Journal of Machine Learning Research 25 (2024) 1-50 Submitted 12/22; Revised 7/24; Published 10/24

Pure Differential Privacy for Functional Summaries with a Laplace-like Process

Haotian Lin hzl435@psu.edu
Department of Statistics,
The Pennsylvania State University,
University Park, PA 16802, USA

Matthew Reimherr mreimherr@psu.edu
Department of Statistics,
The Pennsylvania State University,
University Park, PA 16802, USA

Editor: Jie Peng

Abstract

Many existing mechanisms for achieving differential privacy (DP) on infinite-dimensional functional summaries typically involve embedding these functional summaries into finite-dimensional subspaces and applying traditional multivariate DP techniques. These mechanisms generally treat each dimension uniformly and struggle with complex, structured summaries. This work introduces a novel mechanism to achieve pure DP for functional summaries in a separable infinite-dimensional Hilbert space, named the *Independent Component Laplace Process* (ICLP) mechanism. This mechanism treats the summaries of interest as truly infinite-dimensional functional objects, thereby addressing several limitations of the existing mechanisms. Several statistical estimation problems are considered, and we demonstrate how one can enhance the utility of private summaries by oversmoothing the non-private counterparts. Numerical experiments on synthetic and real datasets demonstrate the effectiveness of the proposed mechanism.

Keywords: Differential Privacy, Functional Data Analysis, Hilbert Space, Reproducing Kernel Hilbert Space, Infinite-Dimensional Statistics.

1. Introduction

Data privacy has garnered critical attention in the last decade as substantial amounts of individualized data are collected. The most widely used paradigm in formal data privacy is *differential privacy* (DP), introduced by Dwork et al. (2006). DP provides a rigorous and interpretable definition of data privacy, as it limits the amount of information attackers can infer from publicly released database queries. Numerous mechanisms have been developed for conventional data settings, such as scalar or vector-valued data. However, advances in technologies enable us to collect and process densely observed data over some temporal or spatial domains, which are coined *functional data* to differentiate them from classic multivariate data (Ramsay et al., 2005; Kokoszka and Reimherr, 2017; Ferraty and Romain, 2011). Although functional data analysis has been proven useful in various fields, such as economics, finance, and genetics, and has been widely researched in the statistical community, there are only a few works concerning privacy preservation within the realm of functional data.

When the statistical summaries are finite-dimensional, additive noise mechanisms are the most commonly used mechanisms to achieve DP, which privatize statistical summaries by adding calibrated noise from predetermined distributions, e.g., Laplace and Gaussian mechanisms (Dwork et al., 2006, 2014). In this paper, we are concerned with establishing an additive noise mechanism for functional summaries, namely infinite-dimensional summaries, to achieve ϵ-DP. Given the challenge that functional summaries are typically infinite-dimensional, most existing mechanisms embed the non-private summaries into a finite-dimensional subspace by using finite basis expansions to approximate summaries and applying classical multivariate privacy tools, such as perturbing the expansion coefficients with i.i.d. noise (Zhang et al., 2012; Wang et al., 2013; Chandrasekaran et al., 2014; Alda and Rubinstein, 2017). This finite-dimensional embedding process is typically unavoidable, as when the summaries are infinite-dimensional, adding i.i.d. noise to each dimension is not even feasible if one wants the private summaries to remain in a specific infinite-dimensional function space. However, these mechanisms have several weaknesses. First, determining the dimension of the subspace is crucial, as it plays a trade-off role between utility and privacy. While data-driven approaches might cause potential privacy leakage, a predetermined dimension will lack adaptation to the data, potentially failing to capture the structure or shape of the functional summaries or injecting excess noise. Second, in multivariate settings, classical privacy tools that add i.i.d. noise to each dimension treat all dimensions equally and allocate the privacy budget over each dimension uniformly, failing to recognize the different levels of importance of coefficients across different dimensions and thus injecting excess noise for "more important" dimensions. This substantially degrades the utility and robustness of private functional summaries. Some previous works have shown that capturing the covariance structure in the data might be able to reduce the amount of noise injected in specific scenarios (Hardt and Talwar, 2010; Awan and Slavković, 2021).

1.1 Our Contributions

To overcome the downsides inherent in DP mechanisms that rely on finite-dimensional embedding, we introduce a mechanism that treats both the functional summary and the privacy noise as truly infinite-dimensional functional objects. Concretely, our contributions can be summarized as follows:

1. We propose an ϵ-DP mechanism by perturbing functional summaries with a random element called the *Independent Component Laplace Process* (ICLP) and name this mechanism *the ICLP mechanism.* We establish the feasibility of the ICLP mechanism (meaning it can achieve DP) in an infinite-dimensional separable Hilbert space, $\mathbb{H}$, by characterizing a subspace of $\mathbb{H}$ and showing that the feasibility holds if and only if the difference between two functional summaries based on adjacent datasets resides in this subspace. We also show how the proposed mechanism applies to the space of continuous functions, even though this space is not a Hilbert space.

2. We provide strategies based on regularized empirical risk minimization (regularized ERM) to obtain qualified functional summaries for the ICLP mechanism. We uncover

the role regularization plays in the trade-off between utility and privacy. Specifically, we show that one can achieve ϵ-DP with a matching order (or even a lower order) of privacy error as the estimation error by slightly oversmoothing the functional summaries in the functional mean protection problem. We also show that the application can go beyond classic functional data settings, as it is also applicable to the realm of more classic non-parametric smoothing problems like kernel density estimation.

3. To obtain privacy-safe regularization parameters in regularized ERM, we propose a privacy-safe selection approach so that choosing the parameters is only tied to the covariance structure of the ICLP noise, thus achieving end-to-end privacy. This approach overcomes the potential privacy leakage in conventional data-driven methods.

The proposed mechanism differentiates itself from existing mechanisms, such as Zhang et al. (2012); Dwork et al. (2014); Alda and Rubinstein (2017), that rely on finite-dimensional embedding and treat each dimension uniformly by adding i.i.d. noise to each dimension in the following senses. First, the ICLP mechanism avoids finite-dimensional subspace embeddings and frees the assumption that every dataset in the database shares the same finite-dimensional subspace. Second, unlike existing mechanisms, the ICLP mechanism treats each dimension heterogeneously, allowing it to achieve a more effective noise injection process while handling truly infinite-dimensional functional summaries and noise.

1.2 Related Works

In the overlap of functional summaries and differential privacy, the landmark paper is Hall et al. (2013), which provided a framework for achieving (ϵ, δ)-DP on infinite-dimensional functional objects but focused on a finite grid of evaluation points. The follow-up work in Mirshani et al. (2019) pushed Hall et al. (2013)'s result forward and established (ϵ, δ)-DP over the full functional path for objects in Banach spaces. In more general spaces, Reimherr and Awan (2019a) considered elliptical perturbations to achieve (ϵ, δ)-DP in locally convex vector spaces, including all Hilbert spaces, Banach spaces, and Fréchet spaces. They also showed the impossibility of achieving ϵ-DP for infinite-dimensional functional objects with elliptical distributions.

Turning to ϵ-DP, a series of works has been proposed by resorting to finite-dimensional representations, such as polynomial bases, trigonometric bases, or Bernstein polynomial bases, to approximate target functional summaries (Wang et al., 2013; Chandrasekaran et al., 2014; Alda and Rubinstein, 2017) and loss functions (Zhang et al., 2012). These mechanisms then perturb the expansion coefficients in $\mathbb{R}^m$ via the m-dimensional i.i.d. Laplace mechanism (Dwork et al., 2014). Privatizing m-dimensional coefficients is feasible through the K-norm mechanism (Hardt and Talwar, 2010; Awan and Slavković, 2021), which encompasses the multivariate i.i.d. Laplace mechanism as a particular instance. However, to the best of our knowledge, no existing literature combines the K-norm mechanism with finite-dimensional embedding techniques to achieve DP for functional summaries. In addition to additive noise mechanisms, Awan et al. (2019) extended the exponential mechanism (McSherry and Talwar, 2007) to arbitrary Hilbert spaces and showed its application to functional principal component analysis. From the robust noise injection perspective, a heterogeneous noise injection scheme (Phan et al., 2019) was proposed by assigning different weighted privacy budgets to each coordinate to further improve the robustness of private summaries.

1.3 Notations and Organization

The following notations are used throughout the rest of this work and follow standard conventions. For asymptotic notations: $f(n) = O(g(n))$ or $f(n) \lesssim O(g(n))$ means for all c there exists $k > 0$ such that $f(n) \leq cg(n)$ for all $n \geq k$; $f(n) \asymp g(n)$ means $f(n) = O(g(n))$ and $g(n) = O(f(n))$; $f(n) = o(g(n))$ means for any $c > 0$ there exists $k > 0$ such that $f(n) \leq cg(n)$ for all $n \geq k$.

The rest of the paper is organized as follows. In Section 2, we provide preliminaries on functional summaries and spaces, differential privacy, and a generalization of finite-dimensional embedding mechanisms. In Section 3, we formally propose the ICLP mechanism and establish its feasibility in separable Hilbert spaces (Theorem 8) and in the space of continuous functions (Theorem 9). In Section 4, we propose approaches for constructing qualified summaries for the ICLP mechanism and apply them to various statistical applications with utility analysis. Implementation of the mechanism is provided in Section 5. We evaluate the performance of the proposed mechanism on both synthetic datasets and real-world applications in Sections 6 and 7. Concluding remarks are given in Section 8. Many technical results, including proofs, lemmas, and propositions, are deferred to the Appendix

2. Preliminaries

2.1 Functional Summaries and Spaces

First, we define the functional summary. Let $\mathbb{H}$ be an infinite-dimensional real separable Hilbert space with inner product $\langle \cdot, \cdot \rangle_{\mathbb{H}}$. For a set $\mathcal{X}$, we define $\mathcal{D} = \mathcal{X}^n$ as the collection of all possible n-unit datasets, and let D be an element of $\mathcal{D}$. This paper considers the functional summary statistic of interest as an element of $\mathbb{H}$, i.e., $f : \mathcal{D} \to \mathbb{H}$.

Next, we briefly introduce the background of random elements in $\mathbb{H}$ and some spaces we will work on in this paper. We refer readers to Hsing and Eubank (2015) for a detailed introduction. A random element $X \in \mathbb{H}$ is said to have mean $\mu \in \mathbb{H}$ and (linear) covariance operator $C : \mathbb{H} \to \mathbb{H}$ if

$$\mathrm{E}[\langle X, h \rangle_{\mathbb{H}}] = \langle \mu, h \rangle_{\mathbb{H}} \qquad \text{and} \qquad \mathrm{Cov}(\langle X, h_1 \rangle_{\mathbb{H}}, \langle X, h_2 \rangle_{\mathbb{H}}) = \langle Ch_1, h_2 \rangle_{\mathbb{H}},$$

for any h, h_1, h_2 in $\mathbb{H}$. When $\mathrm{E}\, \|X - \mu\|_{\mathbb{H}}^2 < \infty$, the covariance operator C exists, is self-adjoint, positive semidefinite, and trace class (thus Hilbert-Schmidt and compact) (Bosq, 2000; Hsing and Eubank, 2015). According to the spectral theorem for self-adjoint compact operators, C admits the eigen decomposition as

$$C(h) = \sum_{j \geq 1} \lambda_j \langle h, \phi_j \rangle_{\mathbb{H}} \phi_j, \quad \forall h \in \mathbb{H},$$

where $\{\lambda_j\}_{j \geq 1}$ and $\{\phi_j\}_{j \geq 1}$ are the eigenvalues and eigenfunctions of C, respectively.

We define two norms associated with C as

$$\|h\|_C = \sqrt{\sum_{j=1}^{\infty} \frac{\langle h, \phi_j \rangle_{\mathbb{H}}^2}{\lambda_j}} \quad \textit{and} \quad \|h\|_{1,C} = \sum_{j=1}^{\infty} \frac{|\langle h, \phi_j \rangle_{\mathbb{H}}|}{\sqrt{\lambda_j}}.$$

We denote the two subspaces of $\mathbb{H}$ induced by $\|\cdot\|_C$ and $\|\cdot\|_{1,C}$ as $\mathcal{H}_C = \{h \in \mathbb{H} : \|h\|_C < \infty\}$ and $\mathcal{H}_{1,C} = \{h \in \mathbb{H} : \|h\|_{1,C} < \infty\}$, respectively. Note $\|\cdot\|_C$ is the classic Cameron-Martin norm induced by C (Bogachev, 1998) and $\mathcal{H}_C$ is called the Cameron-Martin space of C. The $\|\cdot\|_{1,C}$-norm is analogous to a weighted ℓ^1-norm. The space $\mathcal{H}_{1,C}$ is included in $\mathcal{H}_C$, i.e., $h \in \mathcal{H}_{1,C}$ leads to $h \in \mathcal{H}_C$.

We also define the power operator of C. For any $s \geq 0$, the power operator C^s is defined as

$$C^s(h) = \sum_{j \geq 1} \lambda_j^s \langle h, \phi_j \rangle_{\mathbb{H}} \phi_j, \quad \forall h \in \mathbb{H},$$

meaning C^s shares the same eigenfunctions as C while the eigenvalues are raised to the power of s. We define its corresponding power space as

$$\mathcal{H}_{C^s} = \left\{ h \in \mathbb{H} : \|h\|_{C^s} := \sqrt{\sum_{j=1}^{\infty} \frac{\langle h, \phi_j \rangle_{\mathbb{H}}^2}{\lambda_j^s}} < \infty \right\}.$$

The space $\mathcal{H}_{1,C^s}$ and its associated norm $\|\cdot\|_{1,C^s}$ follow a similar definition.

2.2 Differential Privacy

For a given non-private functional summary statistic f_D, we denote its private version as $\tilde{f}_D$, which is a random element of $\mathbb{H}$ indexed by D. We state the definition of differential privacy in terms of conditional distributions (Wasserman and Zhou, 2010).

Definition 1 *Let $\tilde{f}_D$ be the privatized functional summary of f_D. Assume $\{P_D : D \in \mathcal{D}\}$ is the family of probability measures over Ω induced by $\{\tilde{f}_D : D \in \mathcal{D}\}$. We say $\tilde{f}_D$ achieves (ϵ, δ)-DP if for any two adjacent datasets (different in only one record) D and D', and any measurable set $A \in \mathcal{F}$, one has*

$$P_D(A) \leq e^{\epsilon} P_{D'}(A) + \delta. \tag{1}$$

In particular, if $\delta = 0$, we say $\tilde{f}_D$ achieves ϵ-DP.

The definition implies that the summaries of two adjacent datasets should have almost the same probability distribution. The privacy budget ϵ controls how much privacy will be lost while releasing the result, and a small ϵ implies a higher similarity between P_D and $P_{D'}$, and thus increased privacy. Before introducing different DP mechanisms, we first define the global sensitivity of a summary statistic, a central concept of DP (Dwork et al., 2006). For a functional summary $f : \mathcal{D} \to \mathbb{H}$ and a norm $\|\cdot\|$ in the Hilbert space $\mathbb{H}$, the global sensitivity of the summary f_D with respect to the norm $\|\cdot\|$ is given by

$$\Delta = \sup_{D \sim D'} \|f_D - f_{D'}\|,$$

where $D \sim D'$ means D and D' are adjacent datasets. Here, the norm $\|\cdot\|$ is typically tailored based on the employed DP mechanism, e.g., K-norm for the K-norm mechanism (Hardt et al., 2010; Awan and Slavković, 2021), Cameron-Martin space norm for the Gaussian mechanism in separable Banach space (Mirshani et al., 2019). Given that global sensitivity quantifies how much a summary can change with the modification of a single record in the dataset, the additive noise must be calibrated proportionally to the global sensitivity.

2.3 Finite-Dimensional Representation-wise Laplace

We introduce the *F̲inite-Dimensional R̲epresentation-wise L̲aplace* (FRL) mechanism, which generalizes almost all current additive noise mechanisms for ϵ-DP that rely on finite-dimensional representations, such as Wang et al. (2013); Chandrasekaran et al. (2014); Alda and Rubinstein (2017); Zhang et al. (2012), to name a few. Let $\{e_j\}_{j\geq 1}$ be an orthonormal basis in $\mathbb{H}$. Then, one can approximate the summary using M basis functions, i.e.,

$$\hat{f}_D = \sum_{j=1}^{M} f_{Dj} e_j \quad \textit{with} \quad f_{Dj} = \langle f_D, e_j \rangle_{\mathbb{H}}.$$

Expanding the functional summaries via a finite basis facilitates dimension reduction so that the classic multivariate i.i.d. Laplace mechanism (Dwork et al., 2014) can be implemented to privatize the coefficient vector $(f_{D1}, \cdots, f_{DM})$. Specifically, the privatized summary can be expressed as

$$\tilde{f}_D = \sum_{j=1}^{M} (f_{Dj} + Z_j) e_j, \quad \text{with} \quad Z_j \overset{i.i.d.}{\sim} Lap(0, \Delta/\epsilon)$$

where Δ is the global sensitivity in ℓ^1-distance. The FRL mechanism privatizes the functional summary $\hat{f}_D$ without regard to the varying importance of different components. Although some components are more crucial for the estimation, the FRL mechanism treats all components equally during the privatization process, thereby reducing the utility of the privatized summary. Additionally, the truncation level M controls the trade-off among variance, bias, and privacy. This mechanism forces one to either introduce more noise or accept higher bias when more components are required to deal with complex functional summaries.

The m-dimensional coefficient vector can also be privatized using the K-norm mechanism (Hardt and Talwar, 2010; Awan and Slavković, 2021), by perturbing the coefficients through a multivariate random variable in $\mathbb{R}^M$, whose density is proportional to $\exp\{-\epsilon\|\cdot\|_K\}$ for a given K-norm in $\mathbb{R}^M$. Utilizing the ℓ^1-norm results in the aforementioned multivariate i.i.d. Laplace mechanism (the FRL mechanism). However, as noted in Awan and Slavković (2021), when M is large, determining the optimal K-norm is often nontrivial, and sampling these multivariate random variables can be challenging. Due to these challenges and the prevalence of all existing ϵ-DP mechanisms employing finite-dimensional representation with i.i.d. Laplace noise for functional summaries, our discussion primarily focuses on the FRL mechanism. We leave the exploration of using other K-norms with finite-dimensional representation for privatizing functional summaries for future studies.

3. The Independent Component Laplace Process Mechanism

In this section, we first formally define the Independent Component Laplace Process and then propose an additive noise mechanism called *the ICLP mechanism*. Specifically, to achieve ϵ-DP, we will release the privatized summary that takes the form of $\tilde{f}_D = f_D + \sigma Z$, where σ is a positive scalar and Z is an ICLP noise. Initially, we assume f_D lies in a real separable Hilbert space $\mathbb{H}$ and establish the ϵ-DP guarantee. We then show that privacy protection can also hold for the space of continuous functions (which is not a Hilbert space) under certain assumptions on the covariance operator. The proofs of all the theorems can be found in Appendix A.

3.1 Independent Component Laplace Process

The proposed random element is motivated by Mirshani et al. (2019), who achieved (ϵ, δ)-DP for functional summaries in Banach spaces. Formally, their additive noise mechanism can be expressed as $\tilde{f}_D = f_D + \sigma Z$, where $Z \sim GP(0, C)$ is a centered Gaussian process with covariance operator C. There is a dual perspective of this mechanism. By applying the Karhunen-Loéve Theorem (Kosambi, 2016), the mechanism is equivalent to

$$\tilde{f}_D = f_D + \sigma Z = \sum_{j=1}^{\infty} \left(\langle f_D, \phi_j \rangle + \sigma \langle Z, \phi_j \rangle \right) \phi_j, \tag{2}$$

where $\{\lambda_j\}_{j\geq 1}$ and $\{\phi_j\}_{j\geq 1}$ are the eigenvalues and eigenfunctions of C and $\{\langle Z, \phi_j \rangle\}_{j\geq 1}$ are independent Gaussian random variables with zero mean and variance λ_j. This decomposition indicates that the mechanism perturbs each coefficient with independent Gaussian random variables. Unfortunately, the existing Laplace process cannot play a role analogous to the Gaussian process under such a decomposition, as it is an elliptical distribution. It has been proven that no elliptical distribution can achieve ϵ-DP in infinite-dimensional spaces. Specifically, in an infinite-dimensional space, adding any elliptical distribution is equivalent to adding noise from a randomly scaled Gaussian process, which satisfies only the weaker notion of (ϵ, δ)-DP. We refer readers to Theorem 4 of Reimherr and Awan (2019a) for more details. Motivated by this dual perspective via the Karhunen-Loéve expansion and the fact that the most widely used additive noise mechanism for ϵ-DP in the univariate case is the Laplace mechanism, we consider using independent Laplace random variables with heterogeneous variances in the decomposition (2). This is equivalent to perturbing the functional summary with a particular random element defined as follows.

Definition 2 *Let X be a random element in $\mathbb{H}$ with $\mathrm{E}\,\|X\|_{\mathbb{H}}^2 < \infty$ and C be its covariance operator. Denote the eigenvalues and eigenfunctions of C as $\{\lambda_j\}_{j\geq 1}$ and $\{\phi_j\}_{j\geq 1}$. We say X is an Independent Component Laplace Process (ICLP) with mean μ if it admits the following decomposition*

$$X = \mu + \sum_{j=1}^{\infty} \sqrt{\lambda_j} Z_j \phi_j, \tag{3}$$

where Z_j are i.i.d. Laplace random variables with zero mean and variance 1.

Remark 3 *If the objective is to achieve ϵ-DP, alternative i.i.d. sequences of random variables $\{Z_j\}_{j\geq 1}$ with zero mean and unit variance, characterized by a well-constructed density, are technically viable. However, this requires that the density tail of these random variables be calibrated "just right" to comply with ϵ-DP. Typically, their density tails should be as heavy, or closely similar, to the Laplace distribution. Otherwise, using light-tailed random variables, like Gaussian, will cause the probability inequality (1) of ϵ-DP to fail for sets in the tails.*

The collection of square-integrable random elements of $\mathbb{H}$ is itself a Hilbert space with inner product $\mathrm{E}\langle X, Y\rangle_{\mathbb{H}}$. The following theorem states that if a random element X is defined via the infinite sum decomposition in Definition 2, it is still well-defined in $\mathbb{H}$.

Theorem 4 *For a given non-negative decreasing real sequence $\{\lambda_j\}_{j\geq 1}$ that is summable, and an orthonormal basis $\{\phi_j\}_{j\geq 1}$ for $\mathbb{H}$, the random element X defined via (3) is well-defined within $\mathbb{H}$.*

3.2 Feasibility in Separable Hilbert Spaces

When the summary f_D of interest is infinite-dimensional, it turns out that the summary f_D depends heavily on the structure of the random element Z in the additive noise mechanism. Otherwise, it is possible to make the global sensitivity infinite, and thus, no finite amount of noise would be able to achieve DP. For example, in Mirshani et al. (2019), the privatized summary is $\tilde{f}_D = f_D + \sigma Z$, where Z is a zero-mean Gaussian process with covariance operator C. It has been proved that the summary f_D must be "compatible" with the Gaussian process Z, i.e., $f_D - f_{D'}$ lies in the Cameron-Martin space of Z for any adjacent datasets D, D', to achieve (ϵ, δ)-DP, or no finite σ will make the mechanism satisfy (ϵ, δ)-DP. Given that our problem setting is also infinite-dimensional, a similar analysis is needed for the ICLP mechanism. In this section, we will investigate the feasibility of the ICLP mechanism, i.e., identifying the specific conditions under which there always exists a finite σ such that the privatized summary $f_D + \sigma Z$ via the ICLP mechanism achieves ϵ-DP.

To investigate the feasibility of a randomized mechanism for ϵ-DP, one can start with the equivalence or orthogonality of probability measures. As discussed in Awan et al. (2019) and Reimherr and Awan (2019a), the probability measures induced by an ϵ-DP mechanism are necessarily equivalent (though this is not sufficient for DP) in a probabilistic sense; otherwise, it is impossible to achieve DP if the measures are orthogonal. More specifically, if the mechanism produces a private summary $\tilde{f}_D$ that is probabilistically orthogonal to $\tilde{f}_{D'}$, i.e., there exists a $A \in \mathcal{F}$ such that $P_D(A) = 0$ and $P_{D'}(A) = 1$, then the mechanism cannot be DP since f_D and $f_{D'}$ can be distinguished with probability one on A. In the following, we use this perspective to develop the feasibility of the ICLP mechanism. Denote the probability measure family induced by the ICLP mechanism as $\{P_D : D \in \mathcal{D}\}$. In the following theorem, we provide necessary and sufficient conditions for pairwise equivalence in $\{P_D : D \in \mathcal{D}\}$.

Theorem 5 (Equivalence of ICLP probability measures) *Let $D, D' \in \mathcal{D}$ be two adjacent datasets, $\tilde{f}_D, \tilde{f}_{D'}$ be the privatized summaries based on the ICLP mechanism. Denote the corresponding probability measures over $\mathbb{H}$ as P_D and $P_{D'}$, and the covariance operator of ICLP as C. Then P_D and $P_{D'}$ are equivalent if and only if*

$$f_D - f_{D'} \in \mathcal{H}_C = \{h \in \mathbb{H} : \|h\|_C < \infty\}. \tag{4}$$

Theorem 5 shows that if the difference between f_D and $f_{D'}$ resides in the Cameron-Martin space of C, then the probability family will be pairwise equivalent. An analogous result for the equivalence of elliptical distributions appears in Theorem 2 of Reimherr and Awan (2019a), even though the ICLP is not an elliptical distribution. However, it turns out that, unlike elliptical distributions, pairwise equivalence is not enough for the ICLP mechanism to achieve ϵ-DP. To see the reason behind this, one must consider the density of P_D in $\mathbb{H}$. Since there is no common base measure in $\mathbb{H}$ that plays the same role as the Lebesgue measure in $\mathbb{R}^d$, it is more complicated to consider the density in $\mathbb{H}$. Fortunately, we are adding the same type of noise to the functional summaries. Therefore, we only need the density as the Radon-Nikodym derivative of P_D with respect to P_0, where P_0 is the probability measure induced by σZ.

Theorem 6 (Density of ICLP) *Let P_h and P_0 be the probability measures induced by $\{h + \sigma Z\}$ and σZ respectively. Suppose $h \in \mathcal{H}_{1,C}$, then the Radon–Nikodym derivative of P_D with respect to P_0 is given by*

$$\frac{dP_h}{dP_0}(z) = \exp\left\{-\frac{1}{\sigma}\left(\|z-h\|_{1,C} - \|z\|_{1,C}\right)\right\}, \tag{5}$$

P_0 almost everywhere and is unique.

Now, we are ready to show why the condition (4) is insufficient for ϵ-DP. Indeed, even though $f_D - f_{D'} \in \mathcal{H}_C$ guarantees the pairwise equivalence between P_D and $P_{D'}$, it does not guarantee the density in Equation (5) is well-defined and thus cannot be upper bounded, which is a requirement for ϵ-DP however. Meanwhile, $\mathcal{H}_C$ is enough for (ϵ, δ)-DP with Gaussian process with covariance C (Mirshani et al., 2019) since it allows densities to be unbounded up to a set with P_0 measure less than δ. In the following theorem, we will show the appropriate space in which $f_D - f_{D'}$ should reside is the subspace of $\mathcal{H}_{1,C}$.

Theorem 7 (Impossibility of The ICLP Mechanism) *Under the same conditions of Theorem 5, let $\mathcal{H}_{1,C} = \{f \in \mathbb{H} : \|f\|_{1,C} < \infty\}$ be a subspace of $\mathcal{H}_C$ and if*

$$f_{D_1} - f_{D_2} \in \mathcal{H}_C \setminus \mathcal{H}_{1,C},$$

then there is no $\sigma \in \mathbb{R}^+$ such that the ICLP mechanism, $\tilde{f}_D = f_D + \sigma Z$, satisfies ϵ-DP.

Indeed, if f_D resides in the gap between $\mathcal{H}_C$ and $\mathcal{H}_{1,C}$, the sensitivity of f_D will be infinite, and there is no possibility to calibrate the ICLP noise with any finite σ to achieve ϵ-DP. Now, with the proper space in Theorem 7 and the feasible density in Theorem 6, we can establish the ICLP mechanism formally.

Theorem 8 (The ICLP Mechanism) *Let f_D be the functional summary and Z be an ICLP with covariance operator C. Define the global sensitivity of the ICLP mechanism as*

$$\Delta = \sup_{D \sim D'} \|f_D - f_{D'}\|_{1,C} \quad \textit{and} \quad \sigma = \Delta/\epsilon. \tag{6}$$

Then the privatized version of f_D, $\tilde{f}_D = f_D + \sigma Z$, achieves ϵ-DP.

3.3 Example of $\mathcal{H}_C$ and $\mathcal{H}_{1,C}$ for Smooth Functions

In this section, we provide a concrete example of the spaces $\mathcal{H}_C$ and $\mathcal{H}_{1,C}$, in which the smoothness of the elements in these spaces is of primary interest. Consider the Hilbert space $\mathbb{H}$ as $L^2(\mathcal{X})$ where $\mathcal{X} \subseteq \mathbb{R}^d$. For the ICLP covariance function C defined over $\mathcal{X} \times \mathcal{X}$, the eigenfunctions $\{\phi_j\}_{j\geq 1}$ are an orthonormal basis of $L^2(\mathcal{X})$. Suppose the eigenvalues of C decay polynomially with order β, i.e., $\lambda_j \asymp j^{-\beta}$. Then the space $\mathcal{H}_C$ and $\mathcal{H}_{1,C}$ thus can be expressed as

$$\mathcal{H}_C = \left\{ f = \sum_{j=1}^{\infty} f_j \phi_j \in L^2(\mathcal{X}) : \sum_{j=1} j^{2\beta} f_j^2 < \infty \right\},$$

$$\mathcal{H}_{1,C} = \left\{ f = \sum_{j=1}^{\infty} f_j \phi_j \in L^2(\mathcal{X}) : \sum_{j=1} j^{\beta} |f_j| < \infty \right\}.$$

If $\{\phi_j\}_{j\geq 1}$ are the Fourier basis, $\mathcal{H}_C$ corresponds to a Sobolev space of order β, containing functions with weak derivatives up to order β that are L^2-integrable. $\mathcal{H}_{1,C}$ corresponds to a Hölder space of order β, consisting of those functions with continuous derivatives up to order $\lfloor \beta \rfloor$, and the $\lfloor \beta \rfloor$-th derivative is $\beta - \lfloor \beta \rfloor$ Hölder continuous. We refer readers to Section 2.3 of Yang et al. (2017) for further details.

3.4 Extensions To the Space of Continuous Functions

Theorem 8 implies that the ICLP mechanism provides privacy protection for a wide range of infinite-dimensional functional objects in separable Hilbert spaces. The DP *post-processing inequality* (Dwork et al., 2014) is a fundamental property for functional summaries since one may only be practically interested in a few scalar summaries. However, the post-processing inequality only applies to measurable mappings. If $\mathbb{H} = L^2([0,1])$, then this eliminates the possibility of releasing point-wise evaluations of the functional summary since such mappings are not measurable operators in $L^2([0,1])$. Therefore, in this section, we extend the ICLP mechanism to the space of continuous functions, i.e., $\mathcal{C}(T)$ with T a compact set over $\mathbb{R}^d$, where such operators are measurable (and thus protected). We show that the ICLP, under mild conditions, is also in $\mathcal{C}(T)$.

Theorem 9 (Feasibility in the Space of Continuous Functions) *Let $C : T \times T \to \mathbb{R}$ be a symmetric, positive definite, bivariate function over compact domain T. If C is α-Hölder continuous in each coordinate, i.e., there exists a positive constant M_C, and $\alpha \in (0,1]$ such that $|C(t_1,s) - C(t_2,s)| \leq M_C |t_1 - t_2|^{\alpha}$, then there exists an ICLP, Z, with covariance function C and a modification $\tilde{Z} : T \times \Omega \to \mathbb{R}$ of Z that is a continuous process, such that*

1. *$\tilde{Z}$ is sample continuous, i.e., $\forall \omega \in \Omega$, $\tilde{Z}_\omega(t)$ is continuous with respect to $t \in T$;*
2. *For any $t \in T$, $P(\tilde{Z}(t) = Z(t)) = 1$.*

Meaning that there exists a stochastic process in $\mathcal{C}(T)$ equally distributed as the ICLP except on a zero-measure set.

All the ICLP mechanism results for $\mathbb{H}$ in Section 3.2 are now applicable to functional summaries in $\mathcal{C}(T)$. Furthermore, the point-wise evaluation is now a measurable operation and thus is protected. We also note that the proof of Theorem 9 is not just a standard result from stochastic processes but relies heavily on the structure of the ICLP.

4. Qualified Summary Obtainment and Privacy-Preserving Tasks

In this section, we present different approaches to constructing non-private summaries for the ICLP mechanism. Specifically, these constructions must ensure that the difference of functional summaries, $f_D - f_{D'}$, lies in $\mathcal{H}_{1,C}$ for any adjacent datasets D and D', thus qualifying the summaries for the ICLP mechanism. After introducing these approaches, we also apply the ICLP mechanism to achieve privacy protection in different statistical estimation problems with corresponding statistical analysis.

4.1 Generalized Obtainment of Qualified Summaries

Based on Theorems 7 and 8, to achieve ϵ-DP via the ICLP mechanism, the difference of non-private summaries calculated from any adjacent datasets D and D' should reside in $\mathcal{H}_{1,C}$. We call a summary f that satisfies such conditions a qualified summary. However, constructing such f directly is a challenging task. It is easier to address the problem by restricting the individual functional summary f_D residing in $\mathcal{H}_{1,C}$ for any $D \in \mathcal{D}$, which automatically leads to $f_D - f_{D'} \in \mathcal{H}_{1,C}$. Therefore, we leverage regularized ERM to obtain qualified summaries f_D such that $f_D \in \mathcal{H}_{1,C}$.

Formally, let $L(f, D) : \mathbb{H} \times \mathcal{D} \to \mathbb{R}$ be a loss function. The *ICLP with Absolute Regularization* (ICLP-AR) estimator is defined as follows,

$$\text{(ICLP-AR)}: \quad \hat{f}_D = \operatorname*{argmin}_{f \in \mathbb{H}} \left\{ L(f, D) + \psi \, \|f\|_{1,C^\eta} \right\} \quad \textit{for } \eta \geq 1, \tag{7}$$

where C^η is the power kernel of C that shares the same eigenfunctions as C while the eigenvalues are raised to λ_j^η and ψ is the regularization parameter.

The benefits of using a power kernel C^η are twofold. First, the space corresponding to $\|\cdot\|_{1,C^\eta}$ is a subspace of $\mathcal{H}_{1,C}$, guaranteeing that $\hat{f}_D \in \mathcal{H}_{1,C}$. Second, it allows more flexibility to control the regularity (usually smoothness) of the constructed functional summaries. Later on, we will see that even though $\eta = 1$ is a natural setting, setting $\eta > 1$, i.e., constructing a slightly over-smoothing summary, can be helpful for utility and even make privacy error negligible compared to estimation error. However, as we will see in Section 4.2, there are some serious drawbacks to using the $\|\cdot\|_{1,C}$-norm regularization.

Therefore, we consider restricting the functional summary in the power space of $\mathcal{H}_C$, i.e., $\mathcal{H}_{C^\eta}$ and using $\|\cdot\|_{C^\eta}$-norm regularization in regularized ERM as our final strategy, which turns out to work quite well theoretically and practically. Formally, for a given $\eta > 1$, by the Cauchy-Schwarz inequality we have

$$\|h\|_{1,C} = \sum_{j=1}^{\infty} \frac{|h_j|}{\sqrt{\lambda_j}} = \sum_{j=1}^{\infty} \frac{|h_j|}{\lambda_j^{\frac{\eta}{2}}} \lambda_j^{\frac{\eta-1}{2}} \leq \|h\|_{C^\eta} \sqrt{\operatorname{trace}(C^{\eta-1})}.$$

Therefore, by taking $\eta > 1$ such that $C^{\eta-1}$ is a trace-class operator, we obtain the functional summary via the following *ICLP with Quadratic Regularization* (ICLP-QR) strategy,

$$\text{(ICLP-QR)}: \quad \hat{f}_D = \operatorname*{argmin}_{f \in \mathbb{H}} \left\{ L(f, \mathcal{D}) + \psi \, \|f\|_{C^\eta}^2 \right\} \quad \textit{for } \eta > 1. \tag{8}$$

Since the power space $\mathcal{H}_{C^\eta}$ is an RKHS and $\|\cdot\|_{C^\eta}$ takes quadratic form, we name this approach ICLP-QR. Here, η is strictly greater than 1 leads to $\mathcal{H}_{C^\eta} \subseteq \mathcal{H}_{1,C} \subseteq \mathcal{H}_C$, which ensures the feasibility of the ICLP mechanism. Similar to ICLP-AR, the η in ICLP-QR also plays a role in balancing the utility and the privacy of the functional summary.

4.2 Protection for Mean

We consider the problem of privatizing the mean summary. Assume $X_1, \cdots, X_n$ are i.i.d. random elements drawn from an arbitrary real separable Hilbert space $\mathbb{H}$ with mean element $\mathrm{E}\, X_i = \mu_0 \in \mathbb{H}$. Our goal is to release a private estimator for the true mean μ_0 that satisfies ϵ-DP. When using the FRL mechanism, one can start with the sample mean $\hat{\mu}_D = \frac{1}{n}\sum_{i=1}^n X_i$, which is an unbiased estimator of μ_0. For the ICLP mechanism, we use regularized ERM with the square loss to obtain qualified non-private summaries, i.e.,

$$\hat{\mu}_D = \operatorname*{argmin}_{\theta \in \mathbb{H}} \left\{ \frac{1}{n} \sum_{i=1}^{n} \|X_i - \theta\|_{\mathbb{H}}^2 + \psi P(\theta) \right\}, \tag{9}$$

with $P(\theta) = \|\theta\|_{1,C^\eta}$ or $\|\theta\|_{C^\eta}^2$ corresponding to ICLP-AR and ICLP-QR respectively.

To obtain the global sensitivity and utility analysis for the proposed strategies, we first state some standard assumptions in the DP literature regarding the norm of the observed data and the eigenvalue decay rate of C.

Assumption 1 (Boundedness) *Assume for any sample path X, its $\mathbb{H}$-norm is bounded by τ, i.e., $\|X\|_{\mathbb{H}} \le \tau$.*

The bounded norm assumption is commonly used in the DP literature, primarily to ensure finite global sensitivity. This assumption is often adapted to align with specific DP paradigms and mechanisms. For instance, in employing the Gaussian mechanism to attain (ϵ, δ)-DP, it is customary to assume that the ℓ^2-norm of the data is bounded, in line with the global sensitivity being assessed via $\|\cdot\|_{\ell^2}$. Conversely, for the Laplace mechanism, which aims for ϵ-DP, the focus shifts to the ℓ^1-norm, with global sensitivity gauged through $\|\cdot\|_{\ell^1}$. In cases where a more general norm in $\mathbb{R}^d$ is used, such as in the exponential mechanism, the data is presumed to have such a finite general norm. See the mean estimation example in Reimherr and Awan (2019b).

In the context of the ICLP mechanism, where global sensitivity is evaluated under $\|\cdot\|_{1,C}$, it seems logical to posit that $\|X\|_{1,C} \le \tau$ for some finite τ. However, the $\|\cdot\|_{1,C}$-norm is intrinsically linked to C, dependent on the eigenvalues and eigenfunctions of the covariance used in ICLP. This reliance implies that assuming $\|X\|_{1,C} \le \tau$ could narrow the ICLP mechanism's applicability. Practically, it would mean that feasible X_i should be drawn from X whose covariance eigenfunctions align with those of ICLP's covariance C, a condition often unverifiable in practice. Therefore, we consider a more general and relaxed boundedness assumption by assuming $\|X\|_{\mathbb{H}} \le \tau$, which is more likely to be met in most practical applications.

Assumption 2 (Eigenvalue Decay Rate (EDR)) *Suppose the eigenvalue decay rate of C is $\beta > 1$, i.e., there exist constants c_1 and c_2 such that*

$$c_1 j^{-\beta} \le \lambda_j \le c_2 j^{-\beta}, \quad \forall i = 1, 2, \cdots.$$

Note that the eigenvalues λ_j and EDR are only determined by the ICLP covariance C. The polynomial eigenvalue decay rate assumption is standard in the non-parametric literature. For example, if C satisfies the Sacks–Ylvisaker conditions (Sacks and Ylvisaker, 1966, 1968, 1970) of order s, then $\lambda_j \asymp j^{-2(s+1)}$. If setting C equal to the reproducing kernel of the univariate Sobolev space $\mathcal{W}_2^m([0,1])$ results in $\lambda_j \asymp j^{-2m}$, see Micchelli and Wahba (1979) for more instances.

In the following theorems, we derive the closed form of the estimators and provide their global sensitivity analysis for the FRL, ICLP-AR, and ICLP-QR mechanisms.

Theorem 10 (Global Sensitivity Analysis) *Suppose Assumption 1 holds, then*

1. *(FRL) Suppose the functional summary used in FRL is a truncated sample mean function, i.e.,* $\hat{\mu}_D = \sum_{j=1}^M \langle \bar{X}, \phi_j \rangle_{\mathbb{H}} \phi_j$*, then*

$$\Delta = \max_{D,D'} \|\hat{\mu}_D - \hat{\mu}_{D'}\|_{\ell^1} \leq \frac{2M\tau}{n}.$$

2. *(ICLP-AR) The solution of ICLP-AR in (9) is*

$$\hat{\mu}_D = \sum_{j=1}^{\infty} s_{\psi, 2\lambda_j^{\eta/2}} \left(\langle \bar{X}, \phi_j \rangle_{\mathbb{H}} \right) \phi_j, \tag{10}$$

for all $\eta \geq 1$ *and* $s_{a,b}(x) = \operatorname{sgn}(x)\left(|x| - a/b\right)^+$ *is the soft thresholding function with threshold* a/b*. Then, there exists an integer* J^* *such that the global sensitivity of* $\hat{\mu}_D$ *satisfies*

$$\sup_{D \sim D'} \|\hat{\mu}_D - \hat{\mu}_{D'}\|_{1,C} \leq \frac{2\tau}{n} \sum_{j=1}^{J^*} \lambda_j^{-\frac{1}{2}}.$$

3. *(ICLP-QR) The solution of ICLP-QR in (9) is*

$$\hat{\mu}_D = \sum_{j=1}^{\infty} \frac{\lambda_j^{\eta}}{\lambda_j^{\eta} + \psi} \left\langle \bar{X}, \phi_j \right\rangle \phi_j, \tag{11}$$

for all $\eta > 1 + \beta^{-1}$*. Then, the global sensitivity of* $\hat{\mu}_D$ *satisfies*

$$\sup_{D \sim D'} \|\hat{\mu}_D - \hat{\mu}_{D'}\|_{1,C} \leq \frac{2\tau}{n} \sum_{j=1}^{\infty} \left(\frac{\lambda_j^{\eta - \frac{1}{2}}}{\lambda_j^{\eta} + \psi} \right).$$

Remark 11 *The integer* $J^* := \min\{j \geq 1 : \tau \leq \psi / 2\lambda_j^{\eta/2}\}$ *in the ICLP-AR estimator can indeed be viewed as a truncation number as the coefficients after* J^* *will be shrunk to* 0*, i.e., the summation in (10) is indeed finite. The upper bound for global sensitivity is based on the fact that, in the worst-case scenario, the coefficients are not shrunk to zero, and thus, the soft thresholding adjustments are canceled out. Therefore, unfortunately, the ICLP-AR estimator does not produce a better sensitivity than the FRL approach, while the soft thresholding introduces extra bias into the summary. On the other hand, the coefficients of the ICLP-QR estimator (11) will not be shrunk exactly to zero. Hence, one is able to perturb the functional summary with the truly infinite-dimensional ICLP.*

We now analyze the error of the non-private summary $\hat{\mu}_D$ and the private summary $\tilde{\mu}_D$. We call $\mathrm{E}\,\|\hat{\mu}_D - \mu_0\|_{\mathbb{H}}^2$ the estimation error and the quantity $\mathrm{E}\,\|\tilde{\mu}_D - \hat{\mu}_D\|_{\mathbb{H}}^2$ the privacy error. We also call the mean square error (MSE) of $\tilde{\mu}_D$, $\mathrm{E}\,\|\tilde{\mu}_D - \mu_0\|_{\mathbb{H}}^2$, the privacy-estimation error, which represents the amount of error in estimating the population mean by the private summary.

Theorem 12 *Suppose X_i are i.i.d. observations drawn from population X with mean function μ_0, and Assumptions 1 and 2 hold. We also assume $\mu_0 \in \mathcal{H}_{C^\eta}$ for some $\eta > 1 + \beta^{-1}$. Let $\hat{\mu}_D$ be non-private estimators under different approaches and $\tilde{\mu}_D$ as their private version. Then*

$$\mathrm{E}\,\|\tilde{\mu}_D - \mu_0\|_{\mathbb{H}}^2 = \begin{cases} O\left(\dfrac{8M^3\tau^2}{n^2\epsilon^2} + \dfrac{1}{n} + M^{-\eta\beta}\right), & \textit{(FRL)} \\ O\left(\dfrac{4\tau^2}{n^2\epsilon^2}(J^*)^{\beta+2} + \dfrac{1}{n} + \psi\right), & \textit{(ICLP-AR)} \\ O\left(\dfrac{8\tau^2}{n^2\epsilon^2} tr(C^{\eta-1})\psi^{-\frac{2}{\eta}(\frac{\beta+2}{2\beta})} + \dfrac{1}{n} + \psi\right). & \textit{(ICLP-QR)} \end{cases}$$

The first term in the privacy-estimation error is the privacy error, while the last two terms constitute the estimation error. Notably, the optimal estimation error rate for mean function estimation in $\mathbb{H}$ is $O(n^{-1})$. By tuning the regularization parameters in different mechanisms, one can ensure that the privacy error either matches or is of a lower order than the estimation error. This tuning allows the private summary to perform as well as the non-private one in terms of error rate. In the following corollary, we provide the optimal order for these regularization parameters such that the privacy error is $O(n^{-1})$ or $o(n^{-1})$.

Corollary 13 *Under the same conditions as Theorem 12,*

1. *(FRL) Setting $n^{\frac{1}{\eta\beta}} \lesssim M \lesssim n^{\frac{1}{3}}$ leads to*

$$\mathrm{E}\,\|\tilde{\mu}_D - \mu_0\|_{\mathbb{H}}^2 = O(n^{-1}).$$

2. *(ICLP-AR) Setting $(n\epsilon^2)^{-\frac{\eta\beta}{2(\beta+2)}} \lesssim \psi \lesssim n^{-1}$ with $\eta \geq 2(1 + 2\beta^{-1})$ leads to*

$$\mathrm{E}\,\|\tilde{\mu}_D - \mu_0\|_{\mathbb{H}}^2 = O(n^{-1}).$$

In particular, setting $\psi \asymp n^{-1}$ and $\eta > 2(1 + 2\beta^{-1})$,

$$\mathrm{E}\,\|\tilde{\mu}_D - \hat{\mu}_D\|_{\mathbb{H}}^2 = o(n^{-1}).$$

3. *(ICLP-QR) Setting $(n\epsilon^2)^{-\frac{\eta\beta}{\beta+2}} \lesssim \psi \lesssim n^{-1}$ with $\eta \geq 1 + 2\beta^{-1}$ leads to*

$$\mathrm{E}\,\|\tilde{\mu}_D - \mu_0\|_{\mathbb{H}}^2 = O(n^{-1}).$$

In particular, setting $\psi \asymp n^{-1}$ and $\eta > 1 + 2\beta^{-1}$,

$$\mathrm{E}\,\|\tilde{\mu}_D - \hat{\mu}_D\|_{\mathbb{H}}^2 = o(n^{-1}).$$

Corollary 13 suggests that by optimally choosing the finite-dimensional number M in the FRL mechanism or the regularization parameter ψ in ICLP-QR, the privacy error will not dominate the privacy-estimation error, making the privacy-estimation error match the optimal estimation error of $O(n^{-1})$. Additionally, in ICLP-QR, slightly oversmoothing via regularization (choosing $\eta > 1+2\beta^{-1}$) results in the privacy error being a lower order of the estimation error, making it asymptotically negligible. This phenomenon is referred to as "free privacy" since privatizing the summaries will not affect their statistical performance. Similarly, ICLP-AR can also achieve the optimal error rate $O(n^{-1})$ and even gain "free privacy" like ICLP-QR. However, ICLP-AR requires higher oversmoothing to achieve the optimal rate compared to ICLP-QR. This could be attributed to our error analysis using the largest integer J^* (beyond which all coefficients are zero), representing the worst-case scenario. In real-world applications, the actual value of J^* is usually smaller than what we use in our analysis, leading to a performance that might be slightly worse than that of the FRL mechanism, see Section 6. However, it's important to note that the true value of J^* is not analyzable within the current framework.

This section primarily focuses on the mean estimation and privacy protection problem, a common problem in the functional data analysis where the samples X_i are random functions in a function space (Rice and Silverman, 1991; Cai and Yuan, 2011). Although this problem is relatively simple, its approaches and analysis can be methodologically extended to other problems where estimations can be boiled down to estimating the average functional statistic. For example, in functional data analysis, the covariance function estimation is essentially the average of the sample $\{(X_i(t)-\bar{X}(t))(X_i(s)-\bar{X}(s))\}_{i=1}^n$; estimating the coefficient in the function-on-scalar linear regression model is essentially the average of the sample $\{X_iY_i(t)\}_{i=1}^n$. Given the inherently repetitive nature of these problems, we only provide the analysis of the mean protection and refer readers to Appendix C for more detailed discussions on the parallels between these problems.

4.3 Beyond Mean Protection

In this part, we demonstrate how the ICLP mechanism is not only feasible for functional summaries but also can be applied to more general learning problems where the summary of interest is a function.

Kernel Density Estimation. Let $D = \{x_1, \cdots, x_n\} \subseteq T$, where T is a compact set over $\mathbb{R}^d$, be i.i.d. samples from a distribution with density f_0. For any given ICLP covariance kernel K, we adopt the ICLP-QR by picking the density estimation kernel as K^η with $\eta > 1$. For a given $d \times d$ symmetric and positive definite bandwidth matrix $\mathbf{H}$, the kernel density estimator under the ICLP-QR strategy takes the form of

$$\hat{K}_D(x) = \frac{1}{n}\sum_{i=1}^n K^\eta_{\mathbf{H}}(x-x_i) = \frac{1}{n\sqrt{det(\mathbf{H})}}\sum_{i=1}^n K^\eta\left(\mathbf{H}^{-\frac{1}{2}}(x-x_i)\right). \tag{12}$$

We now provide the global sensitivity and utility analysis of $\hat{K}_D(x)$ in the following theorem.

Theorem 14 *Suppose $K^\eta(\cdot,\cdot)$ is point-wise bounded by a constant M_K, then the global sensitivity Δ of $\hat{K}_D(x)$ in (12) satisfies*

$$\Delta = \sup_{D\sim D'} \left\| \hat{K}_D - \hat{K}_{D'} \right\|_{1,K} \le \frac{2M_K}{n\sqrt{det(\mathbf{H})}} \sqrt{tr(K^{\eta-1})}.$$

Furthermore, taking $\mathbf{H}$ to be a diagonal matrix with the same entry, i.e., $\mathbf{H} = h\mathbf{I}$, and assuming $f_0^{''}$ is absolutely continuous, $\int_T (f_0^{'''}(x))^2 dx < \infty$ and $\int_T K^\eta(x) dx = 1$. Then

$$\mathrm{E} \int_T \left(\tilde{f}_D(x) - f_0(x) \right)^2 dx \le O\left(\frac{c_1}{n^2 h^{2d}} + h^4 + \frac{c_2}{nh^d} \right),$$

for some constants c_1 and c_2.

Remark 15 *If h is taken to be $h \asymp n^{\frac{1}{4+d}}$, then $R = O(n^{-\frac{4}{4+d}})$, which matches the optimal kernel density estimation rate (Wasserman, 2006).*

The connection between estimating kernel and the noise kernel also appeared in Hall et al. (2013), where they stated that one could achieve (ϵ, δ)-DP by adding a Gaussian process with its covariance function equal to the kernel used in estimation. For privacy-safe bandwidth, h, we can pick $h \asymp n^{\frac{1}{4+d}}$ to ensure privacy is gained for free. However, a private version of the "rule of thumb", see Rao and Scott (1992) and Hall et al. (2013) is also feasible.

Functionals via Regularized ERM. The functional summaries one desires to release may come from learning algorithms such as regularization-based algorithms. In Section 4.1, we proposed using such algorithms to obtain qualified functional summaries. Here, we generalize the approach to broader scenarios such as non-parametric regression and classification. Let $\mathcal{D} = \{d_1, \cdots, d_n\}$ be the collection of n samples, where d_i is a tuple with finite size. Given a loss function L, we consider the following regularized ERM problem:

$$\hat{f}_D = \operatorname*{argmin}_{f \in \mathcal{H}_{C^\eta}} \left\{ \frac{1}{n} \sum_{i=1}^n L(d_i, f) + \psi \|f\|_{C^\eta} \right\} \quad \text{for some} \quad \eta > 1. \tag{13}$$

When d_i's are couples, i.e., $d_i = (y_i, x_i)$, (13) can be viewed as non-parametric classification (where y_i's take discrete value) or regression (where y_i's take continuous value) problems. The solution of (13) can be expressed as $\hat{f}_D = \sum_{i=1}^n a_i C^\eta(\cdot, d_i)$ by the Representer Theorem (Kimeldorf and Wahba, 1971). However, although the Representer Theorem provides an elegant solution for (13), it is not suitable for calculating the global sensitivity since all the elements in the vector $(a_1, \cdots, a_n)$ change when we swap one individual in the dataset. In the following theorem, we provide a sensitivity analysis for $\hat{f}_D$ under certain regularized conditions.

Theorem 16 *Suppose $\hat{f}_D$ is the solution of (13) and the loss function L in (13) is an M-admissible loss function (Bousquet and Elisseeff, 2002), then the global sensitivity for $\hat{f}_D$ satisfies*

$$\Delta = \sup_{D\sim D'} \left\| \hat{f}_D - \hat{f}_{D'} \right\|_{1,C} \le \frac{M}{\psi n} \sqrt{\sup_x C^\eta(x,x)} \sqrt{tr(C^{\eta-1})}.$$

Remark 17 *One can also prove that the privacy error* $\mathrm{E}\,\|\tilde{f}_D - \hat{f}_D\|_{L^2}^2$ *is bounded by* $c_1(\psi n)^{-2}$*. We do not provide a utility analysis for this case study as the statistical error can vary based on different settings of the problem and is out of the scope of this paper.*

The application scenario is broad since the upper bound for the global sensitivity holds for any convex and locally M-admissible loss function and bounded kernel with finite trace. For example, support vector machines with hinge loss, non-parametric regressions with square loss, and logistic regressions with $log(1+x)$ loss are all applicable settings.

4.4 Privacy-Safe Regularization Parameter Selection

Determining the regularization parameters in different mechanisms, such as the finite-dimensional number M (for subspace embedding mechanisms) and ψ (for the ICLP mechanism), is crucial to ensure the reasonable performance of the private releases. Tuning regularization parameters in statistical modeling has been well studied, and *Cross Validation* (CV), or one of its many variants, is the most widely used approach. However, CV focuses on balancing variance and bias in the estimation error, not the trade-off between privacy and estimation error. To fit CV into the DP framework, *Private Cross Validation* (PCV) was proposed in Mirshani et al. (2019), which aims to find the "sweet spot" between privacy error and estimation error. However, as data-driven approaches, neither CV nor PCV is truly privacy-safe since the regularization parameters may contain information about the data. There are some approaches one can use to obtain end-to-end privacy-guaranteed regularization parameters. For example, one can use out-of-sample public datasets (Zhang et al., 2012) or one can spend extra privacy budget on the tuning process (Chaudhuri et al., 2011; Chaudhuri and Vinterbo, 2013).

Since the ICLP mechanism is tied to a kernel, one can obtain privacy-safe regularization parameters by picking kernels whose eigenvalues decay polynomially, satisfying the conditions in Theorem 12 so that the optimal values for M and ψ in Theorem 12 can be directly used as regularization parameter inputs. We name this approach *"Privacy Safe Selection" (PSS)*. PSS does not degrade the privacy guarantee since the employed regularization parameters rely only on the sample size, the privacy budget, and the additive noise's covariance function. We would like to note that PSS is not a data-driven approach since its acquisition never depends on the data. In practice, the constants for the optimal values in PSS can affect the performance of the ICLP mechanism. In our experiments, we observe that by appropriately normalizing the sample trajectories and the trace of the covariance kernel, setting the constant to 1 usually leads to satisfactory performance.

5. Algorithm and Implementation

Based on the definition of the ICLP, the generic implementation of the mechanism can be achieved using the Karhunen-Loéve expansion.

1. Given any Mercer kernel C, obtain its eigenvalues $\{\lambda_j\}_{j\geq 1}$ and eigenfunctions $\{\phi_j\}_{j\geq 1}$.
2. Generate ICLP noise by $Z = \sum_{j=1}^{\infty} \sqrt{\frac{\lambda_j}{2}} Z_j \phi_j$ where $Z_j \overset{i.i.d.}{\sim} \text{Lap}(1)$.
3. Calibrate Z to desired privacy level by the global sensitivity Δ and privacy budget ϵ.

However, the summation in generating Z cannot be implemented in finite time and usually is terminated at a large integer. Therefore, in practice, we utilize its approximated version, Algorithm 1.

Algorithm 1: Approximated ICLP mechanism

```
Given the covariance kernel $C$ and $K$ different points $(x_1, x_2, \cdots, x_K)$ on the
  compact domain $\mathcal{T}$, calculate the value of $C$ on the grid expanded by
  $(x_1, x_2, \cdots, x_K)$, i.e.,
  $$\hat{C} = \begin{pmatrix} C(x_1, x_1) & \cdots & C(x_1, x_K) \\ \vdots & \ddots & \vdots \\ C(x_K, x_1) & \cdots & C(x_K, x_K) \end{pmatrix}$$
Obtain $K$ estimated eigenvalues $\{\hat{\lambda}_k\}_{k=1}^K$ and eigenfunctions $\{\hat{\phi}_k\}_{k=1}^K$ of $\hat{C}$ by
  eigendecomposition.
for $k$ in 1,2,$\cdots$, $K$ do
    Set $\hat{f}_{Dk} = \langle \hat{f}_D, \hat{\phi}_k \rangle$ and generate $Z_k$ from $\sqrt{\frac{\hat{\lambda}_k}{2}}$ Lap(1)
    $\tilde{f}_{Dk} = \hat{f}_{Dk} + \sigma Z_k$ where $\sigma = \frac{\sqrt{2}\Delta}{\epsilon}$
Return $\tilde{f} = \sum_{k=1}^K \tilde{f}_{Dk} \phi_j$.
```

A natural question about Algorithm 1 is whether all the theoretical analyses still hold if the privacy noise is sampled in a finite approximation manner instead of the "true" infinite sum. The answer is positive as long as the same cutoff K is used both in constructing privacy noise and expressing the original estimate, followed directly by the post-processing inequality. However, a key advantage of our theoretical analyses is that the privacy guarantees will still hold regardless of what K is used. Another problem regarding Algorithm 1 is that even though larger K will lead to more accurate estimates of eigenvalues and eigenfunctions, it also increases the computational burden as the algorithm relies on the Karhunen-Loéve expansion. Next, we investigated how different cutoff values, K, will affect computational time by comparing the average computation time for generating 100 ICLPs to 100 Gaussian Processes. We choose the Gaussian Process as the competitor since it is the stochastic process used to achieve (ϵ, δ)-DP for functional data, and sampling Gaussian processes is nothing more than sampling a multivariate Gaussian with covariance $\hat{C}$. Theoretically, generating one ICLP and one Gaussian process are both in time complexity $O(n^3)$ since both Cholesky decomposition and eigen decomposition are $O(n^3)$. In Table 1, we report the average time to generate 100 ICLPs and 100 Gaussian Processes for different covariance kernels. We found that generating 100 Gaussian Processes is about 30% to 50% faster than generating 100 ICLPs in practice.

6. Experiments

In this section, we numerically evaluate the effectiveness of the ICLP mechanism and other comparable mechanisms, like the FRL and Bernstein mechanisms.

C	K	ICLP	GP	C	K	ICLP	GP
Exponential	100	0.567688	0.273463	Matérn($\nu = \frac{3}{2}$)	100	0.571014	0.270682
	200	2.778597	1.683808		200	2.636642	1.617183
	500	29.98454	20.90210		500	29.38142	20.56870
Gaussian	100	0.553025	0.268379	Matérn($\nu = \frac{5}{2}$)	100	0.551222	0.271477
	200	2.617800	1.610193		200	2.618398	1.615180
	500	29.26284	20.41770		500	29.29077	20.53389

Table 1: Computation time (in seconds) for generating 100 ICLPs and Gaussian Processes under different cutoff values K and covariance kernels C over $[0, 1]$.

6.1 Simulation for Mean Function Protection

In this section, we conduct the simulation for the mean function privacy protection problem discussed in Section 4.2. We use the isotropic Matérn kernel (Cressie and Huang, 1999) as the covariance kernel for the ICLP noise. It takes the form

$$C_\alpha(s,t) = \frac{1}{\Gamma(\nu)2^{\alpha-1}} \left(\frac{\sqrt{2\alpha} d(s,t)}{\rho} \right)^\alpha K_\alpha \left(\frac{\sqrt{2\alpha} d(s,t)}{\rho} \right)$$

where K_α is the modified Bessel function. This is motivated by the fact that the resulting RKHS of C_α ties to a particular Sobolev space, allowing us to control the smoothness directly. Specifically, the RKHS associated with the Matérn kernel C_α over $\mathbb{R}^d$ is norm-equivalent to the Sobolev space $H^{\alpha+\frac{d}{2}}$ and thus the corresponding eigenvalues decay polynomially as $\lambda_j \asymp j^{-2(\alpha+\frac{d}{2})}$. In the following experiments, we set $d = 1$, $\rho = 0.1$, the privacy budget $\epsilon = 1$, and $\alpha = 1.5$ such that $\lambda_j \asymp j^{-4}$. For the functional samples, we consider $\mathbb{H} = L^2([0,1])$ and generate samples as

$$X_i(t) = \mu_0(t) + e_i(t)$$

where $e_i(t)$ are Gaussian processes with zero mean and radial basis function kernel as covariance function. For the true mean function μ_0, we consider the following four different forms:

S-1: $\mu_0(t) = 10t * \exp(-t)$.

S-2: $\mu_0(t) = 0.3 f_{0.3,0.05}(t) + 0.7 * f_{0.8,0.05}(t)$.

S-3: $\mu_0(t) = 0.2 * (f_{0,0.03}(t) + f_{0.2,0.05}(t) + f_{0.5,0.05}(t) - f_{0.75,0.03}(t) + f_{1,0.03}(t))$.

S-4: $\mu_0(t) = \sum_{j=1}^{25} R_{ij}\phi_j(t)$, where $R_{ij} \overset{i.i.d.}{\sim} U[-1,1]$.

Here, $f_{a,b}$ is the probability density function of the normal distribution with mean a and variance b^2. The trajectory complexity of the samples generated from these mean functions rises sequentially. For example, S-1 is a monotonically increasing function, S-2 is a bimodal function, and S-3 and S-4 are functions exhibiting multiple rapid fluctuations. We evaluate these mechanisms via privacy-estimation errors, estimation errors, and privacy errors. These errors are calculated via Monte Carlo by generating 1000 privatized mean estimators.

We also conduct the experiments in different settings. For example, we consider generating the error stochastic process e_i from Gaussian processes with different covariance functions or from basis expansion with heavy-tailed distributions. We also consider the ICLP covariance kernel as the Matérn kernel with $\alpha = 2.5$. We refer to Appendix D for these additional experimental results.

6.1.1 Comparison of PCV and PSS

In Section 4.4, we introduced PSS for parameter selection. Here, we demonstrate its effectiveness by comparing it with the data-driven method PCV. For both the ICLP-AR and ICLP-QR, we set η and ψ_{PSS} to be the values in Theorem 12 such that the privacy error is the same order as the estimation error. For PCV, we obtain ψ_{PCV} by 10-fold PCV within the range of $[0.1\psi_{\text{PSS}}, 10\psi_{\text{PSS}}]$. In the FRL mechanism, the PSS approach is more ambiguous as the PSS values for truncation number such that $\tilde{\mu}$ reaches the optimal rate is a collection of integers, i.e., $\mathcal{M} = \{M_1, \cdots, M_K\}$. We calculate the estimation error for each $M \in \mathcal{M}$ and take the smallest estimation error as the PSS result. For PCV, we consider a wider range of $\mathcal{M}$ by adding and subtracting 3 to its maximum and minimum elements.

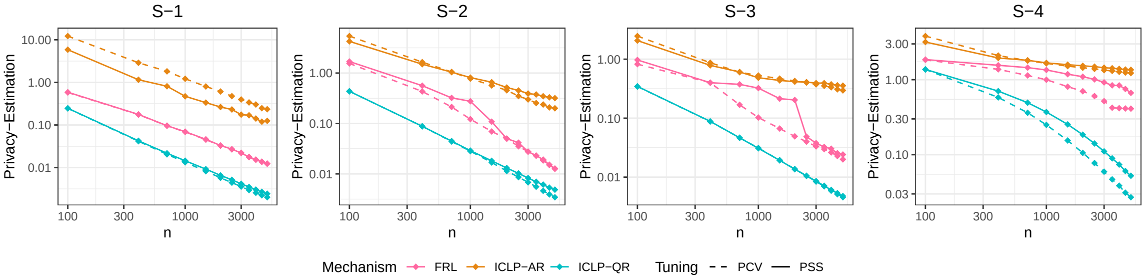


Figure 1: Privacy-estimation error for PSS and PCV approaches to select parameter ψ under different mechanisms, sample size, and true mean functions. Reported values are averages of 100 independent replicated experiments. Both axes are in the base 10 log scale.

In Figure 1, we report the privacy-estimation error for each mechanism under different sample sizes n. For the relatively simple μ_0 in S-1, the error decay curves of the PSS almost line up with the PCV ones for FRL and ICLP-QR. In S-2 and S-3, the ICLP mechanism still has consistent error decay curves between the PSS and PCV, while the PSS curve of the FRL mechanism exhibits a step-down pattern. This pattern occurs because the maximum number in $\mathcal{M}$ is determined by the sample size. Therefore, with complex curves and small sample sizes, the FRL mechanism does not have enough components to estimate the mean function well. As the sample size increases, the availability of more components improves the estimation. In S-4, PCV performs slightly better than the PSS approach for both mechanisms. This is expected since the FRL mechanism requires more dimensions, ICLP-QR requires less regularization, and PCV consistently behaves this way. Since it has been shown that selecting regularization parameters via the PSS approach provides a reasonable and consistent performance compared to PCV, we use the PSS approach in the following experiments to be fully privacy-safe.

6.1.2 Comparison of Different Mechanisms

Under the same settings, we compare the performance of different mechanisms under different sample sizes n, which include the FRL, ICLP-AR, ICLP-QR, and the Bernstein mechanisms[1]. We also include the Gaussian mechanism for achieving (ϵ, δ)-DP (with $\epsilon = 1$ and $\delta = 0.01$) on functional summaries via Gaussian process (Mirshani et al., 2019), and we refer to it as GP-ADP. This provides insight into what is gained by moving from ϵ-DP to (ϵ, δ)-DP.

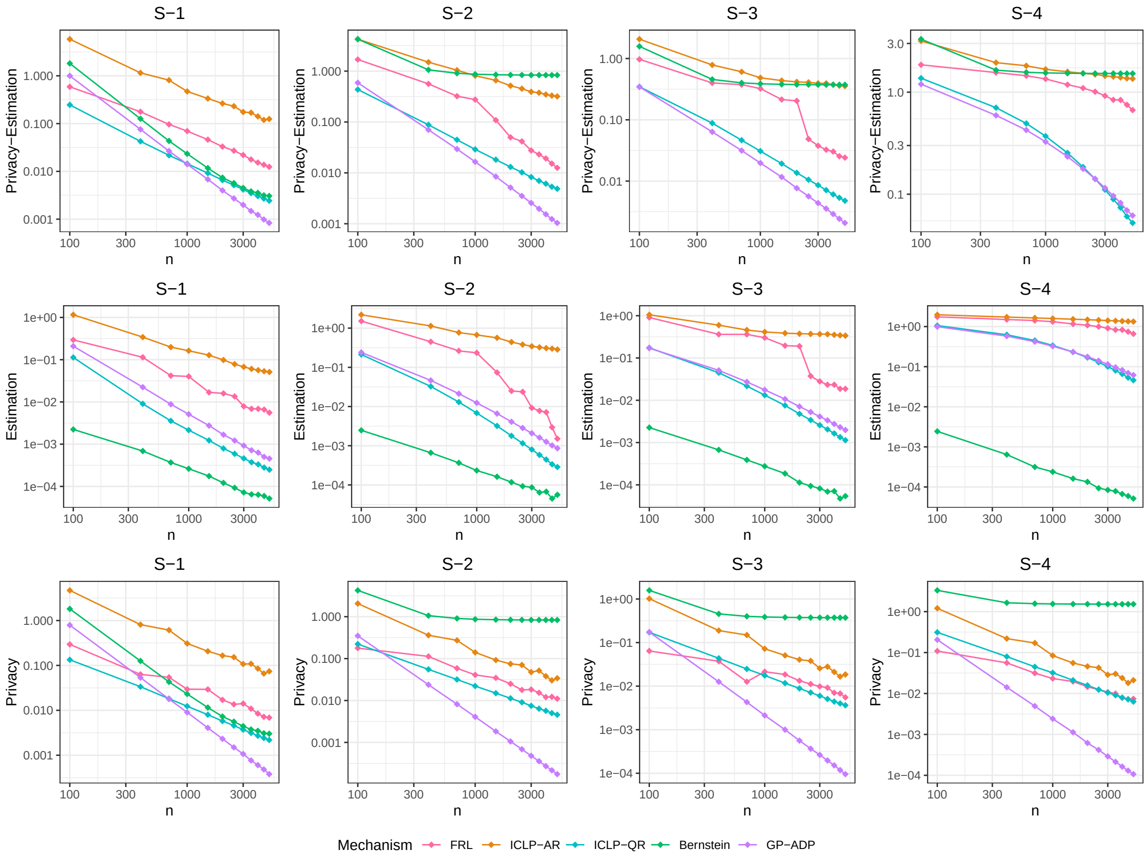


Figure 2: Privacy-estimation, estimation, and privacy errors for different mechanisms under different sample sizes n and true mean functions. The values reported are averages over 100 independent replicated experiments. Both axes are in the base 10 log scale.

Figure 2 illustrates the error decay results of different mechanisms as μ_0 and n are varied. Focusing on the privacy-estimation error, the ICLP-QR consistently outperforms all other ϵ-DP mechanisms across various scenarios. This indicates that the ICLP-QR is effective in releasing privatized summaries with high utility. On the other hand, GP-ADP performs slightly better than the ICLP-QR, especially when n is larger. This is not surprising, as moving to a more relaxed privacy paradigm can lead to private summaries with higher

[1] The implementation of the Bernstein mechanism is based on R package `diffpriv`. We use the sample mean $\bar{X}$ as the non-private summary by setting the cover size parameter as 20.

utility. On the other hand, the ICLP-AR exhibits the worst performance, primarily due to its high estimation error. This poor performance is likely caused by the exact bias introduced by the soft thresholding function, which would require an extremely large sample size to reduce the threshold and eliminate the bias. The FRL and Bernstein mechanisms are close to the ICLP-QR in S-1, showing their effectiveness in simpler mean function scenarios, but they fail to mimic the behaviors of the ICLP-QR when mean functions become more complex, i.e., S-2 to S-4.

Regarding the trade-off between estimation and privacy error, the ICLP-QR mechanism exhibits similar privacy error to the FLR mechanism while significantly outperforming it in estimation error. This confirms the ICLP mechanism's advantages from two perspectives. First, by treating the summary as infinite-dimensional, the ICLP mechanism provides better non-private estimations than the finite-dimensional FLR mechanism, which suffers from worse estimation errors due to fewer components involved. Second, despite adding noise to infinite dimensions, the ICLP mechanism only requires a similar amount of noise as the FLR mechanism, indicating it achieves more effective noise injection by treating different dimensions heterogeneously.

6.2 Simulation for Kernel Density Estimator Protection

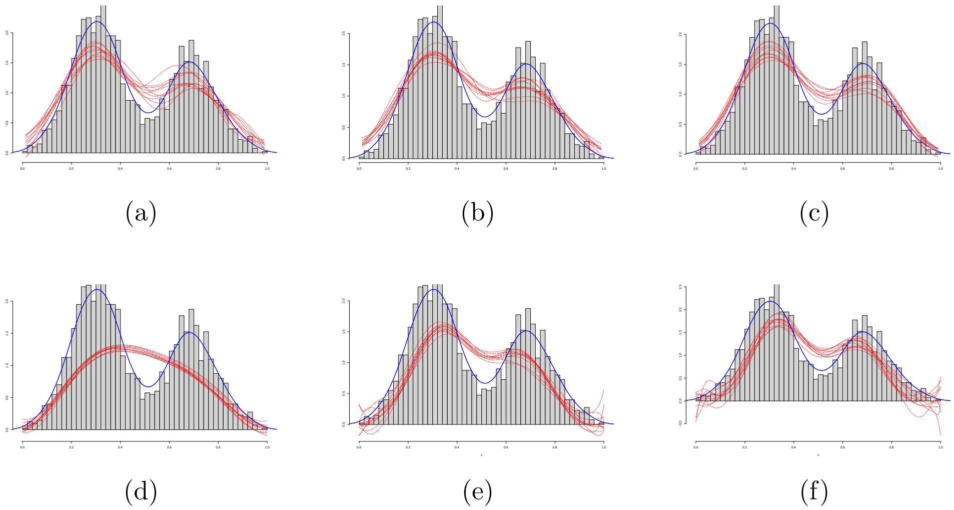
To demonstrate the wide range of application scenarios of the ICLP mechanism, we conduct simulations on kernel density estimations. We consider the setting under $\mathbb{R}$ and $\mathbb{R}^2$ with samples generated from two mixture Gaussian distributions.

1. $\mathbb{R}$ setting:
$$x_i \overset{i.i.d.}{\sim} \sum_{i=1}^{2} p_i \mathcal{N}(\mu_i, 0.1; 0, 1),$$
where $\mathcal{N}(\mu, \sigma; a, b)$ is a truncated normal distribution over $[a, b]$ with $p_1 = 0.6$, $p_2 = 0.4$, $\mu_1 = 0.3$, $\mu_2 = 0.7$.

2. $\mathbb{R}^2$ setting:
$$\mathbf{x}_i \overset{i.i.d.}{\sim} \sum_{i=1}^{2} p_i \mathcal{N}\left(\mu_i, \left(\begin{smallmatrix} 1 & 0.5 \\ 0.5 & 1 \end{smallmatrix}\right); \left(\begin{smallmatrix} 5 \\ -5 \end{smallmatrix}\right), \left(\begin{smallmatrix} 5 \\ -5 \end{smallmatrix}\right)\right),$$
where $\mathcal{N}(\mu, \Sigma; \mathbf{a}, \mathbf{b})$ is a multivariate truncated normal distribution over $[a_1, b_1] \times [a_2, b_2]$ with $p_1 = 0.6$, $p_2 = 0.4$, $\mu_1 = (-3, -3)$, $\mu_2 = (3, 2)$.

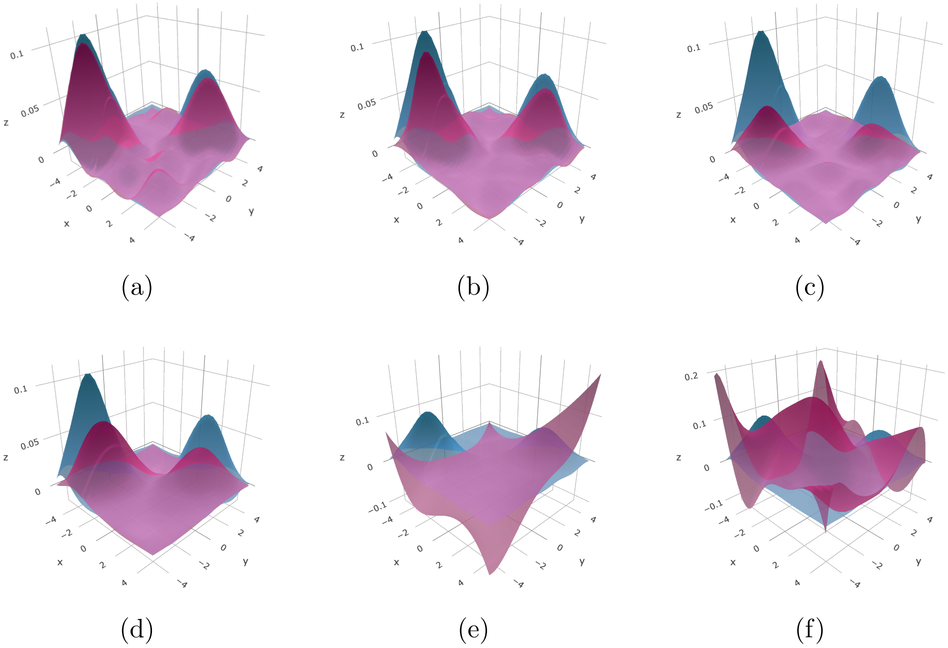
We compare the FRL, ICLP-QR, and Bernstein mechanisms. For the ICLP-QR and the Bernstein mechanism, we pick multiple smoothing parameters η and lattice numbers to demonstrate how they affect private curves and surfaces. For the FRL mechanism, we select the truncated number that provides the best fit under PCV criteria. We use the Gaussian kernel in $\mathbb{R}$ and the exponential kernel in $\mathbb{R}^2$ to build the kernel density estimator. We use $h \asymp n^{1/(4+d)}$ where $d = 1, 2$ to ensure we gain privacy for free and remain privacy safe. The results are reported in Figure 3 and Figure 4.

For the univariate setting, the ICLP-QR performs similarly to the FRL mechanism; a higher η produces less variability in the curves but tends to be over-smoothed. The Bernstein mechanism needs over 30 lattice points in the interval to capture the bimodal pattern but results in producing a messy tail at both ends. A lower lattice number produces better tails but fails to capture the bimodal pattern.


(a) (b) (c)

(d) (e) (f)

Figure 3: Non-private (Blue) and 10 random realization of private (Red) KDEs. (a) ICLP mechanism with $\eta = 1.25$ (b) ICLP mechanism with $\eta = 1.5$ (c) FRL and (d)-(f) Bernstein mechanism with lattice number equal to $10, 30, 50$.



(a) (b) (c)

(d) (e) (f)

Figure 4: 3D plot of non-private (Blue) and private (Red) KDEs over $\mathbb{R}^2$. (a)-(c) ICLP mechanism with $\eta = 1.01, 1.05, 1.2$ (d) FRL and (e)-(f) Bernstein mechanism with lattice number equal to $5, 10$.

For the multivariate setting, one can see that by slightly oversmoothing, the ICLP-QR produces privatized KDEs that are very close to the non-private ones. A smaller η (Figure 4a) is more precise at peaks but will be "noisy" around lower density regions, while a larger η (Figure 4c) produces smooth lower density regions but causes underestimation at peaks. Figure 4b shows there is a clear "sweet point" to trade off the smoothness and underestimation. The FRL mechanism performs similarly to the underestimated ICLP case, but the peaks of the privatized KDE do not fully align with the non-private one. On the other hand, the Bernstein mechanism fails to produce surfaces similar to those of the non-private estimator, even when we increase the number of lattice points.

7. Real Data Applications

This section presents two real data applications of the proposed methods to study the release of functional summaries for different functional data datasets.

7.1 Application on Medical and Energy Usage Functional Data

This application aims to release a private mean function that satisfies ϵ-DP for the following two functional data datasets. The first dataset is the Brain scans Diffusion Tensor Imaging (DTI) dataset[2]. The DTI dataset provides fractional anisotropy (FA) tract profiles for the corpus callosum (CCA) of the right corticospinal tract (RCST) for patients with multiple sclerosis and for controls. Specifically, we study the CCA dataset, which includes 382 patients measured at 93 equally spaced locations of the CCA. The second dataset contains historical electricity demand in Adelaide[3]. The dataset consists of half-hourly electricity demands from Sunday to Saturday in Adelaide between July 6, 1997, and March 31, 2007. Our analysis focuses on Monday specifically, meaning the dataset consists of measurements from 508 days at 48 equally spaced time points.

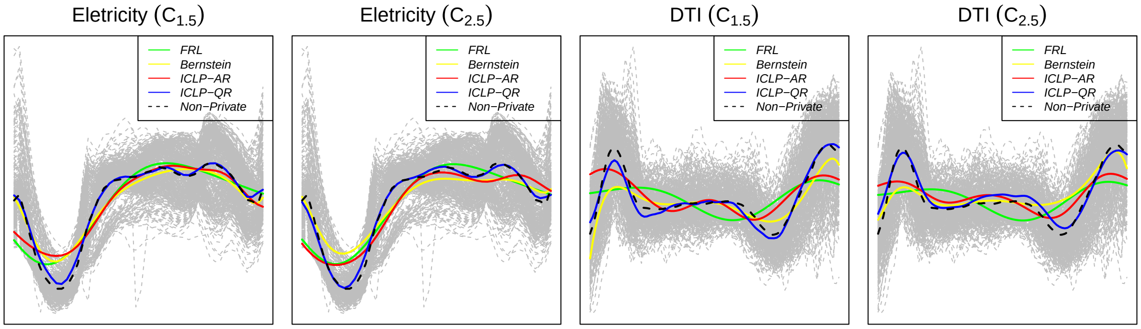


Figure 5: Non-private and private mean functions for different mechanisms with $\epsilon = 1$. The curves in light grey indicate the original samples.

[2] Available in the R package `refund`.

[3] Available in the R package `fds`.

	Eletricity Demand							
	$C_{1.5}$				$C_{2.5}$			
ϵ	FRL	ICLP-AR	ICLP-QR	Bernstein	FRL	ICLP-AR	ICLP-QR	Bernstein
1/8	$0.2828_{0.1981}$	$2.2110_{1.5374}$	$3.9110_{2.9464}$	$1.3726_{1.2935}$	$0.2805_{0.1645}$	$2.2279_{1.5880}$	$4.1614_{3.5764}$	$1.4261_{1.3547}$
1/4	$0.1731_{0.0419}$	$0.6140_{0.3859}$	$1.0391_{0.8066}$	$0.3459_{0.2783}$	$0.1703_{0.0376}$	$0.6053_{0.4227}$	$1.1168_{0.8931}$	$0.3468_{0.2963}$
1/2	$0.1453_{0.0098}$	$0.2117_{0.1276}$	$0.3415_{0.1897}$	$0.0895_{0.0686}$	$0.1427_{0.0094}$	$0.2092_{0.0954}$	$0.3497_{0.2276}$	$0.0920_{0.0692}$
1	$0.1383_{0.0025}$	$0.1106_{0.0352}$	$0.1649_{0.0728}$	$0.0241_{0.0171}$	$0.1360_{0.0026}$	$0.1109_{0.0374}$	$0.1639_{0.0655}$	$0.0251_{0.0168}$
2	$0.1366_{0.0007}$	$0.0858_{0.0151}$	$0.1204_{0.0238}$	$0.0080_{0.0044}$	$0.1342_{0.0006}$	$0.0857_{0.0167}$	$0.1164_{0.0233}$	$0.0089_{0.0049}$
4	$0.1362_{0.0002}$	$0.0793_{0.0074}$	$0.1089_{0.0105}$	$0.0039_{0.0011}$	$0.1338_{0.0002}$	$0.0794_{0.0072}$	$0.1044_{0.0090}$	$0.0048_{0.0012}$

	DTI							
	$C_{1.5}$				$C_{2.5}$			
ϵ	FRL	ICLP-AR	ICLP-QR	Bernstein	FRL	ICLP-AR	ICLP-QR	Bernstein
2	$0.6305_{0.2878}$	$6.2473_{4.2907}$	$4.4245_{3.9617}$	$3.1996_{2.6794}$	$0.6303_{0.2867}$	$6.2155_{4.8556}$	$4.7453_{4.0892}$	$3.0841_{2.7095}$
3	$0.4418_{0.0663}$	$1.6122_{1.0886}$	$1.2691_{1.1059}$	$0.8039_{0.6994}$	$0.4437_{0.0762}$	$1.6275_{1.1514}$	$1.3049_{0.9318}$	$0.7820_{0.7225}$
4	$0.3952_{0.0165}$	$0.4929_{0.3046}$	$0.4701_{0.2522}$	$0.2040_{0.1627}$	$0.3974_{0.0158}$	$0.4907_{0.3274}$	$0.4793_{0.2431}$	$0.2028_{0.1790}$
5	$0.3836_{0.0045}$	$0.2080_{0.0824}$	$0.2747_{0.0756}$	$0.0555_{0.0369}$	$0.3859_{0.0040}$	$0.2073_{0.0890}$	$0.2748_{0.0804}$	$0.0581_{0.0395}$
6	$0.3809_{0.0010}$	$0.1361_{0.0285}$	$0.2265_{0.0265}$	$0.0186_{0.0106}$	$0.3830_{0.0010}$	$0.1366_{0.0287}$	$0.2231_{0.0325}$	$0.0216_{0.0107}$
7	$40.3801_{0.0003}$	$0.1187_{0.0133}$	$0.2143_{0.0133}$	$0.0094_{0.0028}$	$0.3823_{0.0003}$	$0.1187_{0.0137}$	$0.2102_{0.0126}$	$0.0126_{0.0029}$

Table 2: Expected L^2-distance between the private mean function and the sample mean for both electricity demand and DTI(cca) datasets. The numbers in the subscript indicate the standard error ($\times 10^{-3}$).

Since the true mean function is not available in real data applications, we evaluate the performance of each mechanism using the expected L^2-distance between the private summary, $\tilde{\mu}$, and the non-private sample means, $\hat{\mu}$, i.e., $\mathrm{E}\,\|\tilde{\mu}-\hat{\mu}\|_{L^2}^2$. We consider the Matérn kernel with $\alpha = 1.5$ and 2.5 and $\rho = 0.1$. The expected L^2-distance is approximated using Monte Carlo with 1000 generated $\tilde{\mu}$. The results are reported in Table 2, with each value being an average of 100 replicate experiments. We also plot one private mean estimator for each mechanism in Figure 5.

From Table 2, it can be observed that the expected L^2-distance decreases similarly as the privacy budget increases for both datasets. One can see that the expected L^2-distance of the FRL soon stops changing, indicating that most of the errors of the FRL are concentrated on estimation errors. This indicates that in order to avoid adding too much noise to the later components, the FRL mechanism has to compromise on using fewer leading components, leading to higher estimation errors. This can also be seen in Figure 5, where the FRL mechanism only produces a privatized mean function that estimates the overall shape but fails to capture local details. The ICLP-AR and the Bernstein mechanisms have similar performance patterns and much worse results with small privacy budgets. Finally, the ICLP-QR performs the best among all approaches as its privatized mean functions can estimate the shapes precisely and have much smaller expected L^2-distances compared to the non-private mean.

7.2 Application on Human Mortality Data

Publishing the entire age-at-death distribution for a given country or region usually provides more comprehensive information about human lifespan and health status than publishing crude mortality rates. A private version of this distributional summary ensures that an attacker cannot infer information about individuals or groups in a particular age range.

The mortality data for each region are collected from the United Nations World Population Prospects 2019 Databases[4]. The dataset records the number of deaths for each region and age. The goal of this application is to release private mortality distributions for various regions.

	Mechanisms				
Region	ICLP-QR	ICLP-QR	FRL	Bernstein	Bernstein
	$(\eta = 1.01)$	$(\eta = 1.05)$		$(K = 10)$	$(K = 20)$
Eastern Africa	$\mathbf{1.685}_{10.507}$	$5.249_{7.549}$	$5.252_{7.881}$	$3.823_{12.243}$	$2.400_{11.896}$
Middle Africa	$\mathbf{1.122}_{0.451}$	$5.281_{0.410}$	$4.870_{2.152}$	$3.969_{0.313}$	$2.090_{0.327}$
Northern Africa	$\mathbf{2.398}_{0.680}$	$8.873_{0.656}$	$3.991_{1.807}$	$11.315_{0.546}$	$5.710_{0.646}$
Southern Africa	$2.487_{1.491}$	$6.427_{0.919}$	$2.990_{1.809}$	$3.964_{0.666}$	$\mathbf{2.395}_{0.763}$
Western Africa	$\mathbf{1.588}_{13.699}$	$5.775_{9.216}$	$5.752_{11.482}$	$4.347_{15.649}$	$2.609_{15.493}$
Central Asia	$\mathbf{5.804}_{2.715}$	$11.829_{1.547}$	$7.319_{3.739}$	$13.163_{1.346}$	$8.116_{1.808}$
Eastern Asia	$\mathbf{3.551}_{4.414}$	$13.446_{6.872}$	$3.900_{6.958}$	$15.878_{14.110}$	$8.162_{9.175}$
Southern Asia	$\mathbf{2.040}_{2.267}$	$8.619_{3.737}$	$2.968_{2.962}$	$10.044_{8.099}$	$4.962_{4.888}$
South-Eastern Asia	$\mathbf{2.259}_{1.871}$	$9.498_{3.565}$	$2.566_{2.323}$	$10.428_{8.328}$	$5.157_{4.778}$
Western Asia	$\mathbf{2.195}_{0.585}$	$8.689_{0.624}$	$3.768_{2.005}$	$9.380_{0.510}$	$4.640_{0.578}$
Eastern Europe	$\mathbf{3.990}_{3.658}$	$13.501_{4.924}$	$4.817_{5.762}$	$15.542_{9.356}$	$8.438_{6.414}$
Northern Europe	$7.318_{1.626}$	$19.087_{1.187}$	$\mathbf{7.174}_{1.493}$	$22.846_{1.218}$	$13.339_{1.304}$
Southern Europe	$7.457_{0.977}$	$21.076_{0.931}$	$\mathbf{7.137}_{1.083}$	$27.500_{0.852}$	$15.873_{0.927}$
Western Europe	$6.819_{2.184}$	$19.959_{2.782}$	$\mathbf{6.445}_{3.417}$	$25.606_{5.375}$	$14.760_{3.965}$
Caribbean	$6.453_{2.890}$	$11.247_{2.022}$	$\mathbf{5.981}_{3.634}$	$10.714_{1.580}$	$7.231_{2.296}$
Central America	$1.636_{0.548}$	$7.777_{0.685}$	$\mathbf{1.146}_{0.537}$	$5.637_{0.480}$	$2.616_{0.448}$
South America	$\mathbf{2.244}_{2.498}$	$9.461_{4.096}$	$2.886_{3.276}$	$9.294_{6.744}$	$4.700_{4.333}$
Northern America	$2.205_{1.734}$	$10.605_{2.792}$	$\mathbf{1.537}_{2.297}$	$10.028_{5.546}$	$4.746_{3.153}$
Australia/New Zealand	$23.525_{11.299}$	$29.495_{5.040}$	$\mathbf{22.782}_{8.189}$	$28.339_{5.053}$	$23.983_{7.582}$

Table 3: Expected L^2-distance between release KDEs and non-private KDEs for each region with $\epsilon = 1$. The numbers in the subscript indicate the standard error ($\times 10^{-3}$).

We estimate the probability density function for each region and privatize the estimates via the ICLP mechanism and its competitors. The privacy budget is set to 1. We evaluate the performance using $\mathrm{E}\,\|\tilde{f} - \hat{f}\|_{L^2}^2$ where $\tilde{f}$ is the private KDE and $\hat{f}$ is the non-private counterpart. The expectation is approximated using Monte Carlo using 200 private KDEs. The results are reported in Table 3. We also visualized the private KDEs for each region and mechanism in Appendix E.

From Table 3, the ICLP-QR with $\eta = 1.01$ has smaller errors in developing regions, while the FRL mechanism performs better in developed regions. This is reasonable, as developed regions usually have better medical conditions, making the mortality age concentrate between 70 and 80 and their densities unimodal. Thus, a few leading components are sufficient to represent the density function in these regions. Conversely, the situation is different in developing regions, where the infant mortality rates are higher, making their densities multimodal and requiring more components for better estimation.

[4] Available at `https://population.un.org/wpp/Download`.

8. Conclusion

In this paper, we introduce a new mechanism, the ICLP mechanism, to achieve ϵ-DP for infinite-dimensional functional summaries. This mechanism offers a wide range of output privacy protections with more flexible data assumptions and a more effective noise injection process compared to current mechanisms that rely on finite-dimensional embedding. We establish its feasibility in separable Hilbert spaces and spaces of continuous functions. Different approaches are proposed for constructing qualified summaries compatible with the ICLP mechanism, along with parameter selection via PSS, to guarantee end-to-end protection. We also demonstrate that one can balance utility and privacy by controlling the degree of regularization in these strategies. This is demonstrated in the mean protection example by showing that slightly over-smoothing the summary allows the private summary to achieve the optimal rate for non-private mean estimation.

Despite its advantages, the ICLP mechanism has some limitations and presents opportunities for future research. As discussed in Section 5, implementing the ICLP mechanism relies on the Karhunen-Loéve expansion, which can be computationally expensive. Developing a computational approach that does not depend on the Karhunen-Loéve expansion is an important future direction. Additionally, although various experimental results indicate that by appropriately processing the sample trajectories and the ICLP covariance kernel, omitting the constant for PSS values can produce satisfactory performance, we believe a more careful investigation of the constant can further enhance performance.

Acknowledgments

This work was partially supported by the National Science Foundation, NSF SES-1853209.

Appendix A. Proofs for Section 3

A.1 Proof of Theorem 4

Proof Let X be a random element defined via decomposition (3) with covariance operator C, where C belongs to the trace class. To prove that X is well-defined in $\mathbb{H}$, we only need to show $\mathrm{E}\langle X, X\rangle_{\mathbb{H}} < \infty$. Note that

$$\mathrm{E}\langle X, X\rangle_{\mathbb{H}} = \mathrm{E}\sum_{j=1}^{\infty}\sum_{k\geq 1}\sqrt{\lambda_j}\sqrt{\lambda_k}Z_jZ_k\langle\phi_j,\phi_k\rangle_{\mathbb{H}} = \mathrm{E}\left[\sum_{j=1}^{\infty}\lambda_j Z_j^2\right].$$

Since Z_j are i.i.d. Laplace random variables with zero mean and variance 1, we have

$$\mathrm{E}\left[\lambda_j Z_j^2\right] = \sum_{j=1}^{\infty}\lambda_j = \mathrm{Tr}(C) < \infty.$$

Then, by the Fubini theorem,

$$\mathrm{E}\langle X, X\rangle_{\mathbb{H}} = \mathrm{E}\left[\sum_{j=1}^{\infty}\lambda_j Z_j^2\right] = \sum_{j=1}^{\infty}\mathrm{E}\left[\lambda_j Z_j^2\right] < \infty,$$

which proves the existence of X. ■

A.2 Proof of Theorem 5

Proof We first define an isometric isomorphism between the Hilbert space $\mathbb{H}$ and the ℓ^2 space to avoid considering probability measures over $\mathbb{H}$. Given an orthonormal basis $\{\phi_j\}_{j=1}^{\infty}$ of $\mathbb{H}$, define a mapping $\mathcal{T} : \mathbb{H} \to \ell^2$ by $\mathcal{T}(f) = \{\langle f,\phi_j\rangle\}_{j=1}^{\infty}$. The inverse of $\mathcal{T}$ is $\mathcal{T}^{-1}(\{\langle f,\phi_j\rangle)_{j=1}^{\infty}\} = \sum_{j=1}^{\infty}\langle f,\phi_j\rangle\phi_j$ and $\mathcal{T}$ preserves the norms, i.e., $\|\mathcal{T}(f)\|_{\ell^2} = \|f\|_{\mathbb{H}}$. Thus, this mapping is an isometric isomorphism between $\mathbb{H}$ and ℓ^2, and we can consider the probability measure over ℓ^2 rather than over $\mathbb{H}$.

For a Laplace random variable $X \sim \mathrm{Lap}(\mu, b)$ over $\mathbb{R}$, it induces a probability measure over $(\mathbb{R}, \mathcal{B})$, where $\mathcal{B}$ is the Borel set over $\mathbb{R}$, as

$$\gamma_{\mu,b}(dx) = \frac{1}{2b}\exp\left(-\frac{|x-\mu|}{b}\right)dx.$$

Denote $\{\lambda_j\}_{j\geq 1}$ and $\{\phi_j\}_{j\geq 1}$ as the eigenvalues and eigenfunctions of the covariance operator C, respectively, and let $\lambda_j = 2b_j^2$. By the Existence of Product Measures Theorem (Tao, 2011), $P_D \circ \mathcal{T}$ is a unique probability measure defined as $\gamma(f_D, C) := \prod_{j=1}^{\infty}\gamma_{f_{Dj},b_j}$ over $(\mathbb{R}^{\infty}, \mathcal{B}^{\infty}) := (\prod_{j=1}^{\infty}\mathbb{R}_j, \prod_{j=1}^{\infty}\mathcal{B}_j)$. We further restrict $\gamma(f_D, C)$ on $(\ell^2, \sigma(\ell^2))$ and keep denoting it by $\gamma(f_D, C)$. Since $\mathcal{T}$ is an isometric isomorphism mapping between $\mathcal{H}$ and ℓ^2 space, to prove P_D and $P_{D'}$ are equivalent, it is sufficient to prove that the product probability measures $\gamma(f_D, C)$ and $\gamma(f_{D'}, C)$ are equivalent. Now, we prove Theorem 5 by showing $\gamma(h, C)$ and $\gamma(0, C)$ are equivalent if and only if $h \in \mathcal{H}_C$.

For the "if" part, we apply Kakutani's theorem (Kakutani, 1948). The two measures are equivalent if the following sum converges

$$\sum_{j=1}^{\infty} \log \int_R \sqrt{\frac{\gamma_j(h_j, b_j)}{\gamma_j(0, b_j)}} \gamma_j(0, b_j)(dx).$$

This leads to the applicable space for h as

$$\mathcal{H}_C^* = \left\{ h \in \mathbb{H} : \sum_{j=1}^{\infty} \left[\frac{|h_j|}{2b_j} - log(1 + \frac{|h_j|}{2b_j}) \right] < \infty \right\}.$$

To prove $\mathcal{H}_C^* = \mathcal{H}_C$, we only need to show that for a non-negative sequence $\{a_j\}_{j\geq 1}$, the series $\sum_{j=1}^{\infty}[a_n - log(1 + a_n)]$ converges if and only if $\sum_{j=1}^{2} a_n^2$ converges. Let $f(x) = x - log(1+x)$ and $g(x) = x^2$. Besides, note that $\lim_{n\to\infty} f(a_n) = 0$ if and only if $\lim_{n\to\infty} a_n = 0$. Thus,

- If $\sum_{j=1}^{\infty}[a_n - log(1 + a_n)] < \infty$, then $\lim_{n\to\infty} f(a_n) = 0$ and so $\lim_{n\to\infty} a_n = 0$. Then,

  $$\lim_{n\to\infty} \frac{f(a_n)}{g(a_n)} = \lim_{n\to\infty} \frac{a_n - log(1 + a_n)}{a_n^2} = \frac{1}{2}$$

  therefore by the limit comparison test $\sum_{j=1}^{\infty} a_n^2$ converges too.

- If $\sum_{j=1}^{\infty} a_n^2 < \infty$, by the same statement as above also holds and therefore $\sum_{j=1}^{\infty}[a_n - log(1 + a_n)] < \infty$.

For the "only if" part, the proof is the same as Theorem 2 in Reimherr and Awan (2019a). ■

A.3 Proof of Theorem 6

Proof We prove Theorem 6 by showing the form of Radon-Nikodym derivative of $\{P_D : D \in \mathcal{D}\}$ with respect to P_0 is the same as derivative of $\gamma(h, C)$ with respect to $\gamma(0, C)$, and it takes the form of

$$\frac{dP_h}{dP_0}(z) = \exp\left\{ -\frac{1}{\sigma} \left(\|z - h\|_{1,C} - \|z\|_{1,C} \right) \right\}. \tag{14}$$

First, we need to show that the right-hand side of Equation (14) is well-defined when $h \in \mathcal{H}_{1,C}$. Define

$$H_M(z) = \sum_{j=1}^{M} \frac{|z_j - h_j| - |z_j|}{b_j} \quad \textit{and} \quad H(z) = \lim_{M\to\infty} H_M(z).$$

We show there exists a set A with $P_0(A) = 1$ such that $H(z)$ exists and is finite on A. Suppose $z_j \sim \mathrm{Lap}(0, b_j)$, then

$$\begin{aligned}
\mathrm{Var}\left(H_M(z)\right) &= \sum_{j=1}^{M} \mathrm{Var}\left(\frac{|z_j - h_j| - |z_j|}{b_j}\right) \\
&= \sum_{j=1}^{M} 3b_j^2 - b_j^2 \exp\left\{-\frac{2|h_j|}{b_j}\right\} - \exp\left\{-\frac{|h_j|}{b_j}\right\}(b_j^2 - 4b_j|h_j|) \\
&= \sum_{j=1}^{M} -b_j^2\left(1 + \exp\left\{-\frac{|h_j|}{b_j}\right\}\right)^2 + 4b_j^2\left(1 + \frac{|h_j|}{b_j}\exp\left\{-\frac{|h_j|}{b_j}\right\}\right) \\
&= \sum_{j=1}^{M} b_j^2\left(4 - \left(1 + \exp\left\{-\frac{|h_j|}{b_j}\right\}\right)^2\right) + 4b_j^2\frac{|h_j|}{b_j}\exp\left\{-\frac{|h_j|}{b_j}\right\}.
\end{aligned}$$

The second equality is based on the following facts that $E|z_j - h_j| = |h_j| + \exp\{-|h_j|/b_j\}$, $\mathrm{E}\,|z_j| = b_j$ and $\mathrm{Cov}(|z_j - h_j|, |z_j|) = |h_j|b_j + \exp\{-|h_j|/b_j\}\,(b_j^2 + b_j|h_j|/2)$. By Fatou's Lemma and the condition $h \in \mathcal{H}_{1,C}$, we have $\mathrm{Var}(H(z)) < \infty$. The set A is Ω and with P_0-measure 1. Therefore, if $h \in \mathcal{H}_{1,C}$, the right-hand side of Equation (14) exists and is well-defined.

Second, we show that (14) is the Radon-Nikodym derivative of $h + \sigma Z$ with respect to σZ. Define

$$g(x) = \exp\left\{-\frac{\sqrt{2}}{\sigma}\left(\|x - h\|_{1,C} - \|x\|_{1,C}\right)\right\} \quad \text{and} \quad dP_h^*(x) = g(x)dP_0(x).$$

Then, we only need to show that P_h and P_h^* are the same probability measure. We accomplish this by showing they have the same moment-generating function.

$$\begin{aligned}
MGF_{P_h}(t) &= \mathrm{E}_{P_h}\exp\left\{\langle X, t\rangle_{\mathbb{H}}\right\} \\
&= \prod_{j=1}^{\infty}\int_{\mathbb{R}}\exp\left\{x_j t_j\right\} d\gamma_{h_j,b_j}(x_j) \\
&= \prod_{j=1}^{\infty}\frac{\exp\left\{h_j t_j\right\}}{1 - (b_j t_j)^2}.
\end{aligned}$$

where the second inequality comes from the result that P_h is product measure of γ_{h_j,b_j}. For the moment generate function of P_h^*,

$$\begin{aligned} MGF_{P_h^*}(t) &= \mathrm{E}_{P_h^*} \exp\left\{\langle X, t\rangle_{\mathbb{H}}\right\} \\ &= \mathrm{E}_Q\, g(X) \exp\left\{\langle X, t\rangle_{\mathbb{H}}\right\} \\ &= \prod_{j=1}^{\infty} \int_{\mathbb{R}} \exp\left\{-\frac{1}{\sigma}\left(\frac{|x_j - h_j| - |x_j|}{b_j}\right) + x_j t_j\right\} d\gamma_{0,b_j}(x_j) \\ &= \prod_{j=1}^{\infty} \int_{\mathbb{R}} \exp\left\{x_j t_j\right\} d\gamma_{h_j,b_j}(x_j) \\ &= \prod_{j=1}^{\infty} \frac{\exp\left\{h_j t_j\right\}}{1 - (b_j t_j)^2} = MGF_{P_h}(h). \end{aligned}$$

Therefore, P_h and P_h^* have the same moment-generating functions and thus are the same probability measure. ■

A.4 Proof of Theorem 7

Proof We prove the theorem via contradiction. Assume if $f_D \in \mathcal{H}_C \backslash \mathcal{H}_{1,C}$, for any given fixed ϵ, there exists a $\sigma \in \mathbb{R}^+$ such that the ICLP mechanism release $f_D + \sigma Z$ still satisfies ϵ-DP. Then, by the post-processing property of differential privacy, for any transformation $G : \mathcal{H} \to \mathcal{H}$, $G(f_D)$ is also ϵ-DP. Now, for any $J \in \mathbb{N}$, consider G to be a projection mapping into first J components, i.e., $G_J(f_D) = \sum_{j=1}^{J} \langle f_D, \phi_j\rangle \phi_j$. Therefore, for any $J \in \mathbb{N}$ and $f_D \in \mathcal{H}_C \backslash \mathcal{H}_{1,C}$, $G_J(f_D)$ is ϵ-DP, i.e.,

$$\exp\left\{-\frac{\sqrt{2}}{\sigma} \sum_{j=1}^{J} \left(\frac{|\langle f_D - z, \phi_j\rangle|}{\sqrt{2} b_j} - \frac{|\langle z, \phi_j\rangle|}{\sqrt{2} b_j}\right)\right\} \le \exp\left\{\epsilon\right\},$$

except for $z \in A$ where A is zero-measure set.

Define $B_j = \{z_j : |z_j| > |\langle f_D, \phi_j\rangle|\}$ and $S_J = \left\{z \in \ell^2 : z_j \in B_j, \forall 1 \le j \le J \ and \ z_j \in \mathbb{R}, \forall j > J\right\}$. Then for all $z \in S_J$,

$$\exp\left\{-\frac{\sqrt{2}}{\sigma} \sum_{j=1}^{J} \left(\frac{|\langle f_D - z, \phi_j\rangle|}{\sqrt{2} b_j} - \frac{|\langle z, \phi_j\rangle|}{\sqrt{2} b_j}\right)\right\} = \exp\left\{\frac{\sqrt{2}}{\sigma} \sum_{j=1}^{J} \left(\frac{|\langle f_D, \phi_j\rangle|}{\sqrt{2} b_j}\right)\right\} \le \exp\left\{\epsilon\right\}.$$

However, since $h \in \mathcal{H}_C \backslash \mathcal{H}_{1,C}$, for any given fixed ϵ, one can always find an J so that

$$\exp\left\{\frac{1}{\sigma} \sum_{j=1}^{J} \frac{|\langle f_D, \phi_j\rangle|}{b_j}\right\} > \exp\left\{\epsilon\right\}$$

and therefore contradiction holds and no such $\sigma \in \mathbb{R}^+$ exists.

The remaining thing is to prove that S_J is not a zero-measure set. By the Existence of Product Measure in Tao (2011),

$$\gamma_{0,C}(S_J) = \prod_{j=1}^{J} \gamma_{0,\sqrt{2}b_j}(B_j)$$

where the right-hand side of the equation is greater than 0 by definition of B_j. ■

A.5 Proof of Theorem 8

Proof By Theorem 6, the density of $\tilde{f}_D$ with respect to to σZ is

$$\frac{dP_D}{dP_0}(z) = \exp\left\{-\frac{1}{\sigma}\left(\|z - f_D\|_{1,C} - \|z\|_{1,C}\right)\right\}.$$

We aim to show that for any measurable subset $A \subseteq \mathbb{H}$, one has

$$P_D(A) \leq e^{\epsilon} P_{D'}(A),$$

which is equivalent to show

$$P_D(A) = \int_A dP_D(x) = \int_A \frac{dP_D}{dP_{D'}}(x) dP_{D'}(x) \leq e^{\epsilon} \int_A dP_{D'}(x).$$

Notice

$$\begin{aligned}\frac{dP_D}{dP_{D'}}(x) &= \frac{dP_D}{dP_0}(x) / \frac{dP_{D'}}{dP_0}(x) \\ &= \exp\left\{-\frac{1}{\sigma}\left(\|x - f_D\|_{1,C} - \|x - f_{D'}\|_{1,C}\right)\right\} \\ &\leq \exp\left\{\frac{1}{\sigma}\|f_{D'} - f_D\|_{1,C}\right\}.\end{aligned}$$

Recall the global sensitivity for the ICLP mechanism is defined as

$$\Delta = \sup_{D \sim D'} \|f_D - f_{D'}\|_{1,C}.$$

By taking $\sigma = \frac{\Delta}{\epsilon}$, we have $\frac{dP_D}{dP_{D'}}(x) \leq e^{\epsilon}$, $\forall x \in \mathbb{H}$. Thus, the desired inequality holds, i.e.,

$$P_D(A) = \int_A \frac{dP_D}{dP_{D'}}(x) dP_{D'}(x) \leq e^{\epsilon} \int_A dP_{D'}(x).$$

■

A.6 Proof of Theorem 9

Proof First, we decompose the ICLP as $Z(t) - Z(s) = \sum \lambda_j^{1/2} Z_j(\phi_j(t) - \phi_j(s))$, which leads to

$$\mathrm{E}[\exp\{t(Z(t)-Z(s))\}] = \prod_{j=1}^{\infty} \frac{1}{1 - t^2\lambda_j(\phi_j(t) - \phi_j(s))^2} = \exp\left\{ -\sum_{j=1}^{\infty} \log\left(1 - \frac{t^2\lambda_j}{2}(\phi_j(t) - \phi_j(s))^2\right)\right\}$$

with t satisfies $0 \le t^2\lambda_j(\phi_j(t) - \phi_j(s))^2 < 1$ for all j. Let $C_t = C(t, \cdot)$ for any $t \in T$, then by the α-Hölder continuous property of C, we have

$$\begin{aligned}\lambda_j(\phi_j(t) - \phi_j(s))^2 &= \lambda_j \langle C_t - C_s, \phi_j\rangle_C^2 \\ &\le \langle C_t - C_s, C_t - C_s\rangle_C \\ &= C(t,t) - 2C(t,s) + C(s,s) \\ &\le 2M_C|t-s|^{\alpha},\end{aligned}$$

where M_C is the Hölder-continuous constant, and this leads to

$$0 \le t \le (\frac{1}{M_C})^{\frac{1}{2}}|t-s|^{-\frac{\alpha}{2}}.$$

For $x \in [0, 1)$, define function $f(x) = -\log(1-x)$, by the mean value theorem, we have

$$-\log(1-x) = f(x) = f(0) + xf'(\zeta) = \frac{x}{1-\zeta} \le \frac{x}{1-x},$$

for some $\zeta \in (0, x)$. Therefore, applying the above inequality, we have

$$\begin{aligned}-\log(1 - \frac{t^2\lambda_j}{2}(\phi_j(t) - \phi_j(s))^2) &\le \frac{\frac{t^2\lambda_j}{2}(\phi_j(t) - \phi_j(s))^2}{1 - \frac{t^2\lambda_j}{2}(\phi_j(t) - \phi_j(s))^2} \\ &\le \frac{t^2\lambda_j}{2}(\phi_j(t) - \phi_j(s))^2 \max_k \left\{(1 - \frac{t^2\lambda_k}{2}(\phi_k(t) - \phi_k(s))^2)^{-1}\right\}.\end{aligned}$$

Choosing t such that $(1 - \frac{t^2\lambda_k}{2}(\phi_k(t) - \phi_k(s))^2)^{-1} \le M_0 < \infty$, we have

$$\begin{aligned}\mathrm{E}[\exp\{t(Z(t) - Z(s))\}] &\le \exp\left\{M_0\frac{t^2}{2}\sum \lambda_j(\phi_j(t) - \phi_j(s))^2\right\} \\ &= \exp\left\{\frac{M_0 t^2}{2}(C(t,t) - 2C(t,s) + C(s,s))\right\} \\ &\le \exp\left\{M_0 M_C t^2 |t-s|^{\alpha}\right\}.\end{aligned}$$

By the Chernoff bound, we have

$$\begin{aligned}P\left(|Z(t) - Z(s)| \ge a\right) &\le \frac{\mathrm{E}[\exp\{t(Z(t) - Z(s))\}]}{\exp\{ta\}} \\ &\le \exp\left\{M_0 M_C t^2|t-s|^{\alpha} - ta\right\}.\end{aligned}$$

The minimizer of the right-hand with respect to t is $t_0 = \frac{a}{2M_0 M_C |t-s|^\alpha}$. With restriction of t, we get $a \le 2M_0|t-s|^{\frac{1}{2}\alpha}$.

We consider the following two cases:

Case 1 : Suppose $a \le 2M_0|t-s|^{\frac{1}{2}\alpha}$, the minimizer is $t_0 = \frac{a}{2M_0 M_C |t-s|^\alpha}$. We have

$$P\left(|Z(t) - Z(s)| \ge a\right) \le \exp\left\{-\tilde{M_1} a^2 |t-s|^{-\alpha}\right\},$$

for some generic constant $\tilde{M_1}$. Define $a(x) = C|x|^\beta$ with $\beta \ge \frac{1}{2}\alpha$, and two series as

$$\sum_{n=1}^{\infty} a(2^{-n}) = \sum_{n=1}^{\infty} 2^{-n\beta} \quad \textit{and} \quad \sum_{n=1}^{\infty} 2^n \exp\{-\tilde{M_2}|2^n|^{\alpha-2\beta}\}.$$

The first series converges if $\beta < 1$. However, to make the second one converge, we need $\alpha > 2\beta$, which leads to $\alpha > \alpha$ contradiction.

Case 2 : Suppose $a > 2M_0|t-s|^{\frac{1}{2}\alpha}$, then the minimizer is $t_0 = (\frac{1}{M_C})^{\frac{1}{2}}|t-s|^{-\frac{1}{2}\alpha}$, then

$$\exp\left\{M_0 M_C t^2 |t-s|^\alpha - ta\right\} = \exp\left\{M_0 - \left(\frac{1}{M_C}\right)^{\frac{1}{2}} |t-s|^{-\frac{1}{2}\alpha} a\right\}.$$

By picking function $a(x) = 2M_0|x|^\beta > 2M_0|x|^{\frac{1}{2}\alpha}$, with $\beta \in (0, \frac{1}{2}\alpha)$, we have $\sum_{j=1}^{\infty} a(2^{-n}) < \infty$ and

$$\begin{aligned}\sum_{j=1}^{\infty} 2^n b(2^{-n}) &:= \sum_{j=1}^{\infty} 2^n \exp\left\{M_0 - \left(\frac{1}{M_C}\right)^{\frac{1}{2}} |2^{-n}|^{-\frac{1}{2}\alpha} a(2^{-n})\right\} \\ &= \sum_{j=1}^{\infty} 2^n \exp\left\{M_0 - \left(\frac{1}{M_C}\right)^{\frac{1}{2}} |2^n|^{\frac{1}{2}\alpha - \beta}\right\}.\end{aligned}$$

Since $\beta \in (0, \frac{1}{2}\alpha)$, we have $\sum_{j=1}^{\infty} 2^n b(2^{-n}) < \infty$.

Therefore, combining the results from both cases, we have

$$\sum_{n=1}^{\infty} 2^n P\left(|Z(t+2^{-n}) - Z(t)| \ge 2^{-n}\right) < \infty.$$

Finally, the rest of the proof follows the same proof of Theorem 5.2.8 in Lunardi et al. (2015). ■

Appendix B. Proofs for Section 4

B.1 Proof of Theorem 10

Proof For the FRL, ICLP-AR, and ICLP-QR mechanisms, we first derive their exact solution and then conduct the global sensitivity analysis. For two n sample adjacent datasets D and D', we assume they only differ in the first observation, i.e., X_1 and X_1'.

FRL: For FRL, the estimator can be expressed as

$$\hat{f}_D = \sum_{j=1}^{M} \langle \bar{X}, \phi_j \rangle_{\mathbb{H}} \phi_j$$

Its private version can be obtained by applying the multivariable version of the Laplace mechanism to the coefficients $\hat{f}_{D,1:M} = (\langle \bar{X}, \phi_1 \rangle_{\mathbb{H}}, \cdots, \langle \bar{X}, \phi_M \rangle_{\mathbb{H}})$, i.e.,

$$\hat{f}_D = \sum_{j=1}^{M} \left\{ \langle \bar{X}, \phi_j \rangle_{\mathbb{H}} + Z_j \right\} \phi_j, \quad \text{with} \quad Z_j \overset{i.i.d.}{\sim} Lap\left(0, \frac{\Delta}{\epsilon}\right)$$

where $\Delta = \sup_{D \sim D'} \|\hat{f}_{D,1:M} - \hat{f}_{D',1:M}\|_{\ell^1}$. Given the bounded norm in Assumption 1,

$$\Delta = \sup_{D \sim D'} \|\hat{f}_{D,1:M} - \hat{f}_{D',1:M}\|_{\ell^1} \leq \frac{1}{n} \sum_{j=1}^{M} |\langle X_1 - X_1', \phi_j \rangle_{\mathbb{H}}| \leq \frac{2M\tau}{n}.$$

ICLP-AR: To obtain the closed form of the ICLP-AR estimator, we expand $X_i - \theta$ by the eigenfunctions ϕ_j, i.e.

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^{n} \|X_i - \theta\|_{\mathbb{H}}^2 + \psi \|\theta\|_{1,C^{\eta_l}} &= \frac{1}{n} \sum_{i=1}^{n} \left\| \sum_{j=1}^{\infty} \langle X_i - \theta, \phi_j \rangle_{\mathbb{H}} \phi_j \right\|_{\mathbb{H}}^2 + \psi \|\theta\|_{1,C^{\eta_l}} \\ &= \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{\infty} \langle X_i - \theta, \phi_j \rangle_{\mathbb{H}}^2 + \psi \|\theta\|_{1,C^{\eta_l}} \\ &= \sum_{j=1}^{\infty} \left\{ \frac{1}{n} \sum_{i=1}^{n} (X_{ij} - \theta_j)^2 + \psi \frac{|\theta_j|}{\lambda_j^{\eta_l/2}} \right\} \end{aligned} \tag{15}$$

Solving the minimization problem within the bracket, then for each j, we have

$$\hat{\theta}_j = \text{sgn}(\bar{X}_j) \left(\bar{X}_j - \frac{\psi}{2\lambda_j^{\eta_l/2}} \right)^+.$$

This leads to the ICLP-AR estimator as

$$\hat{\mu}_D^l = \sum_{j=1}^{\infty} \hat{\theta}_j \phi_j = \sum_{j=1}^{\infty} s_{\psi, 2\lambda_j^{\eta_l/2}} \left(\langle \bar{X}, \phi_j \rangle_{\mathbb{H}} \right) \phi_j.$$

For the global sensitivity,

$$\begin{aligned} \sup_{D \sim D'} \|\hat{\mu}_D^l - \hat{\mu}_{D'}^l\|_{1,C} &= \sup_{D \sim D'} \sum_{j=1}^{J_\tau} \frac{|s_{\psi, \lambda_j^{\eta_l/2}} \left(\langle \bar{X}_D, \phi_j \rangle_{\mathbb{H}} \right) - s_{\psi, \lambda_j^{\eta_l/2}} \left(\langle \bar{X}_{D'}, \phi_j \rangle_{\mathbb{H}} \right)|}{\lambda_j^{\frac{1}{2}}} \\ &\leq \sup_{D \sim D'} \sum_{j=1}^{J_\tau} \frac{|\langle \bar{X}_D - \bar{X}_{D'}, \phi_j \rangle_{\mathbb{H}}|}{\lambda_j^{\frac{1}{2}}} \leq \frac{2\tau}{n} \sum_{j=1}^{J_\tau} \frac{1}{\lambda_j^{\frac{1}{2}}}, \end{aligned}$$

where the first inequality is based on the fact that, in the worst case, the j-th coefficients based on D and D' will not be shrunk to 0 simultaneously and thus should have the same sensitivity without soft-threshold function.

ICLP-QR: Recall the object function

$$F(\theta) = \frac{1}{n}\sum_{i=1}^{n} \|X_i - \theta\|_{\mathbb{H}}^2 + \psi\|\theta\|_{C^{\eta_r}}^2.$$

and after dropping everything not involving θ, we have

$$\begin{aligned} F(\theta) &= -2\langle \bar{X}, \theta\rangle_{\mathbb{H}} + \langle \theta, \theta\rangle_{\mathbb{H}} + \psi\langle \theta, \theta\rangle_{C^{\eta_r}} \\ &= -2\langle \bar{X}, C^{\eta_r}\theta\rangle_{C^{\eta_r}} + \langle \theta, C^{\eta_r}\theta\rangle_{C^{\eta_r}} + \psi\langle \theta, \theta\rangle_{C^{\eta_r}}. \end{aligned}$$

The second equality is based on Hilbert space's own dual, i.e., $\langle\cdot,\cdot\rangle_{\mathbb{H}} = \langle \cdot, C(\cdot)\rangle_{\mathcal{H}_C}$. Thus the minimizer of the $F(\theta)$ is

$$\begin{aligned} \hat{\mu}_D^r &= \left(C^{\eta_r} + \psi I\right)^{-1} C^{\eta_r}(\bar{X}) \\ &= \sum_{j=1}^{\infty} \frac{\lambda_j^{\eta_r}}{\lambda_j^{\eta_r} + \psi} \left\langle \bar{X}, \phi_j \right\rangle_{\mathbb{H}} \phi_j \end{aligned}$$

where the second equality follow by expansion $\hat{\mu}_D^r$ under the eigenfunction ϕ_j. For the global sensitivity, the upper bound for $\sup_{D\sim D'} \|\hat{\mu}_D^r - \hat{\mu}_{D'}^r\|_{1,C}$ is

$$\begin{aligned} \sup_{D\sim D'} \|\hat{\mu}_D - \hat{\mu}_{D'}\|_{1,C} &= \sup_{D\sim D'} \sum_{j=1}^{\infty} \frac{\lambda_j^{\eta_r - \frac{1}{2}}}{\lambda_j^{\eta_r} + \psi} \left|\left\langle \bar{X} - \bar{X}', \phi_j \right\rangle_{\mathbb{H}}\right| \\ &\leq \frac{2\tau}{n} \sum_{j=1}^{\infty} \frac{\lambda_j^{\eta_r - \frac{1}{2}}}{\lambda_j^{\eta_r} + \psi}, \end{aligned}$$

where the inequality is based on Cauchy-Schwarz inequality and the $\|X_i\|_{\mathbb{H}} \leq \tau$. ■

B.2 Proof of Theorem 12

Proof

FRL: For the privacy error,

$$\mathrm{E}\,\|\tilde{\mu} - \hat{\mu}\|_{\mathbb{H}}^2 = \frac{2\Delta^2}{\epsilon^2} \cdot M \leq \frac{8\tau^2}{\epsilon^2 n^2} M^3.$$

For the estimation error,

$$\begin{aligned} \mathrm{E}\,\|\hat{\mu} - \mu_0\|_{\mathbb{H}}^2 &= \frac{1}{n}\sum_{j=1}^{M} \mathrm{E}\langle X_1 - \mu_0, \phi_j\rangle_{\mathbb{H}}^2 + \sum_{j=M+1}^{\infty} \langle \mu_0, \phi_j\rangle_{\mathbb{H}}^2 \\ &\lesssim \frac{1}{n} + \sum_{j=M+1}^{\infty} \langle \mu_0, \phi_j\rangle_{\mathbb{H}}^2 \\ &\lesssim \frac{1}{n} + \lambda_M^{\eta} \|\mu_0\|_{C^{\eta}}^2 \\ &\lesssim \frac{1}{n} + M^{-\eta\beta}. \end{aligned}$$

The first inequality is based on Assumption 1, while the second one is based on the fact that $\mu_0 \in \mathcal{H}_{C^{\eta}}$.

ICLP-AR: For the privacy error,

$$\begin{aligned} \mathrm{E}\left\|\tilde{\mu}-\hat{\mu}\right\|_{\mathbb{H}}^{2} &= \frac{\Delta^2}{\epsilon^2}\sum_{j=1}^{J^*}\lambda_j \\ &\lesssim \frac{4\tau^2}{\epsilon^2 n^2}\left(\sum_{j=1}^{J^*} j^{\frac{\beta}{2}}\right)^{1}\left(\sum_{j=1}^{\infty} j^{-\beta}\right) \\ &\lesssim \frac{1}{\epsilon^2 n^2}(J^*)^{\beta+2} \end{aligned}$$

Next, we turn to estimation error. Define $\mu_{0,\psi} = \sum_{j=1}^{\infty} f_{\psi,2\lambda_j^{\frac{\eta_l}{2}}}\left(\langle \mu_0, \phi_j\rangle\right)\phi_j$, by triangular inequality

$$\mathrm{E}\left\|\hat{\mu}-\mu_0\right\|_{\mathbb{H}}^{2} \leq \mathrm{E}\left\|\hat{\mu}-\mu_{0,\psi}\right\|_{\mathbb{H}}^{2} + \left\|\mu_{0,\psi}-\mu_0\right\|_{\mathbb{H}}^{2}.$$

For the bias term, let $A = \left\{j : |\mu_j| \geq \frac{\psi}{2\lambda_j^{\eta_l/2}}\right\}$, then

$$\left\|\mu_{0,\psi}-\mu_0\right\|_{\mathbb{H}}^{2} = \sum_{A}\left(\mu_{0j} - f_{\psi,2\lambda_j^{\frac{\eta_l}{2}}}(\mu_{0j})\right)^2 + \sum_{A^c}\left(\mu_{0j} - f_{\psi,2\lambda_j^{\frac{\eta_l}{2}}}(\mu_{0j})\right)^2.$$

Starting with the summation over A, since $\frac{\lambda_j^{-\frac{\eta_l}{2}}}{2} < \frac{|\mu_{0j}|}{\psi}$ we have

$$\sum_{A}\left(\mu_{0j} - f_{\psi,2\lambda_j^{\frac{\eta_l}{2}}}(\mu_{0j})\right)^2 = \sum_{A}\frac{\psi^2}{4\lambda_j^{\eta_l}} \leq \psi\sum_{A}\frac{|\mu_{0j}|}{2\lambda_j^{\frac{\eta_l}{2}}} \leq \frac{\psi}{2}\left\|\mu_0\right\|_{1,C^{\eta_l}}.$$

Turning to summation over A^c, since $|\mu_{0j}| \leq \frac{\psi}{2\lambda_j^{\frac{\eta_l}{2}}}$,

$$\sum_{A^c}\left(\mu_{0j} - f_{\psi,2\lambda_j^{\frac{\eta_l}{2}}}(\mu_{0j})\right)^2 = \sum_{A^c}\mu_{0j}^2 \leq \psi\sum_{A^c}\frac{|\mu_{0j}|}{2\lambda_j^{\frac{\eta_l}{2}}} \leq \frac{\psi}{2}\left\|\mu_0\right\|_{1,C^{\eta_l}}.$$

Therefore, the overall bias is bounded by

$$\left\|\mu_{0,\psi}-\mu_0\right\|_{\mathbb{H}}^{2} \leq \frac{\psi}{2}\left\|\mu_0\right\|_{1,C^{\eta_l}}.$$

Now consider variance term $\mathrm{E}\left\|\hat{\mu}-\mu_{0,\psi}\right\|_{\mathbb{H}}^{2}$,

$$\left\|\hat{\mu}-\mu_0\right\|_{\mathbb{H}}^{2} = \sum_{j=1}^{\infty}\left(f_{\psi,2\lambda_j^{\frac{\eta_l}{2}}}\left(\bar{X}_j\right) - f_{\psi,2\lambda_j^{\frac{\eta_l}{2}}}(\mu_{0j})\right)^2.$$

Similar to the bias part, the summation can be decomposed into the sum of four disjoint pieces

$$\begin{aligned}A_{0,0} &= \{|\bar{X}_j| \leq \psi/2\lambda_j^{\frac{\eta_l}{2}}, |\mu_j| \leq \psi/2\lambda_j^{\frac{\eta_l}{2}}\},\\ A_{0,1} &= \{|\bar{X}_j| \leq \psi/2\lambda_j^{\frac{\eta_l}{2}}, |\mu_j| > \psi/2\lambda_j^{\frac{\eta_l}{2}}\},\\ A_{1,0} &= \{|\bar{X}_j| > \psi/2\lambda_j^{\frac{\eta_l}{2}}, |\mu_j| \leq \psi/2\lambda_j^{\frac{\eta_l}{2}}\},\\ A_{1,1} &= \{|\bar{X}_j| > \psi/2\lambda_j^{\frac{\eta_l}{2}}, |\mu_j| > \psi/2\lambda_j^{\frac{\eta_l}{2}}\}.\end{aligned}$$

When $j \in A_{0,0}$, the summation is zero. Consider $j \in A_{0,1}$, since $|\bar{X}_j| \leq \frac{\psi}{2\lambda_j^{\frac{\eta_l}{2}}}$,

$$\left(f_{\psi, 2\lambda_j^{\frac{\eta_l}{2}}}(\bar{X}_j) - f_{\psi, 2\lambda_j^{\frac{\eta_l}{2}}}(\mu_{0j})\right)^2 = \left(f_{\psi, 2\lambda_j^{\frac{\eta_l}{2}}}(\mu_{0j})\right)^2 = \left(\mu_{0j} - \operatorname{sgn}(\mu_{0j})\frac{\psi}{2\lambda_j^{\frac{\eta_l}{2}}}\right)^2 \leq \left(\mu_{0j} - \bar{X}_j\right)^2.$$

By symmetry, we get the same bound over $A_{1,0}$. So lastly we consider summation over $A_{1,1}$ For $j \in A_{1,1}$ we have

$$\left(f_{\psi, 2\lambda_j^{\frac{\eta_l}{2}}}(\bar{X}_j) - f_{\psi, 2\lambda_j^{\frac{\eta_l}{2}}}(\mu_{0j})\right)^2 = \left(\mu_{0j} - \bar{X}_j - \left(\operatorname{sgn}(\mu_{0j}) - \operatorname{sgn}(\bar{X}_j)\right)\frac{\psi}{2\lambda_j^{\frac{\eta_l}{2}}}\right)^2.$$

If both μ_{0j} and $\bar{X}_j$ have the same sign, then this is just $(\mu_{0j} - \bar{X}_j)^2$. If they have opposite signs, then we have

$$\left|\left(\operatorname{sgn}(\mu_{0j}) - \operatorname{sgn}(\bar{X}_j)\right)\frac{\psi}{2\lambda_j^{\frac{\eta_l}{2}}}\right| \leq \left|\mu_{0j} - \bar{X}_j\right|.$$

Therefore,

$$\left(f_{\psi, 2\lambda_j^{\frac{\eta_l}{2}}}(\bar{X}_j) - f_{\psi, 2\lambda_j^{\frac{\eta_l}{2}}}(\mu_{0j})\right)^2 \leq 4\left(\mu_{0j} - \bar{X}_j\right)^2, \quad \textit{for} \quad j \in A_{1,1}.$$

Finally, the overall variance term is bounded by

$$\mathrm{E}\,\|\hat{\mu} - \mu_0\|_{\mathbb{H}}^2 \leq 4\,\mathrm{E}\left\|\bar{X} - \mu_0\right\|_{\mathbb{H}}^2 \leq \frac{4}{n}\,\mathrm{E}\,\|X_1\|_{\mathbb{H}}^2 \lesssim n^{-1}.$$

ICLP-QR: Recall the global sensitivity,

$$\sup_{D \sim D'} \|\hat{\mu}_D - \hat{\mu}_{D'}\|_{1,C} \leq \frac{2\tau}{n}\sum_{j=1}^{\infty}\frac{\lambda_j^{\eta_r - \frac{1}{2}}}{\lambda_j^{\eta_r} + \psi}.$$

For the summation term, observe that given $\eta > \frac{1}{2} + \frac{1}{\beta}$

$$\begin{aligned}
\sum_{j=1}^{\infty} \frac{\lambda_j^{\eta_r - \frac{1}{2}}}{\lambda_j^{\eta_r} + \psi} &\leq \int_0^{\infty} \frac{C_1 x^{\frac{\beta}{2}}}{C_2 + x^{\eta\beta}} dx \cdot \psi^{-\frac{1}{\eta}\left(\frac{1}{\beta}+\frac{1}{2}\right)} \\
&\leq \left\{ \int_0^{M} \frac{C_1 x^{\frac{\beta}{2}}}{C_2 + x^{\eta\beta}} dx + \int_M^{\infty} x^{\frac{\beta}{2} - \eta\beta} dx \right\} \cdot \psi^{-\frac{1}{\eta}\left(\frac{1}{\beta}+\frac{1}{2}\right)} \\
&\lesssim \psi^{-\frac{1}{\eta}\left(\frac{1}{\beta}+\frac{1}{2}\right)},
\end{aligned}$$

where the first inequality is based on Assumption 2 and change of variable. Therefore, for privacy error, we have

$$\mathrm{E}\, \|\tilde{\mu} - \hat{\mu}\|_{\mathbb{H}}^2 = \frac{\Delta^2}{\epsilon^2} \sum_{j=1}^{\infty} \lambda_j \lesssim (n\epsilon)^{-2} \cdot \psi^{-\frac{2}{\eta}\left(\frac{1}{\beta}+\frac{1}{2}\right)}.$$

For estimation error, the n^{-1} part comes from variance while for bias, we have

$$\|\,\mathrm{E}\,\hat{\mu} - \mu_0\|_{\mathbb{H}}^2 = \sum_{j=1}^{\infty} \left(\frac{\psi}{\lambda_j^{\eta_r} + \psi} \right)^2 \langle \mu_0, \phi_j \rangle^2 \leq \psi \|\mu_0\|_{C^{\eta_r}}^2 \lesssim \psi,$$

where the last inequality is by assuming $\|\mu_0\|_{C^{\eta_r}} < \infty$. Combining privacy error and estimation error, one gets the desired results. ■

B.3 Proof of Theorem 14

Proof Recall that the exact form of the kernel density estimator is

$$\hat{K}_D(x) = \frac{1}{n\sqrt{det(\mathbf{H})}} \sum_{i=1}^{n} K^{\eta}\left(\mathbf{H}^{-\frac{1}{2}}(x - x_i) \right).$$

Then, by the definition of global sensitivity, we have

$$\begin{aligned}
\Delta = \sup_{D \sim D'} \left\| \hat{K}_D - \hat{K}_{D'} \right\|_{1,K} &\leq \frac{1}{n\sqrt{det(\mathbf{H})}} \left\| K^{\eta}(\mathbf{H}^{-\frac{1}{2}} x_n) - K^{\eta}(\mathbf{H}^{-\frac{1}{2}} x_n^{'}) \right\|_{K^{\eta}} \sqrt{tr(K^{\eta-1})} \\
&\leq \frac{1}{n\sqrt{det(\mathbf{H})}} \sqrt{tr(K^{\eta-1})} \sqrt{2\left(K^{\eta}(0) - K^{\eta}(\mathbf{H}^{-\frac{1}{2}}(x_n - x_n^{'})) \right)} \\
&\leq \frac{2M_K}{n\sqrt{det(\mathbf{H})}} \sqrt{tr(K^{\eta-1})}.
\end{aligned}$$

The first inequality is based on the Cauchy–Schwarz inequality, which is also used in deriving the ICLP-QR strategy. The last inequality holds by the assumption that $K^{\eta}(\cdot,\cdot)$ is pointwise bounded.

Turning to the utility, taking **H** to be a diagonal matrix with the same entry, then by the assumptions stated in the Theorem 14 and by the Theorem 6.28 in Wasserman (2006), the risk R satisfies

$$\begin{aligned}
R &= \mathrm{E}\int_T \left(\tilde{f}_D(x) - f_0(x)\right)^2 dx \\
&\leq 2 * \left(\mathrm{E}\int_T \left(\tilde{f}_D(x) - \hat{f}_D(x)\right)^2 dx + \mathrm{E}\int_T \left(\hat{f}_D(x) - f_0(x)\right)^2 dx\right) \\
&\leq O\left(\frac{c_1}{n^2h^{2d}} + h^4 + \frac{c_2}{nh^d}\right).
\end{aligned}$$

for some constants c_1 and c_2. ■

B.4 Proof of Theorem 16

Proof Recall while deriving the ICLP-QR strategy, for a given $\eta > 1$ such that $tr(C^{\eta-1})$ is finite, we have

$$\|h\|_{1,C} \leq \|h\|_{C^\eta}\sqrt{\mathrm{trace}(C^{\eta-1})}.$$

Substituting h by $\hat{f}_D - \hat{f}_{D'}$ leads to

$$\|\hat{f}_D - \hat{f}_{D'}\|_{1,C} \leq \left\|\hat{f}_D - \hat{f}_{D'}\right\|_{C^\eta}\sqrt{\mathrm{trace}(C^{\eta-1})},$$

meaning that we need to found the upper bound for $\|\hat{f}_D - \hat{f}_{D'}\|_{C^\eta}$. Let $t \in [0,1]$, $\delta_{D',D} = \hat{f}_{D'} - \hat{f}_D$ and $L_D(f) = \frac{1}{n}\sum_{i=1}^n L_{d_i,f}$. Note that $\hat{f}_{D'}$ and $\hat{f}_D$ are the minimizers of (9), we have

$$L_D\left(\hat{f}_D\right) + \psi\left\|\hat{f}_D\right\|_{C^\eta}^2 \leq L_D\left(\hat{f}_D + t\delta_{D',D}\right) + \psi\left\|\hat{f}_D + t\delta_{D',D}\right\|_{C^\eta}^2,$$

and

$$L_D\left(\hat{f}_{D'}\right) + \psi\left\|\hat{f}_{D'}\right\|_{C^\eta}^2 \leq L_D\left(\hat{f}_{D'} - t\delta_{D',D}\right) + \psi\left\|\hat{f}_{D'} - t\delta_{D',D}\right\|_{C^\eta}^2.$$

Combining the two inequalities above, we have

$$\begin{aligned}
L_D\left(\hat{f}_D\right) &- L_D\left(\hat{f}_D + t\delta_{D',D}\right) + L_D\left(\hat{f}_{D'}\right) - L_D\left(\hat{f}_{D'} - t\delta_{D',D}\right) \\
&\leq \psi\left(\left\|\hat{f}_D + t\delta_{D',D}\right\|_{C^\eta}^2 - \left\|\hat{f}_D\right\|_{C^\eta}^2 + \left\|\hat{f}_{D'} - t\delta_{D',D}\right\|_{C^\eta}^2 - \left\|\hat{f}_{D'}\right\|_{C^\eta}^2\right).
\end{aligned}$$

Then, using the same proof techniques in Section 4.3 of Hall et al. (2013), we have

$$\|\hat{f}_D - \hat{f}_{D'}\|_{C^\eta} \leq \frac{M}{\psi n}\sqrt{\sup_x C^\eta(x,x)},$$

which completes the proof. ■

Appendix C. Extension of Mean Function

In Section 4.2, we only considered the mean protection. In this section, we demonstrate that many statistical estimation problems in the context of functional data analysis can be reduced down to mean estimation. Therefore, the mean protection technique and the corresponding theoretical analysis derived in Section 4.2 can be applied to these problems as well. Specifically, we consider the estimation and protection of (1) the covariance function and (2) the coefficient function in function-on-scalar linear regression.

C.1 Covariance Function

For a given sample $X_1, \cdots, X_n \in \mathbb{H}$, where $\mathbb{H}$ represents some function spaces. Our goal is to estimate its covariance function, i.e.,

$$C(s,t) = \mathrm{E}\left(X(s) - \mu(s)\right)\left(X(t) - \mu(t)\right)$$

where μ is the mean function, and the empirical sample covariance function is

$$\bar{C}(s,t) = \frac{1}{n}\sum_{i=1}^{n}\left(X_i(s) - \bar{X}(s)\right)\left(X_i(t) - \bar{X}(t)\right).$$

Thus, estimating the covariance function can also be viewed as estimating the mean function of $(X_i(s) - \mu(s))(X_i(t) - \mu(t))$.

To obtain the qualified summary, one can also apply quadratic regularization. For the ICLP covariance function K, denote the tensor product space as

$$\mathcal{H}_{K^\eta \otimes K^\eta} := \mathcal{H}_{K^\eta} \otimes \mathcal{H}_{K^\eta},$$

and RKHS associated with reproducing kernel

$$K^\eta \otimes K^\eta((s_1,t_1),(s_2,t_2)) = K^\eta(s_1,s_2)K^\eta(t_1,t_2).$$

With slight abuse of notation, we let $\otimes$ also denote the tensor product of elements in $\mathbb{H}$, and then the ICLP with quadratic regularization can be expressed as

$$\hat{C} = \underset{C \in \mathcal{H}_{K^\eta \otimes K^\eta}}{\operatorname{argmin}} \left\{ \frac{1}{n}\sum_{i=1}^{n} \left\|(X_i - \bar{X}) \otimes (X_i - \bar{X}) - C\right\|^2_{\mathbb{H}\otimes\mathbb{H}} + \psi\|C\|^2_{\mathcal{H}_{K^\eta \otimes K^\eta}} \right\}$$

which leads to

$$\hat{C} = \left(K^\eta \otimes K^\eta + \psi\mathbf{I}\right)^{-1} K^\eta \otimes K^\eta(\bar{C}).$$

Further assuming $\{\lambda_j\}_{j\geq 1}$ and $\{\phi_j\}_{j\geq 1}$ as the eigenvalues and eigenfunctions of K, then

$$\hat{C}(s,t) = \sum_{j,l\geq 1} \frac{\lambda_l^\eta \lambda_l^\eta}{\lambda_l^\eta \lambda_l^\eta + \psi} \left\langle \bar{C}, \phi_j\phi_l \right\rangle_{\mathbb{H}\otimes\mathbb{H}} \phi_j(s)\phi_l(t).$$

The expression of $\hat{C}$ is analogous to the expression of $\hat{\mu}$ under the quadratic regularization in Section B.1. Therefore, similar global sensitivity and utility analysis can be applied to $\hat{C}$ as well since privatizing the covariance function is using the same ICLP via tensor basis, i.e.,

$$X(s,t) = \sum_{j,l\geq 1} \sqrt{\lambda_k\lambda_l} Z_{kl}\phi_k(s)\phi_l(t)$$

where $Z_{k,l}$ are i.i.d. Laplace random variables with mean 0 and variance 1.

C.2 Function-on-Scalar Linear Regression

We consider the following Function-on-Scalar linear regression, i.e.,

$$Y_i(t) = X_i^T \beta(t) + \epsilon_i(t), \quad \text{for} \quad i = 1, \cdots, n$$

where $Y_i(t)$ and $e_i(t)$ are functional response and error that lies in Hilbert space $\mathbb{H}$, covariates $X_i \in \mathbb{R}^p$ and coefficient function $\beta(t) \in \mathbb{H}^{\otimes p}$ when $\mathbb{H}^{\otimes p}$ denotes p-fold Cartesian product of $\mathbb{H}$. Estimating and privatizing the coefficient functions $\beta(t)$ is of primary interest.

One can obtain the estimation via the classical ordinary least square (OLS) estimator, i.e.,

$$\hat{\beta} = \underset{\beta \in \mathbb{H}^{\otimes p}}{\text{argmin}} \frac{1}{n} \sum_{i=1}^{n} \left\| Y_i - X_i^T \beta \right\|_{\mathbb{H}}^2 .$$

The OLS estimator is then

$$\hat{\beta}(t) = \left(\frac{1}{n} \sum_{i=1}^{n} X_i X_i^T \right)^{-1} \left(\frac{1}{n} \sum_{i=1}^{n} Y_i(t) X_i \right).$$

Noticing the functional components involved in $\hat{\beta}(t)$ are $Y_i(t)$, thus estimating the coefficient function can also be viewed as estimating the mean function of $Y_i(t)X_i$ and the applied the inverse matrix with scalar elements.

In classical simple linear regression, where both response and covariate are scalars, i.e.

$$y_i = x_i \beta + \epsilon_i, \quad \text{for} \quad i = 1, \cdots, n.$$

Let $\mathbf{x} := (x_1, \cdots, x_n)$ and $\mathbf{y} := (y_1, \cdots, y_n)$ To achieve ϵ-DP for the OLS estimator, one typically privatizes the empirical variance of $\mathbf{x}$ and the empirical covariance between $\mathbf{x}$ and $\mathbf{y}$, rather than privatizing the $\hat{\beta}$ directly, see Alabi et al. (2020). Specifically, let $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$, $\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i$, $n\text{cov}(\mathbf{x}, \mathbf{y}) = \langle \mathbf{x} - \bar{x}\mathbf{1}, \mathbf{y} - \bar{y}\mathbf{1} \rangle_{\ell^2}$, and $n\text{var}(\mathbf{x}) = \langle \mathbf{x} - \bar{x}\mathbf{1}, \mathbf{x} - \bar{x}\mathbf{1} \rangle_{\ell^2}$. Assuming the sensitivity of $n\text{cov}(\mathbf{x}, \mathbf{y})$ and $n\text{var}(\mathbf{x})$ are both 1 without losing generality. Then, the private OLS estimator that achieves ϵ-DP is

$$\tilde{\beta} = \frac{n\text{cov}(\mathbf{x}, \mathbf{y}) + Z_1}{n\text{var}(\mathbf{x}) + Z_2}$$

where Z_1 and Z_2 are independent random variables generated from $Lap(0, \frac{1}{\gamma\epsilon})$ and $Lap(0, \frac{1}{(1-\gamma)\epsilon})$ respectively with $\gamma \in (0, 1)$.

Turning back to the Function-on-Scalar regression case, one can privatize the statistic $T_1 := \frac{1}{n} \sum_{i=1}^{n} X_i X_i^T$ and $T_2 := \frac{1}{n} \sum_{i=1}^{n} Y_i(t) X_i$ separately by splitting the privacy budget ϵ in to $\gamma\epsilon$ (for T_1) and $(1-\gamma)\epsilon$ (for T_2). While T_1 is a $p \times p$ matrix with scalar elements, one can privatize it with a classical multivariate privacy tool with budget $\gamma\epsilon$. For the statistic T_2, it is the mean function of $Y_i(t)X_i$ and thus, we can directly apply the mean function protection we develop in Section 4.2.

Appendix D. Additional Results for Mean Protection

This section presents additional experimental results that are conducted in different settings for mean protection. First, we consider generating the error function from a basis expansion instead of a stochastic process. Second, we set the covariance function of the ICLP as the Matérn kernel with $\alpha = 2.5$, resulting in $\beta = 3$. The Matérn kernel has been widely used to control the eigenvalue decay rate in kernel-based methods; see Wang and Jing (2022); Lin and Reimherr (2024a,b)

D.1 Results for Different Error Function

We generate the functional sample curves as

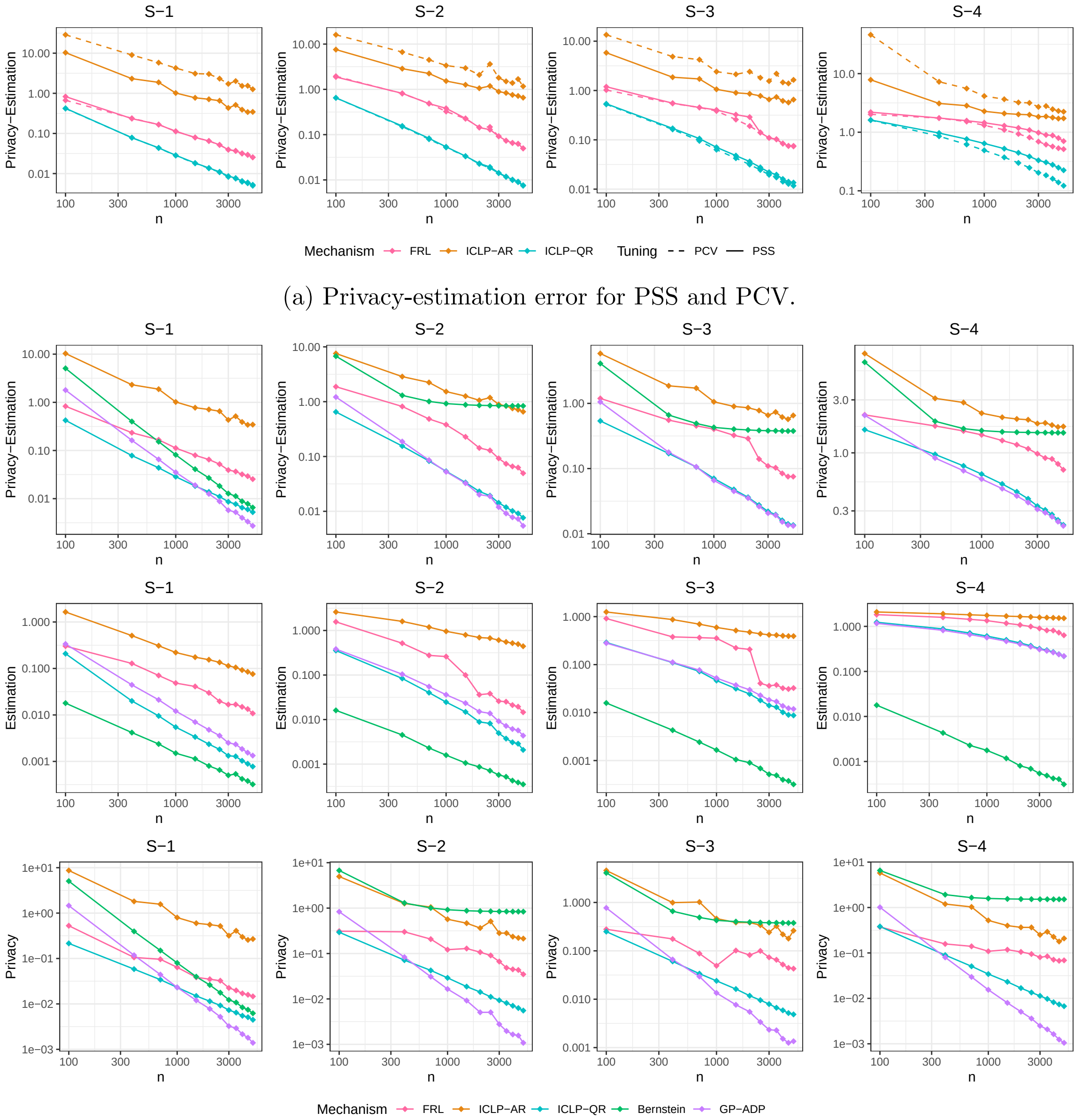
$$X_i(t) = \mu_0(t) + e_i \quad \text{where} \quad e_i(t) = \sum_{j=1}^{100} U_{ij}\phi_j(t)$$

where U_{ij} are i.i.d. random variables from a t distribution with degree of freedom 5. We conduct the same experiments and report the results in Figure 6. One interesting observation is that when e_i are generated from a basis expansion with heavy-tailed coefficients, the PSS and PCV are more aligned with each other in S-2 and S-3.

D.2 Results for Different ICLP Covariance

We set the ICLP covariance function as the Matérn kernel with $\alpha = 2.5$, resulting in $\lambda_j \asymp j^{-6}$. We repeat the comparison between PSS and PCV and the experiments that compare different mechanisms under different n.

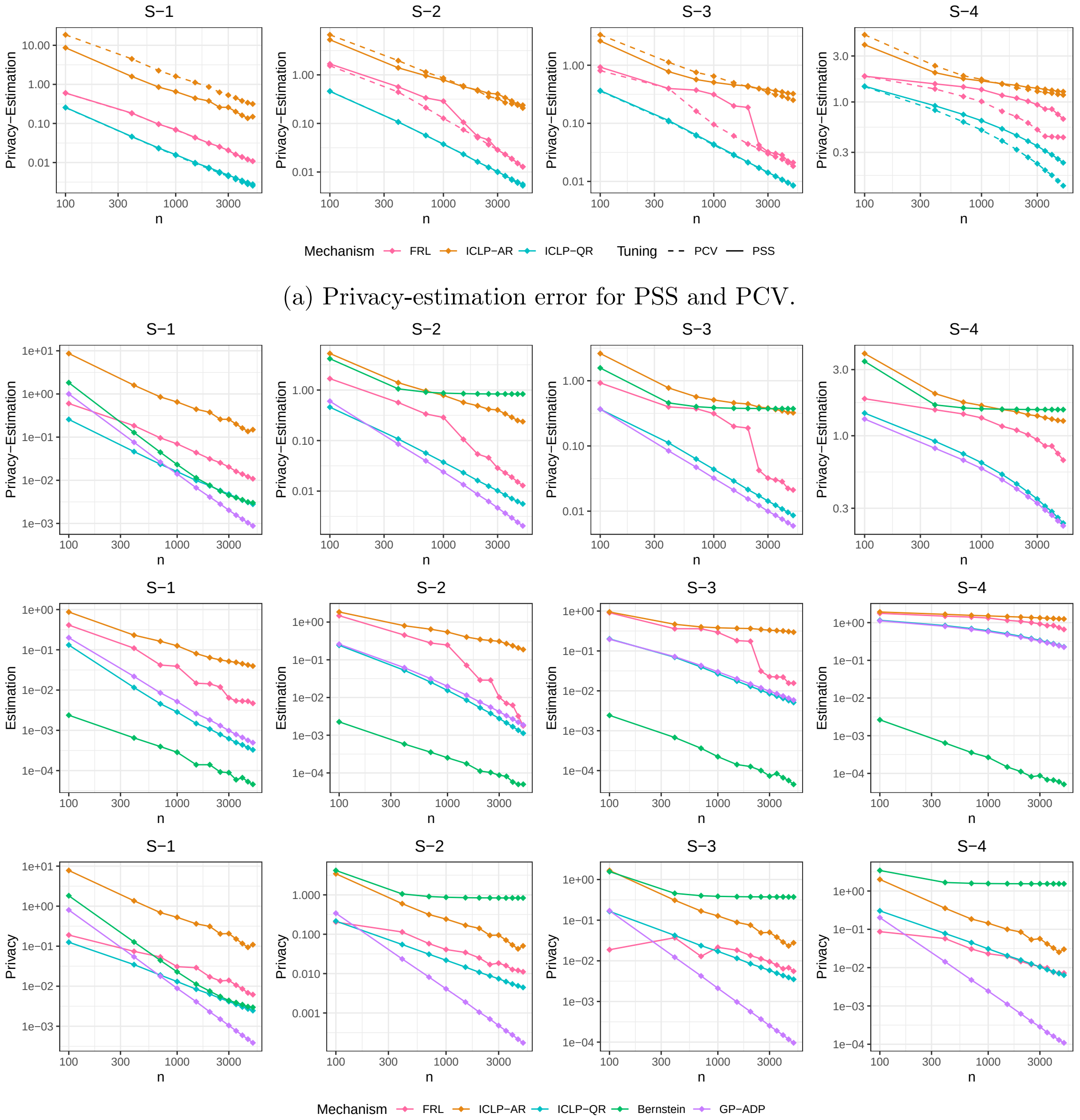
The results are reported in Figure 7 for the error function as a Gaussian stochastic process and Figure 8 for the error function generated from a basis expansion. It can be observed from the figure that the results based on $C_{\frac{5}{2}}$ are almost the same as those based on $C_{\frac{3}{2}}$.



(a) Privacy-estimation error for PSS and PCV.

(b) Privacy-estimation, estimation, and privacy errors with PSS.

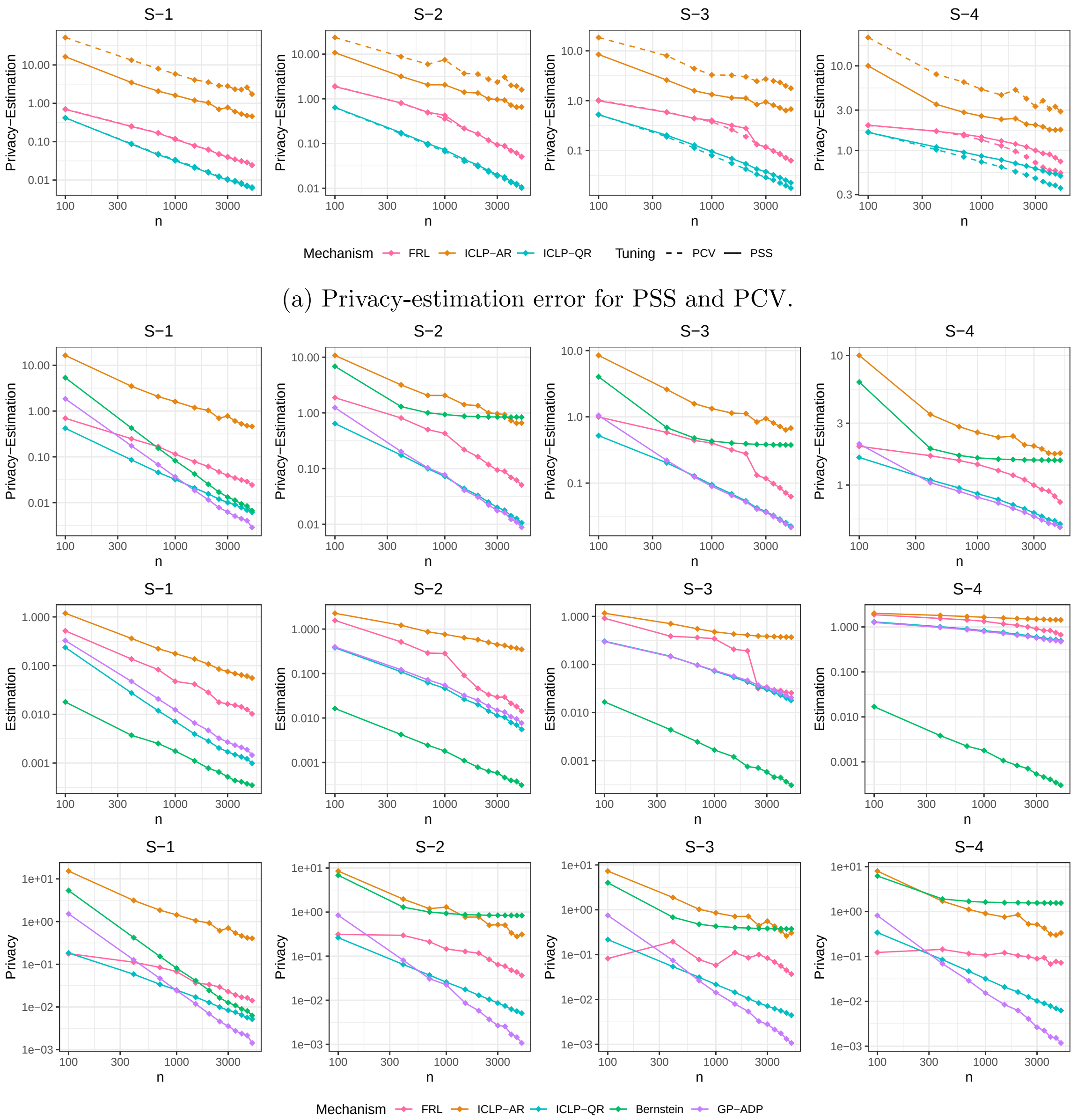
Figure 6: Error decay curves for different mechanisms, sample sizes, and true mean functions. The ICLP covariance function is the Matérn Kernel $C_{\frac{3}{2}}$, and the error function e_i is generated via basis expansion with coefficients randomly drawn from a t distribution.



(a) Privacy-estimation error for PSS and PCV.

(b) Privacy-estimation, estimation, and privacy errors with PSS.

Figure 7: Error decay curves for different mechanisms, sample sizes, and true mean functions. The ICLP covariance function is the Matérn Kernel $C_{\frac{5}{2}}$, and the error function e_i is generated from the Gaussian process with RBF kernels.



(a) Privacy-estimation error for PSS and PCV.

(b) Privacy-estimation, estimation, and privacy errors with PSS.

Figure 8: Error decay curves for different mechanisms, sample sizes, and true mean functions. The ICLP covariance function is the Matérn Kernel $C_{\frac{5}{2}}$, and the error function e_i is generated via basis expansion with coefficients randomly drawn from a t distribution.

Appendix E. Visualization of the Human Mortality Application

In Section 7.2, we only reported the expected L^2-distance of private KDEs to their non-private counterparts. Here, we present the visualization of the comparison between non-private KDEs and private KDEs for different mechanisms in Figure 9.

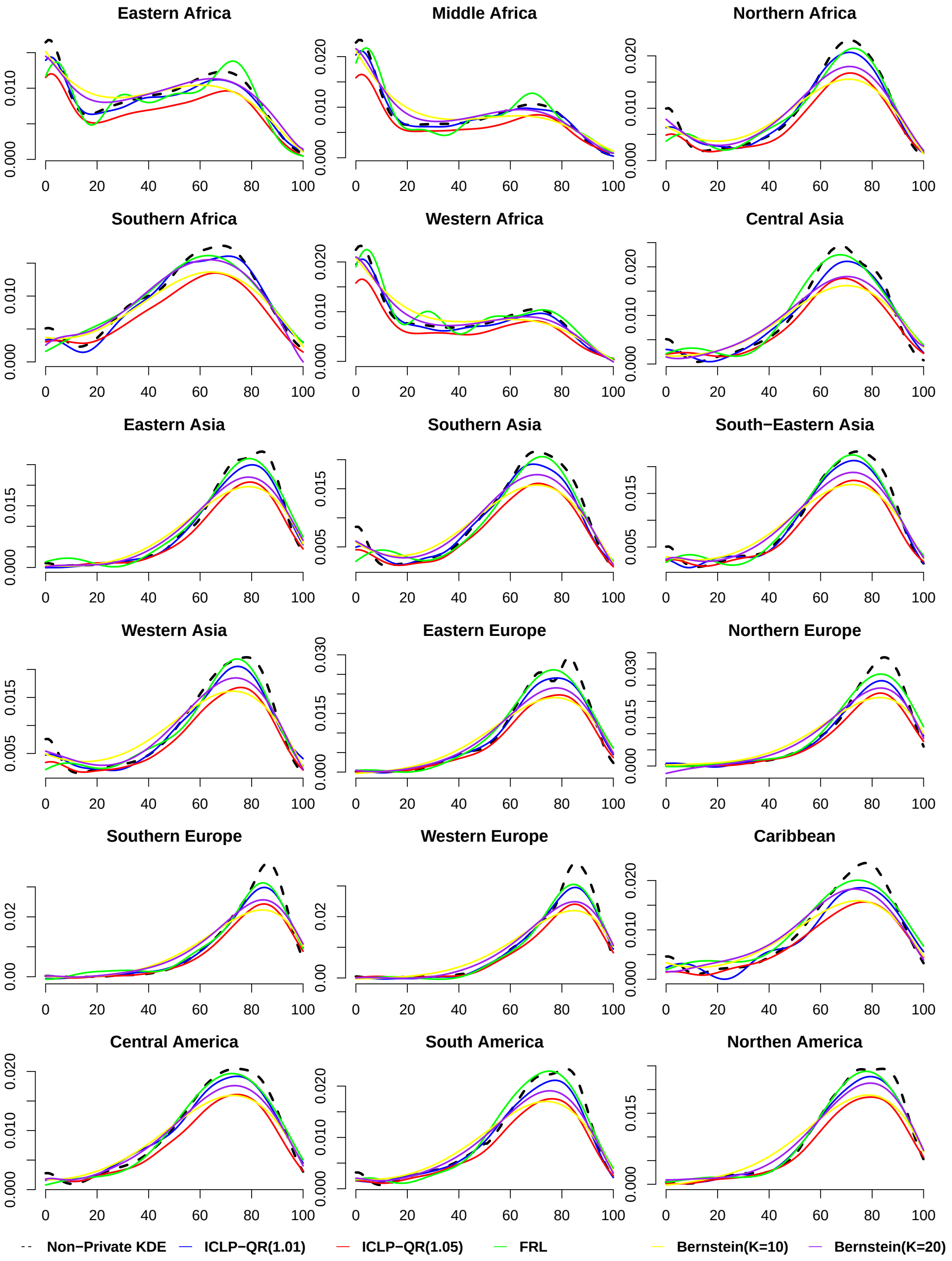


Figure 9: Non-private and private kernel density estimates of age-at-death density in different regions under different mechanisms with $\epsilon = 1$.

References

Daniel Alabi, Audra McMillan, Jayshree Sarathy, Adam Smith, and Salil Vadhan. Differentially private simple linear regression. *arXiv preprint arXiv:2007.05157*, 2020.

Francesco Alda and Benjamin Rubinstein. The bernstein mechanism: Function release under differential privacy. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 31, 2017.

Jordan Awan and Aleksandra Slavković. Structure and sensitivity in differential privacy: Comparing k-norm mechanisms. *Journal of the American Statistical Association*, 116 (534):935–954, 2021.

Jordan Awan, Ana Kenney, Matthew Reimherr, and Aleksandra Slavković. Benefits and pitfalls of the exponential mechanism with applications to hilbert spaces and functional pca. In *International Conference on Machine Learning*, pages 374–384. PMLR, 2019.

Vladimir Igorevich Bogachev. *Gaussian measures*. Number 62. American Mathematical Soc., 1998.

Denis Bosq. *Linear processes in function spaces: theory and applications*, volume 149. Springer Science & Business Media, 2000.

Olivier Bousquet and André Elisseeff. Stability and generalization. *The Journal of Machine Learning Research*, 2:499–526, 2002.

T Tony Cai and Ming Yuan. Optimal estimation of the mean function based on discretely sampled functional data: Phase transition. *The Annals of Statistics*, 39(5):2330–2355, 2011.

Karthekeyan Chandrasekaran, Justin Thaler, Jonathan Ullman, and Andrew Wan. Faster private release of marginals on small databases. In *Proceedings of the 5th conference on Innovations in theoretical computer science*, pages 387–402, 2014.

Kamalika Chaudhuri and Staal A Vinterbo. A stability-based validation procedure for differentially private machine learning. *Advances in Neural Information Processing Systems*, 26:2652–2660, 2013.

Kamalika Chaudhuri, Claire Monteleoni, and Anand D Sarwate. Differentially private empirical risk minimization. *Journal of Machine Learning Research*, 12(3), 2011.

Noel Cressie and Hsin-Cheng Huang. Classes of nonseparable, spatio-temporal stationary covariance functions. *Journal of the American Statistical association*, 94(448):1330–1339, 1999.

Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. In *Theory of cryptography conference*, pages 265–284. Springer, 2006.

Cynthia Dwork, Aaron Roth, et al. The algorithmic foundations of differential privacy. *Found. Trends Theor. Comput. Sci.*, 9(3-4):211–407, 2014.

Frederic Ferraty and Yves Romain. *The Oxford handbook of functional data analaysis*. Oxford University Press, 2011.

Rob Hall, Alessandro Rinaldo, and Larry Wasserman. Differential privacy for functions and functional data. *The Journal of Machine Learning Research*, 14(1):703–727, 2013.

Moritz Hardt and Kunal Talwar. On the geometry of differential privacy. In *Proceedings of the forty-second ACM symposium on Theory of computing*, pages 705–714, 2010.

Moritz Hardt, Katrina Ligett, and Frank McSherry. A simple and practical algorithm for differentially private data release. *arXiv preprint arXiv:1012.4763*, 2010.

Tailen Hsing and Randall Eubank. *Theoretical foundations of functional data analysis, with an introduction to linear operators*, volume 997. John Wiley & Sons, 2015.

Shizuo Kakutani. On equivalence of infinite product measures. *Annals of Mathematics*, pages 214–224, 1948.

George Kimeldorf and Grace Wahba. Some results on tchebycheffian spline functions. *Journal of mathematical analysis and applications*, 33(1):82–95, 1971.

Piotr Kokoszka and Matthew Reimherr. *Introduction to functional data analysis*. Chapman and Hall/CRC, 2017.

DD Kosambi. Statistics in function space. In *DD Kosambi*, pages 115–123. Springer, 2016.

Haotian Lin and Matthew Reimherr. On hypothesis transfer learning of functional linear models. In *Forty-first International Conference on Machine Learning*, 2024a.

Haotian Lin and Matthew Reimherr. Smoothness adaptive hypothesis transfer learning. In *Forty-first International Conference on Machine Learning*, 2024b.

Alessandra Lunardi, Michele Miranda, and Diego Pallara. Infinite dimensional analysis. In *19th Internet Seminar*, volume 2016, 2015.

Frank McSherry and Kunal Talwar. Mechanism design via differential privacy. In *48th Annual IEEE Symposium on Foundations of Computer Science (FOCS'07)*, pages 94–103. IEEE, 2007.

Charles A Micchelli and Grace Wahba. Design problems for optimal surface interpolation. Technical report, WISCONSIN UNIV-MADISON DEPT OF STATISTICS, 1979.

Ardalan Mirshani, Matthew Reimherr, and Aleksandra Slavković. Formal privacy for functional data with gaussian perturbations. In *International Conference on Machine Learning*, pages 4595–4604. PMLR, 2019.

NhatHai Phan, Minh Vu, Yang Liu, Ruoming Jin, Dejing Dou, Xintao Wu, and My T Thai. Heterogeneous gaussian mechanism: Preserving differential privacy in deep learning with provable robustness. *arXiv preprint arXiv:1906.01444*, 2019.

Jim Ramsay, James Ramsay, BW Silverman, et al. *Functional Data Analysis*. Springer Science & Business Media, 2005.

JNK Rao and AJ Scott. A simple method for the analysis of clustered binary data. *Biometrics*, pages 577–585, 1992.

Matthew Reimherr and Jordan Awan. Elliptical perturbations for differential privacy. *Advances in Neural Information Processing Systems*, 32, 2019a.

Matthew Reimherr and Jordan Awan. Kng: The k-norm gradient mechanism. *Advances in neural information processing systems*, 32, 2019b.

John A Rice and Bernard W Silverman. Estimating the mean and covariance structure nonparametrically when the data are curves. *Journal of the Royal Statistical Society: Series B (Methodological)*, 53(1):233–243, 1991.

Jerome Sacks and Donald Ylvisaker. Designs for regression problems with correlated errors. *The Annals of Mathematical Statistics*, 37(1):66–89, 1966.

Jerome Sacks and Donald Ylvisaker. Designs for regression problems with correlated errors: many parameters. *The Annals of Mathematical Statistics*, 39(1):49–69, 1968.

Jerome Sacks and Donald Ylvisaker. Designs for regression problems with correlated errors iii. *The Annals of Mathematical Statistics*, 41(6):2057–2074, 1970.

Terence Tao. *An introduction to measure theory*, volume 126. American Mathematical Society Providence, RI, 2011.

Wenjia Wang and Bing-Yi Jing. Gaussian process regression: Optimality, robustness, and relationship with kernel ridge regression. *Journal of Machine Learning Research*, 23(193): 1–67, 2022.

Ziteng Wang, Kai Fan, Jiaqi Zhang, and Liwei Wang. Efficient algorithm for privately releasing smooth queries. In *NIPS*, pages 782–790. Citeseer, 2013.

Larry Wasserman. *All of nonparametric statistics*. Springer Science & Business Media, 2006.

Larry Wasserman and Shuheng Zhou. A statistical framework for differential privacy. *Journal of the American Statistical Association*, 105(489):375–389, 2010.

Yun Yang, Anirban Bhattacharya, and Debdeep Pati. Frequentist coverage and sup-norm convergence rate in gaussian process regression. *arXiv preprint arXiv:1708.04753*, 2017.

Jun Zhang, Zhenjie Zhang, Xiaokui Xiao, Yin Yang, and Marianne Winslett. Functional mechanism: regression analysis under differential privacy. *arXiv preprint arXiv:1208.0219*, 2012.