Locally Private Gaussian Estimation

Matthew Joseph* Janardhan Kulkarni† Jieming Mao‡ Zhiwei Steven Wu§

November 21, 2018

Abstract

We study a basic private estimation problem: each of n users draws a single i.i.d. sample from an unknown Gaussian distribution, and the goal is to estimate the mean of this Gaussian distribution while satisfying *local differential privacy* for each user. Informally, local differential privacy requires that each data point is individually and independently privatized before it is passed to a learning algorithm. Locally private Gaussian estimation is therefore difficult because the data domain is *unbounded*: users may draw arbitrarily different inputs, but local differential privacy nonetheless mandates that different users have (worst-case) similar privatized output distributions.

We provide both *adaptive* two-round solutions and *nonadaptive* one-round solutions for locally private Gaussian estimation. We then partially match these upper bounds with an information-theoretic lower bound. This lower bound shows that our accuracy guarantees are tight up to logarithmic factors for all sequentially interactive (ε, δ)-locally private protocols.

1 Introduction

Differential privacy is a formal algorithmic guarantee that no single input has a large effect on the output of a computation. Since its introduction [13] over a decade ago, a rich line of work has made differential privacy a compelling privacy guarantee (see Dwork et al. [14] and Vadhan [26] for surveys), and deployments of differential privacy now exist at many organizations, including Apple [3], Google [6, 15], Microsoft [11], Mozilla [4], and the US Census Bureau [1, 22].

Much recent attention, including almost all industrial deployments, has focused on a stronger variant of differential privacy called *local differential privacy* [16, 21, 27]. In the local model private data is distributed across many users, and each user privatizes their data *before* the data is collected by an analyst. Thus, as any locally differentially private computation runs on already-privatized data, data contributors need not worry about compromised data analysts or insecure communication channels. In contrast, (global) differential privacy assumes that the data analyst has trusted access to the unprivatized data. As a result, under global differential privacy any violation of this trust may lead to serious privacy loss for the users contributing the data.

*Computer and Information Science, University of Pennsylvania. majos@cis.upenn.edu. A portion of this work was done while at Microsoft Research Redmond.

†Microsoft Research Redmond. jakul@microsoft.com.

‡Warren Center, University of Pennsylvania. jiemingm@seas.upenn.edu.

§Computer Science and Engineering, University of Minnesota. zsw@umn.edu. A portion of this work was done while at Microsoft Research New York.

However, the stronger privacy guarantees of the local model come at a price: for many problems, "good" solutions under local privacy require far more samples than similarly good solutions under global privacy [21]. Moreover, many problems remain little-understood under local differential privacy. In this paper, we study the simple problem of locally private Gaussian estimation: given n users each holding an i.i.d. draw from an unknown Gaussian distribution $N(\mu, \sigma^2)$, can one accurately estimate the mean μ while guaranteeing local differential privacy for each user?

One challenge of this problem is that, since data is drawn from a Gaussian, there is no a priori (worst-case) bound on the scale of the observations. Naive applications of standard privatization methods like Laplace and Gaussian mechanisms that add noise proportional to the worst-case scale of the data are therefore infeasible. Second, it is desirable to limit the number of *rounds of interaction* between users and the data analyst, as protocols requiring many rounds of user-analyst interaction are difficult to implement.

1.1 Our Contributions

We divide our solution to locally private Gaussian estimation into two cases. In the first case, σ is known to the analyst, and in the second case σ is unknown but bounded in known $\sigma_{\min} \le \sigma \le \sigma_{\max}$. For each case, we provide adaptive two-round and nonadaptive one-round sequentially interactive protocols. Hhere sequential interactivity informally means that no user outputs information more than once (see Section 2 for details). Informal guarantees for these protocols appear below.

Theorem 1.1. *Let $x_1, \dots, x_n \sim_{iid} N(\mu, \sigma^2)$ where $\mu = O(\sigma 2^{n\varepsilon^2/\log(n/\beta)})$ and σ is known. Then*

1. *Adaptive two-round protocol* KVGausstimate *satisfies $(\varepsilon, 0)$-local differential privacy for $x_1, \dots, x_n$ and, with probability at least $1 - \beta$, outputs $\hat{\mu}$ such that*

$$|\hat{\mu} - \mu| = O\left(\frac{\sigma}{\varepsilon}\sqrt{\frac{\log\left(\frac{1}{\beta}\right)}{n}}\right).$$

2. *Nonadaptive one-round protocol* 1RoundKVGausstimate *satisfies $(\varepsilon, 0)$-local differential privacy for $x_1, \dots, x_n$ and, with probability at least $1 - \beta$, outputs $\hat{\mu}$ such that*

$$|\hat{\mu} - \mu| = O\left(\frac{\sigma}{\varepsilon}\sqrt{\frac{\log\left(\frac{1}{\beta}\right)\sqrt{\log(n)}}{n}}\right).$$

Theorem 1.2. *Let $x_1, \dots, x_n \sim_{iid} N(\mu, \sigma^2)$ where $\mu = O\left(2^{\frac{n\varepsilon^2}{\log(n/\beta)}}\right)$ and σ is unknown but bounded in known $0 < \sigma_{\min} \le \sigma \le \sigma_{\max}$ where $\frac{\sigma_{\max}}{\sigma_{\min}} = O\left(2^{\frac{n\varepsilon^2}{\log(n/\beta)}}\right)$. Then*

1. *Adaptive two-round protocol* UVGausstimate *satisfies $(\varepsilon, 0)$-local differential privacy for $x_1, \dots, x_n$ and, with probability at least $1 - \beta$, outputs $\hat{\mu}$ such that*

$$|\hat{\mu} - \mu| = O\left(\frac{\sigma}{\varepsilon}\sqrt{\frac{\log\left(\frac{1}{\beta}\right)\log(n)}{n}}\right).$$

2. *Nonadaptive one-round protocol* 1RoundUVGausstimate *satisfies* $(\varepsilon, 0)$*-local differential privacy for* $x_1, \ldots, x_n$ *and, with probability at least* $1 - \beta$*, outputs* $\hat{\mu}$ *such that*

$$|\hat{\mu} - \mu| = O\left(\frac{\sigma}{\varepsilon}\sqrt{\frac{\log\left(\frac{\sigma_{\max}}{\sigma_{\min}} + 1\right)\log\left(\frac{1}{\beta}\right)\log^{3/2}(n)}{n}}\right).$$

Moreover, we show in the following (informal) information-theoretic lower bound that these upper bounds are tight up to logarithmic factors. Our proof relies on techniques from the strong data-processing inequality literature [7, 23].

Theorem 1.3. *For a given* σ*, there does not exist a sequentially interactive* (ε, δ)*-locally private protocol* $\mathcal{A}$ *such that for any* $\mu = O\left(\frac{\sigma}{\varepsilon}\sqrt{\frac{1}{n}}\right)$*, given* $x_1, \ldots, x_n \sim N(\mu, \sigma^2)$*,* $\mathcal{A}$ *outputs estimate* $\hat{\mu}$ *satisfying* $|\hat{\mu} - \mu| \leq \alpha = O\left(\frac{\sigma}{\varepsilon}\sqrt{\frac{1}{n}}\right)$ *with probability* $\geq 15/16$.

1.2 Related Work

Several works have already studied differentially private versions of various statistical tasks, especially in the global setting. Both Karwa and Vadhan [20] and Kamath et al. [19] are relevant, as they consider similar versions of Gaussian estimation under global differential privacy, respectively in the one-dimensional and high-dimensional cases. For both the known and unknown variance cases, Karwa and Vadhan [20] offer an

$$O\left(\sigma\left[\sqrt{\frac{\log(1/\beta)}{n}} + \frac{\text{poly}\log(1/\beta)}{\varepsilon n}\right]\right)$$

accuracy upper bound for estimating μ. Our upper and lower bounds thus demonstrate that local privacy adds a roughly $\sqrt{n}$ accuracy cost for estimating μ.

In local differential privacy, several recent works have studied related statistical tasks like identity and independence testing [2, 17, 24], albeit restricted to discrete distributions. In concurrent work, Gaboardi et al. [18] also study Gaussian estimation under local differential privacy. They provide an adaptive two-round protocol in the known variance case and an adaptive $O\left(\log\left(\frac{R}{\sigma_{\min}}\right)\right)$-round protocol in the unknown variance case, where R upper bounds both μ and $\sigma_{\max}$ and both protocols are approximately locally private. In our case, R may be as large as $\tilde{\Omega}\left(2^{n\varepsilon^2}\right)$, leading to $\Omega(n\varepsilon^2)$ round complexity for their unknown variance protocol.

In comparison, we construct adaptive two-round and nonadaptive one-round purely locally private protocols improving on these guarantees for both cases: see Figure 1 for a detailed comparison. Moreover, while Gaboardi et al. [18] prove an $\Omega\left(\frac{\sigma}{\varepsilon}\sqrt{\frac{\log(1/\beta)}{n}}\right)$ lower bound for nonadaptive one-round protocols, we prove a logarithmically weaker but also more general $\Omega\left(\frac{\sigma}{\varepsilon}\sqrt{\frac{1}{n}}\right)$ lower bound for adaptive sequentially interactive protocols. Gaboardi et al. [18] also offer extensions to quantile estimation and estimation when σ lacks a known upper bound.

Our lower bounds are structurally similar to existing mutual information-based approaches [5, 12, 25] and build on recent results showing that pure and approximate local differential privacy are "equivalent" [8, 10]. Our lower bound also uses tools from the strong data processing inequality literature [7, 23]; broader application of these techniques to local differential privacy may be of independent interest.

	Gaboardi et al. [18]		This Work	
Setting	Accuracy	Rounds	Accuracy	Rounds
Known σ, adaptive	$O\left(\frac{\sigma}{\varepsilon}\sqrt{\frac{\log\left(\frac{1}{\beta}\right)\log\left(\frac{n}{\beta}\right)\log\left(\frac{1}{\delta}\right)}{n}}\right)$	2	$O\left(\frac{\sigma}{\varepsilon}\sqrt{\frac{\log\left(\frac{1}{\beta}\right)}{n}}\right)$	2
Known σ, nonadaptive	–	–	$O\left(\frac{\sigma}{\varepsilon}\sqrt{\frac{\log\left(\frac{1}{\beta}\right)\sqrt{\log(n)}}{n}}\right)$	1
Unknown σ, adaptive	$O\left(\frac{\sigma}{\varepsilon}\sqrt{\frac{\log\left(\frac{1}{\beta}\right)\log\left(\frac{n}{\beta}\right)\log\left(\frac{1}{\delta}\right)}{n}}\right)$	$O\left(\log\left(\frac{R}{\sigma_{\min}}\right)\right)$	$O\left(\frac{\sigma}{\varepsilon}\sqrt{\frac{\log\left(\frac{1}{\beta}\right)\log(n)}{n}}\right)$	2
Unknown σ, nonadaptive	–	–	$O\left(\frac{\sigma}{\varepsilon}\sqrt{\frac{\log\left(\frac{\sigma_{\max}}{\sigma_{\min}}+1\right)\log\left(\frac{1}{\beta}\right)\log^{3/2}(n)}{n}}\right)$	1

Figure 1: A comparison of upper bounds presented in Gaboardi et al. [18] and our work. In all cases, Gaboardi et al. [18] use (ε, δ)-locally private algorithms while we use $(\varepsilon, 0)$. Here, R denotes an upper bound on both μ and σ. In our setting, $R = \tilde{O}(2^{n\varepsilon^2})$, leading the unknown variance protocol of Gaboardi et al. [18] to round complexity potentially as large as $\tilde{O}(n\varepsilon^2)$.

2 Preliminaries

We consider a setting in which each user $i \in [n]$ has private data consisting of a single i.i.d. draw from an unknown Gaussian distribution, $x_i \sim N(\mu, \sigma^2)$. In our communication protocol, users may exchange messages over public channels with a single (possibly untrusted) central analyst[1]. The analyst's task is to accurately estimate μ while guaranteeing local differential privacy for each user.

We restrict our attention to *sequentially interactive* protocols, where every user sends at most a single message to the analyst in the entire protocol. For simplicity, our definition of sequentially interactive protocols is slightly less general than the one introduced by Duchi et al. [12] (see Section 5 for details). The algorithms we present for our upper bounds all satisfy our more restrictive notion of sequential interactivity, while our lower bounds apply to the more general notion used by Duchi et al. [12].

We also study the *round complexity* of these interactive protocols. Formally, one round of interaction in a protocol consists of the following two steps: 1) the analyst selects a subset of users $S \subseteq [n]$, along with a set of randomizers $\{Q_i \mid i \in S\}$, and 2) each user i in S computes a message $y_i = Q_i(x_i)$ using the assigned function Q_i and sends the message to the analyst.

[1]The notion of a central analyst is a useful simplification but is not intrinsic to the protocol. Technically, as the analyst need not be trusted, any user can fulfill the same role.

2.1 Differential Privacy

Informally, a randomized algorithm is *differentially private* if arbitrarily changing a single input does not change the output distribution "too much". The resulting computation preserves privacy because the output distribution is insensitive to any change of a single user's data. More formally:

Definition 2.1 ((Standard) Differential Privacy)**.** *A randomized algorithm* $\mathcal{A}: X^m \to \mathcal{R}$ *satisfies* (ε, δ)-differential privacy *if for any two databases* $D, D' \in X^m$ *that differ by a single observation, the following holds for any event* $S \subseteq \mathcal{R}$*,*

$$\Pr[\mathcal{A}(D) \in S] \leq e^{\varepsilon} \Pr[\mathcal{A}(D') \in S] + \delta.$$

Here, we study a stronger privacy guarantee called *local differential privacy*. In the local model, each user i computes their message using a *local randomizer*. A local randomizer is a differentially private algorithm taking single-element databases as input. More formally, a randomized function $Q_i: X \to Y$ is an (ε, δ)-local randomizer if, for every pair of observations $x_i, x_i' \in X$ and any $S \subseteq Y$,

$$\Pr[Q_i(x_i) \in S] \leq e^{\varepsilon} \Pr[Q_i(x_i') \in S] + \delta.$$

A sequentially interactive protocol is locally private if every user computes their message using a local randomizer.

Definition 2.2. *A sequentially interactive protocol* $\mathcal{A}$ *is* (ε, δ)-locally private *for private user data* $\{x_1, \dots, x_n\}$ *if, for every user* $i \in [n]$*, the message* Y_i *for every user* i *is computed using an* (ε, δ)-*local randomizer* Q_i*. When* $\delta > 0$*, we say* $\mathcal{A}$ *is* approximately *locally private. If* $\delta = 0$*,* $\mathcal{A}$ *is* purely *locally private.*

3 Known Variance

In this section, we present two solutions for the setting where the variance σ^2 is known (shorthanded "KV"). In Section 3.1, we analyze an *adaptive* protocol KVGausstimate that requires two rounds of analyst-user interaction. In Section 3.2, we analyze a *nonadaptive* protocol 1Round-KVGausstimate achieving a weaker accuracy guarantee in a single round.

3.1 Two-round protocol

We begin with a high-level overview of KVGausstimate before analyzing its components in detail. In KVGausstimate, the analyst splits the n users into halves U_1 and U_2, employing users from U_1 to compute an initial estimate of μ and then users from U_2 to further refine this estimate.

More concretely, the analyst partitions U_1 into $L = \lfloor n/(2k) \rfloor$ subgroups $U_1^1, \dots, U_1^L$ of size $k = \Omega\left(\frac{\log(n/\beta)}{\varepsilon^2}\right)$, where β is the desired failure probability. The analyst then solicits (a privatized version of) $\lfloor x_i/2^j \rfloor \bmod 4$ from each user in subgroup U_1^j. Each user responds by calling RR1, and the analyst aggregates these estimates through KVAgg1. By doing so, the analyst effectively executes a one-round binary search and obtains an initial $O(\sigma)$-accurate estimate $\hat{\mu}_1$ of μ.

The analyst then passes $\hat{\mu}_1$ to users in U_2 and solicits user estimates using (a privatized version of) a de-meaning protocol from the distributed statistical estimation literature [7]. Users in U_2 respond by calls to KVRR2, where each user i de-means their point using $\hat{\mu}_1$, standardizes it using σ, and randomized responds on $\text{sgn}((x_i - \hat{\mu}_1)/\sigma)$. Crucially, this de-meaning relies on knowing an $O(\sigma)$-accurate estimate of $\hat{\mu}$, which necessitates the first estimate $\hat{\mu}_1$. The analyst then uses KVAgg2 to aggregate these responses into an estimate of the CDF of $N(\mu, \sigma^2)$, from which the analyst can finally back out a final estimate $\hat{\mu}_2$. Pseudocode for KVGausstimate appears below. Throughout, we make the following assumptions on our problem parameters, deferring exact constants to the analysis. For neatness, let $L_{\min} = \lfloor \log(\sigma) \rfloor$, $L_{\max} = L + L_{\min} - 1$, and $\mathcal{L} = \{L_{\min}, L_{\min} + 1, \dots, L_{\max}\}$.

Assumption 3.1. $n = \Omega\left(\frac{\log(n/\beta)}{\varepsilon^2}\right)$ *and* $0 \le \mu = O(\sigma 2^{n\varepsilon^2/\log(n/\beta)})$ [2].

Algorithm 1 KVGausstimate

Input: $\varepsilon, k, \mathcal{L}, n, \sigma, U_1, U_2$

```
for j ∈ L do
    for user i ∈ U_1^j do
        User i outputs ỹ_i ← RR1(ε, i, j)
    end for
end for                                                    ▷ End of round 1
Analyst computes Ĥ_1 ← KVAgg1(ε, k, L, U_1)
Analyst computes μ̂_1 ← EstMean1(β, ε, Ĥ_1, k, L)
for user i ∈ U_2 do
    User i outputs ỹ_i ← KVRR2(ε, i, μ̂_1, σ)
end for                                                    ▷ End of round 2
Analyst computes Ĥ_2 ← KVAgg2(ε, n/2, U_2)
Analyst computes T̂ ← √2 · erf^{-1}( 2(−Ĥ_2(−1) + Ĥ_2(1)) / n )
Analyst outputs μ̂_2 ← σT̂ + μ̂_1
```

Output: Analyst estimate $\hat{\mu}_2$ of μ

We start our analysis with a privacy guarantee.

Theorem 3.2. KVGausstimate *satisfies* $(\varepsilon, 0)$*-local differential privacy for* $x_1, \dots, x_n$.

Proof. As KVGausstimate is sequentially interactive, each user only produces one output. It therefore suffices to show that each randomized response routine used in KVGausstimate is $(\varepsilon, 0)$-locally private. In RR1, for any possible inputs x, x' and output y we have

$$\frac{\mathbb{P}[\text{RR1}(x) = y]}{\mathbb{P}[\text{RR1}(x') = y]} \le \frac{e^\varepsilon/(e^\varepsilon + 3)}{1/(e^\varepsilon + 3)} \le e^\varepsilon$$

so RR1 is $(\varepsilon, 0)$-locally private. KVRR2 is $(\varepsilon, 0)$-locally private by similar logic. □

Next, we recall our overall accuracy result for KVGausstimate.

Theorem 3.3. *With probability at least* $1 - \beta$, KVGausstimate *outputs an estimate* $\hat{\mu}_2$ *such that*

$$|\hat{\mu}_2 - \mu| = O\left(\frac{\sigma}{\varepsilon}\sqrt{\frac{\log(1/\beta)}{n}}\right).$$

We prove this result by analyzing the execution of KVGausstimate in sequence below.

[2] While we assume μ is nonnegative, this is largely for convenience – all of our methods extend to negative (but similarly bounded) μ at the expense of constant factors.

3.1.1 Round one

We start with KVGAUSSTIMATE's first round of interaction. First, each user i in group U_1^j runs RR1(ε, i, j) to publish an ε-privatized version $\tilde{y}_i$ of $y_i = \lfloor x_i/2^j \rfloor \bmod 4$. Note that below $p \sim_U X$ denotes a uniform random draw p from set X.

Algorithm 2 RR1

Input: ε, i, j

$y_i \leftarrow \lfloor x_i/2^j \rfloor \bmod 4$
if $p \sim_U [0,1] \leq \frac{e^\varepsilon}{e^\varepsilon+3}$ **then**
    User i publishes $\tilde{y}_i \leftarrow y_i$
else
    User i publishes $\tilde{y}_i \sim_u (\{0,1,2,3\} \setminus \{y_i\})$
end if

Output: Private user estimate $\tilde{y}_i$ of $\mu(j)$

Since these randomized responses contain information about users' local estimates of each bit of μ, the analyst uses KVAGG1$(\varepsilon, k, \mathcal{L}, U_1)$ to aggregate them into histogram $\hat{H}_1$.

Algorithm 3 KVAGG1

Input: $\varepsilon, k, \mathcal{L}, U$

for $j \in \mathcal{L}$ **do**
    for $a \in \{0,1\}$ **do**
        $C^j(a) \leftarrow |\{\tilde{y}_i \mid i \in U^j, \tilde{y}_i = a\}|$
        $\hat{H}^j(a) \leftarrow \frac{e^\varepsilon+3}{e^\varepsilon-1} \cdot \left(C^j(a) - \frac{k}{e^\varepsilon+3}\right)$
    end for
end for
Output $\hat{H}$

Output: Aggregated histogram $\hat{H}$ of private user responses

---

Let H_1 be the "true" histogram, $H_1^j(a) = |\{y_i \mid i \in U_1^j, y_i = a\}|$ for all $a \in \{0,1,2,3\}$ and $j \in \mathcal{L}$. Since the analyst only has access to $\hat{H}_1$, we need to show that $\hat{H}_1$ and H_1 are similar.

Lemma 3.4. *With probability at least* $1 - \beta$, *for all* $j \in \mathcal{L}$,

$$||\hat{H}_1^j - H_1^j||_\infty \leq \left(\frac{\varepsilon+4}{\varepsilon\sqrt{2}}\right) \cdot \sqrt{k \ln(8L/\beta)}.$$

Proof. Choose $a \in \{0,1,2,3\}$ and $j \in \mathcal{L}$. $\mathbb{E}\left[C^j(a)\right] = \frac{H_1^j(a)e^\varepsilon}{e^\varepsilon+3} + \frac{k-H_1^j(a)}{e^\varepsilon+3} = \frac{H_1^j(a)(e^\varepsilon-1)+k}{e^\varepsilon+3}$, so by a pair of Chernoff bounds on the k users in U_1^j, with probability at least $1 - \beta/4L$,

$$|C^j(a) - \tfrac{H_1^j(a)(e^\varepsilon-1)+k}{e^\varepsilon+3}| \leq \sqrt{k \ln(8L/\beta)/2}.$$

Then since $\hat{H}_1^j(a) = \frac{e^\varepsilon+3}{e^\varepsilon-1} \cdot \left(C^j(a) - \frac{k}{e^\varepsilon+3}\right)$, this implies

$$|\hat{H}_1^j(a) - H_1^j(a)| \leq \frac{e^\varepsilon+3}{e^\varepsilon-1} \cdot \sqrt{k \ln(8L/\beta)/2} < \left(\frac{\varepsilon+4}{\varepsilon\sqrt{2}}\right) \cdot \sqrt{k \ln(8L/\beta)}$$

where the last step uses $\frac{e^{\varepsilon}+3}{e^{\varepsilon}-1} < \frac{\varepsilon+4}{\varepsilon}$. Union bounding over $a \in \{0,1,2,3\}$ and all L groups U_1^j completes the proof. □

Next, we show how the analyst uses $\hat{H}_1$ to estimate μ through EstMean1. Intuitively, in subgroup U_1^j when user responses concentrate in a single bin mod 4, this suggests that μ lies in the corresponding bin. In the other direction, when user responses do not concentrate in a single bin, users with points near μ must spread out over multiple bins, suggesting that μ lies near the boundary between bins. We formalize this intuition in EstMean1 and Lemma 3.5.

Algorithm 4 EstMean1

Input: $\beta, \varepsilon, \hat{H}_1, k, \mathcal{L}$
$\psi \leftarrow \left(\frac{\varepsilon+4}{\varepsilon\sqrt{2}}\right) \cdot \sqrt{k \ln(8L/\beta)}$
$j \leftarrow L_{\max}$
$I_j \leftarrow [0, 2^{L_{\max}}]$
while $j \geq L_{\min}$ and $\max_{a \in \{0,1,2,3\}} \hat{H}_1^j(a) \geq 0.52k + \psi$ **do**
  Analyst computes integer c such that $c2^j \in I_j$ and $c \equiv M_1(j) \bmod 4$
  Analyst computes $I_{j-1} \leftarrow [c2^j, (c+1)2^j]$
  $j \leftarrow j - 1$
end while
$j \leftarrow \max(j, L_{\min})$
Analyst computes $M_1(j) \leftarrow \arg\max_{a \in \{0,1,2,3\}} \hat{H}_1^j(a)$
Analyst computes $M_2(j) \leftarrow \arg\max_{a \in \{0,1,2,3\} - \{M_1(j)\}} \hat{H}_1^j(a)$
Analyst computes $c^* \leftarrow$ maximum integer such that $c^* 2^j \in I_j$ and $c^* \equiv M_1(j)$ or $M_2(j) \bmod 4$
Analyst outputs $\hat{\mu}_1 \leftarrow c^* 2^j$
Output: Initial estimate $\hat{\mu}_1$ of μ

Lemma 3.5. *Conditioned on the success of the preceding lemmas, with probability at least* $1 - \beta$, $|\hat{\mu}_1 - \mu| \leq 2\sigma$.

Proof. Recall the definitions of ψ, $M_1(j)$, and $M_2(j)$ from the pseudocode for EstMean1: $\psi = \left(\frac{\varepsilon+4}{\varepsilon\sqrt{2}}\right) \cdot \sqrt{k \ln(8L/\beta)}$, $M_1(j) = \arg\max_{a \in \{0,1,2,3\}} \hat{H}_1^j(a)$, and $M_2(j) = \arg\max_{a \in \{0,1,2,3\} - \{M_1(j)\}} \hat{H}_1^j(a)$. We start by proving two useful claims.

Claim 1: With probability at least $1 - \beta/5$, for all $j \in \mathcal{L}$ where $2^j > \sigma$, if $j' = L_{\max}, L_{\max} - 1, \ldots, j+1$ all have $\hat{H}_1^{j'}(M_1(j)) \geq 0.52k + \psi$, then $\mu \in I_j$.

To see why, suppose $2^j > \sigma$ and let $x \sim N(\mu, \sigma^2)$. Recall the Gaussian CDF $F(x) = \frac{1}{2}\left[1 + \text{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right]$. Then for any $a \not\equiv \lfloor \mu/2^j \rfloor \bmod 4$

$$\mathbb{P}\left[\lfloor x/2^j \rfloor \equiv a \bmod 4\right] \leq \mathbb{P}\left[x \notin [\mu, \mu + 3 \cdot 2^j)\right] < \mathbb{P}\left[x \notin [\mu, \mu + 3\sigma)\right] < 0.51$$

where the second inequality uses $2^j > \sigma$. Thus by a binomial Chernoff bound, the assumption $k > 5000 \ln(5L/\beta)$, and Lemma 3.4, with probability $\geq 1 - \beta/5L$, $\hat{H}_1^j(a) < 0.52k + \psi$. Therefore if for some a we have $\hat{H}_1^j(a) \geq 0.52k + \psi$, $a \equiv \lfloor \mu/2^j \rfloor \bmod 4$. Moreover, if $\mu \in I_j$ then letting c be the (unique) integer such that $c \equiv M_1(j) \bmod 4$ and $c2^j \in I_j$ (since I_j has endpoints $c_1 2^j$ and $(c_1+2)2^j$ for integer c_1) we get $\mu \in [c2^j, (c+1)2^j] = I_j$. As $\mu \in I_{L_{\max}}$ by Assumption 3.1, the claim follows by induction.

Claim 2: Let j be the maximum $j \in \mathcal{L}$ with $\hat{H}_1^j(M_1(j)) < 0.52k + \psi$, and let c^* be the maximum integer such that $c^* 2^j \in I_j$ and $c^* \equiv M_1(j)$ or $M_2(j) \bmod 4$. If $2^j > \sigma$, then with probability at least $1 - 4\beta/5$, $|c^* 2^j - \mu| \le 2\sigma$.

To see why, first note that by Claim 1, $\mu \in I_j$. Let $[c2^j, (c+1)2^j)$ be the subinterval of I_j containing μ for integer c. Then as $2^j > \sigma$, for $x \sim N(\mu, \sigma^2)$, by another application of the Gaussian CDF,

$$\mathbb{P}\left[x \in [c2^j, (c+1)2^j)\right] > \mathbb{P}\left[x \in [\mu, \mu + \sigma)\right] \ge 0.34.$$

Thus by the same method as above, using the assumption $k > 5000 \ln(5/\beta)$, with probability at least $1 - \beta/5$, $\hat{H}_1^j(c \bmod 4) \ge 0.33k - \psi$. By similar logic, since

$$\mathbb{P}\left[\lfloor x/2^j \rfloor \equiv c + 2 \bmod 4\right] < \max_{\lambda \in [0, 2^j]} \mathbb{P}\left[x \notin [\mu - 2^j - \lambda, \mu + 2 \cdot 2^j - \lambda]\right] < \mathbb{P}\left[x \notin [\mu - \sigma, \mu + 2\sigma)\right] \le 0.19$$

with probability at least $1 - \beta/5$, $\hat{H}_1^j(c + 2 \bmod 4) \le 0.2k + \psi$. Next, consider $\hat{H}_1^j(c - 1 \bmod 4)$. If $\mu \ge (c + 0.75)2^j$, then

$$\mathbb{P}\left[x \in [(c-1)2^j, c2^j)\right] \le \mathbb{P}\left[x \notin [\mu - 3\sigma/4, \mu + 9\sigma/4]\right] \le 0.24$$

so with probability at least $1 - \beta/5$

$$\hat{H}_1^j(c - 1 \bmod 4) \le 0.25k + \psi < 0.33k - \psi \le \hat{H}_1^j(c \bmod 4)$$

where the middle inequality uses $k > 625\left(\frac{\varepsilon+4}{\varepsilon\sqrt{2}}\right)^2 \ln(4L/\beta)$. Thus $c \equiv M_1(j)$ or $M_2(j) \bmod 4$; the $\mu \le (c + 0.25)2^j)$ case is symmetric. If instead $\mu \in ((c+0.25)2^j, (c+0.75)2^j)$ then by similar logic with probability at least $1 - \beta/5$

$$\hat{H}_1^j(c \bmod 4) \ge 0.36k - \psi.$$

so by $\psi < 0.08k$ (implied by $k > 40\left(\frac{\varepsilon+4}{\varepsilon\sqrt{2}}\right)^2 \ln(8L/\beta)$) $c \equiv M_1(j)$ or $M_2(j) \bmod 4$. It follows that with probability at least $1 - 3\beta/5$ in all cases $c \equiv M_1(j)$ or $M_2(j) \bmod 4$. Moreover, by a similar application of the Gaussian CDF, one of $c - 1 \bmod 4$ and $c + 1 \bmod 4$ lies in $\{M_1(j), M_2(j)\}$ as well.

Recalling that c^* is the maximum integer such that $c^* 2^j \in I_j$ and $c^* \equiv M_1(j)$ or $M_2(j) \bmod 4$, $c^* - 1 \bmod 4 \in \{M_1(j), M_2(j)\}$ as well. Assume $|c^* 2^j - \mu| > 2\sigma$. By above, $\mu \in [c^* 2^j, (c^*+1)2^j)$ or $[(c^*-1)2^j, (c^* 2^j))$. In the first case,

$$\mathbb{P}\left[\lfloor x/2^j \rfloor \equiv c^* - 1 \bmod 4\right] \le \mathbb{P}\left[x \notin [\mu - 2\sigma, \mu + 2\sigma]\right] \le 0.05$$

so with probability at least $1 - \beta/5$, $\hat{H}_1^j(c^* - 1) \le 0.06k + \psi$, a contradiction of $c^* - 1 \bmod 4 \in \{M_1(j), M_2(j)\}$. In the second case,

$$\mathbb{P}\left[\lfloor x/2^j \rfloor \equiv c^* \bmod 4\right] \le \mathbb{P}\left[x \notin [\mu - 2\sigma, \mu + 2\sigma]\right] \le 0.05$$

and with probability at least $1 - \beta/5$, $\hat{H}_1^j(c^*) \le 0.06k + \psi$, contradicting $c^* \bmod 4 \in \{M_1(j), M_2(j)\}$. Thus $|c^* 2^j - \mu| \le 2\sigma$.

We put these facts together in EstMean1 as follows: let j_1 be the maximum element of $\mathcal{L}$ such that $\hat{H}_1^j(M_1(j)) < 0.52k - \psi$. If $2^{j_1} > \sigma$, then by Fact 2 setting $\hat{\mu}_1 = c^* 2^j$ implies $|\hat{\mu}_1 - \mu| \le 2\sigma$. If instead $2^{j_1} \le \sigma$, then any setting of $\hat{\mu}_1 \in I_j$ (including $\hat{\mu}_1 = c^* 2^j$) guarantees $|\hat{\mu}_1 - \mu| \le 2^{j_1+1} \le 2\sigma$. Thus in all cases, with probability at least $1 - \beta$, $|\hat{\mu}_1 - \mu| \le 2\sigma$. □

3.1.2 Round two

The results above give the analyst an (initial) estimate $\hat{\mu}_1$ such that $|\hat{\mu}_1 - \mu| \le 2\sigma$. Now, the analyst passes this estimate $\hat{\mu}_1$ to users $i \in U_2$, and each user uses $\hat{\mu}_1$ to de-mean their value x_i and randomized respond on the resulting $(x_i - \hat{\mu}_1)/\sigma$ in KVRR2.

Algorithm 5 KVRR2

Input: $\varepsilon, i, \hat{\mu}_1, \sigma$

User i computes $x_i' \leftarrow (x_i - \hat{\mu}_1)/\sigma$
User i computes $y_i \leftarrow \text{sgn}(x_i')$
User i computes $c \sim_U [0,1]$
if $c \le \frac{e^\varepsilon}{e^\varepsilon + 1}$ **then**
    User i publishes $\tilde{y}_i \leftarrow y_i$
else
    User i publishes $\tilde{y}_i \leftarrow -y_i$
end if

Output: Private de-meaned user estimate $\tilde{y}_i$

De-meaning thus effectively transforms the problem of estimating μ into the problem of estimating μ when $|\mu|$ is small. This in turn enables us to use techniques for estimating the CDF near μ (specifically, a private version of Protocol 2 in Braverman et al. [7]).

Algorithm 6 KVAGG2

Input: ε, k, U

for $a \in \{-1, 1\}$ **do**
    $C(a) \leftarrow |\{\tilde{y}_i \mid i \in U, \tilde{y}_i = a\}|$
    $\hat{H}(a) \leftarrow \frac{e^\varepsilon + 1}{e^\varepsilon - 1} \cdot \left(C(a) - \frac{k}{e^\varepsilon + 1}\right)$
end for
Analyst outputs $\hat{H}$

Output: Aggregated histogram $\hat{H}$ of private user responses

We now prove that this de-meaning process results in a more accurate final estimate $\hat{\mu}_2$ of μ.

Lemma 3.6. *Conditioned on the success of the previous lemmas, with probability at least* $1 - \beta$ KVGAUSSTIMATE *outputs* $\hat{\mu}_2$ *such that*

$$|\hat{\mu}_2 - \mu| = O\left(\frac{\sigma}{\varepsilon}\sqrt{\frac{\log(1/\beta)}{n}}\right).$$

Proof. The proof is broadly similar to that of Theorem B.1 in Braverman et al. [7], with some modifications for privacy. First, by Lemma 3.5 $\mu - \hat{\mu}_1 \in [-2\sigma, 2\sigma]$. Letting $\bar{\mu} = (\mu - \hat{\mu}_1)/\sigma$ we get that $x_i' \sim N(\bar{\mu}, 1)$. Next, since $\mathbb{E}[y_i] = 2\mathbb{P}[x_i' \ge 0] - 1$, and in general

$$\Phi_{\mu,\sigma^2}(x) = \frac{1}{2}\left(1 + \text{erf}\left(\frac{x - \mu}{\sigma\sqrt{2}}\right)\right)$$

where Φ_{μ,σ^2} is the CDF of $N(\mu,\sigma^2)$, by $\Phi_{\bar{\mu},1}(0) = \mathbb{P}[x'_i \geq 0]$ we get $\mathbb{E}[y_i] = \text{erf}(\bar{\mu}/\sqrt{2})$ (note that we are analyzing the unprivatized values y_i to start; later, we will use this analysis to prove the analogous result for the privatized values $\tilde{y}_i$).

A Chernoff bound on $[-1,1]$-bounded random variables then shows that, with probability at least $1-\beta/2$, for $y = \frac{2}{n}\sum_{i\in U_2} y_i$ we have

$$|y - \text{erf}(\bar{\mu}/\sqrt{2})| \leq 2\sqrt{\ln(4/\beta)/n}$$

and by $\mathbb{E}[y] = \text{erf}(\bar{\mu}/\sqrt{2})$ we get $|y - \mathbb{E}[y]| \leq 2\sqrt{\ln(4/\beta)/n}$ as well.

Since $\mu - \hat{\mu}_1 \in [-2\sigma, 2\sigma]$, $|\text{erf}(\bar{\mu}/\sqrt{2})| \leq \text{erf}(\sqrt{2})$. Thus $|\mathbb{E}[y]| \leq \text{erf}(\sqrt{2})$, so by $|y - \mathbb{E}[y]| \leq 2\sqrt{\ln(4/\beta)/n}$ we get

$$|y| \leq \text{erf}(\sqrt{2}) + 2\sqrt{\ln(4/\beta)/n}.$$

Using $n > 20000\ln(4/\beta)$ we get $2\sqrt{\ln(4/\beta)/n} < 0.01$ and $\text{erf}(\sqrt{2}) < 0.96$, so $|y| \leq 0.97$ and thus $|y| < \text{erf}(1.6)$. Let M be an upper bound on the Lipschitz constant for erf^{-1} in $[-0.97, 0.97]$,

$$\begin{aligned} M &= \max_{x\in[-0.97,0.97]} \frac{d\text{erf}^{-1}(x)}{dx} \\ &= \max_{x\in[-0.97,0.97]} \frac{\sqrt{\pi}}{2}\exp([\text{erf}^{-1}(x)]^2) \\ &\leq \frac{\sqrt{\pi}}{2}\exp([\text{erf}^{-1}(0.97)]^2) < 10. \end{aligned}$$

Then for any $x, y \in [-0.97, 0.97]$ we have $|\text{erf}^{-1}(x) - \text{erf}^{-1}(y)| \leq M|x-y|$, so setting $T = \sqrt{2}\text{erf}^{-1}(y)$,

$$\begin{aligned} |T - \bar{\mu}| = |\sqrt{2}(\text{erf}^{-1}(y) - \text{erf}^{-1}(\mathbb{E}[y]))| &\leq 10\sqrt{2}|y - \mathbb{E}[y]| \\ &\leq 20\sqrt{2\ln(4/\beta)/n} \end{aligned}$$

using the bound on $|y - \mathbb{E}[y]|$ from above.

It remains to analyze the privatized values $\{\tilde{y}_i\}$ and bound $|T - \hat{T}|$, recalling that we set

$$\hat{T} = \sqrt{2}\cdot\text{erf}^{-1}\left(\frac{2(-\hat{H}_2(-1) + \hat{H}_2(1))}{n}\right)$$

in KVAgg1. By a Chernoff bound analogous to that of Lemma 3.4, with probability at least $1-\beta/2$

$$|T - \hat{T}| \leq \sqrt{2}\left|\text{erf}^{-1}(|y|) - \text{erf}^{-1}\left(|y| + \left[\frac{\varepsilon+2}{\varepsilon}\right]\sqrt{\frac{2\ln(4/\beta)}{n}}\right)\right|.$$

Using $n > 20000\left(\frac{\varepsilon+2}{\varepsilon}\right)^2\ln(4/\beta)$ (which implies $\left[\frac{\varepsilon+2}{\varepsilon}\right]\sqrt{\frac{2\ln(4/\beta)}{n}} \leq 0.01$) and the same derivative trick as above on $[-0.98, 0.98]$, we get

$$|T - \hat{T}| \leq 14\left[\frac{\varepsilon+2}{\varepsilon}\right]\sqrt{\frac{2\ln(4/\beta)}{n}}.$$

Therefore by the triangle inequality

$$|\hat{T} - \bar{\mu}| \leq \left(20 + 14\left[\frac{\varepsilon+2}{\varepsilon}\right]\right)\sqrt{\frac{2\ln(4/\beta)}{n}}$$

and by $\sigma\bar{\mu} = \mu - \hat{\mu}_1$ we get

$$|\sigma\hat{T} - \sigma\bar{\mu}| = |(\sigma\hat{T} + \hat{\mu}_1) - \mu| \leq \sigma\left(20 + 14\left\lceil\frac{\varepsilon+2}{\varepsilon}\right\rceil\right)\sqrt{\frac{2\ln(4/\beta)}{n}}.$$

Thus by taking $\hat{\mu}_2 = \sigma\hat{T} + \hat{\mu}_1$, we get

$$|\hat{\mu}_2 - \mu| = O\left(\frac{\sigma}{\varepsilon}\sqrt{\frac{\log(1/\beta)}{n}}\right).$$

□

3.2 One-round protocol

In this section, we provide a *nonadaptive* version 1RoundKVGausstimate of the protocol above. Recall from the previous section that KVGausstimate uses its first pool of users U_1 to estimate μ, and then passes this estimate $\hat{\mu}_1$ to its second pool of users U_2 to compute a more accurate estimate $\hat{\mu}_2$ of μ. At a high level, 1RoundKVGausstimate executes these two rounds of KVGausstimate *simultaneously* by parallelization. More concretely, 1RoundKVGausstimate splits its second user pool U_2 into $\Theta(\sqrt{\log(n)})$ subgroups and has each subgroup run the second-round protocol from KVGausstimate with different values of $\hat{\mu}_1$. Intuitively, as most users draw points clustered within $O(\sigma\sqrt{\log(n)})$, it suffices that these clustered users de-mean using a "good" guess for μ. Doing this naively would require $\Omega(2^{L_{\max}}/\sigma\sqrt{\log(n)})$ subgroups to ensure that at least one subgroup de-means using $\hat{\mu}$ that is $O(\sigma\sqrt{\log(n)})$ close to μ.

However, we can do better by leveraging the aforementioned clustering and using $\sigma\sqrt{\log(n)}$ as a modulus. We do this by associating with each subgroup U_2^j a set of points $S(j)$ interspersed $O(\sigma\sqrt{\log(n)})$ apart on the real line and having user $i \in U_2^j$ de-mean using the point in $S(j)$ closest to theirs. By defining these sets $S(j)$ carefully, we can guarantee that at least one group has most of its users de-mean using a point near μ.

These two processes come together as folows: at the end of the single round, the analyst aggregates the responses from users in U_1 to compute an estimate $\hat{\mu}_1$ of μ. By comparing $\hat{\mu}_1$ and $S(j)$, the analyst then selects the subgroup U_2^j where, with high probability, most users de-meaned using a value in $S(j)$ close to $\hat{\mu}_1$. This mimics the effect of adaptively passing $\hat{\mu}_1$ to the users in U_2 at the rough cost of a $\log^{1/4}(n)$ factor in accuracy, which results from splitting U_2 into approximately $\sqrt{\log(n)}$ groups. Pseudocode for 1RoundKVGausstimate appears below.

We start with the (slightly stronger) assumptions 1RoundKVGausstimate requires.

Assumption 3.7. *In addition to the assumptions of Assumption 3.1, we have* $n = \Omega\left(\frac{\log(n)\log(1/\beta)}{\varepsilon^2}\right)$.

1RoundKVGausstimate's privacy guarantee follows from the same analysis of randomized response as in KVGausstimate, so we state the guarantee but omit its proof.

Theorem 3.8. 1RoundKVGausstimate *satisfies* $(\varepsilon, 0)$*-local differentially privacy for* $x_1, \ldots, x_n$.

Next, we recall our overall accuracy result for 1RoundKVGausstimate.

Theorem 3.9. *With probability at least* $1-\beta$, 1RoundKVGausstimate *outputs an estimate* $\hat{\mu}_2$ *such that*

$$|\hat{\mu}_2 - \mu| = O\left(\frac{\sigma}{\varepsilon}\sqrt{\frac{\log(1/\beta)\sqrt{\log(n)}}{n}}\right).$$

We define k (here denoted k_1), $\mathcal{L}, U_1$, and U_2 as in KVGausstimate. As 1RoundKVGausstimate's treatment of users in U_1 is identical to that of KVGausstimate, we skip its analysis, instead recalling its final guarantee:

Lemma 3.10. *With probability at least* $1-\beta$, $|\hat{\mu}_1 - \mu| \le 2\sigma$.

Algorithm 7 1RoundKVGausstimate

Input: $\varepsilon, k_1, k_2, \mathcal{L}, n, R, S, \sigma, U_1, U_2$

```
for j ∈ L do
    for user i ∈ U_1^j do
        User i outputs ỹ_i ← RR1(ε, i, j)
    end for
end for
for j ∈ R do
    for user i ∈ U_2^j do
        User i outputs ỹ_i ← 1RoundKVRR2(ε, i, S(j))
    end for
end for                                                ▷ End of round 1
Analyst computes Ĥ_1 ← KVAgg1(ε, k_1, L, U_1)
Analyst computes μ̂_1 ← EstMean1(β, ε, Ĥ_1, k_1, L, )
Analyst computes j* ← argmin_{j∈R} min_{s∈S(j)} |s − μ̂_1|
Analyst computes Ĥ_2 ← KVAgg2(ε, k_2, U_2^{j*})
Analyst computes T̂ ← √2 · erf^{-1}((−Ĥ_2(−1) + Ĥ_2(1)) / k_2)
Analyst outputs μ̂_2 ← σT̂ + argmin_{s∈S(j*)} |s − μ̂_1|
```

Output: Analyst estimate $\hat{\mu}_2$ of μ

This brings us to U_2, and we define new parameters as follows. For neatness, let $\rho = \lceil 2\sqrt{\ln(4n)} \rceil \ge \lceil \sqrt{2\ln(2\sqrt{n})} + 2.1 \rceil$ for $n \ge 32$. We set $R = \{0.2\sigma, 0.4\sigma, \ldots, \rho\sigma\}$ and split U_2 into $|R| = 5\rho$ groups indexed by $j \in R$, each of size $k_2 \ge \lfloor n/2|R| \rfloor \ge \lfloor \frac{n}{20\sqrt{\ln(4n)}} \rfloor = \Omega(n/\sqrt{\log(n)})$, where the last inequality uses $n \ge 25$. Finally, for each $j \in R$ we define $S(j) = \{j + b\rho\sigma \mid b \in \mathbb{Z}\}$.

With this setup, for each $j \in R$ each user $i \in U_2^j$ uses 1RoundKVRR2 to execute a group-specific version of KVRR2: rather than de-meaning by $\hat{\mu}_1$ as in KVRR2, user i now de-means by the nearest point in $S(j)$ (breaking ties arbitrarily).

Algorithm 8 1RoundKVRR2

Input: $\varepsilon, i, S(j)$

User i computes $z_i \leftarrow \arg\min_{z_i \in S(j)} |z_i - x_i|$
User i computes $y_i \leftarrow \text{sgn}((x_i - z_i)/\sigma)$
User i computes $c \sim_U [0,1]$
if $c \le \frac{e^\varepsilon}{e^\varepsilon + 1}$ **then**
  User i publishes $\tilde{y}_i \leftarrow y_i$
else
  User i publishes $\tilde{y}_i \leftarrow -y_i$
end if

Output: Private de-meaned user estimate $\tilde{y}_i$

To analyze 1RoundKVRR2, we first prove that users in each group draw points concentrated around μ.

Lemma 3.11. *With probability at least $1 - \beta$, for all $j \in R$, group U_2^j contains $\le 2\sqrt{k_2}$ users i such that $|x_i - \mu| > \sigma\sqrt{\ln(4n)}$.*

Proof. First, by a Gaussian tail bound, for each user i, $\mathbb{P}\left[|x_i - \mu| \ge \sigma\sqrt{\ln(4n)}\right] \le 1/\sqrt{n}$. Let U_C^j denote the users in group U_2^j such that $|x_i - \mu| > \sigma\sqrt{\ln(4n)}$. Then by a binomial Chernoff bound

$$\mathbb{P}\left[|U^c| > \frac{k_2}{\sqrt{n}} + \sqrt{\frac{3k_2 \ln(|R|/\beta)}{\sqrt{n}}}\right] \le \beta/|R|$$

so using $n \ge 9\ln(|R|/\beta)^2$ and union bounding over $|R| = \Omega(\sqrt{\log(n)})$ groups, the claim follows. ☐

In particular, this implies that for $j^* = \arg\min_{j^* \in R} \min_{s \in S(j^*)} |s - \hat{\mu}_1|$ (i.e., the group with element of $S(j^*)$ closest to $\hat{\mu}_1$), most users draw points in $[\mu - \sigma\sqrt{\ln(4n)}, \mu + \sigma\sqrt{\ln(4n)}]$. Let $s^* = \min_{s \in S(j^*)} |s - \hat{\mu}_1|$. Our final accuracy result will rely on two facts. First, most users in $U_2^{j^*}$ de-mean using s^*. Second, the randomized responses of users who de-mean with s^* are "almost as good" as if they were de-meaned by μ.

Lemma 3.12. *Conditioned on the success of the previous lemmas, with probability at least $1 - \beta$,* 1RoundKVGausstimate *outputs $\hat{\mu}_2$ such that*

$$|\hat{\mu}_2 - \mu| = O\left(\frac{\sigma}{\varepsilon}\sqrt{\frac{\log(1/\beta)\sqrt{\log(n)}}{n}}\right).$$

Proof. Because adjacent points in R are 0.2σ apart, $|s^* - \hat{\mu}_1| \le 0.1\sigma$. Lemma 3.10 and the triangle inequality then imply that $|s^* - \mu| \le 2.1\sigma$. This enables us to mimic the proof of Lemma 3.6, replacing $\mu - \hat{\mu}_1 \in [-2\sigma, 2\sigma]$ with $\mu - s^* \in [-2.1\sigma, 2.1\sigma]$.

We can decompose users in $U_2^{j^*}$ into those with points within $\sigma\rho$ of s^* and those with more distant points. Denote the first set of users by V and the second set by V^c, and recall that the Gaussian CDF is

$$\Phi_{\mu,\sigma^2}(x) = \frac{1}{2}\left(1 + \text{erf}\left(\frac{x - \mu}{\sigma\sqrt{2}}\right)\right).$$

Then, letting $\mathbb{1}$ denote the indicator function,

$$\begin{aligned}\mathbb{E}\left[y_i \cdot \mathbb{1}(i \in V)\right] &= \mathbb{P}\left[y_i = 1, i \in V\right] - \mathbb{P}\left[y_i = -1, i \in V\right] \\ &= \Phi_{\mu,\sigma^2}(s^* + \sigma\rho) + \Phi_{\mu,\sigma^2}(s^* - \sigma\rho) - 2\Phi_{\mu,\sigma^2}(s^*) \\ &= \frac{1}{2}\left[\operatorname{erf}\left(\frac{s^* + \sigma\rho - \mu}{\sigma\sqrt{2}}\right) + \operatorname{erf}\left(\frac{s^* - \sigma\rho - \mu}{\sigma\sqrt{2}}\right)\right] - \operatorname{erf}\left(\frac{s^* - \mu}{\sigma\sqrt{2}}\right) \\ &= \frac{1}{2}\left[\operatorname{erf}\left(\frac{\sigma\rho + s^* - \mu}{\sigma\sqrt{2}}\right) - \operatorname{erf}\left(\frac{\sigma\rho - (s^* - \mu)}{\sigma\sqrt{2}}\right)\right] - \operatorname{erf}\left(\frac{s^* - \mu}{\sigma\sqrt{2}}\right).\end{aligned}$$

where the last step uses the fact that erf is an odd function. Since $\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}dt$ and $|s^* - \mu| \leq 2.1\sigma$,

$$\begin{aligned}\frac{1}{2}\left[\operatorname{erf}\left(\frac{\sigma\rho + s^* - \mu}{\sigma\sqrt{2}}\right) - \operatorname{erf}\left(\frac{\sigma\rho - (s^* - \mu)}{\sigma\sqrt{2}}\right)\right] &\leq \frac{1}{\sqrt{\pi}}\int_{(\sigma\rho - 2.1\sigma)/\sigma\sqrt{2}}^{(\sigma\rho + 2.1\sigma)/\sigma\sqrt{2}} e^{-t^2}dt \\ &< 3e^{-[(\rho - 2.1)/\sqrt{2}]^2} \\ &\leq 3e^{-\ln(4n)/2}\end{aligned}$$

where the second inequality relies on e^{-x} being monotone decreasing and the last step uses $n > 20$, which implies $\rho - 2.1 \geq \sqrt{\ln(4n)}$. Then using $n \geq 3k_2$ we get $3e^{-\ln(4n)/2} \leq \frac{1}{\sqrt{k_2}}$, so

$$\left|\mathbb{E}\left[y_i \cdot \mathbb{1}(i \in V)\right] - \operatorname{erf}\left(\frac{\mu - s^*}{\sigma\sqrt{2}}\right)\right| \leq \frac{1}{\sqrt{k_2}}. \tag{1}$$

Next, as $|s^* - \mu| \leq 2.1\sigma$, users having points within $\sigma\sqrt{2\ln(2\sqrt{n})}$ of μ have points within $\sigma\rho$ of s^*. The Gaussian tail bound from Lemma 3.11 then implies $\mathbb{P}\left[x \in V^c\right] \leq 1/\sqrt{n}$. $\mathbb{E}\left[y_i\right] = \mathbb{E}\left[y_i \cdot \mathbb{1}(i \in V)\right] + \mathbb{E}\left[y_i \cdot \mathbb{1}(i \in V^c)\right]$, and by the above bound on $\mathbb{P}\left[x \in V^c\right]$ and $|y_i| \leq 1$ we get $|\mathbb{E}\left[y_i \cdot \mathbb{1}(i \in V^c)\right]| \leq 1/\sqrt{n}$. Thus

$$|\mathbb{E}\left[y_i \cdot \mathbb{1}(i \in V)\right] - \mathbb{E}\left[y_i\right]| \leq \frac{1}{\sqrt{n}} < \frac{1}{\sqrt{k_2}}. \tag{2}$$

A Chernoff bound on $\{-1, 1\}$-valued random variables then tells us that, for $y = \frac{1}{k_2}\sum_{i \in U_2^{j^*}} y_i$, with probability at least $1 - \beta/2$ we have

$$|y - \mathbb{E}\left[y_i\right]| \leq \sqrt{\frac{2\ln(4/\beta)}{k_2}}. \tag{3}$$

Combining the three numbered equations above with the triangle inequality yields

$$\left|y - \operatorname{erf}\left(\frac{\mu - s^*}{\sigma\sqrt{2}}\right)\right| < \frac{2 + \sqrt{2\ln(4/\beta)}}{\sqrt{k_2}}.$$

Setting $\bar{\mu} = (\mu - s^*)/\sigma$ and using $k_2 \geq (100[2 + \sqrt{2\ln(4/\beta)})])^2$, this rearranges into $|y| \leq \operatorname{erf}(\bar{\mu}/\sqrt{2}) + 0.01$. Since $\bar{\mu} \in [-2.1, 2.1]$, we get

$$|y| < \operatorname{erf}(2.1/\sqrt{2}) + 0.01 < 0.98 < \operatorname{erf}(1.7).$$

Let M be an upper bound on the Lipschitz constant for erf^{-1} in $[-0.98, 0.98]$,

$$\begin{aligned} M &= \max_{x\in[-0.98,0.98]} \frac{d\mathrm{erf}^{-1}(x)}{dx} \\ &= \max_{x\in[-0.98,0.98]} \frac{\sqrt{\pi}}{2}\exp([\mathrm{erf}^{-1}(x)]^2) \\ &\le \frac{\sqrt{\pi}}{2}\exp([\mathrm{erf}^{-1}(0.98)]^2) < 14. \end{aligned}$$

Then for any $x, y \in [-0.98, 0.98]$ we have $|\mathrm{erf}^{-1}(x) - \mathrm{erf}^{-1}(y)| \le M|x-y|$, so for $T = \sqrt{2}\mathrm{erf}^{-1}(y)$,

$$\begin{aligned} |T - \bar{\mu}| = |\sqrt{2}(\mathrm{erf}^{-1}(y) - \mathrm{erf}^{-1}(\mathrm{erf}(\bar{\mu}/\sqrt{2}))| &\le 14\sqrt{2}|y - \mathrm{erf}(\bar{\mu}/\sqrt{2})| \\ &< 28\left(\frac{\sqrt{2}+\sqrt{\ln(4/\beta)}}{k_2}\right). \end{aligned}$$

It remains to bound $|T - \hat{T}|$, where T is the (unknown) aggregation of unprivatized $\{y_i\}$ while $\hat{T}$ is the (known) aggregation of privatized $\{\tilde{y}_i\}$. By a Chernoff bound analogous to that of Lemma 3.4, with probability at least $1 - \beta/2$

$$|T - \hat{T}| \le \sqrt{2}\left|\mathrm{erf}^{-1}(|y|) - \mathrm{erf}^{-1}\left(|y| + \left[\frac{\varepsilon+2}{\varepsilon}\right]\sqrt{\frac{2\ln(4/\beta)}{k_2}}\right)\right|.$$

Using $k_2 > 20000\left(\frac{\varepsilon+2}{\varepsilon}\right)^2\ln(4/\beta)$ (which implies $\left[\frac{\varepsilon+2}{\varepsilon}\right]\sqrt{2\frac{\ln(4/\beta)}{k_2}} \le 0.01$) and the same derivative trick as above on $[-0.99, 0.99]$, we get

$$|T - \hat{T}| \le 25\left[\frac{\varepsilon+2}{\varepsilon}\right]\sqrt{\frac{2\ln(4/\beta)}{k_2}}.$$

Therefore by the triangle inequality

$$|\hat{T} - \bar{\mu}| \le 28\left(\frac{\sqrt{2}+\sqrt{\ln(4/\beta)}}{k_2}\right) + 25\left[\frac{\varepsilon+2}{\varepsilon}\right]\sqrt{\frac{2\ln(4/\beta)}{k_2}} = O\left(\frac{1}{\varepsilon}\sqrt{\frac{\log(1/\beta)}{k_2}}\right)$$

and by $\sigma\bar{\mu} = \mu - s^*$ we get

$$|\sigma\hat{T} - \sigma\bar{\mu}| = |(\sigma\hat{T} + s^*) - \mu| = O\left(\frac{\sigma}{\varepsilon}\sqrt{\frac{\log(1/\beta)}{k_2}}\right).$$

Thus by taking $\hat{\mu}_2 = \sigma\hat{T} + s^*$ and substituting in $k_2 = \Omega(n/\sqrt{\log(n)})$ we get

$$|\hat{\mu}_2 - \mu| = O\left(\frac{\sigma}{\varepsilon}\sqrt{\frac{\log(1/\beta)\sqrt{\log(n)}}{n}}\right).$$

□

4 Unknown Variance

In this section, we consider the more general problem where σ is unknown but bounded by some interval $\sigma_{\min} \le \sigma \le \sigma_{\max}$.

4.1 Two-round protocol

Our adaptive solution, UVGaussTimate, uses two rounds. In round one, the analyst solicits user estimates for σ and μ, and in round two the analyst passes these estimates to another set of users to refine the estimate of μ. Accordingly, our protocol begins by halving the users into groups U_1 (to obtain $O(\sigma)$-accurate estimates of σ and μ in the first round) and U_2 (to refine the initial estimate of μ in the second round).

We start by describing U_1, the group of users estimating σ and μ. We split U_1 into $L_1 = \lfloor n/(2k_1) \rfloor$ subgroups $U_1^1, \dots, U_1^L$ of size $k_1 = \Omega\left(\frac{\log(n/\beta)}{\varepsilon^2}\right)$ (as in Section 3, we defer constants to the analysis). Let $L_{\min} = \lfloor \log(\sigma_{\min}) \rfloor$, $L_{\max} = L_1 + L_{\min} - 1 \ge \lceil \log(\sigma_{\max}) \rceil$, and $\mathcal{L}_1 = \{L_{\min}, L_{\min}+1, \dots, L_{\max}\}$.

Next, we leave U_2 as $n/2$ users without subgroups. All users in U_2 receive (via the analyst) estimate $\hat{\sigma}$ and $\hat{\mu}$ from U_1 and use these estimates to compute a final estimate $\hat{\mu}_2$ with Laplace noise. Roughly, we employ Laplace noise rather than the de-meaning process used in the known variance case because in the unknown variance case we lack the precise ($O(\sigma/\sqrt{n})$-accurate) estimate of σ that correctly de-meaning requires. Throughout, we require the following assumption on our problem parameters.

Assumption 4.1.

$$n = \Omega\left(\frac{\log\left(\frac{\sigma_{\max}}{\sigma_{\min}} + 1\right)\log\left(\frac{n}{\beta}\right)}{\varepsilon^2}\right) \textit{ and } \mu, \frac{\sigma_{\max}}{\sigma_{\min}} = O\left(2^{\frac{n\varepsilon^2}{\log(n/\beta)}}\right).$$

Algorithm 9 UVGaussTimate

Input: $\varepsilon, k_1, \mathcal{L}_1, n, \sigma, U_1, U_2$

for $j \in \mathcal{L}_1$ **do**
  for user $i \in U_1^j$ **do**
    User i outputs $\tilde{y}_i \leftarrow \mathrm{RR1}(\varepsilon, i, j)$
  end for
end for ▷ End of round 1
Analyst computes $\hat{H}_1 \leftarrow \textsc{Agg1}(\varepsilon, \mathcal{L}_1, U_1)$
Analyst computes $\hat{\sigma} \leftarrow \textsc{EstVar}(\beta, \varepsilon, \hat{H}_1, k_1, \mathcal{L}_1)$
Analyst computes $\hat{H}_2 \leftarrow \textsc{KVAgg1}(\varepsilon, k_1, \mathcal{L}_1, U_1)$
Analyst computes $\hat{\mu}_1 \leftarrow \textsc{EstMean1}(\beta, \varepsilon, \hat{H}_1, k_1, \mathcal{L}_1)$
Analyst computes $I \leftarrow [\hat{\mu}_1 - \hat{\sigma}(2 + \sqrt{\ln(4n)}), \hat{\mu}_1 + \hat{\sigma}(2 + \sqrt{\ln(4n)})]$
for user $i \in U_2$ **do**
  User i outputs $\tilde{y}_i \leftarrow \mathrm{UVRR2}(\varepsilon, i, I)$
end for ▷ End of round 1
Analyst outputs $\hat{\mu}_2 \leftarrow \frac{2}{n}\sum_{i \in U_2} \tilde{y}_i$

Output: Analyst estimate $\hat{\mu}_2$ of μ

We begin our analysis with overall privacy and accuracy guarantees.

Theorem 4.2. UVGaussTIMATE *satisfies* $(\varepsilon, 0)$*-local differentially privacy for* $x_1, \ldots, x_n$.

Proof. As we already proved that RR1 is private in Section 3.1, we are left with UVRR2. To prove that UVRR2 is $(\varepsilon, 0)$-locally differentially private as well, we can use a standard Laplace noise privacy guarantee (see e.g. Theorem 3.6 in Dwork et al. [14]): given function f with 1-sensitivity Δf, computing $f(x) + \mathsf{Lap}(\Delta f / \varepsilon)$ satisfies $(\varepsilon, 0)$-differential privacy. □

Theorem 4.3. *With probability at least* $1 - \beta$, UVGaussTIMATE *outputs an estimate* $\hat{\mu}_2$ *such that*

$$|\hat{\mu}_2 - \mu| = O\left(\frac{\sigma}{\varepsilon}\sqrt{\frac{\log(1/\beta)\log(n)}{n}}\right).$$

First, for each $j \in \mathcal{L}_1$ each user i in group U_1^j runs $\mathrm{RR1}(\varepsilon, i, j)$ (originally defined in Section 3.1.1) to publish an ε-privatized version $\tilde{y}_i$ of $y_i = \lfloor x_i / 2^j \rfloor \bmod 4$. Estimating σ relies on similar logic as estimating μ: when $2^j \gg \sigma$, user responses in group U_1^j appear concentrated, and when $2^j \ll \sigma$ then user responses in group U_1^j appear unconcentrated. Examining this transition from concentrated to unconcentrated responses yields an estimate for σ. Estimation of μ, since it does not require knowledge of σ, differs from the subroutine EstMean1 used in KVGaussTIMATE only in the choice of $\mathcal{L}_1$. Here, $\mathcal{L}_1$ ranges from $L_{\min}$ to $L_{\max} \geq \lceil \log(\sigma_{\max}) \rceil$ to account for uncertainty about σ.

The analyst aggregates responses from users in U_1 in two ways. First, the analyst computes a collection of histograms $\hat{H}_1$ using Agg1. $\hat{H}_1$ is an estimate of the "true" histogram collection, $H^j(a) = |\{y_i \mid i \in U_1^j, y_i \in \{a, a+1 \bmod 4\}\}|$ for all $j \in \mathcal{L}_1$. As in Lemma 3.4, we can show that $\hat{H}_1$ and H_1 are similar. As the proof is nearly identical, we omit it.

Lemma 4.4. *With probability at least* $1 - \beta$, *for all* $j \in \mathcal{L}_1$,

$$\|\hat{H}_1^j - H_1^j\|_\infty \leq \left(1 + \tfrac{4}{\varepsilon}\right)\sqrt{2k_1 \ln(8L_1/\beta)}.$$

Algorithm 10 Agg1

Input: $\varepsilon, k, \mathcal{L}, U$

for $j \in \mathcal{L}$ **do**
    for $a = 0, 1, 2, 3$ **do**
        Analyst computes $C^j(a) \leftarrow |\{i \mid i \in U^j, \tilde{y}_i = a\}|$
        Analyst computes $\hat{H}^j(a) \leftarrow \frac{e^\varepsilon + 3}{e^\varepsilon - 1} \cdot \left(C^j(a) - \frac{k}{e^\varepsilon + 3}\right)$
    end for
    for $a = 0, 1, 2, 3$ **do**
        Analyst computes $\hat{H}_1^j(a) \leftarrow \hat{H}^j(a) + \hat{H}^j(a + 1 \bmod 4)$
    end for
end for
Analyst outputs $\hat{H}_1$

Output: Analyst aggregation $\hat{H}_1$ of private user estimates

Next, we show how the analyst uses $\hat{H}_1$ to estimate σ in subroutine EstVar. Here, for neatness we shorthand

$$\tau = \sqrt{2k_1 \ln(2L_1/\beta)} + \left(1 + \tfrac{4}{\varepsilon}\right)\sqrt{2k_1 \ln(8L_1/\beta)}$$

and use the term *concentrated* to denote any histogram $\hat{H}_1^j$ such that $\min_{a\in\{0,1,2,3\}} \hat{H}_1^j(a) \le 0.03k+\tau$ and the term *unconcentrated* to denote $\hat{H}_1^j$ where $\min_{a\in\{0,1,2,3\}} \hat{H}_1^j(a) \ge 0.04k-\tau$. As we show below in Lemma 4.5, when $2^j \gg \sigma$, $\hat{H}_1^j$ is concentrated. Similarly, when $2^j \ll \sigma$, $\hat{H}_1^j$ is unconcentrated. This transition enables the analyst to estimate σ.

Algorithm 11 EstVar

Input: $\beta, \varepsilon, \hat{H}_1, k_1, \mathcal{L}_1$

```
Analyst computes j ← minimum j such that, for all j' ≥ j, Ĥ_1^{j'} is concentrated
if j = ∅ then
    Analyst outputs σ̂ ← 2^{L_max}
else
    Analyst outputs σ̂ ← 2^j
end if
```

Output: Analyst estimate $\hat{\sigma}$ of σ

Lemma 4.5. *Conditioned on the success of the preceding lemmas, with probability at least* $1-\beta$, EstVar *outputs* $\hat{\sigma} \in [\sigma, 8\sigma]$.

Proof. Choose $j \in \mathcal{L}_1$. Below, we reason about two (non-exhaustive) possibilities for j.

Case 1: $2^j \ge 4\sigma$. Then there exists $a \in \{0,1,2,3\}$ and interval I of length $2^{j+1} \ge 8\sigma$ containing $[\mu - 2\sigma, \mu + 2\sigma]$ such that for all $x \in I$, $\lfloor x/2^j \rfloor \bmod 4 \equiv a$ or $a+1 \bmod 4$. By similar application of the Gaussian CDF as in Lemma 3.5, with probability at least $1-\beta/2L_1$,

$$|\{x_i \mid x_i \in I, i \in U_1^j\}| \ge 0.97k_1 - \sqrt{2k_1 \ln(2L_1/\beta)}.$$

Thus by Lemma 4.4, $\hat{H}_1^j(a) \ge 0.97k_1 - \tau$. It follows that $\hat{H}_1^j(a+2) \le 0.03k_1 + \tau$. $2^j \ge 4\sigma$ thus implies that histogram $\hat{H}_1^j$ is concentrated.

Case 2: $2^j \in [\sigma/2, \sigma]$. Choose $a \in \{0,1,2,3\}$. Since $2^j \in [\sigma/2, \sigma]$ there exist at most three subintervals $I_1, I_2, I_3 \subset [\mu - 2\sigma, \mu + 2\sigma]$ such that for all $x \in I = I_1 \cup I_2 \cup I_3$, $\lfloor x/2^j \rfloor \equiv a \bmod 4$, and $|I| \ge \sigma$. Let $x \sim N(\mu, \sigma^2)$. Then by a similar application of the Gaussian CDF as in Lemma 3.5, since

$$\mathbb{P}[x \in I] \ge \mathbb{P}[x \in [\mu - 2\sigma, \mu - \sigma)] \ge 0.13$$

with probability $1-\beta/8L_1$ at least $0.13k - \sqrt{2k_1 \ln(8L_1/\beta)}$ users from U_1^j have points in I. Since this held for arbitrary a, a union bound over all four possibilities of a combined with Lemma 4.4 implies that, with probability at least $1-\beta/2L_1$,

$$\min_{a\in\{0,1,2,3\}} \hat{H}_1^j(a) \ge 0.13k_1 - \tau.$$

$2^j \le \sigma \le 2^{j+1}$ thus implies that histogram $\hat{H}_1^j$ is uniform.

Union bounding both results over $j \in \mathcal{L}_1$, with $k_1 > 800\left(2+\frac{4}{\varepsilon}\right)^2 \ln(8L_1/\beta)$, with probability $1-\beta$ we have $0.13k - \tau > 0.03k + \tau$ for each $j \in \mathcal{L}_1$. Therefore for all $j \in \mathcal{L}_1$ if $2^j \ge 4\sigma$ then $\hat{H}_1^j$ will be concentrated while if $2^{j+1} \ge \sigma \ge 2^j$ then $\hat{H}_1^j$ will be unconcentrated.

Let j be the smallest $j \in \mathcal{L}_1$ such that $\hat{H}_1^j$ is concentrated and for all $j' > j$, $\hat{H}_1^{j'}$ is concentrated as well. If no such j exists, then we know $2^{L_{\max}} \ge \sigma \ge 2^{L_{\max}-2}$, take $\hat{\sigma} = 2^{L_{\max}}$, and we get $\hat{\sigma} \in [\sigma, 4\sigma]$. If not, then by Case 1 above we know $2^j \le 8\sigma$, and by Case 2 we know $2^j \ge \sigma$. Thus taking $\hat{\sigma} = 2^j$, we get $\hat{\sigma} \in [\sigma, 8\sigma]$. ∎

Next, the analyst uses randomized responses from U_1 to compute an initial estimate $\hat{\mu}_1$ of μ. As the process EstMean1 is identical to that used in KVGausstimate up to a different subgroup range $\mathcal{L}_1$, we skip its description and only recall its guarantee:

Lemma 4.6. *Conditioned on the success of the preceding lemmas, with probability at least* $1-\beta$*,* $|\hat{\mu}_1 - \mu| \leq 2\sigma$*.*

From the results above, the analyst obtains an estimate $\hat{\sigma}$ such that $\hat{\sigma} \in [\sigma, 8\sigma]$ and an estimate $\hat{\mu}_1$ such that $|\hat{\mu}_1 - \mu| \leq 2\sigma$. The analyst now uses these to compute interval $I = [\hat{\mu}_1 - \hat{\sigma}(2 + \sqrt{\ln(4n)}), \hat{\mu}_1 + \hat{\sigma}(2 + \sqrt{\ln(4n)})]$, where I is intentionally constructed to (with high probability) contain the points of $\Omega(n)$ users. The analyst then passes I to users in U_2. Users in U_2 respond with noisy responses via independent calls to UVRR2. In UVRR2, each user clips their sample x_i to the interval I and reports a private version $\tilde{y}_i$ using Laplace noise scaled to $|I|$.

At a high level, we employ Laplace noise in this way because Laplace noise requires a small interval I to be a useful privatization method: if I is large, then the $\mathsf{Lap}(|I|/\varepsilon)$ noise required for privacy will be large as well. At the same time, I must be large enough to contain the points of most users. Constructing such an I therefore requires rough estimates of both μ and σ, leading to the two-round approach used here.

Algorithm 12 UVRR2

Input: ε, i, I

1: User i computes $x_i' \leftarrow \arg\min_{x \in I} |x - x_i|$
2: User i outputs $\tilde{y}_i \leftarrow x_i' + \mathsf{Lap}(|I|/\varepsilon)$

Output: Private version of user's point clipped to I

Informally, since I contains the points of an overwhelming fraction of users, the average of noisy points clipped to I will be close to the expected average μ. The analyst can therefore average these user randomized responses to compute a more accurate final estimate of μ.

Lemma 4.7. *Conditioned on the success of the previous lemmas, with probability at least* $1-\beta$*,* $|\hat{\mu}_2 - \mu| = O\left(\frac{\sigma}{\varepsilon}\sqrt{\frac{\log(1/\beta)\log(n)}{n}}\right)$*.*

Proof. There are two sources of error in the analyst's estimate $\hat{\mu}_2 = \frac{2}{n}\sum_i \tilde{y}_i$: error from the unnoised x_i's and error from noise in $\tilde{y}_i$s. Specifically, recalling that $|U_2| = n/2$, we can decompose $\hat{\mu}_2$ as

$$\hat{\mu}_2 = \frac{2}{n}\sum_i \tilde{y}_i = \frac{2}{n}\sum_i (x_i' + \eta_i)$$

where each $\eta_i \sim_{i.i.d.} \mathsf{Lap}(|I|/\varepsilon)$ and $|I| = 2\hat{\sigma}(2 + \sqrt{\ln(4n)})$.

First, using $n > 4\ln(3/\beta)$ by concentration of independent Laplace random variables (see e.g. Lemma 2.8 in Chan et al. [9]) with probability at least $1 - \beta/3$,

$$\left|\frac{2}{n}\sum_i \eta_i\right| \leq \frac{4|I|}{\varepsilon}\sqrt{\frac{2\ln(3/\beta)}{n}} \leq \frac{8\hat{\sigma}(2+\sqrt{\ln(4n)})}{\varepsilon}\sqrt{\frac{2\ln(3/\beta)}{n}} = O\left(\frac{\hat{\sigma}}{\varepsilon}\sqrt{\frac{\log(1/\beta)\log(n)}{n}}\right).$$

This bounds the contribution of Laplace noise to overall error.

It remains to bound $|\frac{2}{n}\sum_i x_i' - \mu|$. Let V denote the set of users with $x_i \in I$ and V^c denote the set of users with $x_i \notin I$. First, by a Gaussian tail bound, for each user i, $\mathbb{P}\left[|x_i - \mu| \geq \sigma\sqrt{\ln(4n)}\right] \leq 1/\sqrt{n}$. Then by a Chernoff bound

$$\mathbb{P}\left[|V^c| > \left(1 + \sqrt{\frac{6\ln(3/\beta)}{n^{3/2}}}\right)\sqrt{n}\right] \leq \beta/3$$

and using $n \geq (6\ln(2/\beta))^{2/3}$ we get $\sqrt{\frac{6\ln(3/\beta)}{n^{3/2}}} \leq 1$, so with probability at least $1 - \beta/3$, $|V^c| \leq 2\sqrt{n}$. Thus

$$\frac{2}{n}\sum_{i \in V^c} |x_i' - \mu| \leq \frac{2}{n}(|V^c| \cdot |I|) \leq \frac{6\hat{\sigma}(2 + \sqrt{\ln(4n)})}{\sqrt{n}} = O\left(\frac{\hat{\sigma}\sqrt{\log(n)}}{\sqrt{n}}\right).$$

This bounds the contribution of error from the (unprivatized) data of users in V^c. Let V denote the set of users in U_2 with points in I. We bound the error contributed by users in V in a similar way. Users in V have $x_i' = x_i$, so by a Chernoff bound on (shifted) $[0, |I|]$-bounded random variables, with probability at least $1 - \beta/3$

$$\frac{2}{n}\sum_{i \in V^c} |x_i' - \mu| = \frac{2}{n}\sum_{i \in V^c} |x_i - \mu| \leq |I|\sqrt{\frac{2\ln(6/\beta)}{n}} \leq \hat{\sigma}(2 + \sqrt{\ln(4n)})\sqrt{\frac{2\ln(6/\beta)}{n}} = O\left(\frac{\hat{\sigma}\sqrt{\log(1/\beta)\log(n)}}{\sqrt{n}}\right).$$

Putting these three bounds together, we get

$$\begin{aligned}
\left|\frac{2}{n}\sum_i \tilde{y}_i - \mu\right| &\leq \frac{2}{n}\sum_i |x_i' + \eta_i - \mu| \\
&\leq \frac{2}{n}\sum_i |x_i' - \mu| + \frac{2}{n}\sum_i |\eta_i| \\
&= \frac{2}{n}\sum_{i \in V} |x_i' - \mu| + \frac{2}{n}\sum_{i \in V^c} |x_i' - \mu| + \frac{2}{n}\sum_i |\eta_i| \\
&= O\left(\frac{\sigma}{\varepsilon}\sqrt{\frac{\log(n)\log(1/\beta)}{n}}\right)
\end{aligned}$$

where the last step uses $\hat{\sigma} \in [\sigma, 8\sigma]$ from Lemma 4.5. □

4.2 One-round protocol

Here we provide a nonadaptive one-round version of UVGAUSSTIMATE, 1ROUNDUVGAUSSTIMATE. As in 1ROUNDKVGAUSSTIMATE, 1ROUNDUVGAUSSTIMATE will simulate the second round of UVGAUSSTIMATE simultaneously with its first round.

More concretely, UVGAUSSTIMATE constructed an interval I based on estimates $\hat{\mu}_1$ of μ and $\hat{\sigma}$ of σ, passed I to users in U_2, and users in U_2 responded with noisy versions of their points clipped to I with $\mathsf{Lap}(|I|/\varepsilon)$ noise. 1ROUNDUVGAUSSTIMATE instead splits U_2 into subgroups, where each subgroup responds using *different* intervals I. As in 1ROUNDKVGAUSSTIMATE, at the end of the single round the analyst obtains estimates $\hat{\mu}_1$ and $\hat{\sigma}$ from users in U_1, constructs an interval I from these estimates, and finds a subgroup of U_2 where most users employed an interval I' similar to I. This similarity guarantees that the responses from that subgroup get the same accuracy as the two-round case up to an $O\left(\sqrt{\log\left(\frac{\sigma_{\max}}{\sigma_{\min}}+1\right)\sqrt{\log(n)}}\right)$ factor. Pseudocode for 1RoundUVGausstimate appears below.

1RoundUVGausstimate's privacy guarantee follows from the same analysis of randomized response and Laplace noise as for UVGausstimate, so we omit its proof.

Theorem 4.8. 1RoundUVGausstimate *satisfies* $(\varepsilon, 0)$*-local differentially privacy for* $x_1, \dots, x_n$.

Next, we recall 1RoundUVGausstimate's accuracy guarantee before proving it below.

Algorithm 13 1RoundUVGausstimate

Input: $\varepsilon, k_1, k_2, \mathcal{L}_1, n, R, \sigma, U_1, U_2,$

Analyst computes $\rho \leftarrow \lceil \sqrt{2\ln(2\sqrt{n})} + 6 \rceil$
for $j \in \mathcal{L}_1$ **do**
  for user $i \in U_1^j$ **do**
    User i outputs $\tilde{y}_i \leftarrow \mathrm{RR1}(\varepsilon, i, j)$
  end for
end for
for $j_1 \in \mathcal{L}_1$ **do**
  for $j_2 \in R_{j_1}$ **do**
    for user $i \in U_2^{j_1, j_2}$ **do**
      User i outputs $\tilde{y}_i \leftarrow$ 1RoundUVRR2$(\varepsilon, i, j_1, j_2, \rho, S)$
    end for
  end for
end for ▷ End of round 1
Analyst computes $\hat{H}_1 \leftarrow$ Agg1$(\varepsilon, k_1, \mathcal{L}_1, U_1)$
Analyst computes $\hat{\sigma} \leftarrow$ EstVar$(\beta, \varepsilon, \hat{H}_1, k_1, \mathcal{L})$
Analyst computes $j_1 \leftarrow \log(\hat{\sigma})$
Analyst computes $\hat{H}_2 \leftarrow$ KVAgg1$(\varepsilon, k_1, \mathcal{L}_1, U_1)$
Analyst computes $\hat{\mu}_1 \leftarrow$ EstMean1$(\beta, \varepsilon, \hat{H}_2, k_1, \mathcal{L}_1)$
Analyst computes $j_2 \leftarrow \arg\min_{j \in R_{j_1}} (\min_{s \in S(j_1, j)} |s - \hat{\mu}_1|)$
Analyst computes $s^* \leftarrow \min_{s \in S(j_1, j_2)} |s - \hat{\mu}_1|$
Analyst outputs $\hat{\mu}_2 \leftarrow s^* + \frac{1}{k_2} \sum_{i \in U_2^{j_1, j_2}} \tilde{y}_i$

Output: Analyst estimate $\hat{\mu}_2$ of μ

Theorem 4.9. *Wih probability at least* $1-\beta$*,* 1RoundUVGausstimate *outputs* $\hat{\mu}_2$ *such that*

$$|\hat{\mu}_2 - \mu| = O\left(\frac{\sigma}{\varepsilon}\sqrt{\frac{\log\left(\frac{\sigma_{\max}}{\sigma_{\min}}+1\right)\log(1/\beta)\log^{3/2}(n)}{n}}\right).$$

We define $k_1, \mathcal{L}_1$, and U_1, as in UVGausstimate and skip the analysis of 1RoundUVGausstimate's treatment of users in U_1 as it is identical to that of UVGausstimate. We recall its collected guarantee:

Lemma 4.10. *With probability at least* $1-\beta$*,* $\hat{\sigma} \in [\sigma, 8\sigma]$ *and* $|\hat{\mu}_1 - \mu| \le 2\sigma$*.*

We again define R and S for U_2, albeit with a few modifications. First, we let $\rho = \lceil \sqrt{\ln(4n)} + 6 \rceil$ for neatness. Then, recalling from Section 4.1 that $\mathcal{L}_1$ ranges over possible values of $\log(\sigma)$, for each $j_a \in \mathcal{L}_1$ we define $R_{j_a} = \{2^{j_a}, 2 \cdot 2^{j_a}, \dots, \rho \cdot 2^{j_a}\}$. Next, for each $j_a \in \mathcal{L}_1$ and $j_b \in R_{j_a}$, we define $S(j_a, j_b) = \{j_b + b\rho 2^{j_a} \mid b \in \mathbb{Z}\}$. Finally, we split U_2 into $L_1 \cdot \rho$ subgroups $U_2^{j_a,j_b}$ of size $k_2 = \Omega\left(\frac{n}{\log\left(\frac{\sigma_{\max}}{\sigma_{\min}}+1\right)\sqrt{\log(n)}}\right)$ for each $j_a \in \mathcal{L}_1$ and $j_b \in R_{j_a}$. As in 1RoundKVGausstimate, we parallelize over these subgroups to simulate the second round of UVGausstimate for different values of (j_a, j_b). The assumptions required are the same as those of Assumption 4.1.

In each subgroup $U_2^{j_a,j_b}$, each user i computes the nearest element $s_i \in S(j_a, j_b)$ to x_i, $s_i = \arg\min_{s \in S(j_a,j_b)} |x_i - s|$ and outputs $x_i - s_i$ plus Laplace noise in 1RoundUVRR2. The analyst then uses estimates $j_1 = \lceil \log(\hat{\sigma}) \rceil$ and $\hat{\mu}_1$ from U_1 to compute $j_2 = \arg\min_{j \in R_{j_1}} (\min_{z \in S(j_1,j)} |z - \hat{\mu}_1|)$. Finally, the analyst aggregates randomized responses from group $U_2^{j_1,\hat{\mu}_2}$ into an estimate $\hat{\mu}_2$.

Algorithm 14 1RoundUVRR2

Input: $\varepsilon, i, j_1, j_2, \rho, S$

User i computes $s_i \leftarrow \min_{s \in S(j_1,j_2)} |s - x_i|$
User i computes $y_i \leftarrow x_i - s_i$
User i outputs $\tilde{y}_i \leftarrow y_i + \mathsf{Lap}\left(2\rho 2^{j_1}/\varepsilon\right)$

Output: Private version of user's point x_i

As in 1RoundKVGausstimate, we start with a concentration result for each $U_2^{j_1,j_2}$. Since its proof is similar to that of Lemma 3.11, we omit it.

Lemma 4.11. *With probability at least $1 - \beta$, for all $j_1 \in \mathcal{L}_1$ and $j_2 \in R_{j_1}$, group $U_2^{j_1,j_2}$ contains $\leq 2\sqrt{k_2}$ users i such that $|x_i - \mu| > \sigma\sqrt{\ln(4n)}$.*

In combination with the previous lemmas, this enables us to prove our final accuracy result.

Lemma 4.12. *Conditioned on the success of the previous lemmas, with probability at least $1 - \beta$, 1RoundUVGausstimate outputs $\hat{\mu}_2$ such that*

$$|\hat{\mu}_2 - \mu| = O\left(\frac{\sigma}{\varepsilon}\sqrt{\frac{\log\left(\frac{\sigma_{\max}}{\sigma_{\min}} + 1\right)\log(1/\beta)\log^{3/2}(n)}{n}}\right).$$

Proof. By Lemma 4.10, $\hat{\sigma} \in [\sigma, 8\sigma]$ and $|\hat{\mu}_1 - \mu| \leq 2\sigma$. Since $j_1 = \log(\hat{\sigma}) \in \mathcal{L}_1$ and $j_2 = \arg\min_{j \in R_{j_1}} (\min_{s \in S(j_1,j)} |s - \hat{\mu}_1|)$, by the definition of $s^* \in S(j_1, j_2)$, $|s^* - \hat{\mu}_1| \leq 0.5\hat{\sigma} < 4\sigma$. Thus $|s^* - \mu| < 6\sigma$.

Consider group $U_2^{j_1,j_2}$. By Lemma 4.11 at most $2\sqrt{k_2}$ users $i \in U_2^{j_1,j_2}$ have $|x_i - \mu| > \sigma\sqrt{\ln(4n)}$. Thus by $|s^* - \mu| < 6\sigma$ and the fact that any two points in $S(j_1, j_2)$ are at least $\hat{\sigma}\rho \geq \sigma(6 + \sqrt{\ln(4n)})$ far apart, we get that at least $k_2 - 2\sqrt{k_2}$ users $i \in U_2^{j_1,j_2}$ set $s_i = s^*$ in their run of 1RoundUVRR2. Denote this subset of users by V, and denote by V^c the set of users $i \in U_2^{j_1,j_2}$ such that $s_i \neq s^*$, and for each user $i \in U_2$ let $y_i = x_i - s_i$.

Let $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp(-(x - \mu)^2/2\sigma^2)$, the density for $N(\mu, \sigma^2)$. Then

$$\int_{\infty}^{\infty} (x - \mu) f(x) dx = \int_{-\infty}^{s^* - \rho\hat{\sigma}} (x - \mu) f(x) dx + \int_{s^* - \rho\hat{\sigma}}^{s^* + \rho\hat{\sigma}} (x - \mu) f(x) dx + \int_{s^* + \rho\hat{\sigma}}^{\infty} (x - \mu) f(x) dx. \quad (4)$$

Let $g(x) = -\frac{\sigma}{\sqrt{2\pi}} \exp(-(x-\mu)^2/2\sigma^2)$, the antiderivative of $(x-\mu)f(x)$. Then

$$\begin{aligned}
\left|\int_{-\infty}^{s^*-\rho\hat{\sigma}} (x-\mu)f(x)dx\right| &= \left|g(s^*-\rho\hat{\sigma}) - \lim_{b\to-\infty} g(b)\right| \\
&= \left|\frac{\sigma}{\sqrt{2\pi}} \cdot \exp\left(-\frac{(s^*-\rho\hat{\sigma}-\mu)^2}{2\sigma^2}\right)\right| \\
&\le \left|\frac{\sigma}{\sqrt{2\pi}} \cdot \exp\left(-\frac{([6-\rho]\sigma)^2}{2\sigma^2}\right)\right| \\
&\le \left|\frac{\sigma}{\sqrt{2\pi}} \cdot \exp\left(-\frac{[6-\rho]^2}{2}\right)\right| \\
&< \frac{\sigma}{\sqrt{2\pi}} \cdot \exp(-\ln(2\sqrt{n})) \\
&< \frac{\sigma}{\sqrt{n}}
\end{aligned}$$

where the first inequality uses $\hat{\sigma} \ge \sigma$ and $|s^* - \mu| < 6\sigma$. Similar logic implies $\left|\int_{s^*+\rho\hat{\sigma}}^{\infty} (x-\mu)f(x)dx\right| \le \sigma/\sqrt{n}$ as well. Therefore by Equation 4 and $\int_{-\infty}^{\infty}(x-\mu)f(x)dx = 0$,

$$\left|\int_{s^*-\rho\hat{\sigma}}^{s^*+\rho\hat{\sigma}} (x-\mu)f(x)dx\right| \le 2\sigma/\sqrt{n}$$

so by $\mathbb{E}[x_i \cdot \mathbb{1}(i \in V)] = \int_{s^*-\rho\hat{\sigma}}^{s^*+\rho\hat{\sigma}} xf(x)dx$, we get

$$\left|\mathbb{E}[x_i \cdot \mathbb{1}(i \in V)] - \mu \int_{s^*-\rho\hat{\sigma}}^{s^*+\rho\hat{\sigma}} f(x)dx\right| \le 2\sigma/\sqrt{n}.$$

Since $\mathbb{E}[x_i \cdot \mathbb{1}(i \in V)]/\mathbb{P}[i \in V] = \mathbb{E}[x_i \mid i \in V]$ and $\mathbb{P}[i \in V] = \int_{s^*-\rho\hat{\sigma}}^{s^*+\rho\hat{\sigma}} f(x)dx$, this means

$$|\mathbb{E}[x_i \mid i \in V] - \mu| \le 2\sigma/\sqrt{n}.$$

By $y_i = x_i - s^*$ for $i \in V$,

$$|\mathbb{E}[y_i \mid i \in V] - (\mu - s^*)| \le 2\sigma/\sqrt{n}.$$

We can therefore decompose

$$\begin{aligned}
\left|\frac{1}{k_2} \sum_{i\in U_2^{j_1,j_2}} y_i - (\mu - s^*)\right| &\le \left|\frac{1}{k_2}\sum_{i\in V}(y_i - (\mu - s^*))\right| + \left|\frac{1}{k_2}\sum_{i\in V^c}(y_i - (\mu - s^*))\right| \\
&\le \left[\frac{2\sigma}{\sqrt{n}} + \rho\hat{\sigma}\sqrt{\frac{2\log(4/\beta)}{k_2}}\right] + \frac{2\rho\hat{\sigma}}{\sqrt{k_2}} \\
&= O\left(\sigma\sqrt{\frac{\log(1/\beta)\log(n)}{k_2}}\right)
\end{aligned}$$

where the the first inequality uses a (with probability at least $1-\beta/2$) Chernoff bound on $\{y_i \mid i \in V\}$ concentrating around $\mathbb{E}[y_i \mid i \in V]$ as well as $|V^c| \le 2\sqrt{k_2}$, and the last step uses $\hat{\sigma} \in [\sigma, 8\sigma]$.

Next, since we can decompose

$$\frac{1}{k_2}\sum_{i\in U_2^{j_1,j_2}}\tilde{y}_i = \frac{1}{k_2}\sum_{i\in U_2^{j_1,j_2}} y_i + \frac{1}{k_2}\sum_{i\in U_2^{j_1,j_2}}\eta_i$$

where each $\eta_i \sim \mathsf{Lap}(\rho\hat{\sigma}/\varepsilon)$, the same concentration of Laplace noise from Lemma 4.7 says that with probability $1-\beta/2$,

$$\left|\frac{1}{k_2}\sum_{i=1}^{k_2}\eta_i\right| = O\left(\frac{\rho\hat{\sigma}}{\varepsilon}\sqrt{\frac{\log(1/\beta)}{k_2}}\right) = O\left(\frac{\sigma}{\varepsilon}\sqrt{\frac{\log(1/\beta)\log(n)}{k_2}}\right).$$

Combining with the bound above and substituting in $k_2 = \Omega\left(\frac{n}{\log\left(\frac{\sigma_{\max}}{\sigma_{\min}}+1\right)\sqrt{\log(n)}}\right)$,

$$\left|\frac{1}{k_2}\sum_{i\in U_2^{j_1,j_2}}\tilde{y}_i - (\mu - s^*)\right| = O\left(\frac{\sigma}{\varepsilon}\sqrt{\frac{\log\left(\frac{\sigma_{\max}}{\sigma_{\min}}\right)\log(1/\beta)\log^{3/2}(n)}{n}}\right).$$

The claim then follows from $\hat{\mu}_2 = s^* + \frac{1}{k_2}\sum_{i\in U_2^{j_1,j_2}}\tilde{y}_i$. □

5 Lower Bound

In this section we prove that the upper bounds proven in Sections 3 and 4 are tight up to log factors for sequentially interactive protocols. Our argument proceeds in three steps: first, in Section 5.1 we show that any protocol achieving good performance on our *estimation* problem achieves good performance on a more easily analyzed *testing* problem. Second, in Section 5.2 we prove a lower bound for $(\varepsilon,0)$-locally private protocols on these testing problems. Finally, we use recent work demonstrating that noninteractive pure and approximate local privacy are "equivalent" for sequentially interactive protocols [8, 10] to generalize our lower bound to protocols satisfying (ε,δ)-local privacy.

For completeness, we start with the more general notion of sequential interactivity used by Duchi et al. [12], which requires that the set of messages $\{Y_i\}$ sent by the users satisfies the following conditional independence structure:

$$\{X_i, Y_1,\dots,Y_{i-1}\}\to Y_i \qquad \text{and,} \qquad Y_i \perp X_j \mid \{X_i,Y_1,\dots,Y_{i-1}\} \text{ for } j\neq i.$$

5.1 From Estimation to Testing

We begin by defining a way to transform an instance of $\mathsf{Estimate}$ into a (formally) easier testing problem Test. We start by defining an instance $\mathsf{Estimate}(n,M,\sigma)$. Here, a protocol receives n samples from a $N(\mu,\sigma^2)$ distribution where σ is known, $\mu\in[0,M]$, and the goal is to estimate μ. Next, define $V\sim_U\{0,1\}$. Consider the following testing problem: for $V=v$, if $v=0$, then each user i draws a sample $x_i\sim_{iid} N(0,\sigma^2)$, while if $v=1$ then each user i draws a sample $x_i\sim_{iid} N(M,\sigma^2)$. The testing problem $\mathsf{Test}(n,M,\sigma)$ is to recover v from $x_1,\dots,x_n$ with high probability.

We will say that a protocol $\mathcal{A}$ (α,β)-estimates $\mathsf{Estimate}(n,M,\sigma)$ if, with probability at least $1-\beta$, $\mathcal{A}(\mathsf{Estimate}(n,M,\sigma)) = \hat{\mu}$ such that $|\hat{\mu}-\mu|<\alpha$. Similarly, we will say that an algorithm $\mathcal{A}$ β-solves $\mathsf{Test}(n,M,\sigma)$ if, with probability at least $1-\beta$, $\mathcal{A}(\mathsf{Test}(n,M,\sigma)) = v$. We now show that $\mathsf{Test}(n,M,\sigma)$ is formally no harder than $\mathsf{Estimate}(n,M,\sigma)$.

Lemma 5.1. *If there exists a sequentially interactive and (ε, δ)-locally private protocol $\mathcal{A}$ that $(M/2, \beta)$-estimates* $\mathsf{Estimate}(n, M, \sigma)$*, then there exists a sequentially interactive and (ε, δ)-locally private protocol $\mathcal{A}'$ that β-solves* $\mathsf{Test}(n, M, \sigma)$*.*

Proof. Let $x_1, \ldots, x_n$ be the samples from an instance of $\mathsf{Test}(n, M, \sigma)$. We define $\mathcal{A}'$ to run $\mathcal{A}(x_1, \ldots, x_n)$ and then output $\arg\min_{\hat{\mu} \in \{0, M\}} |\mathcal{A}(x_1, \ldots, x_n) - \hat{\mu}|$. Since $\mathcal{A}$ $(M/2, \beta)$-estimates $\mathsf{Estimate}(n, M, \sigma)$, with probability at least $1 - \beta$, $|\mathcal{A}(x_1, \ldots, x_n) - \mu| < M/2$. Thus with probability at least $1 - \beta$, $\mathcal{A}'(x_1, \ldots, x_n) = v$. Thus $\mathcal{A}'$ β-solves $\mathsf{Test}(n, M, \sigma)$. As $\mathcal{A}'$ interacted with $x_1, \ldots, x_n$ only through (ε, δ)-locally private $\mathcal{A}$, by preservation of differential privacy under post-processing, $\mathcal{A}'$ is (ε, δ)-locally private as well. Similar logic implies that $\mathcal{A}'$ is also sequentially interactive. ☐

5.2 Lower Bounds for Test

Next, we show that Test is hard for $(\varepsilon, 0)$-locally private protocols. As our result uses some tools from information theory, a brief overview of information theory basics appears in the Appendix.

Lemma 5.2. *Suppose $M \le \sigma / [4(e^{\varepsilon} - 1)\sqrt{2nc}]$,where c is an absolute constant. For any sequentially interactive and $(\varepsilon, 0)$-locally private protocol $\mathcal{A}$ that β-solves* $\mathsf{Test}(n, M, \sigma)$*, $\beta \ge 1/4$.*

Proof. We may express any sequentially interactive $(\varepsilon, 0)$-locally private protocol $\mathcal{A}$ that β-solves $\mathsf{Test}(n, M, \sigma)$ as a Markov chain $V \to X \to Y \to Z$, where V is the random variable selecting v, $X = (x_1, \ldots, x_n)$ is the random variable for users' i.i.d. samples, $Y = (y_1, \ldots, y_n)$ is the random variable for users' $(\varepsilon, 0)$-privatized responses, and $Z = \mathcal{A}(\mathsf{Test}(n, M, \sigma))$. As $V \to X \to Y \to Z$ is a Markov chain (i.e., any two random variables in the chain are conditionally independent given a random variable between them). Thus by a strong data processing inequality for two Gaussians (see e.g. Section 4.1 in Braverman et al. [7] or, for a broader treatment of strong data processing inequalities, Raginsky [23]), there exists absolute constant c such that for each user i, $I(V; Y_i) \le \frac{cM^2}{\sigma^2} I(X_i; Y_i)$, where $I(A; B)$ denotes the mutual information between random variables A and B. Next, since our protocol is $(\varepsilon, 0)$-locally private, by Corollary 1 from Duchi et al. [12], for each user i, $I(X_i; Y_i) \le 4(e^{\varepsilon} - 1)^2$. With the equation above, we get

$$I(V; Y_i) \le \frac{4cM^2(e^{\varepsilon} - 1)^2}{\sigma^2}. \tag{5}$$

Without loss of generality, suppose Z is a deterministic function of Y [3]. From Markov chain

[3]This is without loss of generality because if Z is a random function of Y then it decomposes into a convex combination of deterministic functions of Y.

$V \to X \to Y \to Z$ and the (generic) data processing inequality we get

$$\begin{aligned} I(V;Z) &\le I(V;Y_1,\ldots,Y_n) \\ &= \sum_{i=1}^n I(V;Y_i \mid Y_{i-1},\ldots Y_1) \\ &\le \sum_{i=1}^n I(V,Y_{i-1},\ldots,Y_1;Y_i) \\ &= \sum_{i=1}^n [I(V;Y_i) + I(Y_{i-1},\ldots,Y_1;Y_i|V)] \\ &= \sum_{i=1}^n I(V;Y_i) \end{aligned}$$

where the last step follows from the independence of Y_i and $Y_1,\ldots,Y_{i-1}$ given V. Substituting in Equation 5, $I(V;Z) \le \frac{4ncM^2(e^\varepsilon-1)^2}{\sigma^2}$. Therefore by $M \le \sigma/4(e^\varepsilon - 1)\sqrt{2nc}$ we get $I(V;Z) \le 1/8$.

Define P to be the distribution of Z (over the randomness of V, X, and Y), and let P_0 and P_1 be the distributions for $Z|V = 0$ and $Z|V = 1$ respectively. Then as V is uniform, $P = (P_0 + P_1)/2$, so

$$||P - P_0||_1 = ||P - P_1||_1 = \tfrac{1}{2}||P_0 - P_1||_1.$$

Moreover, by

$$\begin{aligned} \mathbb{P}[Z = V] &= \mathbb{P}[Z = 0, V = 0] + \mathbb{P}[Z = 1, V = 1] \\ &= \frac{1}{2}(P_0(0) + [1 - P_1(0)]) \\ &\le \frac{1}{2}(1 + |P_0(0) - P_1(0)|) \\ &= \frac{1}{2} + \frac{1}{4}||P_0 - P_1||_1 \end{aligned}$$

we get $\mathbb{P}[Z = V] \le \frac{1}{2} + \frac{1}{4}||P_0 - P_1||_1$. Thus

$$\begin{aligned} \frac{||P_0 - P_1||_1^2}{8} &= \frac{1}{4}(||P_0 - P||_1^2 + ||P_1 - P||_1^2) \\ &\le \frac{1}{2}(D_{KL}(P_0||P) + D_{KL}(P_1||P)) \\ &= I(Z;V) \le 1/8 \end{aligned}$$

where the second-to-last inequality uses Pinsker's inequality. It follows that $||P_0 - P_1||_1 \le 1$. Substituting this into $\mathbb{P}[Z = V] \le \frac{1}{2} + \frac{1}{4}||P_0 - P_1||_1$, we get $\mathbb{P}[Z = V] \le \frac{3}{4}$. □

It remains to show that Test is hard for (ε, δ)-locally private protocols. Our result follows almost immediately from existing work [8, 10] showing, roughly, that any sequentially interactive (ε, δ)-locally private protocol $\mathcal{A}$ may be transformed into a sequentially interactive $(O(\varepsilon), 0)$-locally private protocol $\mathcal{A}'$ with similar behavior[4].

[4]While both of these results are stated for noninteractive protocols, it is straightforward to see that their techniques carry over to sequentially interactive protocols. Specifically, both results rely on transforming a single user call to an (ε, δ)-local randomizer into calls to an $(O(\varepsilon), 0)$-local randomizer. Thus, since users in sequentially interactive protocols still only make a single call to a local randomizer, we can apply the same transformations to each single user call and obtain an $(O(\varepsilon), 0)$-locally private sequentially interactive protocol.

Lemma 5.3. *Let $\varepsilon > 0$ and $\delta < \min\left(\frac{\epsilon\beta}{48n\ln(2n/\beta)}, \frac{\beta}{16n\ln(n/\beta)e^{7\varepsilon}}\right)$, and suppose that $\mathcal{A}$ is a sequentially interactive and (ε,δ)-locally private protocol. If $\mathcal{A}$ β-solves $\mathsf{Test}(n,M,\sigma)$, then there exists a sequentially interactive $(10\varepsilon,0)$-locally private $\mathcal{A}'$ that 4β-solves $\mathsf{Test}(n,M,\sigma)$.*

Proof. Our analysis splits into two cases depending on ϵ.

Case 1: $\varepsilon \le 1/4$. In this case, we use a result from Bun et al. [8], included here for completeness.

Lemma 5.4 (Theorem 6.1 in Bun et al. [8] (restated))**.** *Given $\varepsilon \le 1/4$ and $\delta < \epsilon\beta/48n\ln(2n/\beta)$, there exists a $(10\varepsilon,0)$-locally private algorithm $\mathcal{A}'$ such that for every database $U = \{x_1,\dots,x_n\}$, $d_{TV}(\mathcal{A}(U),\mathcal{A}'(U)) \le \beta$, where d_{TV} denotes total variation distance.*

Thus, denoting by $E_{\mathcal{A}}$ the event where $\mathcal{A}$ recovers the correct v on $\mathsf{Test}(n,M,\sigma)$ and $E_{\mathcal{A}'}$ the event where $\mathcal{A}'$ recovers the correct v on $\mathsf{Test}(n,M,\sigma)$, $|\mathbb{P}[E_{\mathcal{A}}] - \mathbb{P}[E_{\mathcal{A}'}]| \le \beta$, where the probabilities are respectively over $\mathcal{A}$ and $\mathcal{A}'$. Thus since $\mathcal{A}$ β-solves $\mathsf{Test}(n,M,\sigma)$, it follows that $\mathcal{A}'$ 2β-solves (and thus also 4β-solves) $\mathsf{Test}(n,M,\sigma)$.

Case 2: $\varepsilon > 1/4$. In this case we use a result from Cheu et al. [10][5]

Lemma 5.5 (Theorem A.1 in Cheu et al. [10] (restated))**.** *Given $\varepsilon > 1/4$ and $\delta < \frac{\beta}{16n\ln(n/\beta)e^{7\varepsilon}}$, there exists an $(8\varepsilon,0)$-locally private protocol $\mathcal{A}'$ such that $\mathcal{A}'$ 4β-solves $\mathsf{Test}(n,M,\sigma)$.*

Our result follows. ☐

5.3 Lower Bound for Estimate

We combine the preceding results to prove a general lower bound for Estimate as follows: for appropriate ε and δ, by Lemma 5.1 any sequentially interactive and $(\frac{\varepsilon}{10},\delta)$-locally private protocol $\mathcal{A}$ that $(M/2,\frac{\beta}{4})$-estimates $\mathsf{Estimate}(n,M,\sigma)$ implies the existence of a sequentially interactive and $(\frac{\varepsilon}{10},\delta)$-locally private protocol $\mathcal{A}'$ that $\frac{\beta}{4}$-solves $\mathsf{Test}(n,M,\sigma)$. Then, Lemma 5.3 implies the existence of a sequentially interactive and $(\varepsilon,0)$-locally private protocol $\mathcal{A}''$ that β-solves $\mathsf{Test}(n,M,\sigma)$. By Lemma 5.2 any such $\mathcal{A}'$ that β-solves $\mathsf{Test}(n,M,\sigma)$ has $\beta \ge 1/4$. We condense this reasoning into the following theorem.

Theorem 5.6. *Let $\varepsilon > 0$ and $\delta < \min\left(\frac{\epsilon\beta}{60n\ln(5n/2\beta)}, \frac{\beta}{16n\ln(n/\beta)e^{7\varepsilon}}\right)$, and let $\mathcal{A}$ be a sequentially interactive (ε,δ)-locally private (α,β)-estimator for $\mathsf{Estimate}(n,M,\sigma)$ where $M = \sigma/4(e^{\varepsilon}-1)\sqrt{2nc}$ and $\beta < 1/16$. Then*

$$\alpha \ge M/2 = \Omega\left(\frac{\sigma}{\varepsilon}\sqrt{\frac{1}{n}}\right).$$

In particular, this implies that the upper bounds of Sections 3 and 4 are tight up to logarithmic factors for *any* sequentially interactive and (ε,δ)-locally private protocol with sufficiently small δ.

6 Acknowledgements

We thank Shahin Jabbari for feedback on an early draft and Akshay Krishnamurthy for useful discussion of the lower bound.

[5] Cheu et al. [10] originally state their result for $\varepsilon > 2/3$, but mildly strengthening their assumed upper bound on δ from $\delta < \frac{\beta}{8n\ln(n/\beta)e^{6\varepsilon}}$ to $\delta < \frac{\beta}{16n\ln(n/\beta)e^{7\varepsilon}}$ yields the result here.

References

[1] John M. Abowd. The challenge of scientific reproducibility and privacy protection for statistical agencies. Technical report, Census Scientific Advisory Committee, 2016.

[2] Jayadev Acharya, Clément L Canonne, Cody Freitag, and Himanshu Tyagi. Test without trust: Optimal locally private distribution testing. *arXiv preprint arXiv:1808.02174*, 2018.

[3] Differential Privacy Team Apple. Learning with privacy at scale. Technical report, Apple, 2017.

[4] Brendan Avent, Aleksandra Korolova, David Zeber, Torgeir Hovden, and Benjamin Livshits. Blender: enabling local search with a hybrid differential privacy model. In *Proc. of the 26th USENIX Security Symposium*, pages 747–764, 2017.

[5] Raef Bassily and Adam Smith. Local, private, efficient protocols for succinct histograms. In *Proceedings of the forty-seventh annual ACM symposium on Theory of computing*, pages 127–135. ACM, 2015.

[6] Andrea Bittau, Úlfar Erlingsson, Petros Maniatis, Ilya Mironov, Ananth Raghunathan, David Lie, Mitch Rudominer, Ushasree Kode, Julien Tinnes, and Bernhard Seefeld. Prochlo: Strong privacy for analytics in the crowd. In *Proceedings of the 26th Symposium on Operating Systems Principles*, SOSP '17, pages 441–459, 2017.

[7] Mark Braverman, Ankit Garg, Tengyu Ma, Huy L Nguyen, and David P Woodruff. Communication lower bounds for statistical estimation problems via a distributed data processing inequality. In *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, pages 1011–1020. ACM, 2016.

[8] Mark Bun, Jelani Nelson, and Uri Stemmer. Heavy hitters and the structure of local privacy. In *Proceedings of the 35th ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems*, pages 435–447. ACM, 2018.

[9] T-H Hubert Chan, Elaine Shi, and Dawn Song. Private and continual release of statistics. *ACM Transactions on Information and System Security (TISSEC)*, 14(3):26, 2011.

[10] Albert Cheu, Adam Smith, Jonathan Ullman, David Zeber, and Maxim Zhilyaev. Distributed differential privacy via mixnets. *arXiv preprint arXiv:1808.01394*, 2018.

[11] Bolin Ding, Janardhan Kulkarni, and Sergey Yekhanin. Collecting telemetry data privately. In *Advances in Neural Information Processing Systems*, pages 3574–3583, 2017.

[12] John C Duchi, Michael I Jordan, and Martin J Wainwright. Local privacy and statistical minimax rates. In *Foundations of Computer Science (FOCS), 2013 IEEE 54th Annual Symposium on*, pages 429–438. IEEE, 2013.

[13] 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.

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

[15] Úlfar Erlingsson, Vasyl Pihur, and Aleksandra Korolova. Rappor: Randomized aggregatable privacy-preserving ordinal response. In *Proceedings of the 2014 ACM SIGSAC conference on computer and communications security*, pages 1054–1067. ACM, 2014.

[16] Alexandre Evfimievski, Johannes Gehrke, and Ramakrishnan 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. ACM, 2003.

[17] Marco Gaboardi and Ryan Rogers. Local private hypothesis testing: Chi-square tests. *arXiv preprint arXiv:1709.07155*, 2017.

[18] Marco Gaboardi, Ryan Rogers, and Or Sheffet. Locally private mean estimation: Z-test and tight confidence intervals. *arXiv preprint arXiv:1810.08054*, 2018.

[19] Gautam Kamath, Jerry Li, Vikrant Singhal, and Jonathan Ullman. Privately learning high-dimensional distributions. *arXiv preprint arXiv:1805.00216*, 2018.

[20] Vishesh Karwa and Salil Vadhan. Finite Sample Differentially Private Confidence Intervals. In Anna R. Karlin, editor, *9th Innovations in Theoretical Computer Science Conference (ITCS 2018)*, pages 44:1–44:9, 2018.

[21] Shiva Prasad Kasiviswanathan, Homin K Lee, Kobbi Nissim, Sofya Raskhodnikova, and Adam Smith. What can we learn privately? *SIAM Journal on Computing*, 40(3):793–826, 2011.

[22] Yu-Hsuan Kuo, Cho-Chun Chiu, Daniel Kifer, Michael Hay, and Ashwin Machanavajjhala. Differentially private hierarchical group size estimation. *arXiv preprint arXiv:1804.00370*, 2018.

[23] Maxim Raginsky. Strong data processing inequalities and ϕ-sobolev inequalities for discrete channels. *IEEE Transactions on Information Theory*, 62(6):3355–3389, 2016.

[24] Or Sheffet. Locally private hypothesis testing. In Jennifer Dy and Andreas Krause, editors, *Proceedings of the 35th International Conference on Machine Learning*, volume 80, pages 4605–4614. PMLR, 2018.

[25] Jonathan Ullman. Tight lower bounds for locally differentially private selection. *arXiv preprint arXiv:1802.02638*, 2018.

[26] Salil Vadhan. The complexity of differential privacy. In *Tutorials on the Foundations of Cryptography*, pages 347–450. Springer, 2017.

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

A Information Theory

We briefly review some standard facts and definitions from information theory, starting with entropy.

Definition A.1. *The* entropy $H(X)$ *of a random variable* X *is*

$$H(X) = \sum_x \mathbb{P}[X = x] \ln\left(\frac{1}{\mathbb{P}[X=x]}\right),$$

and the conditional entropy $H(X|Y)$ *of random variable* X *conditioned on random variable* Y *is*

$$H(X|Y) = \mathbb{E}_y[H(X|Y = y)].$$

Next, we can use entropy to define the mutual information between two random variables. Mutual information between random variables X and Y is roughly the amount by which conditioning on Y reduces the entropy of X (and vice-versa).

Definition A.2. *The* mutual information $I(X;Y)$ *between two random variables* X *and* Y *is*

$$I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X),$$

and the conditional mutual information $I(X;Y|Z)$ *between* X *and* Y *given* Z *is*

$$I(X;Y|Z) = H(X|Z) - H(X|Y,Z) = H(Y|Z) - H(Y|X,Z).$$

We also define the related notion of KL-divergence.

Definition A.3. *The* Kullback-Leibler divergence $D_{KL}(X\|Y)$ *between two random variables* X *and* Y *is*

$$D_{KL}(X\|Y) = \sum_x \mathbb{P}[X = x] \ln\left(\frac{\mathbb{P}[X = x]}{\mathbb{P}[Y = x]}\right),$$

where we often abuse notation and let X *and* Y *denote the distributions associated with* X *and* Y*.*

KL divergence connects to mutual information as follows.

Fact A.4. *For random variables* X*,* Y*, and* Z*,*

$$I(X;Y|Z) = \mathbb{E}_{x,z}\left[D_{KL}\left((Y|X = x, Z = z)\|(Y|Z = z)\right)\right].$$

Finally, we will also use the following connection between KL divergence and $\|\cdot\|_1$ distance.

Lemma A.5 (Pinsker's inequality)**.** *For random variables* X *and* Y*,*

$$\|X - Y\|_1 \leq \sqrt{2D_{KL}(X\|Y)}.$$