully know, the server can attempt at maximizing $\min_{\upsilon(z|x)} \mathbb{E}\{\|x - \hat{x}(y, z)\|_2^2\}$, where $\upsilon(z|x)$ denoting the conditional probability of the side-channel information given the state can vary over restricted set of density functions. This is an interesting avenue for future research.*

Define the relaxed cost function

$$\mathcal{J} := \int_{x\in\mathcal{X}} \mathrm{Tr}(\mathcal{I}(x))p(x)\mathrm{d}x. \quad (10)$$

Since maximizing $\overline{\mathcal{J}}$ is difficult in general, the lower bound $n^2\mathcal{J}^{-1}$ can be maximized to achieve a suboptimal solution; the inequality follows from the application of Proposition 5 in Appendix A. Before stating the problem formulation, it is beneficial to note that $\mathcal{J}$ is in fact a convex function of γ (over a subset of Γ) and, thus, the new problem formulation is more computationally feasible. This proof follows from the convexity of the trace of the Fisher information matrix. Note that the proof of the convexity of the Fisher information for scalar random variables is widely available; see for example [51, pp. 80-81] and [52]. The proof for the multivariate case can also be found in [19]. Define the support of a density function $\gamma(\cdot|x)$ as $\mathrm{supp}(\gamma(\cdot|x)) := \{w \in \mathcal{W} \,|\, \gamma(w|x) > 0\}$.

Assumption 4 *$\gamma(w|x)$ is such that $\mathcal{W}\backslash\mathrm{supp}(\gamma(\cdot|x))$ has a zero Lebesgue measure for all $x \in \mathrm{supp}(p)$.*

Let $\overline{\Gamma} \subseteq \Gamma$ be the set of conditional probability density

functions satisfying Assumptions 2 and 4. The following proposition shows that $\mathcal{J}$ is a convex function over $\overline{\Gamma}$ (which is a convex set) and thus stationarity conditions are necessary and sufficient for finding a minimizer of $\mathcal{J}$ over $\overline{\Gamma}$.

Proposition 2 *$\mathcal{J}$ is a convex function of the density function γ over the set of density functions $\overline{\Gamma}$.*

Proof: The proof is similar to the one in [19] and is thus omitted for the sake of space. ∎

Similarly, Proposition 2 motivates us to search for the minimizer of $\mathcal{J}$ over the set of all density functions $\gamma(\cdot|x)$ that are at most over a measure-zero set equal to zero in $\text{int}(\mathcal{W})$, which satisfies the definition of $\overline{\Gamma}$.

Problem 2 *Find $\gamma^* \in \text{argmin}_{\gamma\in\overline{\Gamma}} \mathcal{J}$.*

As stated earlier, Problems 1 and 2 are both only well-defined for bounded noise support sets $\mathcal{W}$. For unbounded sets $\mathcal{W}$, the solution to both problems is to add a noise with infinite variance. By selecting, for instance, a Gaussian noise with an arbitrarily large variance, the trace of the Fisher information matrix can be pushed towards zero (and the trace of its inverse towards infinity by Proposition 5 in Appendix A). To overcome this problem, the quality of the provided response by the server needs to be balanced with the guaranteed privacy. To do so, a measure of quality of the response can be defined as follows:

$$\begin{aligned}\mathcal{Q} &= \int_{x\in\mathcal{X}} p(x)\mathbb{E}\{\|y - f(x)\|_2^2 \,|\, x\}\mathrm{d}x \\ &= \int_{x\in\mathcal{X}}\int_{w\in\mathcal{W}} w^\top w\gamma(w|x)p(x)\mathrm{d}x.\end{aligned}$$

The smaller $\mathcal{Q}$ is, the better the quality of the provided response to the query is. Note that the measure of quality $\mathcal{Q}$ is a convex function of γ since it is linear in the probability density function. Now, the problem formulation can be revised for unbounded constraint set $\mathcal{W}$.

Problem 3 *Find $\gamma^* \in \text{argmin}_{\gamma\in\overline{\Gamma}} \mathcal{J} + \varrho\mathcal{Q}$, where $\varrho > 0$ is a constant balancing the need for preserving privacy with the quality of the provided response by the server.*

An alternative problem formulation of the following form can be presented in which a hard constraint on the quality of the response is enforced.

Problem 4 *Find $\gamma^* \in \text{argmin}_{\gamma\in\overline{\Gamma}:\mathcal{Q}\leq\vartheta} \mathcal{J}$, where $\vartheta > 0$ denotes the upper bound on performance degradation caused by the additive noise.*

In light of [53], Problems 4 and 3 are equivalent in the sense that, for all $\varrho > 0$, there exists $\vartheta > 0$ such that the solution of Problem 4 is a solution of Problem 3 and *vice versa*.

With the problem formulations at hand, we are ready to calculate the optimal policy of the server. This is the topic of the next section.

3 Privacy-Preserving Policy

In this section, the solutions of the previously-stated problem formulations is presented. We start with Problem 1. In this case, the solution is given in the following the theorem for linear queries of the form $f(x) = Cx$ with $C \in \mathbb{R}^{m\times n}$ and the case where $\gamma(w|x)$ is independent of x, e.g., when measurements of the database are not available. Note that, in this case, the Fisher information matrix can be simplified to

$$\begin{aligned}\mathcal{I}(x) &= \int_{y\in\{Cx\}\oplus\mathcal{W}} \gamma(y - Cx)\left[\frac{\partial\log(\gamma(y-Cx))}{\partial x}\right] \\ &\qquad\qquad \times \left[\frac{\partial\log(\gamma(y-Cx))}{\partial x}\right]^\top \mathrm{d}y \\ &= \int_{w\in\mathcal{W}} \gamma(w)C^\top\left[\frac{\partial\log(\gamma(w))}{\partial w}\right]\left[\frac{\partial\log(\gamma(w))}{\partial w}\right]^\top C\mathrm{d}w.\end{aligned}$$

In this case, $\mathcal{I}(x)$ is no longer a function of x and is thus simply denoted by $\mathcal{I}$. Further, it should be noted that $\overline{\mathcal{J}} = \text{Tr}(\mathcal{I}^{-1})$ and $\mathcal{J} = \text{Tr}(\mathcal{I})$. Therefore, the solution of the problem in this case becomes independent of the choice of $p(x)$. Proving this results for more general cases creates several complications without providing more insight. The subsequent results are however proved for nonlinear queries and general policies. The next theorem presents a necessary condition for the solution of non-convex optimization problem in Problem 1.

Theorem 1 *Let $\gamma^*(w)$ denote a solution of Problem 1 for linear queries of the form $f(x) = Cx$ over the set of probability density functions that are independent of x. Then, it satisfies the following conditions with $u(w) = \sqrt{\gamma^*(w)}$ and some constant $\mu \in \mathbb{R}$:*

$$\begin{cases}\text{Tr}(\mathcal{I}^{-2}C^\top D^2u(w)C) + \mu u(w) = 0, & w \in \mathcal{W}, \\ u(w) = 0, & w \in \partial\mathcal{W}, \\ u(w) \neq 0, & w \in \text{int}\mathcal{W}, \\ \int_{w\in\mathcal{W}} u(w)^2\mathrm{d}w = 1. & \end{cases} \tag{11}$$

Proof: See [1]. ∎

Remark 5 *In Theorem 1, μ denotes the Lagrange multiplier associated with the equality constraint $\int_{w\in\mathcal{W}} u(w)^2\mathrm{d}w = 1$ (to ensure that $\gamma(w) = u(w)^2$ is a*

probability density function). The conditions in (11) *are equivalent to the Karush–Kuhn–Tucker (KKT) conditions for the infinite-dimensional optimization problem in Problem 1. In the rest of the paper, for some specific cases, the value of the multiplier is calculated explicitly. However, in general, the value of the multiplier should be iteratively changed (e.g., using the methods in primal-dual optimization) to find the appropriate value.*

Remark 6 (Complexity of the Solution) *Note that the partial differential equation in* (11) *is nonlinear because* $\mathcal{I}$ *in* $\mathrm{Tr}(\mathcal{I}^{-2}C^\top D^2u(w)C)$ *is a function of* $u(w)$*. Further, Theorem 1 only provides a necessary condition, i.e., the solution of Problem 1 satisfies* (11) *but the reverse does not necessarily hold. As mentioned earlier, these difficulties stem from the complexity of maximizing* $\mathrm{Tr}(\mathcal{I}^{-1})$*, which is a non-concave cost function.*

Following Remark 6, in the remainder of the paper, the relaxed formulation in Problem 2 and its variants are studied. In what follows, $\mathbb{1}_n$ denotes the n-dimensional vector of ones. If the dimension n is clear from the context, $\mathbb{1}$ is used instead of $\mathbb{1}_n$. Further, $F(x)$ denotes the Jacobian of the multivariate function $f(x)$, i.e., a matrix with the element in the i-th row and the j-th column being equal to $\partial f_i(x)/\partial x_j$. The following theorem provides necessary and sufficient conditions for capturing the solution of Problem 2.

Theorem 2 *The solution of Problem 2 is given by* $\gamma^*(w|x)=u(w,x)^2$*, where* $u(w,x)$ *satisfies*

$$\begin{cases} \mathrm{Tr}\Bigg(\begin{bmatrix} F(x)F(x)^\top & F(x) \\ F(x)^\top & I \end{bmatrix} D^2u(w,x)\Bigg) \\ \quad +L(w,x)\begin{bmatrix} \dfrac{\partial\gamma(w|x)}{\partial w} \\ \dfrac{\partial\gamma(w|x)}{\partial x} \end{bmatrix}+\mu(x)u(w,x)=0, & w\in\mathcal{W}, \\ u(w,x)=0, & w\in\partial\mathcal{W}, \\ u(w,x)\neq 0, & w\in\mathrm{int}\mathcal{W}, \\ \int_{w\in\mathcal{W}}u(w,x)^2\mathrm{d}w=1, \end{cases} \tag{12}$$

for some mapping $\mu:\mathcal{X}\to\mathbb{R}$ *and*

$$L(w,x):=\begin{bmatrix} \dfrac{1}{p(x)}\dfrac{\partial p(x)}{\partial x}^\top F(x)^\top+\mathbb{1}^\top D^2f(x) & \dfrac{1}{p(x)}\dfrac{\partial p(x)}{\partial x}^\top \end{bmatrix}. \tag{13}$$

Further, all solutions (if multiple) satisfying (12) *exhibit the same cost.*

Proof: The proofs are moved to the appendices to avoid interrupting the flow of the presentation. See Appendix B. ∎

Remark 7 *Note that the partial differential equation in* (12) *is often classified as a semi-linear equation in the sense that it is linear in the partial derivatives (thus it is a linear differential equation) but the coefficients can be potentially non-linear functions of the independent variable (which makes it "space" varying). Solving these partial differential equations, in general, is a complex task and out of the scope of this paper. In what follows, the partial differential equation in* (12) *is solved for all scalar (potentially nonlinear) queries in Corollary 5.*

In what follows, the partial differential equation in (12) is solved for three special cases explicitly to gain some insight into the structure of the solution of Problem 2.

Example 1 (Smart Meter Privacy) *Let the energy consumption of household* $k\in\mathcal{N}:=\{1,\dots,n\}$ *in a neighbourhood with* $n\in\mathbb{N}$ *houses be denoted by* $x_k\in\mathbb{R}_{\geq 0}$*. These variables can be aggregated into a vector to get* $x=[x_1\ \cdots\ x_n]^\top$*. Assume that the data is stored on an online server so that it can be studied by policy makers and academics. To avoid unintentionally leaking the private details of the customers, the trusted server in possession of the data adds an appropriate noise to the outcome of the queries to which it responds. Note that the additive noise should be somewhat restricted as arbitrary corruptions might render the data useless or unrealistic. For instance, adding a large negative noise can mask possibly wasteful behaviour of the participants or, in extreme, can transform their combined consumption negative, which might be physically impossible if they do not generate any power using renewable energies. In the case where the server decides to publicly release a sanitized version of the consumption data, it must be assumed that* $f(x)=x$ *and* $\mathcal{W}=[\underline{w},\overline{w}]^n$ *with constants* $0<\underline{w}\leq\overline{w}<+\infty$*. The optimal privacy preserving policy for this example is provided in the following corollary.* ♢

Corollary 1 *Let* $n=m$*,* $p(x)=p$*,* $f(x)=x$*, and* $\mathcal{W}=[\underline{w},\overline{w}]^m$ *with* $-\infty<\underline{w}\leq\overline{w}<+\infty$*. The solution of Problem 2 is given by*

$$\gamma^*(w|x)=\left(\frac{2}{\overline{w}-\underline{w}}\right)^m\prod_{i=1}^m\cos^2\left(\frac{\pi}{\overline{w}-\underline{w}}\left(w_i-\frac{\overline{w}+\underline{w}}{2}\right)\right)\mathbb{1}_{w\in\mathcal{W}}.$$

Proof: See Appendix C. ∎

It is expected that, by increasing $\overline{w}-\underline{w}$ in Example 1, it becomes easier to preserve the privacy of the server because the response can be buried deeper within the noise. Although adding more noise can improve the privacy, it also reduces the quality of the provided response by the server. The trade-off between privacy and quality is captured in the following corollary.

Corollary 2 *For the optimal policy in Corollary 1, the following statements hold:*

(i) $\mathbb{E}\{\|x-\hat{x}(y)\|_2^2\} \geq \mathrm{Tr}(\mathcal{I}^{-1}) = \kappa(\overline{w}-\underline{w})^2$ *with a constant* $\kappa > 0$ *for any unbiased estimate of* x *denoted by* $\hat{x}(y)$*;*

(ii) $\mathcal{Q} = m(\overline{w}-\underline{w})^2(2\pi^2-3)/(6\pi^2) \approx 0.2827m(\overline{w}-\underline{w})^2$.

Proof: See Appendix D. ∎

Corollary 2 shows that $\mathcal{Q} = (6\pi^2\kappa/(2\pi^2-3))\mathrm{Tr}(\mathcal{I}^{-1})$ for the optimal policy-preserving policy. Changing any parameter, such as $\overline{w}$ or $\underline{w}$, that may increase the privacy guarantees inevitably degrades the quality of the response. Therefore, as expected, privacy and quality are conflicting criteria.

Example 2 (Computing Weighted Average) *A common query that users want to perform on large databases is to calculate the weighted average of private variables. However, the server in possession of the data does not want the original data to be extracted from the reported weighted average. This operation can be modelled by* $f(x) = Cx$*, where* $C \in \mathbb{R}^{1\times n}$ *is such that* $C\mathbb{1}_n = 1$*. The adversary, for instance, upon using the least square approach to find an estimate of the database, gets* $\hat{x}(y) = C^\dagger y$*, where* $C^\dagger$ *denotes the Moore–Penrose pseudoinverse of* C*, defined as* $C^\top(CC^\top)^{-1}$*. Assumption 3 holds in this example if* C *has a full row rank, which is a reasonable assumption as otherwise the measurements are not independent. Recalling that* C *is not a full column rank matrix (i.e., it is only full row rank and* $m < n$*),* $\hat{x}(y)$ *is not an unbiased estimator as* $\mathbb{E}\{\hat{x}(y)\} = C^\dagger\mathbb{E}\{y\} = C^\dagger Cx$*. For this estimator, it can be deduced that* $\mathbb{E}\{\|x-\hat{x}(y)\|_2^2\} = \|(I-C^\dagger C)x\|_2^2 + \mathbb{E}\{\|C^\dagger w\|_2^2\} \geq \|(I-C^\dagger C)x\|_2^2 + \mathrm{Tr}(\mathcal{I}(x)^{-1}(CC^\top)^{-1})$*, where the last inequality follows from the Cramér-Rao bound in Proposition 1. Finally, it is also worth saying that, in this case,* $\mathrm{Tr}(\mathcal{I}(x)^{-1}(CC^\top)^{-1}) = \mathrm{Tr}(\mathcal{I}(x)^{-1})(CC^\top)^{-1}$ *because* $m = 1$ *(and thus* $(CC^\top)^{-1} \in \mathbb{R}$*). This problem clearly fits the framework presented in this paper. Here, the set* $\mathcal{W}$ *can be either bounded or unbounded based on the situation. The optimal privacy preserving policy for this example is provided in the following corollary.* ♢

Corollary 3 *Let* $n \in \mathbb{N}$*,* $m = 1$*,* $p(x) = p$*,* $f(x) = Cx$*, and* $\mathcal{W} = [\underline{w}, \overline{w}]$ *with* $-\infty < \underline{w} \leq \overline{w} < +\infty$*. If* $C \neq 0$*, the solution of Problem 2 is given by*

$$\gamma^*(w|x) = \frac{2}{\overline{w}-\underline{w}}\cos^2\left(\frac{\pi}{\overline{w}-\underline{w}}\left(w-\frac{\overline{w}+\underline{w}}{2}\right)\right)\mathbb{1}_{w\in\mathcal{W}}. \tag{14}$$

Proof: See Appendix E. ∎

Example 2 evidently fits the criteria of Corollary 3. It is interesting to note that the choice of the weights in C (so long as at least one of them is non-zero) is irrelevant. The next corollary captures the effect of the weight $p(x)$ on the optimal privacy-preserving policy.

Corollary 4 *Let* $n = m = 1$*,* $f(x) = x$*, and* $\mathcal{W} = [\underline{w}, \overline{w}]$ *with* $-\infty < \underline{w} \leq \overline{w} < +\infty$*. The solution of Problem 2 is given by*

$$\gamma^*(w|x) = c(x)\exp\left(-\frac{p'(x)}{p(x)}(w+x)\right) \times \cos^2\left(\frac{\pi}{\overline{w}-\underline{w}}\left(w-\frac{\overline{w}+\underline{w}}{2}\right)\right)\mathbb{1}_{w\in\mathcal{W}}$$

where

$$c(x) = \left[\int_{\underline{w}}^{\overline{w}}\exp\left(-\frac{p'(x)}{p(x)}(w+x)\right) \times \cos^2\left(\frac{\pi}{\overline{w}-\underline{w}}\left(w-\frac{\overline{w}+\underline{w}}{2}\right)\right)\mathrm{d}w\right]^{-1}.$$

Proof: See Appendix F. ∎

Figure 1 illustrates the optimal privacy-preserving policy $\gamma^*(w|x)$ in Corollary 4 versus w and x for the case where the weighting function is $p(x) \propto \exp(-x^2)$ (top) and $p(x) \propto \exp(-x)$ (bottom) when $\underline{w} = 0$ and $\overline{w} = 1$. For any mappings f and g, it is said that $f(x) \propto g(x)$ if there exists constant c such that $f(x) = cg(x)$. For $p(x) = \exp(-x^2)$, the policy is a function of x. However, for $p(x) = \exp(-x)$, the policy is independent of x; this can be attributed to that the ratio $p'(x)/p(x) = -1$ is not a function of x.

Example 3 (Computing Variance) *Another common query, beside the average of a set of private data (see Example 2), is to calculate its statistical variance. This operation can be done by the nonlinear query* $f(x) = [1/(n-1)]\sum_{i=1}^n (x_i - (1/n)\sum_{j=1}^n x_j)^2$*. The optimal privacy preserving policy for this example is provided in the following corollary.* ♢

Corollary 5 *Let* $n \in \mathbb{N}$*,* $m = 1$*,* $p(x) = p$*, and* $\mathcal{W} = [\underline{w}, \overline{w}]$ *with* $-\infty < \underline{w} \leq \overline{w} < +\infty$*. If* $f(x) \neq 0$*, the solution of Problem 2 is given by* (14)*.*

Proof: See Appendix G. ∎

Corollary 5 proves that the for scalar queries, i.e., when $m = 1$, the nonlinearity of the query does not change the distribution of the optimal additive noise. Therefore, the optimal noise distributions for the averaging problem in Example 2 and the variance calculation in Example 3 are the same.

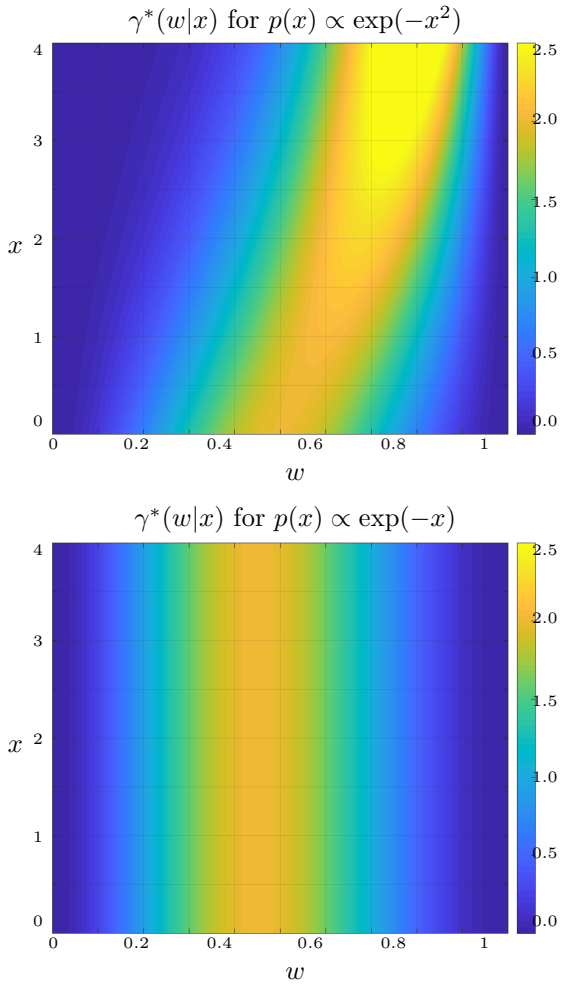

Fig. 1. Optimal privacy-preserving policy for non-uniform weighting functions $p(x) \propto \exp(-x^2)$ (top) and $p(x) \propto \exp(-x)$ (bottom).

Now, we are ready to study the case where the constraint set is unbounded. We start by solving Problem 3 in the next theorem.

Theorem 3 *The solution of Problem 3 is given by* $\gamma^*(w|x) = u(w,x)^2$*, where* $u(w,x)$ *satisfies*

$$\begin{cases} \mathrm{Tr}\left(\begin{bmatrix} F(x)F(x)^\top & F(x)^\top \\ F(x) & I \end{bmatrix} D^2 u(w,x)\right) \\ \qquad +L(w,x)\begin{bmatrix} \dfrac{\partial \gamma(w|x)}{\partial w} \\ \dfrac{\partial \gamma(w|x)}{\partial x} \end{bmatrix} \\ \qquad +(\mu(x)-(\varrho/4)w^\top w)u(w,x) = 0, & w \in \mathcal{W}, \\ u(w,x) = 0, & w \in \partial\mathcal{W}, \\ u(w,x) \neq 0, & w \in \mathrm{int}\mathcal{W}, \\ \int_{w\in\mathcal{W}} u(w,x)^2 \mathrm{d}w = 1, \end{cases} \tag{15}$$

for some mapping $\mu : \mathcal{X} \to \mathbb{R}$ *and* $L(w,x)$ *is defined in* (13)*. Further, all solutions (if multiple) satisfying* (15) *exhibit the same cost.*

Proof: See Appendix H. ∎

An explicit solution of Problems 3 and 4 for the case where $\mathcal{W} = \mathbb{R}^m$ is presented in the following corollary. This case is of special interest as the optimal noise can be compared with the density of the noise suggested in the differential privacy literature, i.e., Laplace noise.

Corollary 6 *Let* $\mathcal{W} = \mathbb{R}^m$ *and* $p(x) = p$*. For linear queries of the form* $f(x) = Cx$ *with full row rank matrix* C*, the solutions of Problems 3 and 4 is given by*

$$\gamma^*(w|x) = \frac{1}{\sqrt{(2\pi)^m \det(\Sigma)}} \exp\left(-\frac{1}{2} w^\top \Sigma^{-1} w\right),$$

where $\Sigma = 2(CC^\top)^{1/2}/\sqrt{\rho}$ *for Problem 3 and* $\Sigma = \vartheta(CC^\top)^{1/2}/\mathrm{Tr}((CC^\top)^{1/2})$ *Problem 4.*

Proof: See Appendix I. ∎

Corollary 6 simply states that, in this case, the optimal noise is a (multivariate) Gaussian random variable with covariance $(CC^\top)^{1/2}/\sqrt{\rho}$. Note that the optimality of the Gaussian noise when minimizing the Fisher information over noises with unbounded support set is not in itself new [54]; however, its application in privacy-preserving policies has not been considered previously. Clearly, as ϱ increases, i.e., the emphasis on the quality of the response to the query increases, the variance of the noise decreases. In this framework, the optimal noise distribution differs from the Laplace distribution, which is a standard choice in the differential privacy literature [12]. Evidently, the weighted averaging setup in Example 2 satisfies the conditions for results of Corollaries 6. For that example, the optimal noise can be simplified into a Gaussian random variable with variance ϑ (due to the scalar nature of the responses). Note that, for this example, the optimal privacy-preserving policy is again independent of the choice of the matrix C.

4 Extensions to Dynamic Estimation

In this section, the problem formulation is extended by considering dynamic estimation problems for linear time-invariant systems. Assume that

$$x[k+1] = Ax[k], \quad x[0] = x_0,$$

where $x[k] \in \mathbb{R}^n$ is the state. The dynamics, for instance, can capture the case where some entries of a database are getting updated in real time or that the database

contains the states of a physical system evolving through time (e.g., position and velocity of a vehicle). Assume that the initial state is deterministic and unknown to the adversary. In each time step $k \in \{0, \dots, T\}$ with T denoting the time horizon, the user can submit a query of the form $Cx[k]$ to the server. The server responds by returning

$$y[k] = Cx[k] + w[k],$$

where $w[k] \in \mathbb{R}^m$ is an additive noise introduced by the server to keep the state private. Here, only the case where the probability density function of $w[k]$ is independent of the state (in the past and the future) is considered. Also consider the case where the support of the noise density is unrestricted. Evidently, the results can be extended to the case where the support set of the noise is constrained following the same line of reasoning as in Section 3.

Define $w_k := [w[0]^\top \; w[1]^\top \; \cdots \; w[k]^\top]^\top$. The policy of the server is the probability density function of w_T denoted by $\gamma : \mathbb{R}^{n(T+1)} \to \mathbb{R}_{\geq 0}$. This is the most general policy that the server can employ. To keep the entries of the database private, the server wants to increase the covariance of the estimation error of the initial condition $\mathbb{E}\{\|\hat{x}_0(y_T) - x_0\|_2^2\}$, where $\hat{x}_0(y_T)$ denotes any unbiased smoothing estimate of the state based on all the received response aggregated into a single vector of the form $y_T = [y[0]^\top \; y[1]^\top \; \cdots \; y[T]^\top]^\top$. Note that $y_T = \Psi_T x_0 + w_T$, where $\Psi_T := [C^\top \; (CA)^\top \; \cdots \; (CA^T)^\top]^\top$. The conditional density of y_T for any x_0 is given by $p(y_T|x_0) = \gamma(y_T - \Psi_T x_0)$. Now, the Cramér-Rao bound (see Proposition 1) can used to show that

$$\mathbb{E}\{\|\hat{x}_0(y_T) - x_0\|_2^2\} \geq \mathrm{Tr}(\mathcal{I}^{-1}) \geq (T+1)^2 n^2 / \mathrm{Tr}(\mathcal{I}),$$

where

$$\begin{aligned}\mathcal{I} &= \int p(y_T|x_0) \left[\frac{\partial \log(p(y_T|x_0))}{\partial x_0}\right] \left[\frac{\partial \log(p(y_T|x_0))}{\partial x_0}\right]^\top \mathrm{d}y_T \\ &= \int \frac{1}{\gamma(w_T)} \Psi_T^\top \left[\frac{\partial \gamma(w_T)}{\partial w_T}\right] \left[\frac{\partial \gamma(w_T)}{\partial w_T}\right]^\top \Psi_T \mathrm{d}w_T\end{aligned}$$

and the second inequality follows from Proposition 5 in Appendix A. The quality of the response can also be measured using

$$\mathcal{Q} = \mathbb{E}\{\|y_T - \Psi_T x_0\|_2^2 \,|\, x_0\} = \int w_T^\top w_T \gamma(w_T) \mathrm{d}w_T.$$

The next theorem provides the optimal policy in the sense of Problem 3 for this case.

Theorem 4 *Let $\Psi_T^\top \Psi_T$ be invertible. The solution of* (3) *for the dynamic estimation setup is given by γ^* such that $\mathbb{P}\{w_T = \Psi_T z\} = 1$ where z is distributed according to the probability density function*

$$p(z) = \frac{1}{\sqrt{(2\pi)^n \det(\Sigma)}} \exp\left(-\frac{1}{2} z^\top \Sigma^{-1} z\right),$$

where $\Sigma = 2(\Psi_T^\top \Psi_T)^{-1/2}/\sqrt{\varrho}$.

Proof: See Appendix J. ∎

Notice that the noise distribution across time is not independently and identically distributed (i.i.d.) across time. The noise at time step k takes the form of $CA^k z$. This means that the server realizes a random variable z based on the probability distribution in Theorem 4. Then it adds z to the initial condition of the system and propagates its effect in all future time steps.

Remark 8 *The matrix $\Psi_T^\top \Psi_T$, which is in fact equal to the observability Gramian over $\{0, \dots, T\}$, is invertible if the pair (A, C) is observable and $T \geq n$. This follows from observability of linear time-invariant systems [55, p. 271] and the Cayley-Hamilton theorem [56, p. 141].*

Example 4 (Traffic Crowd-Sensing with Privacy) *Consider the case where a vehicle is sharing its position with a remote monitoring station in pursuit of estimating the state of the traffic on a road. The house of the vehicle's owner is on this road (which is conveniently modelled by the real line) at $s_0 \in \mathbb{R}$. The vehicle travels on the road with the constant velocity $v_0 \in \mathbb{R}$ after leaving the house. The dynamics of the vehicle is given by*

$$\begin{bmatrix} s[k+1] \\ v[k+1] \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} s[k] \\ v[k] \end{bmatrix}, \quad \begin{bmatrix} s[0] \\ v[0] \end{bmatrix} = \begin{bmatrix} s_0 \\ v_0 \end{bmatrix},$$

where $s[k]$ and $v[k]$, respectively, denote the position and the speed of vehicle over time. The user (i.e., the remote monitoring station) is interested in knowing the position of the vehicle; therefore, it submits a query of the form $Cx[k]$ with $C = [1 \; 0]$. According to Theorem 4, the server (or the vehicle) needs to provide the response $y[k] = Cx[k] + CA^k z$, where, if $T > 2$, z is a zero mean Gaussian random variable with covariance

$$\begin{aligned}\Sigma &= \frac{2}{\sqrt{\rho}} (\Psi_T^\top \Psi_T)^{-1/2} \\ &= \frac{2}{\sqrt{\rho}} \begin{bmatrix} T+1 & T(T+1)/2 \\ T(T+1)/2 & T(T+1)(2T+1)/6 \end{bmatrix}^{-1/2}.\end{aligned}$$

The quality of the response becomes

$$\begin{aligned}\mathcal{Q} = \frac{1}{\sqrt{3\rho}} \Big(&\sqrt{(T+1)(2T^2 + T + 6 - \sqrt{\Delta})} \\ &+ \sqrt{(T+1)(2T^2 + T + 6 + \sqrt{\Delta})} \Big),\end{aligned}$$

where $\Delta = 4T^4 + 4T^3 + 13T^2 - 12T + 36$. *This implies that*[1] $\mathcal{Q} = \mathcal{O}(T\sqrt{T}/\sqrt{\rho})$. *On the other hand,*

$$\mathbb{E}\{\|\hat{x}_0(y_T) - x_0\|_2^2\} = \frac{4\sqrt{3}}{\sqrt{\rho}} \Bigg(\frac{1}{\sqrt{(T+1)(2T^2+T+6-\sqrt{\Delta})}} + \frac{1}{\sqrt{(T+1)(2T^2+T+6+\sqrt{\Delta})}} \Bigg).$$

Thus, $\mathbb{E}\{\|\hat{x}_0(y_T) - x_0\|_2^2\} = \mathcal{O}(1/(T\sqrt{T}\sqrt{\rho}))$. *This shows the privacy guarantee decreases with* T. *Therefore, to keep the privacy guarantee constant, smaller* ρ *should be used for larger horizons* T *(i.e., a lower emphasis on preserving the quality of response must be placed). In fact,* ρ *should be selected such that it scales according to* $1/(T\sqrt{T})$ *with* T. *Doing so, the quality of the response* $\mathcal{Q}$ *depreciates according to* $\mathcal{O}(T^3)$. $\diamondsuit$

5 Discussion

The choice of the Fisher information as a measure of privacy in this paper, motivated by the Cramér-Rao bound, ensures that the privacy guarantees are applicable to a wide range of adversaries in contrast to, e.g., [42] assuming a least mean square error estimator as the adversary. Further, the Cramér-Rao bound provides a clear operational meaning for the measure of privacy. This could be potentially lacking in differential privacy literature [8] and studies using mutual information and entropy as a measure of privacy, e.g., [40]. However, in the presence of a prior distribution for the data and possible correlations between the entries of the database, privacy-preserving methods that do not use this additional information, such as the proposed method in this paper and algorithms relying on differential privacy, can underperform or break down [57], which is not the case for method relying on mutual information.

The use of constrained additive noise in this paper sets it apart from other studies in the literature that use an additive noise whose distribution has an infinite support, such as Laplace or Gaussian [4, 8–10, 12]. In the studies where the optimal noise is investigated, the support of the distribution is most often unrestricted, which again gives rise to the Laplace or Gaussian distributions being the optimal choices [15, 40–42].

In the unconstrained case, the optimal noise distribution minimizing the Fisher information subject to a constraint on the degradation of the quality of the response is proved to be Gaussian in fact[2]. This fact is at odds with the differential privacy literature [8]. Therefore, the provided framework with the guarantees provided by use of the Cramér-Rao bound is weaker than the differential privacy (both in requirements and guarantees). The use of the Gaussian noise is however known to satisfy a weaker variant of the differential privacy, referred to as (ϵ, δ) differential privacy [9, 58]. Further, noting that the privacy-preserving policy in this paper and that of [40] coincide in the unconstrained case, it is easy to see that the presented framework also minimizes the mutual information while having the advantage of providing a better operational meaning for the measure of privacy. This observation can be intuitively explained by the intimate relationship between Fisher information and mutual information [59, 60].

This section is finished by exploring the relationship between differential privacy and the proposed optimal noise in more depth for the weighted averaging setup in Example 2 with $\mathcal{X} = [\underline{x}, \overline{x}]^n$ for some $-\infty < \underline{x} \leq \overline{x} < +\infty$. The corrupted response of the server is ϵ-differentially private if $\mathbb{P}\{Cx + w \in \overline{\mathcal{Y}}\} \leq e^{\epsilon}\mathbb{P}\{Cx' + w \in \overline{\mathcal{Y}}\}$ for all $x, x' \in \mathcal{X}$ that only differ in one element and all Lebesgue-measurable sets $\overline{\mathcal{Y}}$. The following proposition proves that differential privacy can be achieved by an additive Laplace noise.

Proposition 3 *For Example 2 with* $\mathcal{X} = [\underline{x}, \overline{x}]^n$ *for some* $-\infty < \underline{x} \leq \overline{x} < +\infty$, *the corrupted response of the server is* ϵ*-differentially private if* $\gamma(w) = 1/(2b)\exp(-|w|/b)$ *with* $b = \epsilon/[(\overline{x} - \underline{x})\max_i |c_i|]$, *where* c_i *is the* i*-th entry of* C.

Proof: See [61]. ∎

For the differentially private noise in Proposition 3, it can be shown that $\mathcal{Q} = 2b^2$. Therefore, under the constraint $\mathcal{Q} = \vartheta$, one can achieve $[(\overline{x} - \underline{x})\sqrt{\vartheta/2}\max_i c_i]$-differential privacy. Further, it can be shown that for differentially private noise in Proposition 3,

$$\mathcal{I} = 2\int_0^{\infty} \frac{CC^{\top}}{\gamma(w)} \left[\frac{\partial \gamma(w)}{\partial w}\right]^2 \mathrm{d}w = \frac{CC^{\top}}{b^2} = \frac{2CC^{\top}}{\vartheta}.$$

Therefore, for any unbiased estimator of x denoted by $\hat{x}(y)$ under the noise density function in Proposition 3, $\mathbb{E}\{\|x - \hat{x}(y)\|_2^2\} \geq \|(I - C^{\dagger}C)x\|_2^2 + \vartheta/(2(CC^{\top})^2)$. However, for the optimal noise in Corollary 6, it can be shown that $\mathcal{I} = CC^{\top}/\vartheta$. Thus, for any unbiased estimator of x denoted by $\hat{x}(y)$ under the noise in Corollary 6, $\mathbb{E}\{\|x - \hat{x}(y)\|_2^2\} \geq \|(I - C^{\dagger}C)x\|_2^2 + \vartheta/(CC^{\top})^2$. Note that

$$\sup_{x \in [\underline{x}, \overline{x}]^n} \frac{\|(I - C^{\dagger}C)x\|_2^2 + \vartheta/(CC^{\top})^2}{\|(I - C^{\dagger}C)x\|_2^2 + \vartheta/(2(CC^{\top})^2)} = 1 + \kappa > 1,$$

where $\kappa = 1/(1 + 2(CC^{\top})^2 \max_{x \in [\underline{x}, \overline{x}]^n} \|(I - C^{\dagger}C)x\|_2^2/\vartheta)$. Thus, for the same bound on the quality of response,

[1] We say $f(x) = \mathcal{O}(g(x))$ if $\lim_{x \to \infty} |f(x)/g(x)| = c < \infty$.

[2] Note that by changing the measure of the quality of the response (e.g., expection of the norm-1 of the additive noise), one can get other noise density functions.

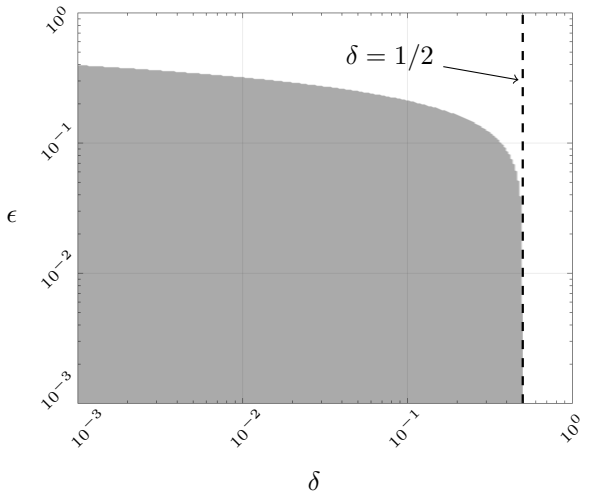

Fig. 2. Let $\vartheta/((\overline{x}-\underline{x})\max_i|c_i|) = 1$. The white and gray areas illustrate the regions for which the condition (16) is satisfied and is not satisfied, respectively. Note that satisfying (16) for $\delta \leq 1/2$ implies that the noise distribution in Corollary 6 is (ϵ, δ)-differentially private.

the optimal noise distribution in Corollary 6 provides a privacy guarantee that is $1+\kappa$ times stronger, albeit in the sense of the error covariance $\mathbb{E}\{\|x-\hat{x}(y)\|_2^2\}$.

For the differentially private noise in Proposition 3, it can be shown that $H(\gamma) = \log_2(e\sqrt{2\vartheta})$. For the optimal noise in Corollary 6, $H(\gamma^*) = \log_2(\sqrt{e2\pi\vartheta})$. Interestingly, $H(\gamma^*) \geq H(\gamma)$, which *intuitively* points to the fact that, for some prior distributions on the entries of the database, the optimal noise in Corollary 6 can potentially provide stronger information theoretic guarantees as well. This is not particularly surprising considering that the optimal noise for privacy in an information theoretic setting is also proved to be Gaussian [40].

The corrupted response of server is (ϵ, δ)-differentially private if $\mathbb{P}\{Cx + w \in \overline{\mathcal{Y}}\} \leq e^{\epsilon}\mathbb{P}\{Cx' + w \in \overline{\mathcal{Y}}\} + \delta$ for all $x, x' \in \mathcal{X}$ that only differ in one element and all Lebesgue-measurable sets $\overline{\mathcal{Y}}$. The following proposition proves that differential privacy can be achieved by an additive Gaussian noise.

Proposition 4 *For Example 2 with $\mathcal{X} = [\underline{x}, \overline{x}]^n$ for some $-\infty < \underline{x} \leq \overline{x} < +\infty$, the corrupted response of the server with noise distribution in Corollary 6 is (ϵ, δ)-differentially private if $\delta \leq 1/2$ and*

$$\vartheta \geq (\overline{x}-\underline{x})\max_i|c_i|\left(\frac{\sqrt{2\ln(1/(2\delta))}}{\epsilon} + \frac{1}{\sqrt{2\epsilon}}\right). \quad (16)$$

Proof: See [61]. ∎

In Figure 2, the white and gray areas illustrate the regions for which the condition (16) is satisfied and is not satisfied, respectively, when $\vartheta/((\overline{x}-\underline{x})\max_i|c_i|) = 1$. Note that satisfying (16) for $\delta \leq 1/2$ (behind the dashed line) implies that the noise distribution in Corollary 6 is (ϵ, δ)-differentially private for the corresponding values of ϵ and δ.

6 Conclusions and Future Work

In this paper, the problem of preserving the privacy of individual entries of a database with constrained additive noise is investigated. A measure of privacy using the Fisher information is developed. The optimal probability density function that maximizes the measure of privacy is computed. It is shown that, in some cases, the privacy-preserving policy of the server could potentially be independent of the entries of the database. Further, for scalar queries when the support set of the additive noise is bounded, the nature of the query, e.g., its linearity or content, does not play a role in the optimal additive noise for preserving the privacy of the database. For unconstrained noises, Gaussian distribution seems to be the optimal privacy-preserving policy if the quality of the provided response is measured using the variance of the additive noise employed by the server. Future work can focus on dynamic nonlinear systems.

References

[1] F. Farokhi and H. Sandberg, "Optimal privacy-preserving policy using constrained additive noise to minimize the Fisher information," in *Proceedings of the 56th IEEE Conference on Decision and Control*, pp. 2692–2697, 2017.

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

[3] B. Hoh, M. Gruteser, H. Xiong, and A. Alrabady, "Enhancing security and privacy in traffic-monitoring systems," *IEEE Pervasive Computing*, vol. 5, no. 4, pp. 38–46, 2006.

[4] Z. Huang, Y. Wang, S. Mitra, and G. E. Dullerud, "On the cost of differential privacy in distributed control systems," in *Proceedings of the 3rd International Conference on High Confidence Networked Systems*, pp. 105–114, ACM, 2014.

[5] S. Pearson and A. Benameur, "Privacy, security and trust issues arising from cloud computing," in *Proceedings of the IEEE Second International Conference on Cloud Computing Technology and Science*, pp. 693–702, 2010.

[6] J. Shao, *Mathematical Statistics*. Springer Texts in Statistics, Springer-Verlag New York, 2003.

[7] C. Dwork, "Differential privacy: A survey of results," in *Theory and Applications of Models of Computation: 5th International Conference, TAMC 2008, Xi'an, China, April 25-29, 2008. Proceedings* (M. Agrawal, D. Du, Z. Duan, and A. Li, eds.), pp. 1–19, Berlin, Heidelberg: Springer Berlin Heidelberg, 2008.

[8] C. Dwork, "Differential privacy," in *Encyclopedia of Cryptography and Security* (H. C. A. van Tilborg and S. Jajodia, eds.), Boston, MA: Springer US, 2011.

[9] J. Le Ny and G. J. Pappas, "Differentially private filtering," *IEEE Transactions on Automatic Control*, vol. 59, no. 2, pp. 341–354, 2014.

[10] S. Han, U. Topcu, and G. J. Pappas, "Differentially private convex optimization with piecewise affine objectives," in *Proceedings of the 53rd IEEE Conference on Decision and Control*, pp. 2160–2166, 2014.

[11] S. Han, U. Topcu, and G. J. Pappas, "Differentially private distributed constrained optimization," *IEEE Transactions on Automatic Control*, vol. 62, no. 1, pp. 50–64, 2017.

[12] C. Dwork and A. Roth, "The algorithmic foundations of differential privacy," *Foundations and Trends in Theoretical Computer Science*, vol. 9, no. 3–4, pp. 211–407, 2014.

[13] J. Soria-Comas and J. Domingo-Ferrer, "Optimal data-independent noise for differential privacy," *Information Sciences*, vol. 250, pp. 200–214, 2013.

[14] Q. Geng and P. Viswanath, "The optimal mechanism in differential privacy," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*, pp. 2371–2375, 2014.

[15] Y. Wang, Z. Huang, S. Mitra, and G. E. Dullerud, "Entropy-minimizing mechanism for differential privacy of discrete-time linear feedback systems," in *Proceedings of the 53rd IEEE Conference on Decision and Control*, pp. 2130–2135, 2014.

[16] M. S. Alvim, M. E. Andrés, K. Chatzikokolakis, P. Degano, and C. Palamidessi, "Differential privacy: On the trade-off between utility and information leakage," in *Formal Aspects of Security and Trust: 8th International Workshop, FAST 2011, Leuven, Belgium, September 12-14, 2011. Revised Selected Papers* (G. Barthe, A. Datta, and S. Etalle, eds.), pp. 39–54, Berlin, Heidelberg: Springer Berlin Heidelberg, 2012.

[17] D. J. Mir, "Information-theoretic foundations of differential privacy," in *Foundations and Practice of Security: 5th International Symposium, FPS 2012, Montreal, QC, Canada, October 25-26, 2012, Revised Selected Papers* (J. Garcia-Alfaro, F. Cuppens, N. Cuppens-Boulahia, A. Miri, and N. Tawbi, eds.), pp. 374–381, Berlin, Heidelberg: Springer Berlin Heidelberg, 2013.

[18] A. Makhdoumi and N. Fawaz, "Privacy-utility tradeoff under statistical uncertainty," in *Proceedings of the 51st Annual Allerton Conference on Communication, Control, and Computing*, pp. 1627–1634, 2013.

[19] 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*, 2017. In Press.

[20] Y. Wang, Z. Huang, S. Mitra, and G. E. Dullerud, "Entropy-minimizing mechanism for differential privacy of discrete-time linear feedback systems," in *Proceedings of the 53rd Annual Conference on Decision and Control*, pp. 2130–2135, 2014.

[21] Z. Huang, S. Mitra, and G. Dullerud, "Differentially private iterative synchronous consensus," in *Proceedings of the 2012 ACM workshop on Privacy in the Electronic Society*, pp. 81–90, 2012.

[22] E. Nozari, P. Tallapragada, and J. Cortés, "Differentially private average consensus with optimal noise selection," in *Proceedings of the 5th IFAC Workshop on Distributed Estimation and Control in Networked Systems*, pp. 203–208, 2015.

[23] V. Katewa, A. Chakrabortty, and V. Gupta, "Protecting privacy of topology in consensus networks," in *Proceedings of the American Control Conference*, pp. 2476–2481, 2015.

[24] Y. Mo and R. M. Murray, "Privacy preserving average consensus," *IEEE Transactions on Automatic Control*, 2016. In Press.

[25] X. Duan, J. He, P. Cheng, Y. Mo, and J. Chen, "Privacy preserving maximum consensus," in *Proceedings of the 54th Annual Conference on Decision and Control*, pp. 4517–4522, 2015.

[26] J. Le Ny, "Privacy-preserving nonlinear observer design using contraction analysis," in *Proceedings of the 54th Annual Conference on Decision and Control*, pp. 4499–4504, 2015.

[27] J. Cortés, G. E. Dullerud, S. Han, J. Le Ny, S. Mitra, and G. J. Pappas, "Differential privacy in control and network systems," in *Proceedings of the 55th Conference on Decision and Control*, pp. 4252–4272, 2016.

[28] M. Hale and M. Egerstedty, "Differentially private cloud-based multi-agent optimization with constraints," in *Proceedings of the American Control Conference*, pp. 1235–1240, IEEE, 2015.

[29] E. Nozari, P. Tallapragada, and J. Cortés, "Differentially private distributed convex optimization via functional perturbation," *IEEE Transactions on Control of Network Systems*, 2016. In Press.

[30] A. Friedman and A. Schuster, "Data mining with differential privacy," in *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 493–502, 2010.

[31] B. I. P. Rubinstein, P. L. Bartlett, L. Huang, and N. Taft, "Learning in a large function space: Privacy-preserving mechanisms for SVM learning," *Journal of Privacy and Confidentiality*, vol. 4, no. 1, pp. 65–100, 2012.

[32] F. McSherry and K. Talwar, "Mechanism design via differential privacy," in *Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science*, pp. 94–103, 2007.

[33] K. Nissim, R. Smorodinsky, and M. Tennenholtz, "Approximately optimal mechanism design via differential privacy," in *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference*, pp. 203–213, ACM, 2012.

[34] R. Dong, W. Krichene, A. M. Bayen, and S. S. Sastry, "Differential privacy of populations in routing games," in *Proceedings of the 54th Annual Conference on Decision and Control*, pp. 2798–2803, 2015.

[35] F. Kargl, A. Friedman, and R. Boreli, "Differential privacy in intelligent transportation systems," in *Proceedings of the 6th ACM Conference on Security and Privacy in Wireless and Mobile Networks*, pp. 107–112, 2013.

[36] G. Kalogridis, C. Efthymiou, S. Z. Denic, T. A. Lewis, and R. Cepeda, "Privacy for smart meters: Towards undetectable appliance load signatures," in *Proceedings of the IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pp. 232–237, 2010.

[37] O. Tan, D. Gunduz, 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.

[38] J. Yao and P. Venkitasubramaniam, "On the privacy-cost tradeoff of an in-home power storage mechanism," in *Communication, Control, and Computing (Allerton), 2013 51st Annual Allerton Conference on*, pp. 115–122, IEEE, 2013.

[39] T. Tanaka and H. Sandberg, "SDP-based joint sensor and controller design for information-regularized optimal LQG control," in *Proceedings of the 54th IEEE Conference on Decision and Control*, pp. 4486–4491, 2015.

[40] E. Akyol, C. Langbort, and T. Basar, "Privacy constrained information processing," in *Proceedings of the 54th IEEE Conference on Decision and Control*, pp. 4511–4516, IEEE, 2015.

[41] F. Farokhi and G. Nair, "Privacy-constrained communication," *IFAC-PapersOnLine*, vol. 49, no. 22, pp. 43–48, 2016.

[42] F. Farokhi, H. Sandberg, I. Shames, and M. Cantoni, "Quadratic Gaussian privacy games," in *Proceedings of the 54th IEEE Conference on Decision and Control*, pp. 4505–4510, 2015.

[43] T. Tanaka, M. Skoglund, H. Sandberg, and K. H. Johansson, "Directed information and privacy loss in cloud-based control," in *Proceedings of the American Control Conference*, 2017.

[44] H. Anderson, "Efficiency versus protection in a general randomized response model," *Scandinavian Journal of Statistics*, pp. 11–19, 1977.

[45] T. K. Nayak and S. A. Adeshiyan, "A unified framework for analysis and comparison of randomized response surveys of binary characteristics," *Journal of Statistical Planning and Inference*, vol. 139, no. 8, pp. 2757–2766, 2009.

[46] X. Zhou, B. Peng, Y. F. Li, Y. Chen, H. Tang, and X. Wang, "To release or not to release: Evaluating information leaks in aggregate human-genome data," in *Computer Security – ESORICS 2011: 16th European Symposium on Research in Computer Security, Leuven, Belgium, September 12-14,2011. Proceedings* (V. Atluri and C. Diaz, eds.), pp. 607–627, Berlin, Heidelberg: Springer Berlin Heidelberg, 2011.

[47] Y. Erlich and A. Narayanan, "Routes for breaching and protecting genetic privacy," *Nature Reviews Genetics*, vol. 15, no. 6, pp. 409–421, 2014.

[48] F. K. Dankar and K. El Emam, "Practicing differential privacy in health care: A review," *Transactions on Data Privacy*, vol. 6, no. 1, pp. 35–67, 2013.

[49] C. Dwork and A. Roth, "The algorithmic foundations of differential privacy," *Foundations and Trends in Theoretical Computer Science*, vol. 9, no. 3-4, pp. 211–407, 2014.

[50] R. Zamir, "A proof of the Fisher information inequality via a data processing argument," *IEEE Transactions on Information Theory*, vol. 44, no. 3, pp. 1246–1250, 1998.

[51] P. J. Huber, *Robust Statistics*. Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, 1981.

[52] E. Uhrmann-Klingen, "Minimal Fisher information distributions with compact-supports," *Sankhyā: The Indian Journal of Statistics, Series A (1961-2002)*, vol. 57, no. 3, pp. 360–374, 1995.

[53] V. Jeyakumar and H. Wolkowicz, "Zero duality gaps in infinite-dimensional programming," *Journal of Optimization Theory and Applications*, vol. 67, no. 1, pp. 87–108, 1990.

[54] P. Stoica and P. Babu, "The Gaussian data assumption leads to the largest Cramér-Rao bound [lecture notes]," *IEEE Signal Processing Magazine*, vol. 28, no. 3, pp. 132–133, 2011.

[55] E. D. Sontag, *Mathematical Control Theory: Deterministic Finite Dimensional Systems*. Texts in Applied Mathematics, Springer New York, 2013.

[56] I. Marshall C. Pease, *Methods of Matrix Algebra*. Mathematics in Science and Engineering, Academic Press, 1965.

[57] D. Kifer and A. Machanavajjhala, "No free lunch in data privacy," in *Proceedings of the 2011 ACM SIGMOD International Conference on Management of Data*, pp. 193–204, 2011.

[58] H. Sandberg, G. Dán, and R. Thobaben, "Differentially private state estimation in distribution networks with smart meters," in *Proceedings of the 54th IEEE Conference on Decision and Control*, pp. 4492–4498, 2015.

[59] B. S. Clarke and A. R. Barron, "Information-theoretic asymptotics of Bayes' methods," *IEEE Transactions on Information Theory*, vol. 36, no. 3, pp. 453–471, 1990.

[60] J. J. Rissanen, "Fisher information and stochastic complexity," *IEEE Transactions on Information Theory*, vol. 42, no. 1, pp. 40–47, 1996.

[61] F. Farokhi and H. Sandberg, "Ensuring privacy with constrained additive noise by minimizing Fisher information." Technical Report, 2018. `arXiv preprint: arXiv:1808.09565 [math.OC] https://arxiv.org/abs/1808.09565`.

[62] R. T. Rockafellar, *Convex Analysis*. Princeton Landmarks in Mathematics and Physics, Princeton University Press, 1997.

[63] C. H. Edwards, *Advanced Calculus of Several Variables*. Academic Press, 1973.

[64] A. S. Mohamed and H. A. Atia, "Separation of the Schrödinger operator with an operator potential in the Hilbert spaces," *Applicable Analysis*, vol. 84, no. 1, pp. 103–110, 2005.

A A Useful Inequality

Proposition 5 $\overline{\mathcal{J}} \geq n^2 \mathcal{J}^{-1}$.

Proof: First, note that

$$\mathrm{Tr}(\mathcal{I}^{-1}) = \sum_{i=1}^{n} \frac{1}{\lambda_i(\mathcal{I})} \geq \frac{n^2}{\sum_{i=1}^{n} \lambda_i(\mathcal{I})} = n^2 \mathrm{Tr}(\mathcal{I})^{-1},$$

where the inequality follows from the Jensen's inequality [62, p. 25] and the facts that the mapping $z \mapsto 1/z$ is convex over $\mathbb{R}_{\geq 0}$ and $\lambda_i(\mathcal{I}) \geq 0$ for all i (since $\mathcal{I}$ is positive semi-definite). Further, it can be shown that

$$\begin{aligned}\int_{x\in\mathcal{X}} \mathrm{Tr}(\mathcal{I}(x)^{-1}) p(x) \mathrm{d}x &\geq n^2 \int_{x\in\mathcal{X}} \mathrm{Tr}(\mathcal{I}(x))^{-1} p(x) \mathrm{d}x \\ &\geq n^2 \left(\int_{x\in\mathcal{X}} \mathrm{Tr}(\mathcal{I}(x)) p(x) \mathrm{d}x \right)^{-1},\end{aligned}$$

where the second inequality follows from the Jensen's inequality and the earlier observation that the mapping $z \mapsto 1/z$ is convex over $\mathbb{R}_{\geq 0}$. ∎

B Proof of Theorem 2

First, noting that the cost function and the constraint set are convex, the stationary condition (that the variational derivative is equal to zero) is sufficient for optimality. Further, if multiple density functions satisfy the

sufficiency conditions, they all exhibit the same cost. In the rest of the proof, this condition is rewritten in a simpler forms. To do so, note that

$$\begin{aligned}\mathcal{I}(x)=&\int_{y\in\{f(x)\}\oplus\mathcal{W}}\gamma(y-f(x)|x)\left[\frac{\partial\log(\gamma(y-f(x)|x))}{\partial x}\right]\\&\times\left[\frac{\partial\log(\gamma(y-f(x)|x))}{\partial x}\right]^\top\mathrm{d}y\\=&\int_{w\in\mathcal{W}}\frac{1}{\gamma(w|x)}\left[F(x)^\top\frac{\partial\gamma(w|x)}{\partial w}+\frac{\partial\gamma(w|x)}{\partial x}\right]\\&\times\left[F(x)^\top\frac{\partial\gamma(w|x)}{\partial w}+\frac{\partial\gamma(w|x)}{\partial x}\right]^\top\mathrm{d}w.\end{aligned}$$

Thus,

$$\begin{aligned}\mathrm{Tr}(\mathcal{I}(x))=&\int_{w\in\mathcal{W}}\frac{1}{\gamma(w|x)}\left[F(x)^\top\frac{\partial\gamma(w|x)}{\partial w}+\frac{\partial\gamma(w|x)}{\partial x}\right]^\top\\&\times\left[F(x)^\top\frac{\partial\gamma(w|x)}{\partial w}+\frac{\partial\gamma(w|x)}{\partial x}\right]\mathrm{d}w.\end{aligned}$$

As a result,

$$\begin{aligned}\mathcal{J}=&\int_{x\in\mathcal{X}}\int_{w\in\mathcal{W}}\frac{p(x)}{\gamma(w|x)}\left[F(x)^\top\frac{\partial\gamma(w|x)}{\partial w}+\frac{\partial\gamma(w|x)}{\partial x}\right]^\top\\&\times\left[F(x)^\top\frac{\partial\gamma(w|x)}{\partial w}+\frac{\partial\gamma(w|x)}{\partial x}\right]\mathrm{d}w\mathrm{d}x.\end{aligned}$$

Following the result of [53], the Lagrangian can be constructed as

$$\begin{aligned}\mathcal{L}=&\int_{x\in\mathcal{X}}\int_{w\in\mathcal{W}}\frac{p(x)}{\gamma(w|x)}\left[F(x)^\top\frac{\partial\gamma(w|x)}{\partial w}+\frac{\partial\gamma(w|x)}{\partial x}\right]^\top\\&\times\left[F(x)^\top\frac{\partial\gamma(w|x)}{\partial w}+\frac{\partial\gamma(w|x)}{\partial x}\right]\mathrm{d}w\mathrm{d}x\\&-\int_{x\in\mathcal{X}}p(x)\mu(x)\bigg(\int_{w\in\mathcal{W}}\gamma(w|x)\mathrm{d}w-1\bigg)\mathrm{d}x\\=&\int_{x\in\mathcal{X}}\int_{w\in\mathcal{W}}p(x)\bigg(-\mu(x)\gamma(w|x)\\&+\frac{1}{\gamma(w|x)}\left[F(x)^\top\frac{\partial\gamma(w|x)}{\partial w}+\frac{\partial\gamma(w|x)}{\partial x}\right]^\top\\&\times\left[F(x)^\top\frac{\partial\gamma(w|x)}{\partial w}+\frac{\partial\gamma(w|x)}{\partial x}\right]\bigg)\mathrm{d}w\mathrm{d}x\\&+\int_{x\in\mathcal{X}}\mu(x)p(x)\mathrm{d}x,\end{aligned}$$

where $\mu:\mathcal{X}\to\mathbb{R}$ is the Lagrange multiplier corresponding to the equality constraint $\int_{w\in\mathcal{W}}\gamma(w|x)\mathrm{d}w=1$ for all $x\in\mathrm{supp}(p)$. Using Theorem 5.3 in [63, p. 440], it can be seen that the extrema must satisfy (B.1) on top of the next page. Introducing the change of variable $\gamma(w|x)=u(w,x)^2$ results in

$$\begin{aligned}&\mu(x)+\frac{4}{u(x,w)}\mathrm{Tr}\left(\begin{bmatrix}F(x)F(x)^\top & F(x)\\F(x)^\top & I\end{bmatrix}D^2u(w,x)\right)\\&\quad+\frac{4}{u(x,w)}\mathbb{1}^\top D^2f(x)\frac{\partial u(w,x)}{\partial w}\\&\quad+\frac{4}{u(x,w)}\frac{1}{p(x)}\frac{\partial p(x)}{\partial x}^\top\begin{bmatrix}F(x)^\top & I\end{bmatrix}\begin{bmatrix}\dfrac{\partial\gamma(w|x)}{\partial w}\\[2mm]\dfrac{\partial\gamma(w|x)}{\partial x}\end{bmatrix}=0,\end{aligned}$$

for all $w\in\mathrm{int}(\mathcal{W})$ and $x\in\mathrm{supp}(p)$. Again, if $u(w,x)\neq0$ for all $w\in\mathrm{int}(\mathcal{W})$ and $x\in\mathrm{supp}(p)$, it can be deduced that

$$\begin{aligned}&\bar{\mu}(x)u(w,x)+\mathrm{Tr}\bigg(\begin{bmatrix}F(x)F(x)^\top & F(x)\\F(x)^\top & I\end{bmatrix}D^2u(w,x)\bigg)\\&\quad+\mathbb{1}^\top D^2f(x)\frac{\partial u(w,x)}{\partial w}\\&\quad+\frac{1}{p(x)}\frac{\partial p(x)}{\partial x}^\top\begin{bmatrix}F(x)^\top & I\end{bmatrix}\begin{bmatrix}\dfrac{\partial\gamma(w|x)}{\partial w}\\[2mm]\dfrac{\partial\gamma(w|x)}{\partial x}\end{bmatrix}=0,\end{aligned}$$

where $\bar{\mu}(x)=\mu(x)/4$. However, if $u(w,x)=0$ for some $w\in\mathrm{int}(\mathcal{W})$ and $x\in\mathrm{supp}(p)$, the equality cannot be satisfied with any $\mu\in\mathbb{R}$.

C Proof of Corollary 1

In this proof, a solution of the form $u(w,x)=u(w)$ and $\mu(x)=\mu$ is sought for the partial differential equation in (12). In this case, the partial differential equation in (12) becomes

$$\begin{cases}\nabla^2u(w)+\mu u(w)=0, & w\in\mathcal{W},\\u(w)=0, & w\in\partial\mathcal{W}.\end{cases}\tag{C.1}$$

This is a special case of the time-independent Schrödinger equation. This knowledge can be used to solve the partial differential equation explicitly. Following [64], the solution of (C.1) is unique. The rest easily follows from showing that the provided density function satisfies the partial differential equation and its boundary conditions.

D Proof of Corollary 2

First, we prove part (i). Allow $u^{(\underline{w},\overline{w})}(w)$ be such that $\gamma(w)=u^{(\underline{w},\overline{w})}(w)^2$ with $\gamma(w)$ denoting the solution of Problem 2 in Corollary 1 for $\mathcal{W}=[\underline{w},\overline{w}]$. To emphasize

$$\frac{p(x)}{\gamma(w|x)^2}\begin{bmatrix}\frac{\partial\gamma(w|x)}{\partial w}\\ \frac{\partial\gamma(w|x)}{\partial x}\end{bmatrix}^\top\begin{bmatrix}F(x)F(x)^\top & F(x)\\ F(x)^\top & I\end{bmatrix}\begin{bmatrix}\frac{\partial\gamma(w|x)}{\partial w}\\ \frac{\partial\gamma(w|x)}{\partial x}\end{bmatrix}+2\sum_{i=1}^{m}\frac{\partial}{\partial w_i}\left(\frac{p(x)}{\gamma(w|x)}e_i^\top\begin{bmatrix}F(x)F(x)^\top & F(x)\\ F(x)^\top & I\end{bmatrix}\begin{bmatrix}\frac{\partial\gamma(w|x)}{\partial w}\\ \frac{\partial\gamma(w|x)}{\partial x}\end{bmatrix}\right)$$
$$+2p(x)\sum_{i=m+1}^{n+m}\frac{\partial}{\partial x_i}\left(\frac{1}{\gamma(w|x)}e_i^\top\begin{bmatrix}F(x)F(x)^\top & F(x)\\ F(x)^\top & I\end{bmatrix}\begin{bmatrix}\frac{\partial\gamma(w|x)}{\partial w}\\ \frac{\partial\gamma(w|x)}{\partial x}\end{bmatrix}\right)$$
$$+2\sum_{i=m+1}^{n+m}\frac{\partial p(x)}{\partial x_i}\left(\frac{1}{\gamma(w|x)}e_i^\top\begin{bmatrix}F(x)F(x)^\top & F(x)\\ F(x)^\top & I\end{bmatrix}\begin{bmatrix}\frac{\partial\gamma(w|x)}{\partial w}\\ \frac{\partial\gamma(w|x)}{\partial x}\end{bmatrix}\right)+p(x)\mu(x)=0. \tag{B.1}$$

the fact that $\mathcal{W}$ is a function of $\underline{w}$ and $\overline{w}$, in this proof, the notation $\mathcal{W}^{(\underline{w},\overline{w})}$ is used. Therefore,

$$\begin{aligned}\mathcal{I}&=\int_{w\in\mathcal{W}^{(\underline{w},\overline{w})}}\left[\frac{\partial u^{(\underline{w},\overline{w})}(w)}{\partial w}\right]\left[\frac{\partial u^{(\underline{w},\overline{w})}(w)}{\partial w}\right]^\top\mathrm{d}w\\&=\int_{w\in\mathcal{W}^{(\underline{w},\overline{w})}}\frac{1}{(\overline{w}-\underline{w})^{m+2}}\left[\frac{\partial u^{(0,1)}(w')}{\partial w'}\right]_{w'=(w-\underline{w})/(\overline{w}-\underline{w})}\\&\qquad\times\left[\frac{\partial u^{(0,1)}(w')}{\partial w'}\right]^\top_{w'=(w-\underline{w})/(\overline{w}-\underline{w})}\mathrm{d}w\\&=\frac{1}{(\overline{w}-\underline{w})^2}\underbrace{\int_{w'\in\mathcal{W}^{(0,1)}}\left[\frac{\partial u^{(0,1)}(w')}{\partial w'}\right]\left[\frac{\partial u^{(0,1)}(w')}{\partial w'}\right]^\top\mathrm{d}w'}_{:=\kappa}.\end{aligned}$$

Now, we prove part (ii). To do so, note that

$$\begin{aligned}\mathcal{Q}&=\mathrm{Tr}(\mathbb{E}\{ww^\top\})\\&=\mathrm{Tr}(\mathbb{E}\{(w-\mathbb{E}\{w\})(w-\mathbb{E}\{w\})^\top\})+m(\overline{w}+\underline{w})^2/4,\end{aligned}$$

where the second equality follows from that $\mathbb{E}\{w\}^\top\mathbb{E}\{w\}=m(\overline{w}+\underline{w})^2/4$. Noting that w_i is independent of w_j for the optimal policy in Corollary 1 results in

$$\begin{aligned}&\mathbb{E}\{(w-\mathbb{E}\{w\})(w-\mathbb{E}\{w\})^\top\}\\&\quad=\mathrm{diag}(\mathbb{E}\{(w_1-\mathbb{E}\{w_1\})^2\},\cdots,\mathbb{E}\{(w_m-\mathbb{E}\{w_m\})^2\}).\end{aligned}$$

Note that

$$\begin{aligned}\mathbb{E}\{(w_i-\mathbb{E}\{w_i\})^2\}=&\left(\frac{2}{\overline{w}-\underline{w}}\right)\int_{\underline{w}}^{\overline{w}}\left(w_i-\frac{\overline{w}+\underline{w}}{2}\right)^2\\&\times\cos^2\left(\frac{\pi}{\overline{w}-\underline{w}}\left(w_i-\frac{\overline{w}+\underline{w}}{2}\right)\right)\mathrm{d}w_i\\=&\frac{(\pi^2-6)(\overline{w}-\underline{w})^2}{12\pi^2}.\end{aligned}$$

E Proof of Corollary 3

In this proof, a solution of the form $u(w,x)=u(w)$ and $\mu(x)=\mu$ is sought for the partial differential equation in (12). In this case, it can be shown that $\mathrm{Tr}(C^\top D^2u(w)C)=u''(w)\mathrm{Tr}(C^\top C)$. Therefore, the partial differential equation in (12) becomes the ordinary differential equation $u''(w)+\bar{\mu}u(w)=0$ for $w\in\mathcal{W}$ with the boundary condition that $u(w)=0$ for all $w\in\partial\mathcal{W}$, where $\bar{\mu}=\mu/\mathrm{Tr}(C^\top C)$. The differential equation $u''(w)+\bar{\mu}u(w)=0$ admits a solution of the form $u(w)=\alpha\cos(\sqrt{\bar{\mu}}(w-\beta))$ where $\alpha,\beta\in\mathbb{R}$ are constants depending on the boundary conditions. It should be ensured that $u(\underline{w})=u(\overline{w})=0$ since $\gamma(w)=u(w)^2$ for $w\in\partial\mathcal{W}$. Thus $\bar{\mu}>0$. Two distinct situations may occur:

- $\nexists q\in\mathbb{Z}$ such that $\sqrt{\bar{\mu}}=(2q+1)\pi/(\overline{w}-\underline{w})$: To be able to satisfy $u(\overline{w})=u(\underline{w})=0$, it must be that $\alpha=0$. In this case, $\int_{w\in\mathcal{W}}u(w)^2\mathrm{d}w=0$, which contradicts the requirement that $\int_{w\in\mathcal{W}}\gamma(w)\mathrm{d}w=1$.
- $\exists q\in\mathbb{Z}$ such that $\sqrt{\bar{\mu}}=(2q+1)\pi/(\overline{w}-\underline{w})$: In this case, $\beta=(\overline{w}+\underline{w})/2$. To ensure that $u(w)\neq0$ for all $w\in\mathrm{int}(\mathcal{W})$, select $q=0$. Finally, to be able to satisfy the equality constraint $\int_{w\in\mathcal{W}}\gamma(w)\mathrm{d}w=1$, pick $\alpha=\sqrt{2/(\overline{w}-\underline{w})}$.

Finally, note that since the cost function and the constraint set are convex, all the extrema are minimizers.

F Proof of Corollary 4

It can be shown that (12) becomes

$$\mathrm{Tr}\left(\begin{bmatrix}1&1\\1&1\end{bmatrix}D^2u(w,x)\right)+\frac{p'(x)}{p(x)}\left[\frac{\partial u(w,x)}{\partial w}+\frac{\partial u(w,x)}{\partial x}\right]+\mu(x)u(w,x)=0.$$

Introducing the change of variable $v = x + w$ results in

$$\frac{\partial^2 u(v,x)}{\partial v^2} + \frac{p'(x)}{p(x)}\frac{\partial u(v,x)}{\partial v} + \mu(x)u(v,x) = 0. \quad \text{(F.1)}$$

Note that v must belong to $[x+\underline{w}, x+\overline{w}]$. To ensure that $u(v,x) = 0$ on $x+\underline{w}$ and $x+\overline{w}$ (because $u(w,x) = 0$ on $\partial\mathcal{W}$), $\mu(x)$ must be selected such that $[p'(x)/p(x)]^2 - 4\mu(x) < 0, \forall x \in \text{supp}(x)$. Under this condition, the solution of the partial differential equation in (F.1) becomes

$$\begin{aligned} u(v,x) =& \alpha \exp\left(-\frac{p'(x)}{2p(x)}v\right) \\ & \times \cos\left(\sqrt{\left[\frac{p'(x)}{2p(x)}\right]^2 - \mu(x)}(v-\beta)\right). \end{aligned}$$

Following the same line of reasoning as in the proof of Corollary 3, it can be inferred that $\beta = ((x+\underline{w}) + (x+\overline{w}))/2 = x + (\underline{w}+\overline{w})/2$ and $\sqrt{[p'(x)/(2p(x))]^2 - \mu(x)} = \pi/(\overline{w}-\underline{w})$. This concludes the proof.

G Proof of Corollary 5

The proof follows from the selection of $u(w,x) = u(w)$ and $\mu(x) = \text{Tr}(F(x)F(x)^\top)\mu$ and following the same line of reasoning as in Corollary 3.

H Proof of Theorem 3

Similarly, because the cost function and the constraint set are convex, the stationarity condition is sufficient for optimality. Further, if multiple density functions satisfy the conditions, they all exhibit the same cost. In this case, the Lagrangian can be constructed as

$$\begin{aligned} \mathcal{L} = \int_{x\in\mathcal{X}}\int_{w\in\mathcal{W}} p(x)\Big(&-\mu(x)\gamma(w|x) + \varrho w^\top w\gamma(w|x) \\ &+ \frac{1}{\gamma(w|x)}\left[F(x)^\top\frac{\partial\gamma(w|x)}{\partial w} + \frac{\partial\gamma(w|x)}{\partial x}\right]^\top \\ &\times\left[F(x)^\top\frac{\partial\gamma(w|x)}{\partial w} + \frac{\partial\gamma(w|x)}{\partial x}\right]\Big)\mathrm{d}w\mathrm{d}x \\ &+ \int_{x\in\mathcal{X}}\mu(x)p(x)\mathrm{d}x. \end{aligned}$$

The rest of the proof follows the same line of reasoning as in Theorem 2.

I Proof of Corollary 6

First, we present the solution of Problem 3. In this proof, a solution of the form $u(w,x) = u(w)$ and $\mu(x) = \mu$ is sought for the partial differential equation in (15). Note that $u(w) = \alpha\exp(-w^\top\Sigma^{-1}w/4)$ satisfies the partial differential equation (15) with $\mu = \text{Tr}(C^\top\Sigma^{-1}C)/2$ and $\Sigma = 2(CC^\top)^{1/2}/\sqrt{\rho}$. Further to ensure that $\int_{w\in\mathcal{W}} u(w)^2 = 1$, select $\alpha = 1/(\sqrt{(2\pi)^m\det(\Sigma)})$. Now, we present the solution of Problem 3. This follows from that for Problem 4 the duality gap is zero [53]. Therefore, the constraint on the variance can be added to the cost function using a Lagrange multiplier, which transforms the problem into that of Problem 3. Therefore, following Corollary 6, the solution is equal to the density function of a zero-mean Gaussian random variable. Finally, for Gaussian random variables, the Fisher information is a decreasing function of the variance. Therefore, the Lagrange multiplier is set so that the inequality constraint on the variance becomes active. That means $\text{Tr}(2(CC^\top)^{1/2}/\sqrt{\rho}) = \vartheta$. As a result, $\sqrt{\rho} = 2\text{Tr}((CC^\top)^{1/2})/\vartheta$.

J Proof of Theorem 4

Applying the result of Theorem 3 for linear query functions and probability density functions that are independent of the content of the server, the sufficient condition of optimality comes from the solution to the partial differential equation $\text{Tr}(\Psi_T^\top D^2u(w_T)\Psi_T) + (\mu - (\varrho/4)w_T^\top w_T)u(w_T) = 0$. Introduce the change of variable $w_T = \Psi_T z$ for $z \in \mathbb{R}^n$ (recall that only one solution for the partial differential equation needs to be calculated). It can be shown that $D^2\bar{u}(z) = D^2u(\Psi_T z) = \Psi_T^\top D^2u(w_T)\big|_{w_T=\Psi_T z}\Psi_T$. Therefore, $\text{Tr}(D^2u(\Psi_T z)) + (\mu - (\varrho/4)z^\top\Psi_T^\top\Psi_T z)u(\Psi_T z) = 0$. Define $\bar{u}(z) = u(\Psi_T z)$. Thus, $\text{Tr}(D^2\bar{u}(z)) + (\mu - (\varrho/4)z^\top\Psi_T^\top\Psi_T z)\bar{u}(z) = 0$. The rest follows from that $\bar{u}(z) = \alpha\exp(-z^\top\Sigma^{-1}z/4)$ satisfies this partial differential equation. To do so, note that $\text{Tr}(D^2\bar{u}(z)) = \bar{u}(z)\text{Tr}(\Sigma^{-1}zz^\top\Sigma^{-1}) - \text{Tr}(\Sigma^{-1})/2$. Therefore, $\mu = \text{Tr}(\Sigma^{-1})/2$, $\Sigma = 2(\Psi_T^\top\Psi_T)^{-1/2}/\sqrt{\varrho}$, and $\alpha = 1/\sqrt[4]{(2\pi)^m\det(\Sigma)}$.