Enhancing Feature-Specific Data Protection via Bayesian Coordinate Differential Privacy [*]

Maryam Aliakbarpour [†]   Syomantak Chaudhuri [‡]   Thomas A. Courtade [‡]

Alireza Fallah [‡]   Michael I. Jordan [§]

Abstract

Local Differential Privacy (LDP) offers strong privacy guarantees without requiring users to trust external parties. However, LDP applies uniform protection to all data features, including less sensitive ones, which degrades performance of downstream tasks. To overcome this limitation, we propose a Bayesian framework, Bayesian Coordinate Differential Privacy (BCDP), that enables feature-specific privacy quantification. This more nuanced approach complements LDP by adjusting privacy protection according to the sensitivity of each feature, enabling improved performance of downstream tasks without compromising privacy. We characterize the properties of BCDP and articulate its connections with standard non-Bayesian privacy frameworks. We further apply our BCDP framework to the problems of private mean estimation and ordinary least-squares regression. The BCDP-based approach obtains improved accuracy compared to a purely LDP-based approach, without compromising on privacy.

1 Introduction

The rapid expansion of machine learning applications across various sectors has fueled an unprecedented demand for data collection. With the increase in data accumulation, user privacy concerns also intensify. Differential privacy (DP) has emerged as a prominent framework that enables algorithms to utilize data while preserving the privacy of individuals who provide it Warner [1965], Evfimievski et al. [2003], Dwork et al. [2006, 2014], Vadhan [2017], Desfontaines [2020]. A variant of DP, local differential privacy (LDP), suitable for distributed settings where users do not trust any central authority has been extensively studied [Warner, 1965, Evfimievski et al., 2003, Beimel et al., 2008, Kasiviswanathan et al., 2011, Chan et al., 2012]. In this framework,

---

[*]The authors are listed in alphabetical order.

[†]Department of Computer Science & Ken Kennedy Institute, Rice University

[‡]Department of Electrical Engineering and Computer Sciences, University of California, Berkeley

[§]Department of Electrical Engineering and Computer Sciences & Department of Statistics, University of California, Berkeley; INRIA, Paris

each user shares an obfuscated version of their datapoint, meaningful in aggregate to the central server yet concealing individual details. Leading companies, including Google [Erlingsson et al., 2014], Microsoft [Ding et al., 2017], and Apple [Differential Privacy Team, Apple, 2017, Thakurta et al., 2017], have employed this approach for data collection.

In the LDP framework, a privacy parameter must be set—typically by the data collector—that determines the level of protection for user information. This decision is particularly challenging in high-dimensional settings, where user data includes multiple attributes such as name, date of birth, and social security number, each with differing levels of sensitivity. One common strategy is to allow the most sensitive data feature to dictate the level of privacy, providing the highest level of protection. Alternatively, separate private mechanisms can be applied to each feature according to its required privacy level. However, this approach fails when there is correlation between features. For instance, setting a lower level of privacy for a feature that is consistently aligned with a more sensitive one risks compromising the privacy of the sensitive feature.

This complexity raises a compelling question: Can we tailor privacy protection to specific data features if the correlation among data coordinates is controlled, without defaulting to the privacy level of the most sensitive coordinate? In other words, can we leverage users' lower sensitivity preferences for certain features to reduce the error in our inference algorithms, provided the correlation with sensitive coordinates is weak?

Our contributions: Our main contribution is addressing the challenge of feature-specific data protection from a Bayesian perspective. We present a novel Bayesian privacy framework, which we refer to as *Bayesian Coordinate Differential Privacy* (BCDP). BCDP complements LDP and allows a tailored privacy demand per feature. In essence, BCDP ensures that the odds of inferring a specific feature (also referred to as a coordinate) remain nearly unchanged before and after observing the outcome of the private mechanism. The framework's parameters control the extent to which the odds remain unchanged, enabling varying levels of privacy per feature. As part of developing this new framework, we explore its formal relationship with other related DP notions and examine standard DP properties, such as *post-processing* and *composition*.

To demonstrate the efficacy of our framework, we study two fundamental problems in machine learning under a baseline level of LDP privacy, coupled with more stringent per-coordinate BCDP constraints. First, we address *multivariate mean estimation*. Our proposed algorithm satisfies BCDP constraints while outperforming a standard LDP algorithm, where the privacy level is determined by the most sensitive coordinate. Our experimental results confirm this advantage. Second, we present an algorithm for *ordinary least squares regression* under both LDP and BCDP constraints.

A key technical component of our algorithm for these two problems is a mechanism that builds on *any* off-the-shelf LDP mechanism and can be applied to various problems. Our algorithm involves making parallel queries to the LDP mechanism with a series of carefully selected privacy parameters where the i-th query omits the $i - 1$ most sensitive coordinates. The outputs are then aggregated to produce the most accurate estimator for each coordinate.

Bayesian Feature-Specific Protection: We now give a high-level discussion of privacy in the Bayesian framework to motivate our approach. Formal definitions are in Section 2. Consider a set of users, each possessing a high-dimensional data point x with varying privacy protection needs across different features. For instance, x may contain genetic markers for diseases, and the i-th coordinate x_i could reveal Alzheimer's markers that the user wishes to keep strictly confidential. In a local privacy setting, users apply a locally private mechanism M to obtain obfuscated data $M(x)$, for public aggregation and analysis. Our focus is on understanding what can be inferred about x, especially the sensitive feature x_i, from the obfuscated data.

Now, imagine an adversary placing a bet on a probabilistic event related to x. Without extra information, their chances of guessing correctly are based on prior knowledge of the population. However, access to $M(x)$ could prove advantageous for inference. Privacy in the Bayesian model is measured by how much $M(x)$ improves the adversary's chances. In other words, the privacy parameter limits how much their odds can shift.

The standard Bayesian interpretation of LDP ensures uniform privacy across all data features. In contrast, our framework BCDP enhances LDP by enabling feature-specific privacy controls. That is, BCDP restricts adversarial inference of events concerning individual coordinates. For example, BCDP ensures that the advantage an adversary gains from observing $M(x)$ when betting on whether a person has the Alzheimer's trait (i.e., $x_i = 1$) is smaller than what would be achieved under a less stringent LDP constraint.

We remark that BCDP does not replace LDP, since it lacks global data protection. Also, correlation plays a crucial role in BCDP's effectiveness: if a sensitive feature is highly correlated with less sensitive features, increasing protection for the sensitive feature imposes stricter protection for other features as well. As the correlation between features increases, BCDP's guarantees converge to the uniform guarantees of LDP.

Related Work: Statistical inference with LDP guarantees has been extensively studied. For comprehensive surveys see Xiong et al. [2020], Yang et al. [2023]. Below, we discuss the results that are most relevant to our setup.

Several papers consider the uneven protection of privacy across features. For instance, Ghazi et al. [2022] and Acharya et al. [2020] focus on settings where the privacy parameters vary across different neighboring datasets. Other models arise from altering the concept of neighboring datasets [Kifer and Machanavajjhala, 2014, He et al., 2014], or they emphasize indistinguishability guarantees within a redefined metric space for the datasets or data points Chatzikokolakis et al. [2013], Alvim et al. [2018], Imola et al. [2022], Andrés et al. [2013]. Another relevant notion for classification tasks is *Label-DP* which considers labels of training data as sensitive, but not the features themselves [Chaudhuri and Hsu, 2011, Beimel et al., 2013, Wang and Xu, 2019, Ghazi et al., 2021]. A recent generalization of Label-DP, proposed by Mahloujifar et al. [2023], is Feature-DP, which allows specific features of the data to be disclosed by the algorithm while ensuring the privacy of the remaining information. A line of work, e.g., Kifer and Machanavajjhala [2011],

Kenthapadi et al. [2013], studies attribute DP, which limits changes in the output distribution of the privacy mechanism when a single coordinate of the data is modified.

A common shortcoming in the aforementioned approaches is that the heterogeneous privacy protection across features, without accounting for their correlations, can lead to unaddressed privacy leakage—an issue our framework is specifically designed to mitigate. Another key distinction of our work is the integration of the BCDP framework with the LDP framework, enabling simultaneous protection at both the feature level and the global level.

A Bayesian approach to Central-DP is considered in Triastcyn and Faltings [2020] where the notation of neighboring dataset are taken over pairs that are drawn from the same distribution. This is similar in spirit to our framework but we operate in the Local-DP regime so we work with a prior instead. Xiao and Devadas [2023] consider a Bayesian approach to privacy where the goal is to prevent an observer from being able to *nearly* guess the input to a mechanism under a given notion of distance and error probability. Another Bayesian interpretation of DP is presented in Kasiviswanathan and Smith [2014] and show that DP ensures that the posterior distribution after observing the output of a DP algorithm is close to the prior in Total-Variation distance. None of these results explicitly focus on varying privacy features across coordinates.

Private mean estimation and private least-squares regression are well studied problems in literature in both central [Biswas et al., 2020, Karwa and Vadhan, 2017, Kifer et al., 2012, Vu and Slavkovic, 2009] and local [Asi et al., 2022, Feldman and Talwar, 2021, Zheng et al., 2017] DP regimes. We depart from these works as we consider a BCDP constraint in addition to an LDP constraint.

Organization: We present a formal treatment of our framework in Section 2. Some key properties of the framework are presented in Section 3. We consider the problem of *mean estimation* in Section 4 and *ordinary least-squares regression* in Section 5.

2 Problem Formulation

Let $\mathcal{X}$ denote the data domain represented as the Cartesian product $\mathcal{X} := \times_{j=1}^{d} \mathcal{X}_j$. One can consider this as a general decomposition where each of d canonical data coordinates can be categorical or a real value, or a combination of both, as the setting dictates. For a vector $\boldsymbol{x} \in \mathcal{X}$, we let x_i denote the i-th coordinate, and $\boldsymbol{x}_{-i}$ the vector with $(d-1)$ coordinates obtained by removing the i-th coordinate x_i from $\boldsymbol{x}$. A private mechanism M is a randomized algorithm that maps $\boldsymbol{x}$ to $M(\boldsymbol{x})$ in an output space $\mathcal{Y}$.

Let $(\mathcal{X}, \mathcal{F}_{\mathcal{X}})$ and $(\mathcal{Y}, \mathcal{F}_{\mathcal{Y}})$ be measurable spaces. Let each coordinate space $\mathcal{X}_i$ be equipped with σ-algebra $\mathcal{F}_{\mathcal{X}_i}$, and take $\mathcal{F}_{\mathcal{X}}$ equal to the product σ-algebra $\times_{i=1}^{d} \mathcal{F}_{\mathcal{X}_i}$. A randomized mechanism M from $\mathcal{X}$ to $\mathcal{Y}$ can be represented by a Markov kernel $\mu : (\mathcal{X}, \mathcal{F}_{\mathcal{X}}) \to (\mathcal{Y}, \mathcal{F}_{\mathcal{Y}})$ where for an input $\boldsymbol{x} \in \mathcal{X}$, the output $M(\boldsymbol{x})$ is a $\mathcal{Y}$-valued random variable with law $\mu_{\boldsymbol{x}}(\cdot)$.

Local differential privacy   We first revisit the definition of Local Differential Privacy (LDP).

Definition 1 (Local DP [Kasiviswanathan et al., 2011]). A mechanism M is ε-LDP if, for all $R \in \mathcal{F}_\mathcal{Y}$ and all $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{X}$, we have:

$$\mu_{\boldsymbol{x}}(R) \leq e^{\varepsilon} \mu_{\boldsymbol{x}'}(R). \tag{1}$$

With the understanding that $M(\boldsymbol{x}) \sim \mu_{\boldsymbol{x}}(\cdot)$, Definition 1 can be rewritten in the more familiar form $\Pr\{M(\boldsymbol{x}) \in R\} \leq e^{\varepsilon}\Pr\{M(\boldsymbol{x}') \in R\} \ \forall \boldsymbol{x}, \boldsymbol{x}' \in \mathcal{X}, R \in \mathcal{F}_\mathcal{Y}$.

Next, we focus on a *Bayesian interpretation* of this setting. Suppose we have an underlying data distribution π. That is, we equip $(\mathcal{X}, \mathcal{F}_\mathcal{X})$ with the probability measure π which we call the *prior*. This induces a probability space $(\mathcal{X} \times \mathcal{Y}, \mathcal{F}_\mathcal{X} \times \mathcal{F}_\mathcal{Y}, P)$ where the measure P is characterized via

$$P(S \times R) = \int_S \mu_{\boldsymbol{x}}(R)d\pi(\boldsymbol{x}), \ S \in \mathcal{F}_\mathcal{X}, R \in \mathcal{F}_\mathcal{Y}. \tag{2}$$

Note that the mechanism M defining μ is precisely represented on this space by the random variable $M : (\boldsymbol{x}, y) \in \mathcal{X} \times \mathcal{Y} \mapsto y$. By a slight abuse of notation, we continue to denote the (randomized) output of the mechanism by $M(\boldsymbol{x})$. We first define the sets on $\mathcal{X}$ which have positive probability under π

$$\mathcal{F}_\mathcal{X}^+ = \{S \in \mathcal{F}_\mathcal{X} | \pi(S) > 0\}. \tag{3}$$

For $S \in \mathcal{F}_\mathcal{X}^+$ and $R \in \mathcal{F}_\mathcal{Y}$, note that

$$P\{M(\boldsymbol{x}) \in R | \boldsymbol{x} \in S\} = \frac{P(S \times R)}{\pi(S)}. \tag{4}$$

Bayesian differential privacy: We now introduce a Bayesian formulation of the LDP definition which involves both the mechanism M *and* the prior π. This can be seen as the local version of the definitions proposed in earlier works, e.g., Kifer and Machanavajjhala [2014], Triastcyn and Faltings [2020].

Definition 2 (Bayesian DP). The pair (π, M) is ε-BDP if for all $R \in \mathcal{F}_\mathcal{Y}$ and all $S, S' \in \mathcal{F}_\mathcal{X}^+$, we have:

$$P\{M(\boldsymbol{x}) \in R | \boldsymbol{x} \in S\} \leq e^{\varepsilon} P\{M(\boldsymbol{x}) \in R | \boldsymbol{x} \in S'\}. \tag{5}$$

We show in Proposition 1 that ε-LDP and ε-BDP can be viewed as equivalent, modulo some technicalities.

Proposition 1. *If M is ε-LDP then (π, M) is ε-BDP. Conversely, if (π, M) is ε-BDP, then (assuming $\mathcal{Y}$ is a Polish space) there exists a mechanism M' which is ε-LDP such that $P\{M(\boldsymbol{x}) = M'(\boldsymbol{x})\} = 1$.*

The reverse direction of the equivalence above relies on $\mathcal{Y}$ being a Polish space, which should capture all settings of practical interest. The proof can be found in Appendix A.1.

As discussed in Kifer and Machanavajjhala [2014], we can interpret the BDP definition in terms of odds ratios. The *prior odds* of two disjoint events S and S' is given by $\pi(S)/\pi(S')$. Now, for a set $R \in \mathcal{F}_\mathcal{Y}$ with positive measure, the *posterior odds*, after observing mechanism output $M(\boldsymbol{x}) \in R$, is equal to

$$\frac{P\{\boldsymbol{x} \in S | M(\boldsymbol{x}) \in R\}}{P\{\boldsymbol{x} \in S' | M(\boldsymbol{x}) \in R\}}. \tag{6}$$

One can easily verify that the BDP definition implies that the ratio of posterior and prior odds, also known as the *odds ratio*, is in the interval $[e^{-\varepsilon}, e^\varepsilon]$. In other words, the BDP definition limits how much the output of the mechanism changes the odds, which in turn limits our ability to distinguish whether the data $\boldsymbol{x}$ was in S or S'.

Coordinate differential privacy: We wish to measure the privacy loss of each coordinate incurred by a mechanism M. Initially, one might hypothesize that by altering the definition of LDP to reflect changes in specific coordinates, we could achieve privacy protection for certain features of the user's data. More precisely, consider the following definition of Coordinate DP (CDP)[1] where $\boldsymbol{c}$ represents a vector of the desired levels of privacy for each coordinate.

Definition 3 (Coordinate DP). Let $\boldsymbol{c} \in \mathbb{R}^d_{\geq 0}$. A mechanism M is $\boldsymbol{c}$-CDP if, for all $R \in \mathcal{F}_\mathcal{Y}$ and all $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{X}$, we have the implication:

$$\boldsymbol{x}_{-i} = \boldsymbol{x}'_{-i} \implies \mu_{\boldsymbol{x}}(R) \leq e^{c_i} \mu_{\boldsymbol{x}'}(R). \tag{7}$$

A special case of CDP, where all c_i's are equal, is referred to as attribute DP in the literature Kifer and Machanavajjhala [2011], Kenthapadi et al. [2013], Ghazi et al. [2022]. Additionally, CDP itself can be viewed as a special case of Partial DP, introduced in Ghazi et al. [2022].

The drawback of this definition is that it does not account for potential correlation among the data coordinates and, therefore, does not provide the type of guarantee we hope for regarding sensitive parts of the data. In other words, in $\boldsymbol{c}$-CDP, stating that the privacy level of coordinate i is c_i could be misleading. For instance, having $c_i = 0$ does not necessarily ensure that no information is leaked about x_i, as the other (potentially less-well protected) coordinates $\boldsymbol{x}_{-i}$ may be correlated with it.

Our formulation: Bayesian coordinate differential privacy: In order to capture the correlation among the data coordinates, we adopt a Bayesian interpretation of differential privacy. As seen in Proposition 1, BDP and LDP are closely related and this serves as the main motivation for our definition.

For $i \in [d]$, define

$$\mathcal{F}^+_{\mathcal{X}_i} = \{S_i \in \mathcal{F}_{\mathcal{X}_i} | \pi\{x_i \in S_i\} > 0\}. \tag{8}$$

---

[1]The acronym CDP is not to be confused with Concentrated DP, which is not considered in this work.

For $S_i \in \mathcal{F}_{\mathcal{X}_i}^+$, we have

$$P\{M(\boldsymbol{x}) \in R | x_i \in S_i\} = \frac{P((S_i \times \mathcal{X}_{-i}) \times R)}{\pi(S_i \times \mathcal{X}_{-i})}, \tag{9}$$

where $\mathcal{X}_{-i} = \times_{j \in [d], j \neq i} \mathcal{X}_j$. Thus, similar to Definition 2, we have a natural version of privacy for each coordinate which we shall refer to as Bayesian Coordinate DP (BCDP).

Definition 4 (Bayesian Coordinate DP). Let $\boldsymbol{\delta} \in \mathbb{R}_{\geq 0}^d$. A pair (π, M) is $\boldsymbol{\delta}$-BCDP if, for each $i \in [d]$, and for all $R \in \mathcal{F}_{\mathcal{Y}}$, $S_i, S_i' \in \mathcal{F}_{\mathcal{X}_i}^+$, we have:

$$P\{M(\boldsymbol{x}) \in R | x_i \in S_i\} \leq e^{\delta_i} P\{M(\boldsymbol{x}) \in R | x_i \in S_i'\}. \tag{10}$$

Unlike the Coordinate-DP definition, this definition accounts for potential correlation between coordinates through incorporation of the prior π.

It is worth noting that, similar to the BDP definition, we can interpret the BCDP definition in terms of the odds ratio, with the difference that we consider the two events $\{x_i \in S_i\}$ and $\{x_i \in S_i'\}$. In other words, having a small δ_i guarantees that the mechanism's output does not provide much useful information for discriminating between the possibilities that the underlying data has coordinate x_i in S_i or S_i'. This aligns with our goal of proposing a new formulation that characterizes how much information is revealed about each particular coordinate. We further discuss this interpretation in the next section.

3   Properties of BCDP

In this section, we discuss a number of properties following from the BCDP definition. Figure 1 compiles some of the results from Section 2 and Section 3 on how the different notions of Local DP (LDP), Coordinate-DP (CDP), Bayesian DP (BDP), and Bayesian Coordinate-DP (BCDP) are interrelated.

3.1   BCDP through the Lens of Hypothesis Testing

In this subsection, we present an interpretation of what the privacy vector $\boldsymbol{\delta}$ quantifies in $\boldsymbol{\delta}$-BCDP. We begin by revisiting the connection between hypothesis testing and the definition of differential privacy, and demonstrate how this connection extends to the $\boldsymbol{\delta}$-BCDP definition. For an ε-LDP mechanism M represented by Markov kernel μ, suppose an adversary observes a realization of the output Y and is tasked with determining whether the input was $\boldsymbol{x}$ or $\boldsymbol{x}'$. This scenario can be formulated as the following binary hypothesis problem:

$$H_0 : Y \sim \mu_{\boldsymbol{x}}(\cdot),$$
$$H_1 : Y \sim \mu_{\boldsymbol{x}'}(\cdot).$$

Let the rejection region be $R \subseteq \mathcal{Y}$. Denote the type I error (false alarm) rate as $\alpha(R, M, \boldsymbol{x}, \boldsymbol{x}')$

and the type II error (missed detection) rate as $\beta(R, M, \boldsymbol{x}, \boldsymbol{x}')$. We restate a known result below [Kairouz et al., 2015, Wasserman and Zhou, 2010].

Proposition 2. *A mechanism $M(\cdot)$ is ε-LDP iff the following condition holds for any $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{X}$ and any $R \in \mathcal{F}_{\mathcal{Y}}$:*

$$e^{\varepsilon} \alpha(R, M, \boldsymbol{x}, \boldsymbol{x}') + \beta(R, M, \boldsymbol{x}, \boldsymbol{x}') \geq 1. \tag{11}$$

We note that the criterion (11) can be equivalently replaced by the criterion

$$\alpha(R, M, \boldsymbol{x}, \boldsymbol{x}') + e^{\varepsilon} \beta(R, M, \boldsymbol{x}, \boldsymbol{x}') \geq 1.$$

As noted in Kairouz et al. [2015], this operational perspective shows that it is impossible to force type I and type II error rates to simultaneously vanish for a DP algorithm, and the reverse implication also holds.

For BCDP: Now, consider a $\boldsymbol{\delta}$-BCDP mechanism M. Suppose an adversary has knowledge of the prior π and again observes the algorithm's output Y. For any $i \in [d]$, the adversary aims to distinguish whether the output was generated from data where the i-th coordinate belongs to the set S_i or S'_i. In other words, we consider the following binary hypothesis testing problem:

$$H_0^i : Y \sim P\{\cdot \,|x_i \in S_i\},$$
$$H_1^i : Y \sim P\{\cdot \,|x_i \in S'_i\}.$$

Let the rejection region for the i-th hypothesis test be $R_i \in \mathcal{F}_{\mathcal{Y}}$. For the i-th test, denote the type I and type II error rates as $\alpha(R_i, M, S_i, S'_i)$ and $\beta(R_i, M, S_i, S'_i)$. Then, the following result holds:

Theorem 1. *A pair (π, M) is $\boldsymbol{\delta}$-BCDP iff the following condition holds for every coordinate $i \in [d]$:*

$$e^{\delta_i} \alpha(R_i, M, S_i, S'_i) + \beta(R_i, M, S_i, S'_i) \geq 1, \tag{12}$$

for any $S_i, S'_i \in \mathcal{F}_{\mathcal{X}_i}^+$ and $R_i \in \mathcal{F}_{\mathcal{Y}}$.

We note that the criteria (12) can be equivalently replaced by the criteria

$$\alpha(R_i, M, S_i, S'_i) + e^{\delta_i} \beta(R_i, M, S_i, S'_i) \geq 1.$$

The proof of Theorem 1 is presented in Appendix A.2. Thus, the theorem shows that the vector $\boldsymbol{\delta}$ in the $\boldsymbol{\delta}$-BCDP formulation captures the privacy loss for each of the coordinates in the same way ε in ε-DP captures the privacy loss of the entire data. In contrast, the vector $\boldsymbol{c}$ in $\boldsymbol{c}$-CDP does not have such an interpretation of privacy loss across the coordinates.
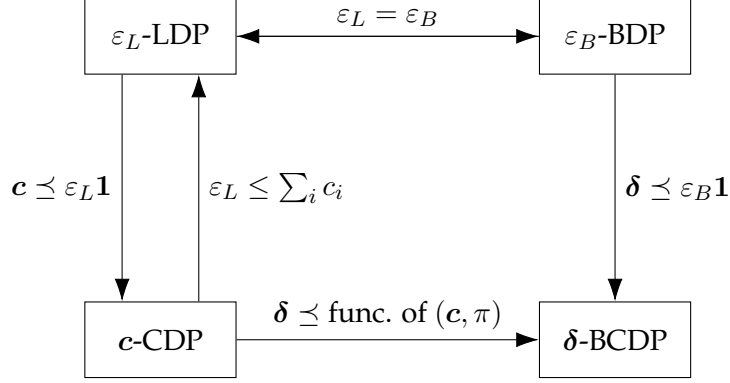
Figure 1: General relation of LDP (Definition 1), BDP (Definition 2), CDP (Definition 3) and BCDP (Definition 4). The condition $\boldsymbol{a} \preceq \boldsymbol{b}$ is to be read as $a_i \leq b_i \; \forall i$. The implication arrows and the described transformation of the parameters are to be interpreted as sufficient condition. For example, an ε_L-DP mechanism is guaranteed to be $\boldsymbol{c}$-CDP for $\boldsymbol{c} \preceq \varepsilon_L \mathbf{1}$. The function that translates $\boldsymbol{c}$-CDP under the prior π to BCDP is presented in the Proposition 4.

3.2 BCDP can be achieved via LDP

We begin by a simple result that follows directly from ε-BDP definition. If a mechanism is ε-LDP, then it is also ε-BDP. Noticing that ε-BDP ensures a ratio of e^ε for each coordinate as well in (10), we see that each coordinate has a privacy level of ε as well. Proposition 3 formalizes this observation and the proof can be found in Appendix A.3.

Proposition 3. (LDP to BCDP) *An ε-LDP mechanism satisfies $\varepsilon\mathbf{1}$-BCDP.*

Thus, a baseline approach to implement $\boldsymbol{\delta}$-BCDP is to use a $(\min_i \delta_i)$-LDP mechanism. Nevertheless, this would mean that we provide the most conservative privacy guarantee required for one coordinate to all coordinates. However, as we will show, we can maintain an $\varepsilon > \min \boldsymbol{\delta}$ for the overall privacy guarantee while providing the $\boldsymbol{\delta}$-BCDP guarantee, therefore achieving lower error for downstream tasks.

3.3 BCDP does not imply LDP

It should be noted that BCDP does not ensure LDP, BDP, or CDP. Example 1 below demonstrates a mechanism that is BCDP, but not LDP.

Example 1. Consider data $\boldsymbol{x} = (x_1, x_2)$ where x_1, x_2 are i.i.d. Bern$(\frac{1}{2})$, and let $\mathcal{Y} = \{0, 1\}$. Let M be defined by the table below, with parameters $a, b, c \in (0, 1]$. Mechanism M is not LDP (and hence, neither BDP nor CDP). However, for BCDP, the terms $P\{M(\boldsymbol{x}) = y | x_i = 0\}$ and $P\{M(\boldsymbol{x}) = y | x_i = 1\}$ have a finite multiplicative ranges in $y = 0, 1$, for both $i = 1, 2$. For instance, if $a = b = c = 1/2$, the resulting M is $(\log(2), \log(2))$-BCDP.

$P\{M(\boldsymbol{x}) = 1\|\boldsymbol{x}\}$		x_1	
		0	1
x_2	0	a	b
	1	0	c

The takeaway from the above example is that there can be priors and mechanisms such that the prior-mechanism pair is δ-BCDP but not ε-DP for any value of $\varepsilon < \infty$. Intuitively, BCDP quantifies privacy loss for each coordinate. In Example 1, one cannot distinguish much about any particular coordinate when the output is observed due to the BCDP guarantee. However, one may still be able to make conclusions over the coordinates jointly. For instance, if the mechanism output in Example 1 is 1, then the input could not have been $\boldsymbol{x} = (0, 1)$.

As this result suggests, just ensuring BCDP is not sufficient to guarantee overall privacy. In other words, the right way to think of BCDP is that it is a tool that allows us to quantify the privacy offered by an algorithm for each coordinate. The user can demand overall ε-LDP along with δ-BCDP where the value δ_i could be significantly smaller than ε for those coordinates that the user feels more concerned about their privacy.

3.4 Achieving BCDP via CDP

As we discussed earlier, CDP does not imply the BCDP, as it does not take into account the correlation among the features. Therefore, a natural question arises: could CDP, along with some bound on the dependence among the features, lead to a bound with respect to the BCDP definition? The next result provides an answer to this question. We denote the total variation distance between two distributions P and Q as $\mathsf{TV}(P, Q)$.

Proposition 4. (CDP to BCDP) *Suppose mechanism M is $\boldsymbol{c}$-CDP and there exists $q_1, \ldots, q_d$ such that*

$$\mathsf{TV}(\pi_{\boldsymbol{x}_{-i}|x_i \in S_i}, \pi_{\boldsymbol{x}_{-i}|x_i \in S_i'}) \leq q_i \ \forall \ S_i, S_i' \in \mathcal{F}_{\mathcal{X}_i}^+. \tag{13}$$

Then:

1. *The mechanism is $\sum_{i=1}^{n} c_i$-LDP and $\boldsymbol{\delta}$-BCDP with $\delta_i = c_i + \log(1 + q_i e^{\sum_{j \neq i} c_j} - q_i)$.*

2. *In case that the mechanism is ε-LDP for some $\varepsilon \geq 0$, then it is $\boldsymbol{\delta}'$-BCDP with $\delta_i' = \min\{c_i + \log(1 + q_i e^\varepsilon - q_i), \varepsilon\}$.*

For the proof, see Appendix A.4. To better highlight this result, consider a simple example in which that the input's d coordinates are redundant copies of the same value and we have access to a $\boldsymbol{c}$-CDP mechanism M. Then, as Proposition 4 suggests, and given that we have $q_i = 1$ here, this mechanism is $\boldsymbol{\delta}$-BCDP with $\delta_i = \sum_j c_j$. This is, of course, intuitive, as all coordinates reveal

information about the same value. It also reinforces that the BCDP is sensitive to correlations, unlike CDP. The first part of Proposition 4 provides an immediate method for constructing a BCDP mechanism using mechanisms that offer CDP guarantees, such as applying off-the-shelf LDP mechanisms per coordinate. This can also be done without requiring full knowledge of the prior. The second part of Proposition 4 suggests that in cases where we design an algorithm that has a tighter LDP guarantee than just $\sum_{i=1}^{n} c_i$, we can further improve the BCDP bound. One example of such a construction is provided in our private mean estimation algorithm in Section 4. One challenge with this approach is translating the constraint $\boldsymbol{\delta}$ to a vector $\boldsymbol{c}$. In Appendix A.5, we provide a tractable approach to do so by imposing a linear constraint on $\boldsymbol{c}$.

3.5 Do Composition and Post-Prossesing Hold for BCDP?

Composition is a widely used and well-known property of LDP. Nonetheless, as the following example highlights, the composition, as traditionally applied in differential privacy settings, does not hold in the BCDP setting.

Example 2. Let $\boldsymbol{x} \in \mathbb{R}^3$ be a Bernoulli vector where coordinates' distributions are independent and each one is $\mathrm{Ber}(1/2)$. Consider the mechanism

$$M(\boldsymbol{x}) = \begin{cases} [x_1 \oplus x_2, x_3]^\top & \text{with probability } \frac{1}{2} \\ [x_2, x_1 \oplus x_3]^\top & \text{with probability } \frac{1}{2}, \end{cases} \tag{14}$$

where $\oplus$ denotes the sum in $\mathbb{Z}_2$. It is straightforward to verify that $M(\cdot)$ is in fact $\delta_1 = 0$-BCDP with respect to the first coordinate. However, having two independent copies of $M(\cdot)$ is not BCDP with respect to the first coordinate. To see this, consider the case where $x_1 = 1$ and, in two copies of $M(\boldsymbol{x})$, one adds x_1 to x_2 and the other adds it to x_3. By observing the two copies, we can determine $x_1 = 1$.

On the other hand, the post-processing property holds for BCDP (see Appendix A.6 for the proof).

Proposition 5. (Post processing for BCDP) *Let $M(\cdot)$ be a $\boldsymbol{\delta}$-BCDP mechanism and $K : \mathcal{Y} \to \mathcal{Z}$ be a measurable function. Then, $K \circ M$ is also $\boldsymbol{\delta}$-BCDP.*

4 Private Mean Estimation

To demonstrate the practical use of the BCDP framework, we consider the private mean estimation problem.

Problem Setup: Let there be n users submitting their data to a central server via Local DP mechanisms. Let user j's data be $\boldsymbol{x}^{(j)} \in [-1, 1]^d$, where $\boldsymbol{x}^{(j)}$ is drawn from prior $\pi^{(j)}$ on $[-1, 1]^d$; the server and users are not assumed to know the priors (which may be different for each individual

Mechanism 1 Proposed locally private mechanism $M_{\mathsf{mean}}(\boldsymbol{x}, \boldsymbol{c})$

Input: user data $\boldsymbol{x} \in [-1, 1]^d$ and non-decreasing vector $\boldsymbol{c}$.
$c_0 \leftarrow 0$
For $k \in [d]$, set $w_k = \frac{(c_k - c_{k-1})^2}{(d-k+1)}$
for $k \in [d]$ **do**
$\quad \boldsymbol{Y}^k \leftarrow M_{\mathsf{LDP}}(\boldsymbol{x}_{k:d}, c_k - c_{k-1}, \sqrt{d-k+1}) \in \mathbb{R}^{d-k+1}$
end for
$\hat{\nu}_i \leftarrow (\sum_{k=1}^{i} Y_{i+1-k}^k w_k)/(\sum_{k=1}^{i} w_k)$ for $i \in [d]$
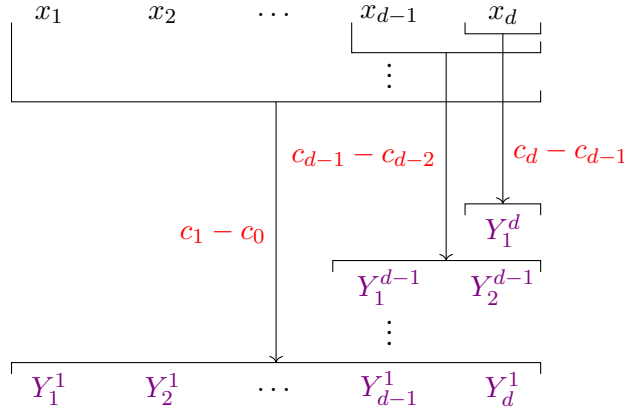Return $\hat{\boldsymbol{\nu}}$

Figure 2: The figure illustrates how the mechanism M_{mean} in Mechanism 1 obtains the estimate $\hat{\nu}$. For example, the vector Y^k is obtained by taking the vector $(x_k, x_{k+1}, \ldots, x_d)$ and using the LDP channel M_{LDP} with privacy parameter $c_k - c_{k-1}$. $\hat{\nu}_i$ is obtained by taking a linear combination of the component of $\{Y^k\}_{k \in [d]}$ corresponding to x_i, i.e., directly below x_i in the figure.

user). Our goal is to estimate the empirical mean $\bar{\boldsymbol{x}} = \sum_{j=1}^{n} \frac{\boldsymbol{x}^{(j)}}{n} \in [-1, 1]^d$ under ε-LDP and $\boldsymbol{\delta}$-BCDP. We consider l_2-norm as our error metric. Without loss of generality, we assume that $\boldsymbol{\delta}$ is in non-decreasing order. We make the following assumption regarding the priors.

Assumption 1. For every $i \in [d-1]$ and every $j \in [n]$, we have

$$\mathsf{TV}(\pi^{(j)}_{x_{-i}|x_i \in S_i}, \pi^{(j)}_{x_{-i}|x_i \in S'_i}) \leq q \quad \forall S_i, S'_i \in \mathcal{F}^+_{\mathcal{X}_i}. \tag{15}$$

Here q can be interpreted as a measure of how much different coordinates are interdependent. For instance, in the case where coordinates are independent we have $q = 0$, and in the case where they are fully equal, we have $q = 1$. Note that we do not need (15) to hold for the last coordinate ($i = d$). While this slight relaxation does not play a role in this section, it will be useful when we move to the private optimization application.

It is also worth emphasizing that our analysis extends to cases where we have heterogeneous bounds across different coordinates (similar to (13)) and different across users. However, we opt for a universal bound for the sake of simplifying the presentation of results. We assume that the upper bound q is known to the designer of the local channels, which can be either the end user

Mechanism 2 Local DP channel for data in $\mathcal{B}_2(r)$ proposed by Duchi et al. [2013a]

Input: user data $\boldsymbol{v}$, Local-DP level $\alpha \geq 0$, and bound r such that $\boldsymbol{v} \in \mathcal{B}_2(r)$.

$d \leftarrow \text{dimension}(\boldsymbol{v})$

$B \leftarrow \sqrt{\pi} dr \frac{e^{\alpha}+1}{e^{\alpha}-1} \frac{\Gamma(\frac{d+1}{2})}{\Gamma(\frac{d}{2}+1)}$

$K \sim \text{Bern}(\frac{e^{\alpha}}{e^{\alpha}+1})$

$S \sim \text{Bern}(\frac{1}{2} + \frac{\|\boldsymbol{v}\|_2}{2r})$

$\tilde{\boldsymbol{v}} \leftarrow (2S-1)\frac{r\boldsymbol{v}}{\|\boldsymbol{v}\|_2}$

Return $Z \sim \text{Uniform}(\boldsymbol{z} \in \mathbb{R}^d : (2K-1)\boldsymbol{z}^T\tilde{\boldsymbol{v}} > 0, \|\boldsymbol{z}\|_2 = B)$

---

themselves or the central server.

Our proposed privacy mechanism M_{mean} (Mechanism 1) builds upon any LDP mechanism M_{LDP}, capable of handling different dimensional inputs, that can be used for mean estimation as a black-box tool. While Algorithm 1 satisfies the privacy constraints for any black-box LDP mechanism M_{LDP}, we shall focus on the mechanism presented by Duchi et al. [2013a] for proving error bounds. The mechanism of Duchi et al. [2013a], which is known to be minimax optimal for mean estimation, is outlined in Mechanism 2 for completeness. A figure explaining M_{mean} is presented in Figure 2.

Theorem 2 shows that our proposed mechanism satisfies the desired privacy constraints and provides a bound on the mean-squared error (MSE). The notation $a \wedge b$ refers to $\min\{a, b\}$. later in this section we present a refined result with a more interpretable MSE bound in Corollary 1.

Theorem 2. *Suppose Assumption 1 holds and let $\tilde{\delta}_i := \min\{\delta_i, \varepsilon\}$ for any $i \in [d]$. Then, M_{mean}, i.e., Mechanism 1, with parameters*

$$c_d = \min\left\{\log\left(\frac{e^{\zeta\tilde{\delta}_1} + q - 1}{q}\right), \tilde{\delta}_d\right\}, \; c_0 = 0, \tag{16}$$

$$\forall i < d, \; c_i = \begin{cases} c_d & \text{if } c_d \leq \tilde{\delta}_i, \\ \tilde{\delta}_i - \log(1 + qe^{c_d} - q) & \text{otherwise}, \end{cases}$$

where $\zeta \in (0, 1]$ is a free parameter, is $\boldsymbol{\delta}$-BCDP and ε-LDP. Moreover, using Mechanism 2 as M_{LDP}, we obtain the following mean-squared error bound

$$\mathbb{E}\left[\left\|\bar{\boldsymbol{x}} - \Pi\left(\frac{1}{n}\sum_{j=1}^{n} M_{\text{mean}}(\boldsymbol{x}^{(j)}, \boldsymbol{c})\right)\right\|_2^2 \middle| \boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(n)}\right] \lesssim \frac{1}{n}\sum_{i=1}^{d}\left(\frac{1}{\sum_{k=1}^{i}\frac{(c_k - c_{k-1})^2}{d-k+1}}\right) \wedge d, \tag{17}$$

where $\Pi(\cdot)$ denotes projection into $[-1, 1]^d$ and the expectation is taken over the randomness in $M_{\text{mean}}(\cdot)$.

The proof is presented in Appendix B.2. Here, we provide a proof sketch for the privacy guarantee. We first establish that M_{mean} is $\boldsymbol{c}$-CDP and c_d-LDP. The latter implies that M_{mean} also satisfies $c_d\mathbf{1}$-BCDP. Thus, if $\tilde{\delta}_i \geq c_d$, we have already shown that M_{mean} is δ_i-BCDP with respect to

13

the i-th coordinate. For the case where $c_d > \tilde{\delta}_i$, we use Proposition 4, which essentially implies that M_{mean} is $c_i + \log(1 + qe^{c_d} - q)$-BCDP with respect to the i-th coordinate. Substituting the value of c_i from the statement of Theorem 2 gives us the desired δ_i-BCDP guarantee for coordinate i. Finally, the choice of c_d in (16) ensures that the proposed c_i's are all non-negative and valid.

We next make a few remarks regarding this result. *First,* note that while this bound applies to empirical mean estimation, when the data points $\{x^{(j)}\}_{j=1}^n$ are independent and identically distributed, the result can also be used to derive an out-of-sample error bound for the statistical private mean estimation problem. In this case, an additional $\frac{d}{n}$ term would be added to the error bound. *Secondly,* note that in the special case where there is no coordinate-level privacy requirement (i.e., $\delta_i = \infty$), setting $\zeta = 1$ implies $c_1 = \ldots = c_d = \varepsilon$, which yields the error bound $\mathcal{O}(d^2/(n\varepsilon^2))$. This matches the ε-LDP lower bound provided in Duchi et al. [2013a]. *Lastly,* we would like to discuss the role of parameter ζ. Note that when the correlation parameter q is sufficiently high or the most sensitive privacy requirement δ_1 is low enough such that c_d is determined by the first term, $\log((e^{\zeta\tilde{\delta}_1} + q - 1)/q)$, in (16), one can verify that $c_1 = (1 - \zeta)\delta_1$. As ζ increases, c_1 decreases, leading to a worse estimate for the first coordinate. However, c_d increases, indicating that, for the less-sensitive coordinates, we are increasing the privacy parameter, potentially resulting in lower estimation error for those coordinates. In other words, the parameter ζ allows us to control how we allocate the privacy budget of the most sensitive coordinate, δ_1, between the estimation of the first coordinate and the privacy cost imposed by its correlation with other coordinates.

We next present a corollary that studies a special case where the coordinates are divided into more sensitive and less sensitive groups. In this case, users requests a general ε-LDP guarantee for all coordinates and a δ-BCDP guarantee specifically for the more sensitive coordinates. Under this scenario, we further simplify the error bound which allows us to better highlight implications of Theorem 2.

Corollary 1. *Suppose we are under the premise of Theorem 2, with $\delta_i = \delta$ for $1 \le i \le k$ and $\delta_i = \varepsilon > 2\delta$ for $d \ge i > k$. Then, by choosing $\zeta = 0.5$, the mean-squared error upper bound in (17) simplifies to*

$$\mathcal{O}(1)\frac{1}{n} \cdot \left(\frac{dk}{\delta^2} + \frac{(d-k)^2}{\varepsilon^2} \right), \tag{18}$$

for the case $q \le \frac{e^{\delta/2}-1}{e^\varepsilon-1}$, and to

$$\mathcal{O}(1)\frac{1}{n} \cdot \left(\frac{dk}{\delta^2} + \frac{(d-k)^2}{\frac{d-k}{d}\delta^2 + (c_d - \delta/2)^2} \right), \tag{19}$$

otherwise, with $c_d = \log(\frac{e^{\delta/2}+q-1}{q})$ decreasing from ε to $\delta/2$ as q increases from $(e^{\delta/2} - 1)/(e^\varepsilon - 1)$ to 1.

Under the special structure of δ and ε in Corollary 1, the MSE can be thought of as sum of MSE of the more sensitive and the less sensitive coordinates. For low levels of correlation, the MSE of less private coordinates behave like the MSE of an ε-LDP mechanism on $d - k$ dimensional vector but the MSE of the more sensitive part *depends on the dimension d of the whole vector,* not

just the sensitive part, showing how the more sensitive part of the data affects the whole vector. As the correlation increases $q \to 1$, we have $c_d \to \delta/2$ and the MSE matches that of a $\min \boldsymbol{\delta}$-LDP mechanism. A supporting experiment is deferred to Appendix B.1.

5   Private Least-Squares Regression

We focus on how the BCDP framework can be employed on least-squares problems. Consider a setting in which n datapoints are distributed among n users, where $\boldsymbol{x}^{(i)} = [\boldsymbol{z}^{(i)^\top} \, l^{(i)}]^\top \in \mathbb{R}^d$ represents the data of user i, with $\boldsymbol{z}^{(i)}$ as the feature vector for the i-th data point and $l^{(i)}$ as its label. Our goal is to solve the following empirical risk minimization problem privately over the data of n users:

$$\min_{\boldsymbol{\theta} \in \Theta} f(\boldsymbol{\theta}) := \frac{1}{n} \sum_{i=1}^{n} \ell(\boldsymbol{\theta}, \boldsymbol{x}^{(i)}), \tag{20}$$

where $\ell(\boldsymbol{\theta}, \boldsymbol{x}^{(i)}) := \frac{1}{2}(\boldsymbol{\theta}^\top \boldsymbol{z}^{(i)} - l^{(i)})^2$ and $\Theta \subset \mathbb{R}^{d-1}$ is a compact convex set. We denote the solution of (20) by f^*, achieved at some point $\boldsymbol{\theta}^*$.

The customary approach in private convex optimization is to use a gradient-based method, perturbing the gradient at each iteration to satisfy the privacy requirements. However, this approach is not suitable for cases in which coordinate-based privacy guarantees are needed. The difficulty is that each coordinate of the gradient is typically influenced by all coordinates of the data. As a result, and it becomes challenging to distinguish the level of privacy offered to different coordinates of the data. Instead, we first perturb the users' data locally and then compute the gradient at the perturbed point to update the model at the server.

However, computing the gradient based on perturbed data introduces its own challenges. In particular, the gradient of $\ell(\boldsymbol{\theta}, \boldsymbol{x})$ with $\boldsymbol{x} = [\boldsymbol{z}^\top l]^\top$, given by

$$\nabla \ell(\boldsymbol{\theta}, \boldsymbol{x}) = \boldsymbol{z}\boldsymbol{z}^\top \boldsymbol{\theta} - l\boldsymbol{z}, \tag{21}$$

is not a linear function of the data. As a result, although the privacy mechanism outputs an unbiased estimate of the true data, this non-linearity introduces an error in the gradient with a non-zero mean, preventing the algorithm from converging to the optimal solution. To overcome this challenge, we observe that the aforementioned non-linearity in the data arises from the product terms $\boldsymbol{z}\boldsymbol{z}^\top$ and $l\boldsymbol{z}$. Therefore, if we create two private copies of the data and replace each term of the product with one of the copies, we obtain a stochastic unbiased estimate of the gradient at the server. The catch, of course, is that we need to divide our privacy budget in half.

A summary of the algorithm is presented in Algorithm 3. There, $\text{OPT}(g, \delta)$ refers to any convex optimization algorithm of choice, e.g., gradient descent, that finds the minimizer of the function $g(\cdot)$ over the set Θ and up to an error of δ. We make the assumption that the data is bounded.

Assumption 2. $\boldsymbol{x}^{(i)} \in [-1, 1]^d$ for all $i \in [n]$.

We also need Assumption 1, which bounds the dependence of each coordinate on the other

Algorithm 3 BCDP and LDP least-squares regression

Input: Users' data $\{\boldsymbol{x}^{(i)}\}_{i=1}^n$ and vector $\boldsymbol{c}$
for $i \in [n]$ **do**
    //user i sends locally perturbed data to the server
    $\hat{\boldsymbol{x}}^{(i,1)} \leftarrow M_{\mathsf{mean}}(\boldsymbol{x}^{(i)}, \boldsymbol{c}/2)$
    $\hat{\boldsymbol{x}}^{(i,2)} \leftarrow M_{\mathsf{mean}}(\boldsymbol{x}^{(i)}, \boldsymbol{c}/2)$
end for
//server optimizes with the perturbed data
$\hat{f}(\boldsymbol{\theta}) := \frac{1}{2n}\sum_{i=1}^n \left(\boldsymbol{\theta}^\top \hat{\boldsymbol{z}}^{(i,1)} \hat{\boldsymbol{z}}^{(i,2)^\top} \boldsymbol{\theta} - 2\boldsymbol{\theta}^\top \hat{l}^{(i,1)} \hat{\boldsymbol{z}}^{(i,2)} \right)$
$\hat{\boldsymbol{\theta}} \leftarrow \mathrm{OPT}(\hat{f}, \frac{1}{n})$
Return $\hat{\boldsymbol{\theta}}$

---

coordinates of the data. Recall that the last coordinate is exempt from this assumption, which is why we position the label as the last coordinate of the data. It is not reasonable to assume the feature vector is weakly correlated with the label, as that would violate the premise of regression. We next state the following result on Algorithm 3.

Theorem 3. *Suppose Assumptions 1 and 2 hold. Then, Algorithm 3, with $\boldsymbol{c}$ set similar to Theorem 2, is δ-BCDP and ε-LDP. Moreover, using Mechanism 2 as M_{LDP}, we obtain the following error rate*

$$\mathbb{E}\left[f(\hat{\boldsymbol{\theta}}) - f^* \middle| \boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(n)} \right] \lesssim \max_{\boldsymbol{\theta} \in \Theta}(\|\boldsymbol{\theta}\|^2 + \|\boldsymbol{\theta}\|) \left(\frac{\log d}{\sqrt{n}}(r^2 + rd) \wedge d \right), \tag{22}$$

with $r^2 = \sum_{i=1}^d 1/(\sum_{k=1}^i \frac{(c_k - c_{k-1})^2}{d-k+1})$.

The proof of this theorem can be found in Appendix C.1. It is worth noting that r^2 appears in the mean-squared error bound in Theorem 2. That said, if we consider special cases similar to the setting in Corollary 1, the term r^2 would simplify in a similar manner.

We conclude this section by noting that, while we focus on linear regression here, our approach can provide an unbiased gradient and guarantee privacy in any setting where the loss function $\ell(\boldsymbol{\theta}, \boldsymbol{x})$ has a quadratic (or even low-degree polynomial) dependence on $\boldsymbol{x}$, such as neural networks with ℓ_2 loss. However, the optimization error in such cases may require further considerations, as non-convex loss functions could arise.

6 Acknowledgements

References

J. Acharya, K. Bonawitz, P. Kairouz, D. Ramage, and Z. Sun. Context aware local differential privacy. In *Proceedings of the 37th International Conference on Machine Learning*, pages 52–62. PMLR, 2020.

M. Alvim, K. Chatzikokolakis, C. Palamidessi, and A. Pazii. Invited paper: Local differential privacy on metric spaces: Optimizing the trade-off with utility. In *2018 IEEE 31st Computer Security Foundations Symposium (CSF)*, pages 262–267, 2018.

M. E. Andrés, N. E. Bordenabe, K. Chatzikokolakis, and C. Palamidessi. Geo-indistinguishability: Differential privacy for location-based systems. In *Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security*, pages 901–914, 2013.

H. Asi, V. Feldman, and K. Talwar. Optimal algorithms for mean estimation under local differential privacy. In *International Conference on Machine Learning*, pages 1046–1056. PMLR, 2022.

A. Beimel, K. Nissim, and E. Omri. Distributed private data analysis: Simultaneously solving how and what. In *Advances in Cryptology–CRYPTO 2008: 28th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 17-21, 2008. Proceedings 28*, pages 451–468. Springer, 2008.

A. Beimel, K. Nissim, and U. Stemmer. Private learning and sanitization: Pure vs. approximate differential privacy. In *International Workshop on Approximation Algorithms for Combinatorial Optimization*, pages 363–378. Springer, 2013.

S. Biswas, Y. Dong, G. Kamath, and J. Ullman. Coinpress: Practical private mean and covariance estimation. *Advances in Neural Information Processing Systems*, 33:14475–14485, 2020.

T.-H. H. Chan, E. Shi, and D. Song. Optimal lower bound for differentially private multi-party aggregation. In *Proceedings of the 20th Annual European Conference on Algorithms*, ESA'12, page 277–288. Springer-Verlag, 2012. ISBN 9783642330896.

K. Chatzikokolakis, M. E. Andrés, N. E. Bordenabe, and C. Palamidessi. Broadening the scope of differential privacy using metrics. In *Privacy Enhancing Technologies: 13th International Symposium, PETS 2013, Bloomington, IN, USA, July 10-12, 2013. Proceedings 13*, pages 82–102. Springer, 2013.

K. Chaudhuri and D. Hsu. Sample complexity bounds for differentially private learning. In *Proceedings of the 24th Annual Conference on Learning Theory*, pages 155–186. JMLR Workshop and Conference Proceedings, 2011.

D. Desfontaines. A list of real-world uses of differential privacy. https://desfontain.es/blog/real-world-differential-privacy.html, 2020. Accessed: 2024-05-22.

Differential Privacy Team, Apple. Learning with privacy at scale. https://machinelearning.apple.com/research/learning-with-privacy-at-scale, 2017. Accessed: 2024-05-22.

B. Ding, J. Kulkarni, and S. Yekhanin. Collecting telemetry data privately. *Advances in Neural Information Processing Systems*, 30, 2017.

J. C. Duchi, M. I. Jordan, and M. J. Wainwright. Local privacy, data processing inequalities, and statistical minimax rates. *arXiv preprint arXiv:1302.3203*, 2013a.

J. C. Duchi, M. I. Jordan, and M. J. Wainwright. Local privacy and statistical minimax rates. In *2013 IEEE 54th Annual Symposium on Foundations of Computer Science*, pages 429–438. IEEE, 2013b.

C. Dwork, F. McSherry, K. Nissim, and A. Smith. Calibrating noise to sensitivity in private data analysis. In *Theory of Cryptography: Third Theory of Cryptography Conference, TCC 2006, New York, NY, USA, March 4-7, 2006. Proceedings 3*, pages 265–284. Springer, 2006.

C. Dwork, A. Roth, et al. The algorithmic foundations of differential privacy. *Foundations and Trends® in Theoretical Computer Science*, 9(3–4):211–407, 2014.

Ú. Erlingsson, V. Pihur, and A. Korolova. Rappor: Randomized aggregatable privacy-preserving ordinal response. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*, pages 1054–1067, 2014.

A. Evfimievski, J. Gehrke, and R. Srikant. Limiting privacy breaches in privacy preserving data mining. In *Proceedings of the Twenty-second ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*, pages 211–222, 2003.

V. Feldman and K. Talwar. Lossless compression of efficient private local randomizers. In *International Conference on Machine Learning*, pages 3208–3219. PMLR, 2021.

B. Ghazi, N. Golowich, R. Kumar, P. Manurangsi, and C. Zhang. Deep learning with label differential privacy. *Advances in Neural Information Processing Systems*, 34:27131–27145, 2021.

B. Ghazi, R. Kumar, P. Manurangsi, and T. Steinke. Algorithms with more granular differential privacy guarantees. *arXiv preprint arXiv:2209.04053*, 2022.

X. He, A. Machanavajjhala, and B. Ding. Blowfish privacy: tuning privacy-utility trade-offs using policies. In *Proceedings of the 2014 ACM SIGMOD International Conference on Management of Data*, SIGMOD '14, page 1447–1458. Association for Computing Machinery, 2014. ISBN 9781450323765.

J. Imola, S. Kasiviswanathan, S. White, A. Aggarwal, and N. Teissier. Balancing utility and scalability in metric differential privacy. In *Uncertainty in Artificial Intelligence*, pages 885–894. PMLR, 2022.

P. Kairouz, S. Oh, and P. Viswanath. The composition theorem for differential privacy. In F. Bach and D. Blei, editors, *Proceedings of the 32nd International Conference on Machine Learning*, volume 37 of *Proceedings of Machine Learning Research*, pages 1376–1385, Lille, France, 07–09 Jul 2015. PMLR.

V. Karwa and S. Vadhan. Finite sample differentially private confidence intervals. *arXiv preprint arXiv:1711.03908*, 2017.

S. P. Kasiviswanathan and A. Smith. On the 'semantics' of differential privacy: A Bayesian formulation. *Journal of Privacy and Confidentiality*, 6(1), 2014.

S. P. Kasiviswanathan, H. K. Lee, K. Nissim, S. Raskhodnikova, and A. Smith. What can we learn privately? *SIAM Journal on Computing*, 40(3):793–826, 2011.

K. Kenthapadi, A. Korolova, I. Mironov, and N. Mishra. Privacy via the johnson-lindenstrauss transform. *Journal of Privacy and Confidentiality*, 5(1), Aug. 2013. doi: 10.29012/jpc.v5i1.625.

D. Kifer and A. Machanavajjhala. No free lunch in data privacy. In *Proceedings of the 2011 ACM SIGMOD International Conference on Management of data*, pages 193–204, 2011.

D. Kifer and A. Machanavajjhala. Pufferfish: A framework for mathematical privacy definitions. *ACM Transactions on Database Systems (TODS)*, 39(1):1–36, 2014.

D. Kifer, A. Smith, and A. Thakurta. Private convex empirical risk minimization and high-dimensional regression. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 25.1–25.40, Edinburgh, Scotland, 25–27 Jun 2012. PMLR. URL https://proceedings.mlr.press/v23/kifer12.html.

S. Mahloujifar, C. Guo, G. E. Suh, and K. Chaudhuri. Machine learning with feature differential privacy. In *Federated Learning and Analytics in Practice: Algorithms, Systems, Applications, and Opportunities*, 2023.

A. G. Thakurta, A. H. Vyrros, U. S. Vaishampayan, G. Kapoor, J. Freudiger, V. R. Sridhar, and D. Davidson. Learning new words, 2017.

A. Triastcyn and B. Faltings. Bayesian differential privacy for machine learning. In *International Conference on Machine Learning*, pages 9583–9592. PMLR, 2020.

J. A. Tropp. The expected norm of a sum of independent random matrices: An elementary approach. In *High Dimensional Probability VII: The Cargese Volume*, pages 173–202. Springer, 2016.

S. Vadhan. The complexity of differential privacy. *Tutorials on the Foundations of Cryptography: Dedicated to Oded Goldreich*, pages 347–450, 2017.

D. Vu and A. Slavkovic. Differential privacy for clinical trial data: Preliminary evaluations. In *2009 IEEE International Conference on Data Mining Workshops*, pages 138–143, 2009. doi: 10.1109/ ICDMW.2009.52.

D. Wang and J. Xu. On sparse linear regression in the local differential privacy model. In *International Conference on Machine Learning*, pages 6628–6637. PMLR, 2019.

S. L. Warner. Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, 60(309):63–69, 1965.

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

H. Xiao and S. Devadas. Pac privacy: Automatic privacy measurement and control of data processing. In *Annual International Cryptology Conference*, pages 611–644. Springer, 2023.

X. Xiong, S. Liu, D. Li, Z. Cai, and X. Niu. A comprehensive survey on local differential privacy. *Security and Communication Networks*, 2020:1–29, 2020.

M. Yang, T. Guo, T. Zhu, I. Tjuawinata, J. Zhao, and K.-Y. Lam. Local differential privacy and its applications: A comprehensive survey. *Computer Standards & Interfaces*, page 103827, 2023.

K. Zheng, W. Mou, and L. Wang. Collect at once, use effectively: Making non-interactive locally private learning possible. In *International Conference on Machine Learning*, pages 4130–4139. PMLR, 2017.

A Proofs for Section 2 and Section 3

A.1 Proof of Proposition 1

Proposition 1. *If M is ε-LDP then (π, M) is ε-BDP. Conversely, if (π, M) is ε-BDP, then (assuming $\mathcal{Y}$ is a Polish space) there exists a mechanism M' which is ε-LDP such that $P\{M(\boldsymbol{x}) = M'(\boldsymbol{x})\} = 1$.*

Proof. (A) Forward direction:

For any $R \in \mathcal{F}_{\mathcal{Y}}$ and $S, S' \in \mathcal{F}_{\mathcal{X}}^+$, we have

$$\mu_{\boldsymbol{x}}(R) \leq e^{\varepsilon} \mu_{\boldsymbol{x}'}(R) \tag{23}$$

$$\implies \int_{\boldsymbol{x} \in S} \mu_{\boldsymbol{x}}(R) d\pi(\boldsymbol{x}) \leq e^{\varepsilon} \int_{\boldsymbol{x} \in S} \mu_{\boldsymbol{x}'}(R) d\pi(\boldsymbol{x}) \tag{24}$$

$$\implies P\{M(\boldsymbol{x}) \in R, \boldsymbol{x} \in S\} \leq e^{\varepsilon} \pi(S) \mu_{\boldsymbol{x}'}(R), \tag{25}$$

$$\implies P\{M(\boldsymbol{x}) \in R | \boldsymbol{x} \in S\} \leq e^{\varepsilon} \mu_{\boldsymbol{x}'}(R). \tag{26}$$

Similarly integrate again $\int_{\boldsymbol{x}' \in S'} \cdot d\pi(\boldsymbol{x}')$ to get the desired result.

(B) Converse:

Fix $R \in \mathcal{F}_{\mathcal{Y}}$ and let

$$U(R) = \mathsf{ess\,sup}_{\boldsymbol{x} \in \mathcal{X}} \mu_{\boldsymbol{x}}(R), \tag{27}$$

$$L(R) = \mathsf{ess\,inf}_{\boldsymbol{x} \in \mathcal{X}} \mu_{\boldsymbol{x}}(R), \tag{28}$$

where the essential supremum and infimum are with respect to $(\mathcal{X}, \mathcal{F}_{\mathcal{X}}, \pi)$. Fix any $\delta > 0$ and define

$$S = \{\boldsymbol{x} \in \mathcal{X} : \mu_{\boldsymbol{x}}(R) > U(R) - \delta\}, \; S' = \{\boldsymbol{x} \in \mathcal{X} : \mu_{\boldsymbol{x}}(R) < L(R) + \delta\}. \tag{29}$$

By definition of essential supremum and infimum and using the fact that μ is a Markov kernel, $S, S' \in \mathcal{F}_{\mathcal{X}}^+$.

Using the ε-BDP condition, we can obtain

$$U(R) - \delta \leq \frac{1}{\pi(S)} \int_{\boldsymbol{x} \in S} \mu_{\boldsymbol{x}}(R) d\pi(\boldsymbol{x}) \tag{30}$$

$$\leq \frac{e^{\varepsilon}}{\pi(S')} \int_{\boldsymbol{x} \in S'} \mu_{\boldsymbol{x}}(R) d\pi(\boldsymbol{x}) \tag{31}$$

$$\leq e^{\varepsilon}(L(R) + \delta). \tag{32}$$

Since $\delta > 0$ is arbitrary, we get $U(R) \leq e^{\varepsilon} L(R) \; \forall R \in \mathcal{F}_{\mathcal{Y}}$.

Now, for R, define

$$E(R) := \{\boldsymbol{x} \in \mathcal{X} : \mu_{\boldsymbol{x}}(R) \notin [L(R), U(R)]\}, \tag{33}$$

and note that $\pi(E(R)) = 0$ by definition of essential supremum and infimum and using the fact that μ is a Markov kernel.

We shall now modify M on a judiciously chosen π-null set to obtain an ε-LDP mechanism M'. Since $\mathcal{Y}$ is second countable as it is a Polish space, there exists a countable collection of open sets $\mathcal{V} = \{V_i\}_{i \geq 1} \subset \mathcal{F}_\mathcal{Y}$ such that every open set $R \in \mathcal{F}_\mathcal{Y}$ can be written as a disjoint union of some countable subset of $\mathcal{V}$.

Define $E = \cup_{i \geq 1} E(V_i)$ and note that $\pi(E) = 0$ by countable sub-additivity. Fix any $\bar{\boldsymbol{x}} \in E^c$ and define the modification M' of M via

$$M'(\boldsymbol{x}) = \begin{cases} M(\bar{\boldsymbol{x}}) & \text{if } \boldsymbol{x} \in E \\ M(\boldsymbol{x}) & \text{otherwise.} \end{cases} \tag{34}$$

Equivalently, M' is represented by the Markov kernel $\mu' : \mathcal{X} \times \mathcal{F}_\mathcal{Y} \to [0,1]$ defined by

$$\mu'_{\boldsymbol{x}}(R) = \begin{cases} \mu_{\bar{\boldsymbol{x}}}(R) & \text{if } \boldsymbol{x} \in E \\ \mu_{\boldsymbol{x}}(R) & \text{otherwise.} \end{cases} \tag{35}$$

Since $\pi(E) = 0$, it is easy to see that $P\{M(\boldsymbol{x}) = M'(\boldsymbol{x})\} = 1$. By construction, since $E = \cup_{i \geq 1} E(V_i)$,

$$\mu'_{\boldsymbol{x}}(V_i) \leq U(V_i) \leq e^\varepsilon L(V_i) \leq e^\varepsilon \mu'_{\boldsymbol{x}'}(V_i) \ \forall \boldsymbol{x}, \boldsymbol{x}' \in \mathcal{X}, V_i \in \mathcal{V}. \tag{36}$$

By second countability, and σ-additivity, we conclude,

$$\mu'_{\boldsymbol{x}}(R) \leq e^\varepsilon \mu'_{\boldsymbol{x}'}(R) \ \forall \boldsymbol{x}, \boldsymbol{x}' \in \mathcal{X}, \forall \text{open } R \in \mathcal{F}_\mathcal{Y}. \tag{37}$$

Every second countable space is Radon, so by outer regularity of both measures $\mu_{\boldsymbol{x}}$ and $\mu'_{\boldsymbol{x}}$, we finally obtain

$$\mu'_{\boldsymbol{x}}(R) \leq e^\varepsilon \mu'_{\boldsymbol{x}'}(R) \ \forall \boldsymbol{x}, \boldsymbol{x}' \in \mathcal{X}, \forall R \in \mathcal{F}_\mathcal{Y}. \tag{38}$$

Hence, M' is ε-LDP as desired. $\qquad\square$

A.2 Proof of Theorem 1

Theorem 1. *A pair (π, M) is δ-BCDP iff the following condition holds for every coordinate $i \in [d]$:*

$$e^{\delta_i} \alpha(R_i, M, S_i, S'_i) + \beta(R_i, M, S_i, S'_i) \geq 1, \tag{12}$$

for any $S_i, S'_i \in \mathcal{F}^+_{\mathcal{X}_i}$ and $R_i \in \mathcal{F}_\mathcal{Y}$.

Proof. By definition $\alpha(R_i, M, S_i, S'_i) = P\{M(\boldsymbol{x}) \in R_i | \boldsymbol{x} \in S_i\}$ and

$$P\{M(\boldsymbol{x}) \in T_i | x_i \in S_i\} \leq e^{\delta_i} P\{M(\boldsymbol{x}) \in T_i | x_i \in S'_i\} \ \forall S_i, S'_i \in \mathcal{F}^+_{\mathcal{X}_i}, T_i \in \mathcal{F}_\mathcal{Y} \tag{39}$$

Setting $T_i = R_i^C$, get $\alpha(R_i, M, S_i, S'_i) + e^{\delta_i} \beta(R_i, M, S_i, S'_i) \geq 1$. One can also switch S_i, S'_i above and set $T_i = R_i$ to get $e^{\delta_i} \alpha(R_i, M, S_i, S'_i) + \beta(R_i, M, S_i, S'_i) \geq 1$.

The converse is straightforward as well.

$$e^{\delta_i}\alpha(R_i, M, S_i, S'_i) + \beta(R_i, M, S_i, S'_i) \geq 1 \ \forall S_i, S'_i \in \mathcal{F}^+_{\mathcal{X}_i}, R_i \in \mathcal{F}_{\mathcal{Y}} \tag{40}$$

$$\Leftrightarrow P\{M(\boldsymbol{x}) \in R_i | x_i \in S'_i\} \leq e^{\delta_i} P\{M(\boldsymbol{x}) \in R_i | x_i \in S_i\} \ \forall S_i, S'_i \in \mathcal{F}^+_{\mathcal{X}_i}, R_i \in \mathcal{F}_{\mathcal{Y}}. \tag{41}$$

$\square$

A.3 Proof of Proposition 3

Proposition 3. (LDP to BCDP) *An ε-LDP mechanism satisfies $\varepsilon\mathbf{1}$-BCDP.*

Proof. By Proposition 1, an ε-DP algorithm is ε-BDP as well. Recall $P\{M(\boldsymbol{x}) \in R | \boldsymbol{x} \in S\} = \frac{P(S \times R)}{\pi(S)}$. Thus, ε-BDP guarntee ensures that $\forall S, S' \in \mathcal{F}^+_{\mathcal{X}}$,

$$\frac{P(S \times R)}{\pi(S)} \leq e^{\varepsilon} \frac{P(S' \times R)}{\pi(S')}.$$

Restricting to the sets of form $S = S_i \times \mathcal{X}_{-i}$ and $S' = S'_i \times \mathcal{X}_{-i}$ for $S_i, S'_i \in \mathcal{F}^+_{\mathcal{X}_i}$ above yields the desired BCDP result.

$\square$

A.4 Proof of Proposition 4

Proposition 4. (CDP to BCDP) *Suppose mechanism M is $\boldsymbol{c}$-CDP and there exists $q_1, \ldots, q_d$ such that*

$$\mathsf{TV}(\pi_{\boldsymbol{x}_{-i}|x_i \in S_i}, \pi_{\boldsymbol{x}_{-i}|x_i \in S'_i}) \leq q_i \ \forall \ S_i, S'_i \in \mathcal{F}^+_{\mathcal{X}_i}. \tag{13}$$

Then:

1. *The mechanism is $\sum_{i=1}^n c_i$-LDP and $\boldsymbol{\delta}$-BCDP with $\delta_i = c_i + \log(1 + q_i e^{\sum_{j \neq i} c_j} - q_i)$.*

2. *In case that the mechanism is ε-LDP for some $\varepsilon \geq 0$, then it is $\boldsymbol{\delta}'$-BCDP with $\delta'_i = \min\{c_i + \log(1 + q_i e^{\varepsilon} - q_i), \varepsilon\}$.*

Proof. Consider the sets $R \in \mathcal{F}_{\mathcal{Y}}$ such that $P\{M(\boldsymbol{x}) \in R\} = 0$, then for $S_i, S'_i \in \mathcal{F}^+_{\mathcal{X}_i}$, we have $P\{M(\boldsymbol{x}) \in R | x_i \in S_i\} = P\{M(\boldsymbol{x}) \in R | x_i \in S'_i\} = 0$. Thus, to bound the $\boldsymbol{\delta}$ in $\boldsymbol{\delta}$-BCDP guarantee, we can restrict to R in the set $\mathcal{F}^+_{\mathcal{Y}} := \{R \in \mathcal{F}_{\mathcal{Y}} | P\{M(\boldsymbol{x}) \in R\} > 0\}$.

For $S_i, S'_i \in \mathcal{F}^+_{\mathcal{X}_i}, R \in \mathcal{F}^+_{\mathcal{Y}}$,

$$\frac{P\{M(\boldsymbol{x}) \in R | x_i \in S_i\}}{P\{M(\boldsymbol{x}) \in R | x_i \in S'_i\}} = \frac{P\{M(\boldsymbol{x}) \in R, x_i \in S_i\}}{\pi\{x_i \in S_i\}} \frac{\pi\{x_i \in S'_i\}}{P\{M(\boldsymbol{x}) \in R, x_i \in S'_i\}}. \tag{42}$$

For the set R under consideration, define

$$l(\boldsymbol{x}_{-i}) = \inf_{x_i \in \mathcal{X}_i} \mu_{x_i,\boldsymbol{x}_{-i}}(R), \tag{43}$$

$$u(\boldsymbol{x}_{-i}) = \sup_{x_i \in \mathcal{X}_i} \mu_{x_i,\boldsymbol{x}_{-i}}(R). \tag{44}$$

By CDP property, we have $u(\boldsymbol{x}_{-i}) \le e^{c_i} l(\boldsymbol{x}_{-i})$. Now note that

$$\frac{P\{M(\boldsymbol{x}) \in R, x_i \in S_i\}}{\pi\{x_i \in S_i\}} = \frac{1}{\pi\{x_i \in S_i\}} \int_{\boldsymbol{x}_{-i} \in \mathcal{X}_{-i}, x_i \in S_i} \mu_{x_i,\boldsymbol{x}_{-i}}(R) d\pi(x_i, \boldsymbol{x}_{-i}) \tag{45}$$

$$\le \frac{1}{\pi\{x_i \in S_i\}} \int_{\boldsymbol{x}_{-i} \in \mathcal{X}_{-i}, x_i \in S_i} u(\boldsymbol{x}_{-i}) d\pi(x_i, \boldsymbol{x}_{-i}) \tag{46}$$

$$= \int_{\boldsymbol{x}_{-i} \in \mathcal{X}_{-i}} u(\boldsymbol{x}_{-i}) d\pi_{\boldsymbol{x}_{-i}|x_i \in S_i}(\boldsymbol{x}_{-i}). \tag{47}$$

Similarly, we have

$$\frac{P\{M(\boldsymbol{x}) \in R, x_i \in S_i'\}}{P\{x_i \in S_i'\}} \ge \int_{\boldsymbol{x}_{-i} \in \mathcal{X}_{-i}} l(\boldsymbol{x}_{-i}) d\pi_{\boldsymbol{x}_{-i}|x_i \in S_i'}(\boldsymbol{x}_{-i}). \tag{48}$$

$\square$

Using (47) and (48) in (42), we obtain

$$\frac{P\{M(\boldsymbol{x}) \in R | x_i \in S_i\}}{P\{M(\boldsymbol{x}) \in R | x_i \in S_i'\}} \le \frac{\int_{\boldsymbol{x}_{-i} \in \mathcal{X}_{-i}} u(\boldsymbol{x}_{-i}) d\pi_{\boldsymbol{x}_{-i}|x_i \in S_i}(\boldsymbol{x}_{-i})}{\int_{\boldsymbol{x}_{-i} \in \mathcal{X}_{-i}} l(\boldsymbol{x}_{-i}) d\pi_{\boldsymbol{x}_{-i}|x_i \in S_i'}(\boldsymbol{x}_{-i})} \tag{49}$$

$$\le e^{c_i} \frac{\int_{\boldsymbol{x}_{-i} \in \mathcal{X}_{-i}} l(\boldsymbol{x}_{-i}) d\pi_{\boldsymbol{x}_{-i}|x_i \in S_i}(\boldsymbol{x}_{-i})}{\int_{\boldsymbol{x}_{-i} \in \mathcal{X}_{-i}} l(\boldsymbol{x}_{-i}) d\pi_{\boldsymbol{x}_{-i}|x_i \in S_i'}(\boldsymbol{x}_{-i})} \tag{50}$$

$$\le e^{c_i} \left(1 + \mathsf{TV}(\pi_{\boldsymbol{x}_{-i}|x_i \in S_i}, \pi_{\boldsymbol{x}_{-i}|X_i \in S_i'}) \frac{\max_{\boldsymbol{x}_{-i}} l(\boldsymbol{x}_{-i}) - \min_{\boldsymbol{x}_{-i}} l(\boldsymbol{x}_{-i})}{\min_{\boldsymbol{x}_{-i}} l(\boldsymbol{x}_{-i})} \right) \tag{51}$$

$$\le e^{c_i} \left(1 + \mathsf{TV}(\pi_{\boldsymbol{x}_{-i}|x_i \in S_i}, \pi_{\boldsymbol{x}_{-i}|x_i \in S_i'})(e^{\sum_{j \ne i} c_j} - 1) \right), \tag{52}$$

where we used $\frac{\max_{\boldsymbol{x}_{-i}} l(\boldsymbol{x}_{-i})}{\min_{\boldsymbol{x}_{-i}} l(\boldsymbol{x}_{-i})} \le e^{\sum_{j \ne i} c_j}$ by CDP property.

The $\sum_i c_i$-LDP property follows by straightforward composition.

Finally, notice that, in case that we know the mechanism is ε-LDP, we can bound $\frac{\max_{\boldsymbol{x}_{-i}} l(\boldsymbol{x}_{-i})}{\min_{\boldsymbol{x}_{-i}} l(\boldsymbol{x}_{-i})}$ by e^ε. In addition, by Proposition 3, we know $\delta_i' \le \varepsilon$. Thus, we get $\delta_i' \le \min\{c_i + \log(1 + q_i e^\varepsilon - q_i), \varepsilon\}$

A.5 Additional Results for Section 3.4

Proposition 6. *Let*

$$A_{\boldsymbol{\delta}}[i, j] = \begin{cases} \frac{1}{\delta_i} & i = j, \\ \frac{1}{\log\left(1 + \frac{e^{\delta_i} - 1}{q_i}\right)} & else. \end{cases} \tag{53}$$

Then, the condition $A_{\boldsymbol{\delta}}\boldsymbol{c} \leq \mathbf{1}$ implies $\delta_i \geq c_i + \log(1 + q_i e^{\sum_{j \neq i} c_j} - q_i)$ for all $i \in [d]$.

Proof. Without loss of generality, we establish the proof for $i = 1$. For a given value of c_1, we have the constraint that

$$\sum_{j=2}^{d} c_j \leq \log\left(1 + \frac{e^{\delta_1 - c_1} - 1}{q_1}\right). \tag{54}$$

Note that $g(x) = \log\left(1 + \frac{e^{\delta_1 - x} - 1}{q_1}\right)$ is concave and $g(0) = \log\left(1 + \frac{e^{\delta_1} - 1}{q_1}\right)$ and $g(\delta_1) = 0$. Thus, $g(x) \geq \left(1 - \frac{x}{\delta_1}\right)\log\left(1 + \frac{e^{\delta_1} - 1}{q_1}\right)$. Therefore, we can ensure (54) by imposing

$$\sum_{j=2}^{d} c_j \leq \left(1 - \frac{c_1}{\delta_1}\right)\log\left(1 + \frac{e^{\delta_1} - 1}{q_1}\right). \tag{55}$$

$\square$

A.6 Proof of Proposition 5

Proposition 5. (Post processing for BCDP) *Let $M(\cdot)$ be a $\boldsymbol{\delta}$-BCDP mechanism and $K : \mathcal{Y} \to \mathcal{Z}$ be a measurable function. Then, $K \circ M$ is also $\boldsymbol{\delta}$-BCDP.*

Proof. Let the σ-field on $\mathcal{Z}$ be denoted by $\mathcal{F}_{\mathcal{Z}}$. For $T \in \mathcal{F}_{\mathcal{Z}}, Y(T) := \{y \in \mathcal{Y} | K(y) \in T\} \in \mathcal{F}_{\mathcal{Y}}$ since K is a measurable function. Thus, we have the following for any $i \in [d], T \in \mathcal{F}_{\mathcal{Z}}$ and $S_i, S_i' \in \mathcal{F}_{\mathcal{X}_i}^{+}$ by BCDP guarantees on M

$$P\{K(M(\boldsymbol{x})) \in T | x_i \in S_i\} = P\{M(\boldsymbol{x}) \in Y(T) | x_i \in S_i\} \tag{56}$$

$$\leq e^{\delta_i} P\{M(\boldsymbol{x}) \in Y(T) | x_i \in S_i'\} \tag{57}$$

$$= e^{\delta_i} P\{K(M(\boldsymbol{x})) \in T | x_i \in S_i'\}. \tag{58}$$

$\square$

B Private Mean Estimation

B.1 Numerical Experiment

Here, we provide a simple numerical experiment to validate the performance of our proposed algorithm. For baseline comparison, we consider the M_{LDP} privacy mechanism from Duchi et al. [2013b] with privacy parameter $\min \boldsymbol{\delta}$. For our proposed privacy mechanism M_{mean}, we have a degree of freedom in choosing ζ. We note that, as $q \to 1$, we would want $\zeta \to 1$ since in this case, the best algorithm is to do $\min \boldsymbol{\delta}$-LDP. As $q \to 0$, we would want $c_i = \tilde{\delta}_i$ as CDP is ideal in this case. Thus, we settle with a heuristic choice $\zeta = (1 + q)/2$.
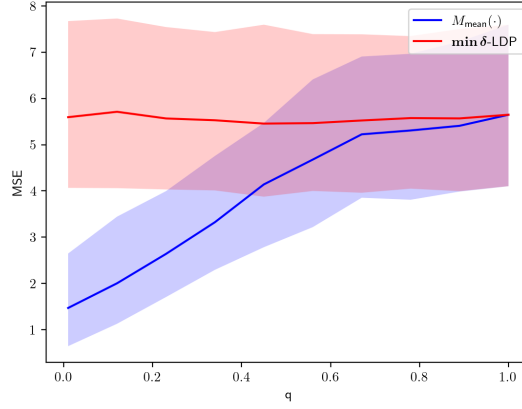
Figure 3: MSE as q is varied keeping $\boldsymbol{\delta}$ and ε constant for M_{mean} and $\min \boldsymbol{\delta}$-LDP.

We first construct a distribution for which Assumption 1 holds. We consider the distribution where

$$Z \sim \mathsf{Bern}\left(\frac{1}{2}\right), \tag{59}$$

$$(x_1, \ldots, x_d) = \begin{cases} (Z, \ldots, Z) & \text{w.p. } q \\ \text{i.i.d. } \mathsf{Bern}\left(\frac{1}{2}\right) & \text{else.} \end{cases} \tag{60}$$

It is easy to verify that $\mathsf{TV}(\pi_{\boldsymbol{x}_{-i}|x_i=1}, \pi_{\boldsymbol{x}_{-i}|x_i=0}) = q \quad \forall i$ in this setting. We fix $d = 10$, $\boldsymbol{\delta} = (0.2, 0.2, \varepsilon, \varepsilon, \ldots, \varepsilon)$, and $\varepsilon = 2$. This coordinate-wise privacy demand captures the impact of requiring more privacy protection for the first two-coordinates on the algorithm's error.

We vary the correlation q and plot the median MSE values, along with the 25-th and 75-th quantiles, for our proposed algorithm BCDP and the baseline algorithm in Figure 3. We set $n = 10000$, and sample user data i.i.d. $2 \times \mathsf{Bern}(\frac{1}{2}) - 1$. The experiment is run over 1000 trials.

When $q \approx 1$, there is a high degree of correlation between the coordinates and our algorithm resembles $\min \boldsymbol{\delta}$-LDP. For other values of q, the proposed BCDP algorithm can leverage the heterogeneous nature of the privacy demand and obtain better performance.

B.2   Proofs

Theorem 2. *Suppose Assumption 1 holds and let $\tilde{\delta}_i := \min\{\delta_i, \varepsilon\}$ for any $i \in [d]$. Then, M_{mean}, i.e., Mechanism 1, with parameters*

$$c_d = \min\left\{\log\left(\frac{e^{\zeta \tilde{\delta}_1} + q - 1}{q}\right), \tilde{\delta}_d\right\}, \quad c_0 = 0, \tag{16}$$

$$\forall i < d, \; c_i = \begin{cases} c_d & \text{if } c_d \leq \tilde{\delta}_i, \\ \tilde{\delta}_i - \log(1 + qe^{c_d} - q) & \text{otherwise,} \end{cases}$$

where $\zeta \in (0,1]$ is a free parameter, is $\boldsymbol{\delta}$-BCDP and ε-LDP. Moreover, using Mechanism 2 as M_{LDP}, we obtain the following mean-squared error bound

$$\mathbb{E}\left[\left\| \bar{\boldsymbol{x}} - \Pi\left(\frac{1}{n}\sum_{j=1}^{n} M_{\text{mean}}(\boldsymbol{x}^{(j)}, \boldsymbol{c}) \right) \right\|_2^2 \middle| \boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(n)} \right] \lesssim \frac{1}{n}\sum_{i=1}^{d}\left(\frac{1}{\sum_{k=1}^{i} \frac{(c_k - c_{k-1})^2}{d-k+1}} \right) \wedge d, \qquad (17)$$

where $\Pi(\cdot)$ denotes projection into $[-1,1]^d$ and the expectation is taken over the randomness in $M_{\text{mean}}(\cdot)$.

Proof. Our proof has two main parts that are presented below: proof of privacy guarantee, and proof of the error bounds.

Proof of privacy: Note that we use the channel $M_{\text{LDP}}(.)$ a total of d times with varying size inputs in Mechanism 1. Focusing on any particular data point $\boldsymbol{x}$, let

$$\tilde{M}(\boldsymbol{x}) := \left(M_{\text{LDP}}(\boldsymbol{x}_{i:d}, c_i - c_{i-1}, \sqrt{d-i+1}) \right)_{i=1}^{d}. \qquad (61)$$

First, notice that, by composition, $\tilde{M}$ is c_d-LDP. Given that $c_d \leq \tilde{\delta}_d \leq \varepsilon$, this immediately implies that our algorithm is ε-LDP. To show the BCDP bound, given Proposition 5, it suffices to show $\tilde{M}$ is $\boldsymbol{\delta}$-BCDP. First, notice that Proposition 3, along with choice of $c_d \leq \tilde{\delta}_d$, implies that this algorithm is δ_d-BCDP with respect to coordinate d. In fact, for any i for which $c_d \leq \tilde{\delta}_i$, a similar argument implies that $\tilde{M}$ is $\tilde{\delta}_i$-BCDP with respect to coordinate i. Hence, we could only focus on other coordinates $i < d$ for which $c_d > \tilde{\delta}_i$.

Next, we claim that $\tilde{M}(\boldsymbol{x})$ is $\boldsymbol{c}$-CDP. To see this, notice that, by perturbing the k-th coordinate of $\boldsymbol{x}$, the terms affected in $\tilde{M}(\boldsymbol{x})$ are $M_{\text{LDP}}(\boldsymbol{x}_{i:d}, c_i - c_{i-1}, \sqrt{d-i+1})$ for $i \leq k$; in Figure 2, this corresponds to the bottom i rows of the Y(s), i.e., $\{Y^k\}_{k \in [i]}$. Thus, it suffices to show

$$\tilde{M}_k(\boldsymbol{x}) := \left(M_{\text{LDP}}(\boldsymbol{x}_{i:d}, c_i - c_{i-1}, \sqrt{d-i+1}) \right)_{i=1}^{k} \qquad (62)$$

is c_k-CDP with respect to coordinate k. However, this immediately follows as $\tilde{M}_k$ is c_k-LDP, implying it is also c_k-CDP with respect to the k-th coordinate.

As a result, by the second part of Proposition 4, and given that $\tilde{M}$ is c_d-LDP and Assumption 1 holds, we obtain that $\tilde{M}$ is $c_i + \log(1 + qe^{c_d} - q)$-BCDP with respect to coordinate i. It remains to show that $c_i + \log(1 + qe^{c_d} - q) \leq \delta_i$ for any $i < d$. This also follows from the choice of c_i for $i < d$ and $c_d > \tilde{\delta}_i$, which are the conditions we are operating under. It is worth noting that the condition

$$c_d \leq \log\left(\frac{e^{\zeta\tilde{\delta}_1} + q - 1}{q} \right) \qquad (63)$$

guarantees that the chosen c_i's are nonnegative, given $\zeta \in (0,1)$.

Proof of the mean-squared error bound:

Duchi et al. [2013a] show that the channel $M_{\mathsf{LDP}}(\cdot)$'s output is unbiased and it holds that for any $\boldsymbol{x} \in [-1,1]^d$, we have

$$\mathbb{E}[\|M_{\mathsf{LDP}}(\boldsymbol{x}, \varepsilon, d) - \boldsymbol{x}\|_2^2 | \boldsymbol{x}] \leq \mathcal{O}(d) \left(\frac{d}{\varepsilon^2} \wedge 1 \right). \tag{64}$$

Note that, due to the symmetry of Algorithm 2, the error is identically distributed across each coordinate in expectation. Hence, for any vector $\boldsymbol{x}$ and any $k \leq i$, the random variable Y_{i-k+1}^k is an unbiased estimators of $\boldsymbol{x}_i$ with its variance bounded by $\mathcal{O}(1) \left(\frac{d-i+1}{(c_i-c_{i-1})^2} \wedge 1 \right)$.

Note that, the errors from different runs of the algorithm are independent. Therefore, using the Cauchy–Schwarz inequality, we can see that the minimum variance in estimating $\boldsymbol{x}_i$ through a linear combination of $Y^{(i,j)}$ is achieved by choosing weights proportional to the inverse of the variances which would result in the following estimator

$$\frac{\sum_{k=1}^{i} Y_{i+1-k}^k w_k}{\sum_{k=1}^{i} w_k} \tag{65}$$

with the error rate given by

$$\mathcal{O}(1) \frac{1}{\sum_{k=1}^{i} \frac{(c_k-c_{k-1})^2}{d-k+1}} \wedge 1. \tag{66}$$

Also, note that, the algorithm's runs over different users are independent. Therefore, taking average of $M_{\mathsf{mean}}(\boldsymbol{x}^{(j)}, \boldsymbol{c})$ over j, i.e., different users' data, gives us an estimator of

$$\frac{1}{n} \sum_{j=1}^{n} \boldsymbol{x}_i^{(j)} \tag{67}$$

with the error rate

$$\mathcal{O}(1) \frac{1}{n} \cdot \frac{1}{\sum_{k=1}^{i} \frac{(c_k-c_{k-1})^2}{d-k+1}} \wedge 1. \tag{68}$$

Summing this over the coordinates gives us the following upper bound:

$$\mathcal{O}(1) \frac{1}{n} \sum_{i=1}^{d} \frac{1}{\sum_{k=1}^{i} \frac{(c_k-c_{k-1})^2}{d-k+1}} \wedge d. \tag{69}$$

$\square$

Corollary 1. *Suppose we are under the premise of Theorem 2, with $\delta_i = \delta$ for $1 \leq i \leq k$ and $\delta_i = \varepsilon > 2\delta$ for $d \geq i > k$. Then, by choosing $\zeta = 0.5$, the mean-squared error upper bound in (17) simplifies to*

$$\mathcal{O}(1) \frac{1}{n} \cdot \left(\frac{dk}{\delta^2} + \frac{(d-k)^2}{\varepsilon^2} \right), \tag{18}$$

for the case $q \le \frac{e^{\delta/2}-1}{e^\varepsilon -1}$, *and to*

$$\mathcal{O}(1)\frac{1}{n} \cdot \left(\frac{dk}{\delta^2} + \frac{(d-k)^2}{\frac{d-k}{d}\delta^2 + (c_d - \delta/2)^2} \right), \tag{19}$$

otherwise, with $c_d = \log(\frac{e^{\delta/2}+q-1}{q})$ *decreasing from* ε *to* $\delta/2$ *as* q *increases from* $(e^{\delta/2} - 1)/(e^\varepsilon - 1)$ *to 1.*

Proof. For convenience, we set $\zeta = 0.5$. We do not focus on refining the constants in the Mean-Squared Error (MSE) bound with a potentially better choice of ζ. Now consider two cases for the correlation bound q

1. $q \in [0, \frac{e^{\delta/2}-1}{e^\varepsilon -1}]$: in this case, we have $c_{k+1} = c_{k+2} = \ldots = c_d = \varepsilon$ and $c_1 = c_2 = \ldots = c_k = \delta - \log(1 + qe^\varepsilon - q)$. Using the range of q, we get $\delta - \log(1 + qe^\varepsilon - q) \in [\delta/2, \delta]$. Thus, replacing the c_1 to c_k values by the worst-case of $\delta/2$ we get an error of

$$\mathsf{MSE} \lesssim \frac{d}{n} \left(\frac{k}{\delta^2} + \frac{d-k}{\delta^2 + \frac{d}{d-k}(\varepsilon - \delta/2)^2} \right) \wedge d \tag{70}$$

Using $\varepsilon \ge 2\delta$, we get

$$\mathsf{MSE} \lesssim \left(\frac{dk}{n\delta^2} + \frac{(d-k)^2}{n\varepsilon^2} \right) \wedge d \tag{71}$$

2. $q \in (\frac{e^{\delta/2}-1}{e^\varepsilon -1}, 1]$: in this case we have $c_{k+1} = \ldots = c_d = \log\frac{e^{\delta/2}+q-1}{q}$ and $c_1 = \ldots = c_k = \delta - \log(1 + qe^{c_d} - q) = \delta/2$. Thus, we get an error of

$$\mathsf{MSE} \lesssim \left(\frac{dk}{n\delta^2} + \frac{(d-k)d}{n\left(\delta^2 + \frac{d}{d-k}(c_d - \delta/2)^2\right)} \right) \wedge d \tag{72}$$

$\square$

C Least-Squares Regression

C.1 Proof of Theorem 3

Theorem 3. *Suppose Assumptions 1 and 2 hold. Then, Algorithm 3, with $\boldsymbol{c}$ set similar to Theorem 2, is δ-BCDP and ε-LDP. Moreover, using Mechanism 2 as M_{LDP}, we obtain the following error rate*

$$\mathbb{E}\left[f(\hat{\boldsymbol{\theta}}) - f^* \middle| \boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(n)} \right] \lesssim \max_{\boldsymbol{\theta}\in\Theta}(\|\boldsymbol{\theta}\|^2 + \|\boldsymbol{\theta}\|)\left(\frac{\log d}{\sqrt{n}}(r^2 + rd) \wedge d \right), \tag{22}$$

with $r^2 = \sum_{i=1}^{d} 1/(\sum_{k=1}^{i} \frac{(c_k - c_{k-1})^2}{d-k+1})$.

Proof. **Proof of privacy:** The proof of the privacy guarantee follows a similar structure to the proof of Theorem 2. Note that while composition does not hold for BCDP, it does hold for LDP and

CDP. This is why we divide c by two, rather than dividing δ by two. Additionally, since the OPT algorithm only observes the private copies of the data, by the post processing inequality, it does not matter how the algorithm finds the minimizer (e.g., how many passes it makes over each private data point), as the privacy guarantee remains unchanged.

Proof of error rate: Without loss of generality, we can ignore the constant term in $f(\cdot)$ that does not depend on $\boldsymbol{\theta}$, and hence, with a slight abuse of notation, we can rewrite $f(\cdot)$ as

$$f(\boldsymbol{\theta}) := \frac{1}{2n} \sum_{i=1}^{n} \left(\boldsymbol{\theta}^{\top} \boldsymbol{z}^{(i)} \boldsymbol{z}^{(i)^{\top}} \boldsymbol{\theta} - 2\boldsymbol{\theta}^{\top} l^{(i)} \boldsymbol{z}^{(i)} \right). \tag{73}$$

Also, recall that $\hat{f}(\cdot)$ is given by

$$\hat{f}(\boldsymbol{\theta}) := \frac{1}{2n} \sum_{i=1}^{n} \left(\boldsymbol{\theta}^{\top} \hat{\boldsymbol{z}}^{(i,1)} \hat{\boldsymbol{z}}^{(i,2)^{\top}} \boldsymbol{\theta} - 2\boldsymbol{\theta}^{\top} \hat{l}^{(i,1)} \hat{\boldsymbol{z}}^{(i,2)} \right). \tag{74}$$

Denote the minimizer of $\hat{f}(\boldsymbol{\theta})$ over Θ by $\boldsymbol{\theta}^*(\{\hat{\boldsymbol{x}}^{(i,1)}, \hat{\boldsymbol{x}}^{(i,2)}\}_{i=1}^{n})$. Also, we denote the output of Algorithm 3 by $\hat{\boldsymbol{\theta}}(\{\hat{\boldsymbol{x}}^{(i,1)}, \hat{\boldsymbol{x}}^{(i,2)}\}_{i=1}^{n})$. Finally, recall that $\boldsymbol{\theta}^*$ denotes the minimizer of function f.

For convenience and brevity, we drop the conditioning on $\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(n)}$ from the notation, but all expectations henceforth are conditioned on $\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(n)}$.

Next, note that

1. By the definition of OPT, we have

$$\hat{f}\left(\hat{\boldsymbol{\theta}}(\{\hat{\boldsymbol{x}}^{(i,1)}, \hat{\boldsymbol{x}}^{(i,2)}\}_{i=1}^{n})\right) - \hat{f}\left(\boldsymbol{\theta}^*(\{\hat{\boldsymbol{x}}^{(i,1)}, \hat{\boldsymbol{x}}^{(i,2)}\}_{i=1}^{n})\right) \leq \frac{1}{n}. \tag{75}$$

2. By the definition of $\boldsymbol{\theta}^*(\{\hat{\boldsymbol{x}}^{(i,1)}, \hat{\boldsymbol{x}}^{(i,2)}\}_{i=1}^{n})$, we have

$$\hat{f}\left(\boldsymbol{\theta}^*(\{\hat{\boldsymbol{x}}^{(i,1)}, \hat{\boldsymbol{x}}^{(i,2)}\}_{i=1}^{n})\right) - \hat{f}(\boldsymbol{\theta}^*) \leq 0. \tag{76}$$

3. Given the construction of $\hat{f}$, it is an unbiased estimator of f for any fixed $\boldsymbol{\theta}$. Hence, we have

$$\mathbb{E}[\hat{f}(\boldsymbol{\theta}^*) - f(\boldsymbol{\theta}^*)] = 0. \tag{77}$$

Summing all the three, given us

$$\mathbb{E}\left[\hat{f}\left(\hat{\boldsymbol{\theta}}(\{\hat{\boldsymbol{x}}^{(i,1)}, \hat{\boldsymbol{x}}^{(i,2)}\}_{i=1}^{n})\right) - f(\boldsymbol{\theta}^*)\right] \leq \frac{1}{n}. \tag{78}$$

However, we need to bound

$$\mathbb{E}\left[f\left(\hat{\boldsymbol{\theta}}(\{\hat{\boldsymbol{x}}^{(i,1)}, \hat{\boldsymbol{x}}^{(i,2)}\}_{i=1}^{n})\right) - f(\boldsymbol{\theta}^*)\right]. \tag{79}$$

Therefore, it suffices to bound

$$\mathbb{E}\left[f\left(\hat{\boldsymbol{\theta}}(\{\hat{\boldsymbol{x}}^{(i,1)}, \hat{\boldsymbol{x}}^{(i,2)}\}_{i=1}^n)\right) - \hat{f}\left(\hat{\boldsymbol{\theta}}(\{\hat{\boldsymbol{x}}^{(i,1)}, \hat{\boldsymbol{x}}^{(i,2)}\}_{i=1}^n)\right)\right]. \tag{80}$$

Observe that this term is upper bounded by

$$\mathbb{E}\left[\sup_{\boldsymbol{\theta}}(f(\boldsymbol{\theta}) - \hat{f}(\boldsymbol{\theta}))\right]. \tag{81}$$

Note that (81) is upper bounded by

$$\sup_{\boldsymbol{\theta} \in \Theta} \|\theta\|^2 \mathbb{E}\left[\left\|\frac{1}{n}\sum_{i=1}^n \boldsymbol{z}^{(i)}\boldsymbol{z}^{(i)\top} - \hat{\boldsymbol{z}}^{(i,1)}\hat{\boldsymbol{z}}^{(i,2)\top}\right\|\right] + \sup_{\boldsymbol{\theta} \in \Theta} \|\theta\| \mathbb{E}\left[\left\|\frac{1}{n}\sum_{i=1}^n l^{(i)}\boldsymbol{z}^{(i)} - \hat{l}^{(i,1)}\hat{\boldsymbol{z}}^{(i,2)}\right\|\right]. \tag{82}$$

We bound the first term above. The second term can be bounded similarly. To do so, first note that we can decompose $\hat{\boldsymbol{z}}^{(i,1)}\hat{\boldsymbol{z}}^{(i,2)\top} - \boldsymbol{z}^{(i)}\boldsymbol{z}^{(i)\top}$ as

$$(\hat{\boldsymbol{z}}^{(i,1)} - \boldsymbol{z}^{(i)})\boldsymbol{z}^{(i)\top} + \boldsymbol{z}^{(i)}(\hat{\boldsymbol{z}}^{(i,2)} - \boldsymbol{z}^{(i)})^\top + (\hat{\boldsymbol{z}}^{(i,1)} - \boldsymbol{z}^{(i)})(\hat{\boldsymbol{z}}^{(i,2)} - \boldsymbol{z}^{(i)})^\top. \tag{83}$$

Hence, it suffices to bound the expected norm of the average of each term separately, and all of them can be handled similarly. We will focus on the last term in particular, i.e., we will bound

$$\mathbb{E}\left[\left\|\frac{1}{n}\sum_{i=1}^n A^i\right\|\right] \tag{84}$$

with $A^i := (\hat{\boldsymbol{z}}^{(i,1)} - \boldsymbol{z}^{(i)})(\hat{\boldsymbol{z}}^{(i,2)} - \boldsymbol{z}^{(i)})^\top$. Next, we make the following claim regarding matrices $\{A^i\}_{i=1}^n$, proved in Appendix C.2.

Lemma 1. *Let*

$$r^2 := \sup_{\boldsymbol{x} \in [-1,1]^d} \mathbb{E}[\|M_{\mathsf{mean}}(\boldsymbol{x}, \boldsymbol{c}/2) - \boldsymbol{x}\|^2 | \boldsymbol{x}]. \tag{85}$$

Then, the following two results hold:

$$\left\|\sum_i \mathbb{E}[A^i A^{i\top}]\right\| \leq nr^4, \tag{86}$$

$$\mathbb{E}[\max_i \|A^i\|^2] \leq nr^4. \tag{87}$$

Let us first show how this lemma completes the proof. Using Theorem 1 in Tropp [2016], Jensen's inequality, along with this lemma, we can immediately upper bound (84) by

$$\mathcal{O}(1)\log(d)\frac{r^2}{\sqrt{n}}. \tag{88}$$

Similarly, the terms $\|(\hat{\boldsymbol{z}}^{(i,1)} - \boldsymbol{z}^{(i)})\boldsymbol{z}^{(i)\top}\|$ and $\|\boldsymbol{z}^{(i)}(\hat{\boldsymbol{z}}^{(i,2)} - \boldsymbol{z}^{(i)})^\top\|$ can be upper bounded by $\mathcal{O}(\frac{\log(d)rd}{\sqrt{n}})$

using similar steps as proof of Lemma 1.

Recall $r^2 = \mathcal{O}(1) \sum_{i=1}^d \frac{1}{\sum_{k=1}^i \frac{(c_k - c_{k-1})^2}{d-k+1}}$. Note that we use M_{mean} without the truncation to $[-1, 1]^d$ to preserve unbaisedness, and hence, the minimum with d does not appear in r^2, unlike Theorem 2. Thus we obtain the bound

$$\mathbb{E}\left[\sup_{\boldsymbol{\theta}}(f(\boldsymbol{\theta}) - \hat{f}(\boldsymbol{\theta}))\right] \lesssim \max_{\boldsymbol{\theta} \in \Theta}(\|\boldsymbol{\theta}\|^2 + \|\boldsymbol{\theta}\|)\left(\frac{r^2 \log d}{\sqrt{n}} + \frac{rd \log d}{\sqrt{n}}\right). \tag{89}$$

We can also project the estimate $\hat{\theta}$ to the set Θ to obtain the bound $\max_{\boldsymbol{\theta} \in \Theta}(\|\boldsymbol{\theta}\|^2 + \|\boldsymbol{\theta}\|)d$. Taking minimum of the terms, we get Theorem 3.

C.2 Proof of Lemma 1

First, note that

$$A^i A^{i\top} = \|\hat{\boldsymbol{z}}^{(i,2)} - \boldsymbol{z}^{(i)}\|^2 (\hat{\boldsymbol{z}}^{(i,1)} - \boldsymbol{z}^{(i)})(\hat{\boldsymbol{z}}^{(i,1)} - \boldsymbol{z}^{(i)})^\top, \tag{90}$$

which yields

$$\|\mathbb{E}[A^i A^{i\top}]\| \le r^2 \mathbb{E}[\|(\hat{\boldsymbol{z}}^{(i,1)} - \boldsymbol{z}^{(i)})(\hat{\boldsymbol{z}}^{(i,1)} - \boldsymbol{z}^{(i)})^\top\|]. \tag{91}$$

Using the result that $\|uu^\top\| = \|u\|^2$ for any vector u completes the proof of the first part of the lemma. To prove the second part, note that

$$\mathbb{E}[\max_i \|A^i\|^2] \le \mathbb{E}[\sum_{i=1}^n \|A^i\|^2] = \sum_{i=1}^n \mathbb{E}[\|A^i\|^2]. \tag{92}$$

Using the result that $\|uv^\top\| \le \|u\|\|v\|$ for any two vectors u and v implies

$$\mathbb{E}[\|A^i\|^2] \le \mathbb{E}\left[\|\hat{\boldsymbol{z}}^{(i,1)} - \boldsymbol{z}^{(i)}\|^2 \|\hat{\boldsymbol{z}}^{(i,2)} - \boldsymbol{z}^{(i)}\|^2\right]. \tag{93}$$

Finally, the conditional independence of $\hat{\boldsymbol{z}}^{(i,1)}$ and $\hat{\boldsymbol{z}}^{(i,2)}$ given $z^{(i)}$ completes the proof of lemma. $\square$