Contents

A iDP-SignRP Under Individual Differential Privacy (iDP)

A.1 Relaxation: Individual Differential Privacy (iDP)

Many extensions or relaxation of DP have been proposed to improve the utility of DP mechanisms. Examples include Concentrated Differential Privacy [4], Rényi Differential Privacy [10], and Gaussian Differential Privacy [3]. These alternatives provide better composition properties than the composition theorems of DP, thus reducing the noise needed [6]. Another possible direction to elevate the empirical performance of DP is to relax the DP definition by constraining the scope of neighboring datasets depending on the specific use case of DP [12; 2]. In this paper, we consider the concept called "individual differential privacy" (iDP), also known as "data-centric DP", as follows.

Definition A.1 (Individual DP [12])**.** *Given a dataset U, an algorithm $\mathcal{M}$ satisfies (ϵ, δ)-iDP for U if for any dataset U' that is adjacent to U, it holds that*

$$Pr[\mathcal{M}(U) \in O] \leq e^{\epsilon} Pr[\mathcal{M}(U') \in O] + \delta,$$
$$Pr[\mathcal{M}(U') \in O] \leq e^{\epsilon} Pr[\mathcal{M}(U) \in O] + \delta.$$

We should emphasize that, individual DP does not satisfy the rigorous DP definition, as iDP only focuses on the "point-wise" guarantee of privacy. It protects the neighborhood of a specific dataset of interest, instead of fulfilling DP requirements for all possible adjacent databases. While iDP does not provide the same level of privacy protection as the "worst-case" standard DP, it might be sufficient in certain application scenarios, e.g., data publishing/release, when the procedure is non-interactive and the released dataset is indeed the target that one is interested in privatizing. We discuss it in our work as iDP may provide another direction/option for balancing the trade-off between privacy and utility in practice, based on specific applications.

The intuition of iDP is that, while the standard DP (Definition 2.1) requires indistinguishability between any pair of neighboring databases, in some practical scenarios, the data custodian only holds one "ground truth" database U that needs to be protected. Limiting the scope of the neighborhood could be reasonable in certain practical scenarios. The "indistinguihability" requirement is only cast on U and its neighbors specifically, instead of on any possible dataset. iDP has achieved excellent utility for computing robust statistics at small ϵ [12].

For the DP algorithms that have been discussed previously in this paper, we first note that, for DP-RP and DP-OPORP, the local sensitivity at any $u \in \mathcal{U}$ equals the global sensitivity. In other words, iDP does not help improve DP-RP and DP-OPORP. Also, we will soon discuss the reason why SignRP can be much better than SignOPORP under iDP. Therefore, we will mainly investigate the SignRP algorithms under iDP. Because the "indistinguihability" requirement of DP is only for U and its neighbors locally, operationally, for SignRP, iDP essentially follows the local flipping probability (Section 4.2 and Figure 1) when computing the perturbation level, which can be much smaller than that required by the standard DP.

We propose two iDP-SignRP methods, based on noise addition and sign flipping, respectively. Both approaches share the same key idea of iDP, that is, many signs of the projected values do not need perturbations. This can be seen from Figure 1, where the "local flipping probability" is non-zero only in the regime when the projected data is near 0 (i.e., $L = 1$ in Algorithm 4). Since in other cases the local flip probability is zero, perturbation is not needed. As a result, out of k projections, only a fraction of the projected values needs to be perturbed. This significantly reduces the noise injected to SignRP and boosts the utility by a very large margin.

A.2 iDP-SignRP-G by Gaussian Noise Addition

In Algorithm 6, we present the iDP-SignRP-G method for one data vector u. We use the "local flipping probability" (e.g., in Figure 1) to choose which projections are perturbed before taking signs. After applying random projection to get k projected values, we do the following steps:

1. We compute noise-indicators $(I_1, ..., I_k)$ for each projected value in $x = \frac{1}{\sqrt{k}} W^T u$ using Algorithm 7. Denote $\mathcal{A} = \{I_j : I_j = 1, j = 1, ..., k\}$ and $N_+ = |\mathcal{A}|$. This is the maximal number of different signs of x and $x' = W^T u', \forall u' \in Nb(u)$.

Algorithm 6: iDP-SignRP-G (DP-SignRP with Gaussian noise)

Input: Data $u \in [-1, 1]^p$; Privacy parameters $\epsilon > 0$, $\delta \in (0, 1)$; Number of projections k
Output: Differentially private sign random projections
Apply RP by $x = \frac{1}{\sqrt{k}} W^T u$, where $W \in \mathbb{R}^{p \times k}$ is a random Rademacher matrix
For every projected value in x, compute $(I_1, ..., I_k)$ by Algorithm 7
Let $\mathcal{A} = \{I_j : I_j = 1, j = 1, ..., k\}$ and $\tilde{N}_+ = |\mathcal{A}|$
Compute sensitivity $\triangle_2 = \beta\sqrt{\frac{\tilde{N}_+}{k}}$
Compute σ by Theorem 3.2 with $\triangle_2$ and privacy budget ϵ and δ
Compute $\tilde{s}_j = \begin{cases} sign(x_j), & j \notin \mathcal{A} \\ sign(x_j + G), & j \in \mathcal{A} \end{cases}$, where $G \sim N(0, \sigma^2)$ is iid Gaussian noise
Return $\tilde{s} = [\tilde{s}_1, ..., \tilde{s}_k]$

Algorithm 7: Compute noise-indicator of iDP-SignRP-G for one projection

Input: Data $u \in [-1, 1]^p$; one projected value z; adjacency parameter β
Output: Indicator I w.r.t. projection w for data vector u
$I = 0$
If $\beta/\sqrt{k} \geq |z|$
    $I = 1$
End If

2. We compute the sensitivity $\triangle_2 = \beta \max_{i=1,...,p} \|W_{[i,\mathcal{A}]}\|$, where $W_{[i,\mathcal{A}]}$ denotes the i-th row of W indexed at $\mathcal{A}$, which is an N_+-dimensional vector.
3. We use the optimal Gaussian mechanism (Theorem 3.2) to compute σ, with $\triangle_2$ computed above and privacy parameters (ϵ, δ).
4. For $j = 1, .., k$, if $j \notin \mathcal{A}$, we take $\tilde{s}_j = sign(x_j)$; if $j \in \mathcal{A}$, we take $\tilde{s}_j = sign(x_j + G)$ where $G \sim N(0, \sigma^2)$ is a Gaussian noise. Finally we output $\tilde{s} = [\tilde{s}_1, ..., \tilde{s}_k]$.

Let's explain the intuition behind DP-SignRP-G. Since a neighboring data vector u' only differs from u in one dimension by at most β, for each single projection w, when $\beta \max_{i=1,...,p} |w_i| \leq |w^T u|$, there is no neighbor u' of u that may change the sign of the projected value of u, i.e., $sign(w^T u') \neq sign(w^T u)$. In other words, when $\beta \max_{i=1,...,p} |w_i| \leq |w^T u|$, no noise is needed for this projected value to attain iDP. This is the reason why we call the output of Algorithm 7 a "noise-indicator". Consequently, in step 4 of iDP-SignRP-G it suffices to add Gaussian noise only to those projected values x_j with $j \in \mathcal{A}$, instead of to all k projections as in DP-RP-G-OPT.

Theorem A.1 (iDP-SignRP-G)**.** *Algorithm 6 is (ϵ, δ)-iDP for data u.*

Proof. For a data vector u, let $Nb(u)$ be its neighbor set with vector that differs from u by at most β in one dimension. Denote $x = \frac{1}{\sqrt{k}} W^T u$ and $x' = \frac{1}{\sqrt{k}} W^T u'$. Let $(I_1, ..., I_k)$ be the noise-indicators from Algorithm 7 and $\mathcal{A} = \{i : I_j = 1\}$, $\tilde{N}_+ = |\mathcal{A}|$. Consider the two sets separately:

- For $j \in [k] \setminus \mathcal{A}$, by the condition $\beta/\sqrt{k} \leq |z|$, we know that $\forall u' \in Nb(u)$, it holds that $sign(x_i) = sign(x'_i)$.
- For $j \in \mathcal{A}$, consider the sub-vector $x_{\mathcal{A}}$. Adding iid Gaussian noise to $x_{\mathcal{A}}$ according to Theorem 3.2 with $\triangle_2 = \beta\sqrt{\frac{\tilde{N}_+}{k}}$ ensures the (ϵ, δ)-DP of $x_{\mathcal{A}}$. By the post processing property of DP, we know that $sign(x_{\mathcal{A}})$ is also (ϵ, δ)-DP. Thus, for any $Q \in \{-1, 1\}^{N_+}$, we have $Pr(sign(x_{\mathcal{A}}) = Q) - e^{\epsilon} Pr(sign(x'_{\mathcal{A}}) = Q) \leq \delta, \forall u' \in Nb(u)$.

Combining two parts, we have for any $Q \in \{-1, 1\}^k$,

$$Pr(sign(x) = Q) - e^{\epsilon} Pr(sign(x') = Q) = Pr(sign(x_{\mathcal{A}}) = Q) - e^{\epsilon} Pr(sign(x'_{\mathcal{A}}) = Q) \leq \delta,$$

for all $u' \in Nb(u)$. By the symmetry of DP (on the sub-vector $x_{\mathcal{A}}$), we also know that $Pr(sign(x') = Q) - e^{\epsilon} Pr(sign(x) = Q) \leq \delta$. This proves the (ϵ, δ)-iDP by Definition A.1. ☐

A.3 iDP-SignRP-RR by Randomized Response

Algorithm 8: iDP-SignRP-RR

Input: Data $u \in [-1, 1]^p$, privacy parameters $\epsilon > 0$, $0 < \delta < 1$, number of projections k

Output: Differentially private sign random projections

Apply RP by $x = \frac{1}{\sqrt{k}} W^T u$, where $W \in \mathbb{R}^{p \times k}$ is a random Rademacher matrix

For every column in W, compute $(I_1, ..., I_k)$ by Algorithm 7

Let $\mathcal{A} = \{I_j : I_j = 1, j = 1, ..., k\}$ and $\tilde{N}_+ = |\mathcal{A}|$

Compute $\tilde{s}_j = \begin{cases} sign(x_j), & j \notin \mathcal{A} \\ sign(x_j), & j \in \mathcal{A} \text{ with prob. } \frac{e^{\epsilon'}}{e^{\epsilon'}+1} \\ -sign(x_j), & j \in \mathcal{A} \text{ with prob. } \frac{1}{e^{\epsilon'}+1} \end{cases}$ for $j = 1, ..., k$, with $\epsilon' = \epsilon / \tilde{N}_+$

Return $\tilde{s}$ as the DP-SignRP of u

Similar to Section 4, we also have an iDP-SignRP-RR method with pure ϵ-DP guarantee by randomly flipping the signs after SignRP, as summarized in Algorithm 3. After we apply random projection $x = \frac{1}{\sqrt{k}} W^T u$, we call the same procedure as in iDP-SignRP-G to determine set $\mathcal{A}$ representing the projected values that needs perturbation for iDP. For $j \notin \mathcal{A}$, we use the original $\tilde{s}_j = sign(x_j)$. For $j \in \mathcal{A}$, we keep $sign(x_j)$ with probability $\frac{e^{\epsilon'}}{e^{\epsilon'}+1}$ and flip the sign otherwise, where $e^{\epsilon'} = \epsilon / \tilde{N}_+$ with $\tilde{N}_+ = |\mathcal{A}|$.

Theorem A.2. *Algorithm 3 achieves ϵ-iDP for data u.*

Proof. The high-level proof idea is similar to that of Theorem A.1. For $u \in [-1, 1]^p$ let u' be an β-neighboring data. Let $s = sign(W^T u) \in \{-1, +1\}^k$, $s' = sign(W^T u') \in \{-1, +1\}^k$, and denote $\tilde{s}$ and $\tilde{s}'$ as the randomized output of s and s' by Algorithm 3, respectively. Consider $\mathcal{A}$ in Algorithm 3. By Algorithm 7, we know that for $j \notin \mathcal{A}$, $Pr(\tilde{s}_j = \tilde{s}'_j) = Pr(s_j = s'_j) = 1$, $\forall u' \in Nb(u)$. For projections in $\mathcal{A}$, denote $S = \{j \in \mathcal{A} : s_j \neq s'_j\}$ and $S^c = \mathcal{A} \setminus S$. For any vector $y \in \{-1, +1\}^k$, we further define $S_0 = \{j \in S : s_j = y_j\}$, $S_1 = \{j \in S : s_j \neq y_j\}$, $S_0^c = \{j \in S^c : s_j = y_j\}$ and $S_1^c = \{j \in S^c : s_j \neq y_j\}$. Since the k projections are independent, by composition we have

$$\log \frac{Pr(\tilde{s} = y)}{Pr(\tilde{s}' = y)} = \log \frac{\prod_{j \notin \mathcal{A}} Pr(\tilde{s}_j = y_j) \prod_{j \in S_0^c} \frac{e^{\epsilon'}}{e^{\epsilon'}+1} \prod_{j \in S_1^c} \frac{1}{e^{\epsilon'}+1} \prod_{j \in S_0} \frac{e^{\epsilon'}}{e^{\epsilon'}+1} \prod_{j \in S_1} \frac{1}{e^{\epsilon'}+1}}{\prod_{j \notin \mathcal{A}} Pr(\tilde{s}'_j = y_j) \prod_{j \in S_0^c} \frac{e^{\epsilon'}}{e^{\epsilon'}+1} \prod_{j \in S_1^c} \frac{1}{e^{\epsilon'}+1} \prod_{j \in S_0} \frac{1}{e^{\epsilon'}+1} \prod_{j \in S_1} \frac{e^{\epsilon'}}{e^{\epsilon'}+1}}$$

$$\leq \log \frac{\prod_{j \in S} \frac{e^{\epsilon'}}{e^{\epsilon'}+1}}{\prod_{j \in S} \frac{1}{e^{\epsilon'}+1}} = |S| \epsilon' \leq \tilde{N}_+ \epsilon' = \epsilon,$$

which proves the ϵ-iDP according to Definition A.1. ☐

The number of projections that requires noise addition $\tilde{N}_+$ is also tightly related to the $P_+(\|u\|, p)$ (Proposition 4.4 and (5)). Particularly, $\tilde{N}_+$ would be small when the data has relatively large norm compared with the change in neighboring data β. Therefore, both iDP-SignRP methods would have better utility when the data norm is large.

The reduction from k to $\tilde{N}_+$ in iDP not only waives the need to add noise to many projected values, but also requires smaller Gaussian noise or smaller flipping probability for the values that need to be perturbed. Specifically, note that in Algorithm 6, the optimal Gaussian mechanism is deployed with sensitivity $\triangle_2 = \beta\sqrt{\frac{\tilde{N}_+}{k}}$, instead of $\triangle_2 = \beta$ as in (3) for DP-RP-G-OPT.

iDP-SignOPORP. Similarly, we can also apply iDP to the SignOPORP method. Basically, we only need to replace x in Line 3 in both Algorithm 6 and Algorithm 8 by the OPORP of u. However, we note that this iDP-SignOPORP procedure is considerably worse than iDP-SignRP in performance. This is because, by the binning step in OPORP, the average scale of each projected value becomes much smaller. This implies that in Algorithm 7, the magnitude of z would be much smaller, so a lot more projected values will require perturbation, which leads to a utility loss. This illustrates the superiority of SignRP under iDP: since each RP aggregates the whole data vector, SignRP is more robust to a small change in the data. Hence, less noise is needed.

A.4 Empirical Results on iDP

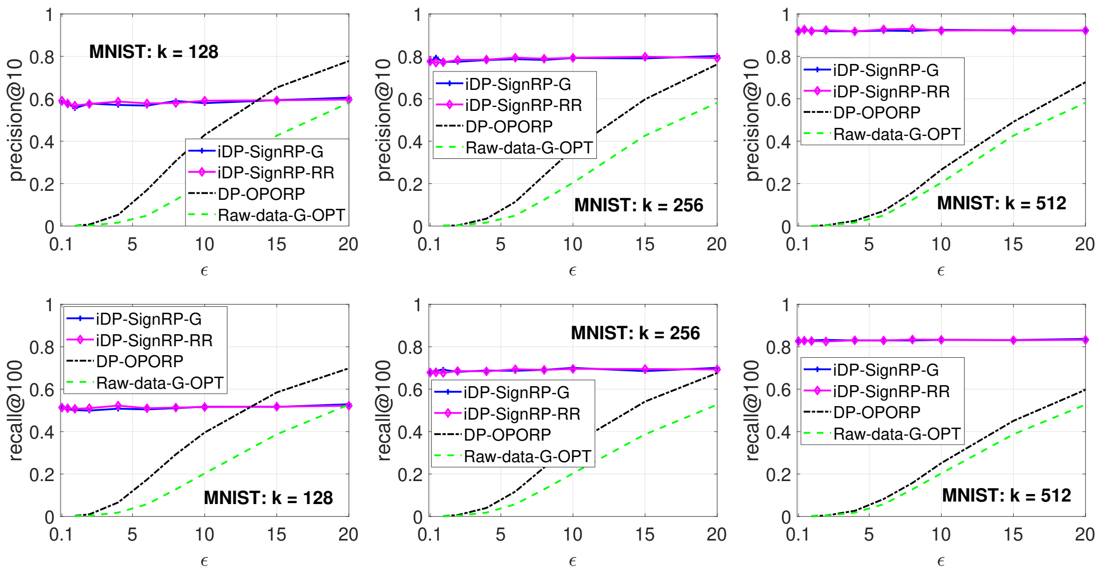

Figure 5: Retrieval on MNIST with iDP-SignRP, $\beta = 1$, $\delta = 10^{-6}$.

To demonstrate the empirical gain in utility of iDP-SignRP, we conduct the same set of experiments as in Section 5. Figure 5 reports the precision and recall on MNIST, and Figure 6 presents the SVM test accuracy on WEBSPAM. As we can see, iDP-SignRP achieves very high utility even when $\epsilon < 0.1$. We see that the curves of iDP-SignRP are almost flat. This is because only a small fraction of projected values are perturbed, so the untouched projected values already provides rich information for search and classification. In other words, the experimental results illustrate that the SignRP itself is already very strong in protecting the individual differential privacy. In other words, SignRP itself is already a strong method to protect the privacy of each specific dataset with respect to the individual DP, e.g., in non-interactive data publishing tasks.

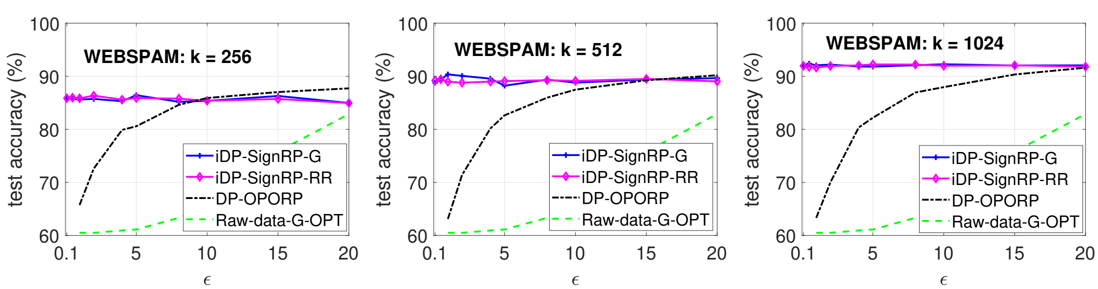

Figure 6: SVM on WEBSPAM with iDP-SignRP, $\beta = 1$, $\delta = 10^{-6}$.

B Comparison of Different Projection Matrices and the Benefits of Rademacher RP

Besides Gaussian random projection, we can also adopt other types of projection matrices which might even work better for DP. The following distributions of w_{ij} are popular:

- The uniform distribution, $\sqrt{3} \times unif[-1, 1]$. The $\sqrt{3}$ factor is placed here to have $\mathbb{E}(w_{ij}^2) = 1$ by following the convention in the practice of random projections.
- The "very sparse" distribution, as used in [8]:

$$w_{ij} = \sqrt{s} \times \begin{cases} -1 & \text{with prob.} \quad 1/(2s) \\ 0 & \text{with prob.} \quad 1 - 1/s, \\ +1 & \text{with prob.} \quad 1/(2s) \end{cases} \tag{8}$$

  which generalizes [1] (for $s = 1$ and $s = 3$). Note that when $s = 1$, it is also called the "symmetric Bernoulli" distribution or the "Rademacher" distribution.

Next, we compare these various types of projection matrices and show that Rademacher (symmetric Bernoulli) random projection is superior to Gaussian random projection for both DP-RP and DP-SignRP in that less perturbation is required to achieve the same privacy level.

B.1 Rademacher Projection for DP-RP

From Theorem 3.1 and Theorem 3.2, it is clear that the noise magnitude of Gaussian noise in DP-RP directly depends on the l_2-sensitivity $\triangle_2$, which, according to (3), equals the largest row norm of the projection matrix W. Among the above mentioned distributions, the dense Rademacher projection ($s = 1$ in (8)) has $\triangle_2 = \frac{1}{\sqrt{k}}\beta \times \sqrt{k} = \beta$ which is independent of p. This could be much smaller than the dense Gaussian projection (i.e., DP-RP-G-OPT).

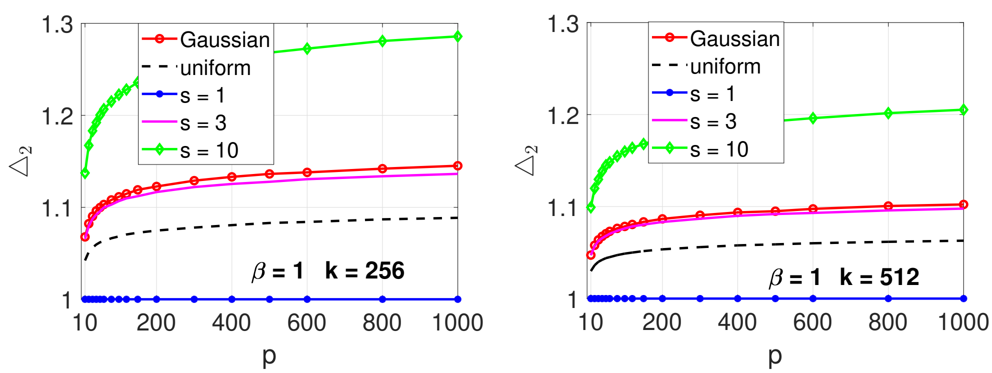

Figure 7: The l_2-sensitivity $\triangle_2$ (3) for different types of random projection matrices against the data dimensionality p, at $k = 256$ and $k = 512$, respectively. $\beta = 1$.

In Figure 7, we numerically simulate the $\triangle_2$ of different projection matrices, which shows that the Rademacher projection produces the smallest sensitivity. This, when plugged into the optimal Gaussian mechanism (Theorem 3.2), leads to smaller Gaussian noise variance needed.

B.2 Rademacher Projection for DP-SignRP

From our analysis, it is clear that the flipping probability of DP-SignRP (both DP-SignRP-RR and DP-SignRP-RR-smooth) essentially depends on how concentrated the projected data is around zero. Particularly, N_+ in Algorithm 3, as given in Proposition 4.4, is a high probability upper bound on a Binomial random variable with success probability $Pr(\beta \max_{i=1,...,p} |w_i| \geq |w^T u|)$ with $w \sim N(0, 1)$. In Algorithm 4, $L_j = \lceil \frac{|w_j^T u|}{\beta \max_{i=1,...,p} |W_{ij}|} \rceil$. For both quantities, a smaller value leads to a smaller sign flipping probability and thus better utility.

N_+ **in DP-SignRP-RR.** We first consider the N_+ in Algorithm 3, which determines the flipping probability $\frac{1}{e^{\epsilon/N_+}+1}$. Particularly, N_+ in Algorithm 3, as given in Proposition 4.4, is a high probability upper bound on a Binomial random variable with success probability

$$P_+ = Pr\left(\beta \max_{i=1,...,p} |w_i| \geq |w^T u|\right), \tag{9}$$

where w is the p-dimensional projection vector. When w_i is sampled from the Rademacher distribution, i.e., $w_i \in \{-1, +1\}$ with equal probabilities, the probability calculation can be simplified:

$$P_{+,b} = Pr\left(\beta \max_{i=1,...,p} |w_i| \geq |\sum_{i=1}^{p} w_i u_i|\right) = Pr\left(\beta \geq |\sum_{i=1}^{p} w_i u_i|\right) \approx 2\Phi\left(\frac{\beta}{\|u\|}\right) - 1. \tag{10}$$

Based on the central limit theorem, the normal approximation (10) is accurate unless p is very small. Recall that, when w_i's are sampled from the Gaussian distribution, we can calculate an upper bound in (21), which is re-written as below:

$$P_{+,g} = Pr\left(\beta \max_{i=1,...,p} |w_i| \geq |\sum_{i=1}^{p} w_i u_i|\right) \leq \int_0^\infty 2p[2\Phi(t)-1]^{p-1}[2\Phi(\beta t/\|u\|)-1]\phi(t)dt. \tag{11}$$

Next, we provide a simulation study to justify the approximation and compare different distributions in terms of their impact on the probability (9), for $\beta = 1$ as well as $\beta = 0.1$. For simplicity, we simulate the data as a p-dimensional vector of uniform random numbers sampled from $unif[-1, 1]$. We experiment with five different choices of w: the standard Gaussian, the uniform, the "very sparse" distribution (8) with $s = 1$, $s = 3$, and $s = 10$. We vary p from 10 to 1000. For each case, we repeat the simulations 10^7 times to ensure sufficient accuracy. Figure 8 verifies that the two approximations (10) and (11) are accurate. In Figure 9, we provide the curves for more types of projection matrices. From both figures, we clearly see that using the Rademacher projection can considerably reduce (9) compared with Gaussian (and other) projections, leading to smaller N_+ value. This typically implies better utility.

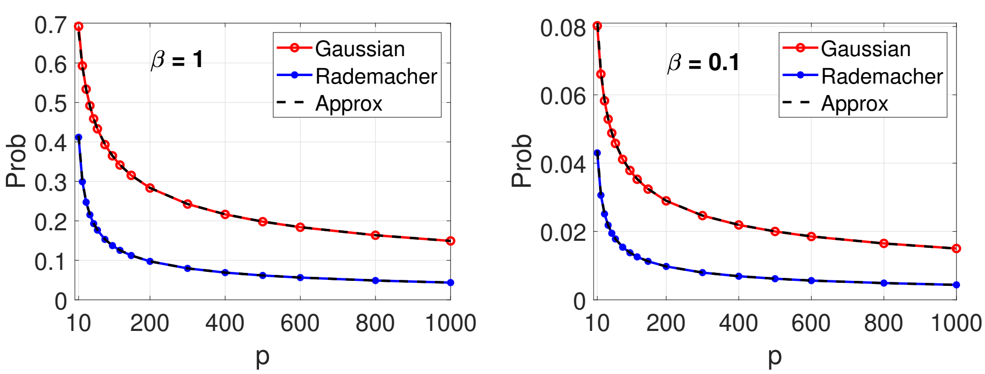

Figure 8: Simulations for evaluating (10) and (11), using two choices for w: the Gaussian distribution and the Rademacher distribution (i.e., (8) with $s = 1$). We plot the two upper bounds (10) and (11) as black dashed curves, which both overlap with their corresponding simulations.

L_j **in DP-SignRP-RR-smooth.** Similarly, we numerically evaluate the L_j in Algorithm 4. We run Algorithm 4 with $k = 512$, which gives 512 L_j values. In Figure 10, we plot the proportion (or the approximated distribution) of the values of L_j among k projections. As we see, Rademacher projection produces least number of small L_j values, and largest number of higher L_j values. As the smooth flipping probability equals $\frac{1}{\exp(\frac{L_j}{k}\epsilon)+1}$, larger L_j leads to smaller probability of sign flipping. Hence, Rademacher is again the best choice for the projection matrix.

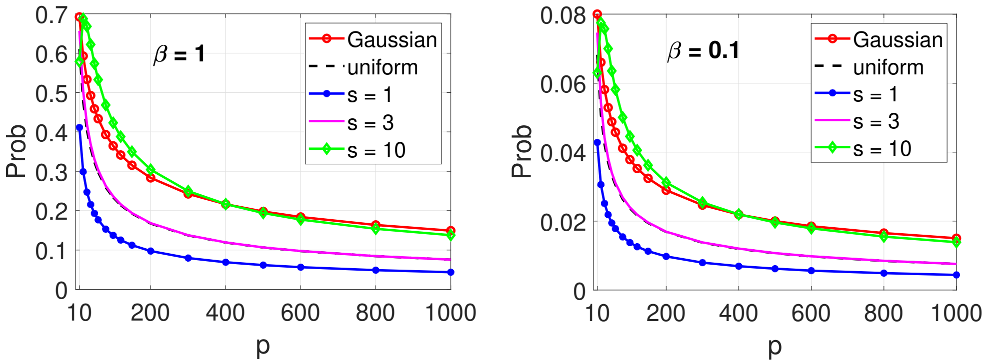

Figure 9: Simulations (same as in Figure 9) for evaluating (9), using five different choices for w: the Gaussian, the uniform, the "very sparse" distribution (8) with $s = 1, 3$ and 10. $s = 1$ is the Rademacher distribution. The data vector is simulated by sampling each entry from $unif[-1, 1]$.

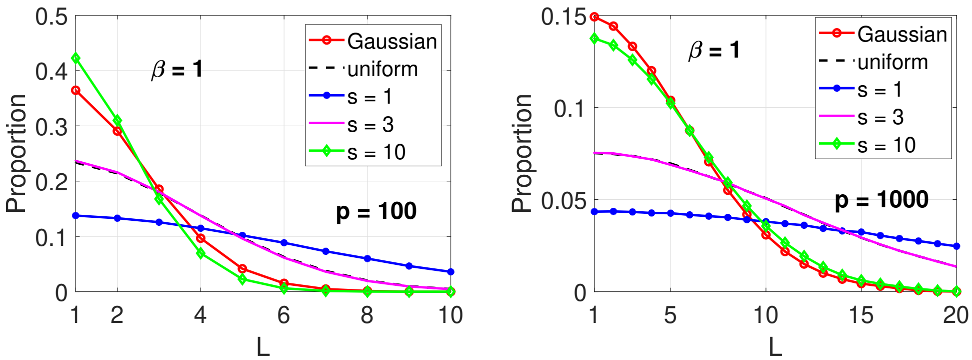

Figure 10: Simulations for evaluating L_j in Algorithm 4, using different choices for w. $s = 1$ is the Rademacher distribution. Left: $p = 100$, right: $p = 1000$. The y-axis is the proportion (normalized histogram) of the values of all the L_j, $j = 1, ..., k$ computed using $k = 512$ projected samples.

C Comparison of DP-RP and DP-OPORP on Inner Product Estimation

In this section, we theoretically compare DP-RP and DP-OPORP. In Section B, we have shown that Rademacher projection requires lowest noise magnitude. Thus, we consider DP-RP with Rademacher projections here. For clarity, we summarize the algorithm in Algorithm 9. We name it "DP-RP-G-OPT-B", where "G" stands for the **G**aussian noise mechanism and "B" stands for "symmetric **B**ernoulli" projections.

Algorithm 9: DP-RP-G-OPT-B

Input: Data $u \in [-1, 1]^p$, privacy parameters $\epsilon > 0$, $\delta \in (0, 1)$, number of projections k

Output: (ϵ, δ)-differentially private random projections $\tilde{x} \in \mathbb{R}^k$

Apply RP $x = \frac{1}{\sqrt{k}} W^T u$, where $W \in \mathbb{R}^{p \times k}$ is a random Rademacher matrix

Generate iid random noise vector $G \in \mathbb{R}^k$ following $N(0, \sigma^2)$ where σ is obtained by Theorem 3.2 with $\triangle_2 = \beta$

Return $\tilde{x} = x + G$

In our analysis, for simplicity we assume the data are normalized, i.e., the data vector has l_2 norm equal to 1. In this case, the inner product is also the cosine. The baseline method is the most straightforward: we add optimal Gaussian noise to each dimension of the original data (Raw-data-G-OPT). For this strategy, the sensitivity is also $\triangle_2 = \beta$. This means, when we compare all three methods: Raw-data-G-OPT, DP-RP-G-OPT-B, and DP-OPORP, the noise level σ is the same. This makes it convenience to conduct the comparisons, from which we can gain valuable insights.

Theorem C.1 (Raw-data-G-OPT, i.e., adding optimal Gaussian noise on raw data)**.** *Let σ be the solution to (4) with $\triangle_2 = \beta$. For any $u, v \in \mathcal{U}$, let $\tilde{u}_i = u_i + a_i$ and $\tilde{v}_i = v_i + b_i$ be the DP noisy vectors, with $a_i, b_i \sim N(0, \sigma^2)$ i.i.d. Then, denote $\hat{g}_{org} = \sum_{i=1}^{p} \tilde{u}_i \tilde{v}_i$. we have*

$$\mathbb{E}[\hat{g}_{org}] = \sum_{i=1}^{p} u_i v_i, \qquad Var\left(\hat{g}_{org}\right) = \sigma^2 \sum_{i=1}^{p} \left(u_i^2 + v_i^2\right) + p\sigma^4. \tag{12}$$

Proof. To add Gaussian noise to the original data, it suffices to find the sensitivity, which, by Definition 2.2, is $\triangle_2 = \beta$. Thus, the approach is (ϵ, δ)-DP according to the optimal Gaussian mechanism (Theorem 3.2). To compute the mean and variance, consider some $i \in [p]$. We have

$$\mathbb{E}\left[(u_i + a_i)(v_i + b_i)\right] = \mathbb{E}[u_i v_i + a_i v_i + b_i u_i + a_i b_i] = u_i v_i.$$

Thus, taking the sum implies $\mathbb{E}[\hat{g}_{org}] = \sum_{i=1}^{p} u_i v_i$. For the variance,

$$\mathbb{E}\left[(u_i + a_i)(v_i + b_i)\right]^2 = \mathbb{E}[u_i v_i + a_i v_i + b_i u_i + a_i b_i]^2 = u_i^2 v_i^2 + \sigma^2\left(u_i^2 + v_i^2\right) + \sigma^4,$$

which leads to

$$Var\left((u_i + a_i)(v_i + b_i)\right) = \sigma^2\left(u_i^2 + v_i^2\right) + \sigma^4.$$

Therefore, by independence,

$$Var\left(\hat{g}_{org}\right) = Var\left(\sum_{i=1}^{p} (u_i + a_i)(v_i + b_i)\right) = \sigma^2 \sum_{i=1}^{p} \left(u_i^2 + v_i^2\right) + p\sigma^4,$$

which proves the claim. □

For DP-RP-G-OPT-B and DP-OPORP, we have the following results.

Theorem C.2 (DP-RP-G-OPT-B inner product estimation)**.** *Let σ be the solution to (4) with $\triangle_2 = \beta$. In Algorithm 9, let $W \in \{-1, 1\}^{p \times k}$ be a Rademacher random matrix. Denote $x = \frac{1}{\sqrt{k}} W^T u$, $y = \frac{1}{\sqrt{k}} W^T v$, and a, b are two random Gaussian noise vectors following $N(0, \sigma^2)$. Let $\hat{g}_{rp} = \sum_{j=1}^{k} (x_j + a_j)(y_j + b_j)$. Then, $\mathbb{E}[\hat{g}_{rp}] = \sum_{i=1}^{p} u_i v_i$, and*

$$Var\left(\hat{g}_{rp}\right) = \sigma^2 \sum_{i=1}^{p} \left(u_i^2 + v_i^2\right) + k\sigma^4 + \frac{1}{k}\left(\sum_{i=1}^{p} u_i^2 \sum_{i=1}^{p} v_i^2 + \left(\sum_{i=1}^{p} u_i v_i\right)^2 - 2\sum_{i=1}^{p} u_i^2 v_i^2\right). \tag{13}$$

Proof. The conditional mean and variance can be computed as

$$\mathbb{E}\left[\sum_{j=1}^{k} (x_j + a_j)(y_j + b_j) | x_j, y_j, j = 1, ..., k\right] = \sum_{j=1}^{k} x_j y_j,$$

$$Var\left(\sum_{j=1}^{k} (x_j + a_j)(y_j + b_j) | x_j, y_j, j = 1, ..., k\right) = \sigma^2 \sum_{j=1}^{k} (x_j^2 + y_j^2) + k\sigma^4,$$

where the variance calculation follows from Theorem C.1. Hence, we have

$$\mathbb{E}\left[\sum_{j=1}^{k} (x_j + a_j)(y_j + b_j)\right] = \mathbb{E}\left[\sum_{j=1}^{k} x_j y_j\right] = \sum_{i=1}^{p} u_i v_i,$$

$$\begin{aligned} Var\left(\hat{g}_{rp}\right) =& \mathbb{E}\left[\sigma^2 \sum_{j=1}^{k} (x_j^2 + y_j^2) + k\sigma^4\right] + Var\left(\sum_{j=1}^{k} x_j y_j\right) \\ =& \sigma^2 \sum_{i=1}^{p} \left(u_i^2 + v_i^2\right) + k\sigma^4 + \frac{1}{k}\left(\sum_{i=1}^{p} u_i^2 \sum_{i=1}^{p} v_i^2 + \left(\sum_{i=1}^{p} u_i v_i\right)^2 - 2\sum_{i=1}^{p} u_i^2 v_i^2\right). \end{aligned} \tag{14}$$

In the above calculation, the formula of $Var\left(\sum_{j=1}^{k} x_j y_j\right)$ is from the result in [8] with $s = 1$ for Rademacher distribution. □

Algorithm 10: DP-OPORP

Input: Data $u \in [-1,1]^p$, privacy parameters $\epsilon > 0$, $\delta \in (0,1)$, number of projections k
Output: Differentially private OPORP
Apply Algorithm 2 with a random Rademacher projection vector to obtain the OPORP x
Set sensitivity $\Delta_2 = \beta$
Generate iid random vector $G \in \mathbb{R}^k$ following $N(0, \sigma^2)$ where σ is computed by Theorem 3.2
Return $\tilde{x} = x + G$

Theorem C.3 (DP-OPORP inner product estimation)**.** *Let σ be the solution to (4) with $\triangle_2 = \beta$. Let $w \in \{-1,1\}^p$ be a Rademacher random vector. In Algorithm 10, let x and y be the OPORP of u and v, and a, b be two random Gaussian noise vectors following $N(0,\sigma^2)$. Denote $\hat{g}_{oporp} = \sum_{j=1}^k (x_j + a_j)(y_j + b_j)$. Then, $\mathbb{E}[\hat{g}_{oporp}] = \sum_{i=1}^p u_i v_i$, and*

$$Var\left(\hat{g}_{oporp}\right) = \sigma^2 \sum_{i=1}^p \left(u_i^2 + v_i^2\right) + k\sigma^4 + \frac{1}{k}\left(\sum_{i=1}^p u_i^2 \sum_{i=1}^p v_i^2 + \left(\sum_{i=1}^p u_i v_i\right)^2 - 2\sum_{i=1}^p u_i^2 v_i^2\right)\frac{p-k}{p-1}. \tag{15}$$

Proof. The proof is similar to that of Theorem C.2, with the help of the result in [9]. □

The variance reduction factor $\frac{p-k}{p-1}$ can be quite beneficial when p is not very large. Also, see [9] for the normalized estimators for both OPORP and VSRP (very sparse random projections). The normalization steps can substantially reduce the estimation variance.

Comparison. For the convenience of comparison, let us assume that the data are row-normalized, i.e., $\|u\|^2 = 1$ for all $u \in \mathcal{U}$. Let $\rho = \sum_{i=1}^p u_i v_i$. We have

$$\begin{aligned}
Var\left(\hat{g}_{org}\right) &= 2\sigma^2 + p\sigma^4, \\
Var\left(\hat{g}_{rp}\right) &= 2\sigma^2 + k\sigma^4 + \frac{1}{k}\left(1 + \rho^2 - 2\sum_{i=1}^p u_i^2 v_i^2\right), \\
Var\left(\hat{g}_{oporp}\right) &= 2\sigma^2 + k\sigma^4 + \frac{1}{k}\left(1 + \rho^2 - 2\sum_{i=1}^p u_i^2 v_i^2\right)\frac{p-k}{p-1}.
\end{aligned}$$

For high-dimensional data (large p), we see that $\hat{g}_{rp}$ and $\hat{g}_{oporp}$ has roughly the same variance, approximately $2\sigma^2 + k\sigma^4 + \frac{1}{k}$. We would like to compare this with $Var\left(\hat{g}_{org}\right) = 2\sigma^2 + p\sigma^4$ the variance for adding noise directly to the original data.

Let's define the ratio of the variances:

$$R = \frac{2\sigma^2 + p\sigma^4}{2\sigma^2 + k\sigma^4 + \frac{1}{k}} \sim \frac{p\sigma^4}{k\sigma^4} = \frac{p}{k} \quad \text{(if } p \text{ is large or } \sigma \text{ is high)} \tag{16}$$

to illustrate the benefit of RP-type algorithms (DP-RP and DP-OPORP) in protecting the privacy of the (high-dimensional) data. If $\frac{p}{k} = 100$, then it is possible that the ratio of the variances can be roughly 100. This would be a huge advantage. Figure 11 plots the ratio R for $p = 1000$ and $p = 10000$ as well as a series of k/p values, with respect to σ.

Figure 11 also illustrates when it might be a good strategy to directly add noise to the original data. For example, when $p = 1000$, the ratio can be below 1 if $\sigma < 0.1$. One can numerically verify that, (in Figure 12) in order for $\sigma < 0.1$ at $\Delta_2 = \beta = 1$, we need $\epsilon > 100$. In other words, adding noise to the raw data might be plausible when $\epsilon > 100$. In the literature, however, many DP applications typically require a much smaller ϵ, such as $\epsilon \in [0.1, 20]$ (e.g., [5; 7]). Therefore, DP-RP and DP-OPORP is much better (i.e., has much smaller inner product estimation variance) than adding noise to the raw data in common privacy regimes.

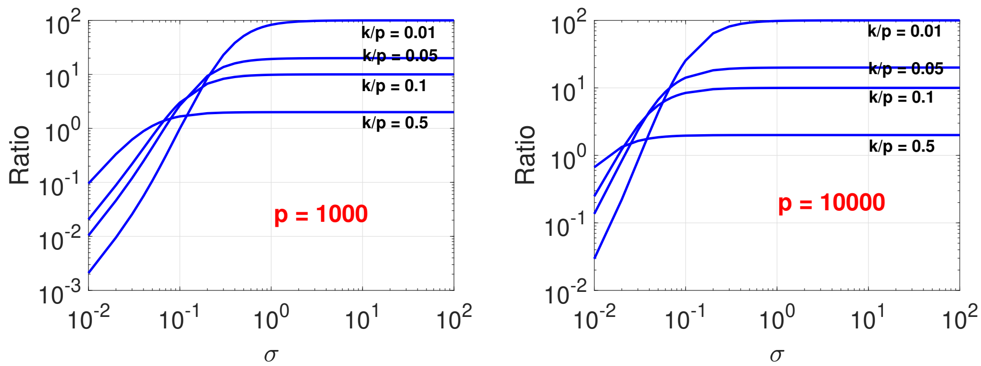

Figure 11: We plot the ratio of variances in (16) for $p = 1000$ and $p = 10000$. We choose k values with $k/p \in \{0.01, 0.05, 0.1, 0.5\}$. Then for any σ value, we are able to compute the ratio R. For larger σ, we have $R \sim \frac{p}{k}$ as expected. See Figure 12 for the relationship among σ, Δ, and ϵ (and δ).

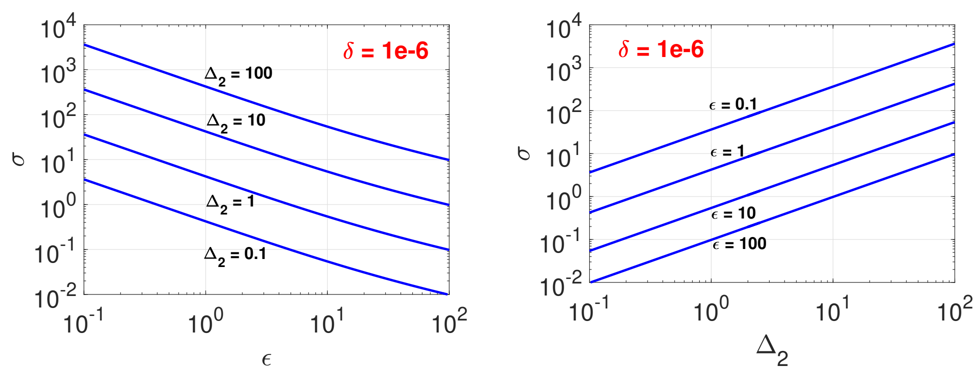

Figure 12: Left panel: the optimal Gaussian noise σ versus ϵ for a series of Δ_2 values, by solving the nonlinear equation (4) in Theorem 3.2, for $\delta = 10^{-6}$. Right panel: the optimal Gaussian noise σ versus Δ_2 for a series of ϵ values.

D More Experiment Results

We provide the complete set of plots of our experimental results. In Figure 13 and Figure 14, we report the precision@10 and recall@100 curves of DP-RP variants and DP-OPORP on MNIST and CIFAR, respectively. In Figure 15, we report the test accuracy on the Webspam dataset of these methods. From all plots, we see that DP-RP-G-OPT-B and DP-OPORP perform equally the best on all the tasks, significanly better than the strategy of adding Gaussian noise to the raw data.

In Figure 16, we report the recall@100 metric of DP-SignOPORP methods in addition to the precision@10 metric shown in the main paper. For completeness, we also include the SVM test accuracy in Figure 17, which is the same as Figure 4 in the main context.

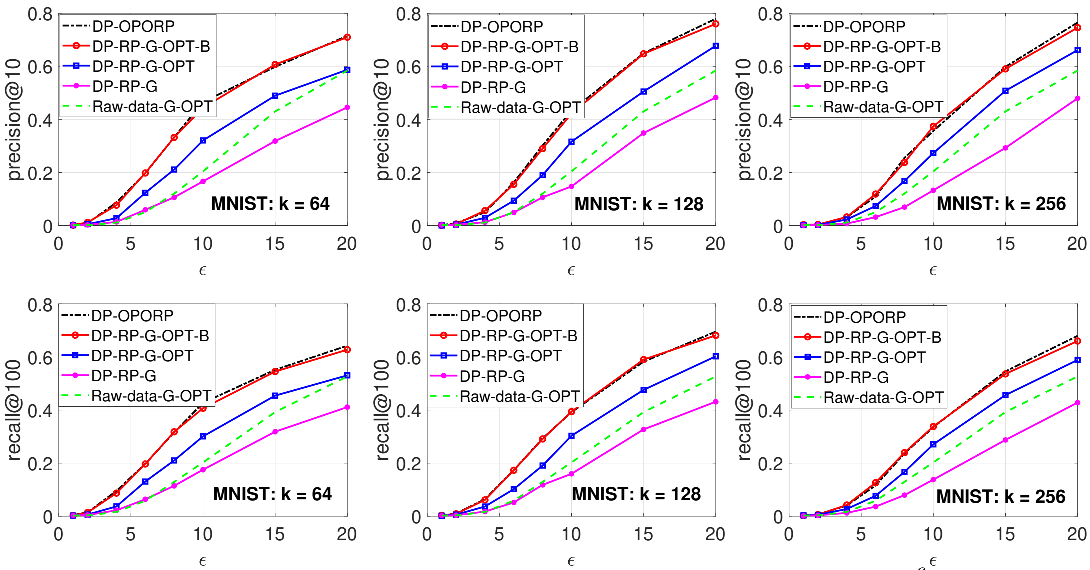

Figure 13: Retrieval recall and precision on MNIST, $\beta = 1$, $\delta = 10^{-6}$.

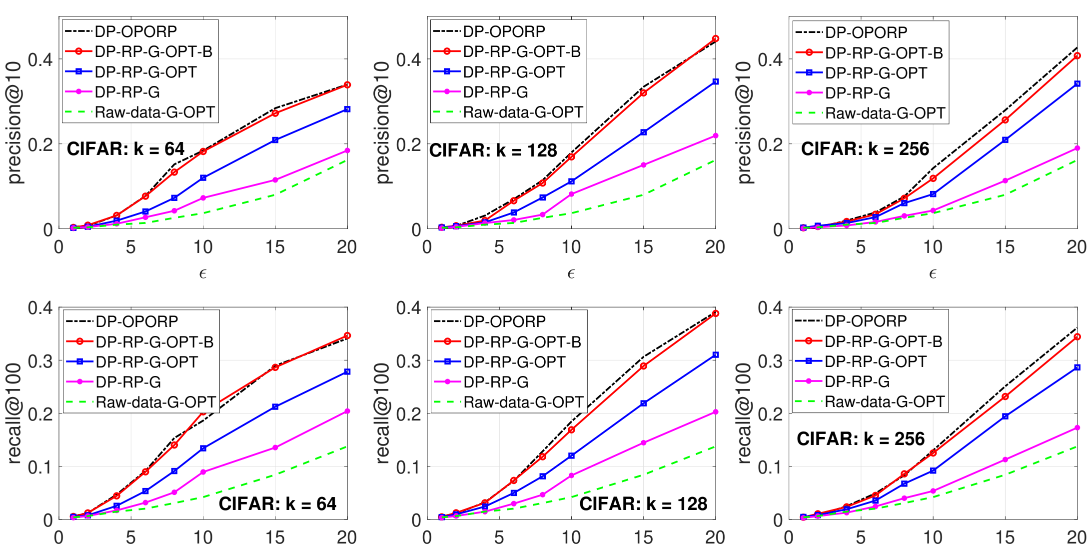

Figure 14: Retrieval recall and precision on CIFAR, $\beta = 1$, $\delta = 10^{-6}$.

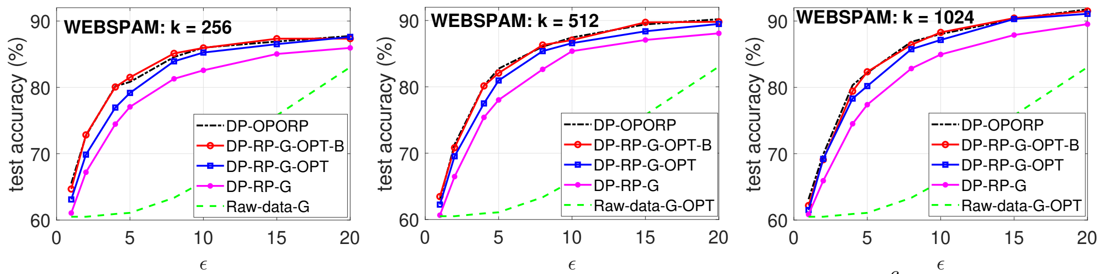

Figure 15: SVM classification on WEBSPAM, $\beta = 1$, $\delta = 10^{-6}$.

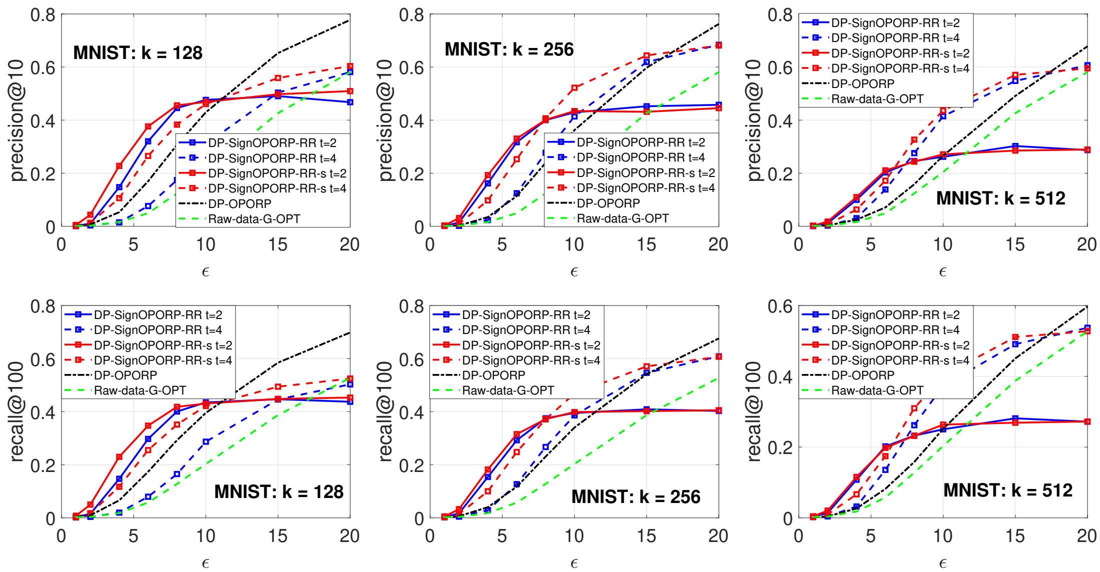

Figure 16: Retrieval on MNIST with DP-SignOPORP-RR and DP-SignOPORP-RR-smooth (in the caption, "-s" stands for "-smooth". For DP-OPORP and Raw-data-G-OPT, we let $\delta = 10^{-6}$.

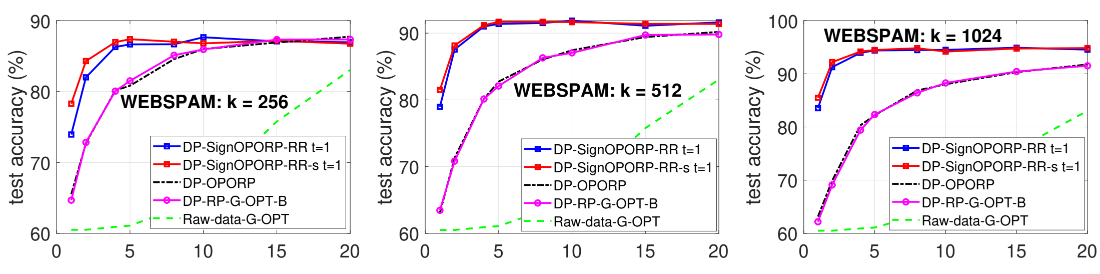

Figure 17: SVM classification on Webspam with DP-SignOPORP-RR and DP-SignOPORP-RR-s. For DP-OPORP and Raw-data-G-OPT, we let $\delta = 10^{-6}$.

E Deferred Proofs

The following lemma on Gaussian random variables will be used in our proof for Lemma 4.2, and may also be of independent interest.

Lemma E.1. *Let* $\begin{pmatrix} X \\ Y \end{pmatrix} \sim N\begin{pmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{pmatrix}$. *Denote* $r = \sigma_x/\sigma_y$. *Then we have:*

1. $Pr(|X| > |Y|) = \frac{1}{\pi}\left[\tan^{-1}\left(\frac{r-\rho}{\sqrt{1-\rho^2}}\right) + \tan^{-1}\left(\frac{r+\rho}{\sqrt{1-\rho^2}}\right)\right]$. *When* $r \leq 1$, *the maximum is achieved at* $\rho = 0$, *i.e.,* $\max_\rho Pr(|X| < |Y|) = \frac{2}{\pi}\tan^{-1}(r)$.

2. The conditional expectation:

$$\mathbb{E}\left[|X|\big||X| > |Y|\right] = \sigma_x\sqrt{\frac{\pi}{2}} \cdot \frac{\frac{r-\rho}{\sqrt{1+r^2-2r\rho}} + \frac{r+\rho}{\sqrt{1+r^2+2r\rho}}}{\tan^{-1}\left(\frac{r-\rho}{\sqrt{1-\rho^2}}\right) + \tan^{-1}\left(\frac{r+\rho}{\sqrt{1-\rho^2}}\right)}.$$

3. The conditional tail probability: for any $r > 0$, $\rho \in (-1, 1)$, *for any* $t > 0$,

$$Pr(|X| > t\big||X| > |Y|) \leq \exp\left(-\frac{t^2}{2\sigma_x^2}\right).$$

Proof. The bivariate normal density function is

$$\begin{aligned} f(x,y) &= \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}\exp\left(-\frac{\frac{x^2}{\sigma_x^2} - \frac{2\rho xy}{\sigma_x\sigma_y} + \frac{y^2}{\sigma_y^2}}{2(1-\rho^2)}\right) \\ &= \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}\exp\left(-\frac{x^2}{2\sigma_x^2}\right)\exp\left(-\frac{(\frac{y}{\sigma_y} - \rho\frac{x}{\sigma_x})^2}{2(1-\rho^2)}\right). \end{aligned}$$

Therefore, we have $\mathbb{E}\left[|X|\big||X| > |Y|\right] = \frac{A}{P}$, with

$$\begin{aligned} A &= \int_{-\infty}^{\infty}\frac{|x|}{\sqrt{2\pi}\sigma_x\sqrt{1-\rho^2}}\exp\left(-\frac{x^2}{2\sigma_x^2}\right)dx\int_{-|x|}^{|x|}\frac{1}{\sqrt{2\pi}\sigma_y}\exp\left(-\frac{(\frac{y}{\sigma_y} - \rho\frac{x}{\sigma_x})^2}{2(1-\rho^2)}\right)dy, \\ P &= \int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma_x\sqrt{1-\rho^2}}\exp\left(-\frac{x^2}{2\sigma_x^2}\right)dx\int_{-|x|}^{|x|}\frac{1}{\sqrt{2\pi}\sigma_y}\exp\left(-\frac{(\frac{y}{\sigma_y} - \rho\frac{x}{\sigma_x})^2}{2(1-\rho^2)}\right)dy. \end{aligned}$$

Note that $P = Pr(|X| > |Y|)$ in the first statement of the theorem. Our calculation will use the following two identities involving the Gaussian functions [11]:

$$\int_0^\infty \phi(ax)\Phi(bx)dx = \frac{1}{2\pi|a|}\left(\frac{\pi}{2} + \tan^{-1}\left(\frac{b}{|a|}\right)\right), \tag{17}$$

$$\int_0^\infty x\phi(ax)\Phi(bx)dx = \frac{1}{2\sqrt{2\pi}}\left(1 + \frac{b}{\sqrt{1+b^2}}\right), \tag{18}$$

where $\phi(x)$ and $\Phi(x)$ are the pdf and cdf of the standard Gaussian distribution.

With a proper change of random variables, we can compute A as

$$\begin{aligned} A &= \int_{-\infty}^{\infty}\frac{|x|}{\sqrt{2\pi}\sigma_x\sqrt{1-\rho^2}}\exp\left(-\frac{x^2}{2\sigma_x^2}\right)dx\int_{\frac{\frac{-|x|}{\sigma_y} - \rho\frac{x}{\sigma_x}}{\sqrt{1-\rho^2}}}^{\frac{\frac{|x|}{\sigma_y} - \rho\frac{x}{\sigma_x}}{\sqrt{1-\rho^2}}}\sqrt{1-\rho^2}\frac{1}{\sqrt{2\pi}}e^{-s^2}ds \\ &= \int_{-\infty}^{\infty}\frac{|x|}{\sqrt{2\pi}\sigma_x}\exp\left(-\frac{x^2}{2\sigma_x^2}\right)\left[\Phi\left(\frac{\frac{|x|}{\sigma_y} - \rho\frac{x}{\sigma_x}}{\sqrt{1-\rho^2}}\right) - \Phi\left(\frac{\frac{-|x|}{\sigma_y} - \rho\frac{x}{\sigma_x}}{\sqrt{1-\rho^2}}\right)\right]dx \\ &= \int_{-\infty}^{\infty}\frac{\sigma_x|t|}{\sqrt{2\pi}}\exp\left(-\frac{t^2}{2}\right)\left[\Phi\left(\frac{\frac{\sigma_x}{\sigma_y}|t| - \rho t}{\sqrt{1-\rho^2}}\right) - \Phi\left(-\frac{\frac{\sigma_x}{\sigma_y}|t| - \rho t}{\sqrt{1-\rho^2}}\right)\right]dt \\ &:= A_1 - A_2. \end{aligned}$$

For the first term we have

$$
\begin{aligned}
A_1 &= \sigma_x \left[\int_0^\infty \frac{t}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) \Phi\left(\frac{\frac{\sigma_x}{\sigma_y} - \rho}{\sqrt{1-\rho^2}} t \right) dt + \int_{-\infty}^0 \frac{-t}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) \Phi\left(-\frac{\frac{\sigma_x}{\sigma_y} + \rho}{\sqrt{1-\rho^2}} t \right) dt \right] \\
&= \sigma_x \left[\int_0^\infty \frac{t}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) \Phi\left(\frac{\frac{\sigma_x}{\sigma_y} - \rho}{\sqrt{1-\rho^2}} t \right) dt + \int_0^\infty \frac{s}{\sqrt{2\pi}} \exp\left(-\frac{s^2}{2}\right) \Phi\left(\frac{\frac{\sigma_x}{\sigma_y} + \rho}{\sqrt{1-\rho^2}} s \right) ds \right] \\
&= \sigma_x \left[\frac{1}{2\sqrt{2\pi}} \left(1 + \frac{\frac{\frac{\sigma_x}{\sigma_y} - \rho}{\sqrt{1-\rho^2}}}{\sqrt{1 + \frac{(\frac{\sigma_x}{\sigma_y} - \rho)^2}{1-\rho^2}}} \right) + \frac{1}{2\sqrt{2\pi}} \left(1 + \frac{\frac{\frac{\sigma_x}{\sigma_y} + \rho}{\sqrt{1-\rho^2}}}{\sqrt{1 + \frac{(\frac{\sigma_x}{\sigma_y} + \rho)^2}{1-\rho^2}}} \right) \right] \\
&= \sigma_x \left[\frac{1}{\sqrt{2\pi}} + \frac{1}{2\sqrt{2\pi}} \left(\frac{r-\rho}{\sqrt{1+r^2-2r\rho}} + \frac{r+\rho}{\sqrt{1+r^2+2r\rho}} \right) \right],
\end{aligned}
$$

where we denote $r = \frac{\sigma_x}{\sigma_y}$ and use (18). Similarly, we have that

$$
\begin{aligned}
A_2 &= \sigma_x \left[\int_0^\infty \frac{t}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) \Phi\left(-\frac{r+\rho}{\sqrt{1-\rho^2}} t \right) dt + \int_{-\infty}^0 \frac{-t}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) \Phi\left(\frac{r-\rho}{\sqrt{1-\rho^2}} t \right) dt \right] \\
&= \sigma_x \left[\int_0^\infty \frac{t}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) \Phi\left(-\frac{r+\rho}{\sqrt{1-\rho^2}} t \right) dt + \int_0^\infty \frac{s}{\sqrt{2\pi}} \exp\left(-\frac{s^2}{2}\right) \Phi\left(-\frac{r-\rho}{\sqrt{1-\rho^2}} s \right) ds \right] \\
&= \sigma_x \left[\frac{1}{\sqrt{2\pi}} - \frac{1}{2\sqrt{2\pi}} \left(\frac{r-\rho}{\sqrt{1+r^2-2r\rho}} + \frac{r+\rho}{\sqrt{1+r^2+2r\rho}} \right) \right].
\end{aligned}
$$

Therefore, we obtain

$$
A(\rho, r) = A_1 - A_2 = \frac{\sigma_x}{\sqrt{2\pi}} \left(\frac{r-\rho}{\sqrt{1+r^2-2r\rho}} + \frac{r+\rho}{\sqrt{1+r^2+2r\rho}} \right). \tag{19}
$$

To compute P, by doing a similar change of variables, we have

$$
\begin{aligned}
P &= \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) \left[\Phi\left(\frac{r|t| - \rho t}{\sqrt{1-\rho^2}} \right) - \Phi\left(-\frac{r|t| - \rho t}{\sqrt{1-\rho^2}} \right) \right] dt \\
&:= P_1 - P_2.
\end{aligned}
$$

Using (17), we obtain

$$
\begin{aligned}
P_1 &= \int_0^\infty \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) \Phi\left(\frac{r-\rho}{\sqrt{1-\rho^2}} t \right) dt + \int_{-\infty}^0 \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) \Phi\left(-\frac{r+\rho}{\sqrt{1-\rho^2}} t \right) dt \\
&= \int_0^\infty \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) \Phi\left(\frac{r-\rho}{\sqrt{1-\rho^2}} t \right) dt + \int_0^\infty \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{s^2}{2}\right) \Phi\left(\frac{r+\rho}{\sqrt{1-\rho^2}} s \right) ds \\
&= \frac{1}{2\pi} \left(\frac{\pi}{2} + \tan^{-1}\left(\frac{r-\rho}{\sqrt{1-\rho^2}} \right) \right) + \frac{1}{2\pi} \left(\frac{\pi}{2} + \tan^{-1}\left(\frac{r+\rho}{\sqrt{1-\rho^2}} \right) \right) \\
&= \frac{1}{2} + \frac{1}{2\pi} \left[\tan^{-1}\left(\frac{r-\rho}{\sqrt{1-\rho^2}} \right) + \tan^{-1}\left(\frac{r+\rho}{\sqrt{1-\rho^2}} \right) \right], \\
P_2 &= \frac{1}{2} - \frac{1}{2\pi} \left[\tan^{-1}\left(\frac{r-\rho}{\sqrt{1-\rho^2}} \right) + \tan^{-1}\left(\frac{r+\rho}{\sqrt{1-\rho^2}} \right) \right],
\end{aligned}
$$

which leads to

$$
P(\rho, r) = P_1 - P_2 = \frac{1}{\pi} \left[\tan^{-1}\left(\frac{r-\rho}{\sqrt{1-\rho^2}} \right) + \tan^{-1}\left(\frac{r+\rho}{\sqrt{1-\rho^2}} \right) \right] \tag{20}
$$

Therefore, we know that

$$\mathbb{E}\left[|X|\big||X|>|Y|\right]=\frac{A(\rho,r)}{P(\rho,r)}=\sigma_x\sqrt{\frac{\pi}{2}}\cdot\frac{\frac{r-\rho}{\sqrt{1+r^2-2r\rho}}+\frac{r+\rho}{\sqrt{1+r^2+2r\rho}}}{\tan^{-1}\left(\frac{r-\rho}{\sqrt{1-\rho^2}}\right)+\tan^{-1}\left(\frac{r+\rho}{\sqrt{1-\rho^2}}\right)},$$

with $r=\sigma_x/\sigma_y$. We now investigate the derivative of P. By some algebra, we can show that

$$\frac{\partial P(\rho,r)}{\partial\rho}=\frac{2r\rho(r^2-1)}{(1+r^2-2r\rho)(1+r^2+2r\rho)\sqrt{1-\rho^2}}.$$

When $0<r\leq 1$, $\frac{\partial P(\rho,r)}{\partial\rho}\geq 0$ when $\rho\leq 0$ and $\frac{\partial P(\rho,r)}{\partial\rho}\leq 0$ when $\rho>0$. Therefore, $\max_\rho P(\rho,r)=P(0,r)=\frac{2}{\pi}\tan^{-1}(r)$.

Tail bound. By our previous calculations, the conditional distribution of X given $|X|>|Y|$ is

$$f(x\big||X|>|Y|)=\frac{\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)\left[\Phi\left(\frac{r|x|-\rho x}{\sqrt{1-\rho^2}}\right)-\Phi\left(-\frac{r|x|-\rho x}{\sqrt{1-\rho^2}}\right)\right]}{P},\quad x\in\mathbb{R},$$

with $P=Pr(|X|>|Y|)$ in (20) the normalizing constant to make the integral equal to 1.

The conditional tail probability can be computed as follows. For some $t>0$, by symmetry,

$$\begin{aligned}&Pr(|X|>t,|X|>|Y|)\\=&2\int_t^\infty\frac{1}{\sqrt{2\pi}\sigma_x}\exp\left(-\frac{x^2}{2\sigma_x^2}\right)\left[\Phi\left(\frac{r|x|-\rho x}{\sqrt{1-\rho^2}}\right)-\Phi\left(-\frac{r|x|-\rho x}{\sqrt{1-\rho^2}}\right)\right]dx\\=&2\int_{\frac{t}{\sigma_x}}^\infty\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)\left[\Phi\left(\frac{r|x|-\rho x}{\sqrt{1-\rho^2}}\right)-\Phi\left(-\frac{r|x|-\rho x}{\sqrt{1-\rho^2}}\right)\right]dx\\:=&2(\tilde{P}_1-\tilde{P}_2).\end{aligned}$$

For $\tilde{P}_1$, using polar coordinates we have

$$\begin{aligned}\tilde{P}_1=&\frac{1}{2\pi}\int_{\frac{t}{\sigma_x}}^\infty e^{-\frac{x^2}{2}}dx\int_{-\infty}^{\frac{r-\rho}{\sqrt{1-\rho^2}}x}e^{-\frac{y^2}{2}}dy\\=&\frac{1}{2\pi}\int_{-\frac{\pi}{2}}^{\tan^{-1}(\frac{r-\rho}{\sqrt{1-\rho^2}})}d\theta\int_{\frac{t}{\sigma_x\cos(\theta)}}^\infty e^{-\frac{r^2}{2}}rdr\\=&\frac{1}{2\pi}\int_{-\frac{\pi}{2}}^{\tan^{-1}(\frac{r-\rho}{\sqrt{1-\rho^2}})}\exp\left(-\frac{t^2}{2\sigma_x^2\cos^2(\theta)}\right)d\theta.\end{aligned}$$

Similarly,

$$\tilde{P}_2=\frac{1}{2\pi}\int_{-\frac{\pi}{2}}^{\tan^{-1}(-\frac{r+\rho}{\sqrt{1-\rho^2}})}\exp\left(-\frac{t^2}{2\sigma_x^2\cos^2(\theta)}\right)d\theta.$$

Therefore, we obtain

$$\begin{aligned}Pr(|X|>t,|X|>|Y|)&=\frac{1}{\pi}\int_{\tan^{-1}(-\frac{r+\rho}{\sqrt{1-\rho^2}})}^{\tan^{-1}(\frac{r-\rho}{\sqrt{1-\rho^2}})}\exp\left(-\frac{t^2}{2\sigma_x^2\cos^2(\theta)}\right)d\theta\\&=\frac{1}{\pi}\int_{-\tan^{-1}(\frac{r+\rho}{\sqrt{1-\rho^2}})}^{\tan^{-1}(\frac{r-\rho}{\sqrt{1-\rho^2}})}\exp\left(-\frac{t^2}{2\sigma_x^2\cos^2(\theta)}\right)d\theta\\&\leq e^{-\frac{t^2}{2\sigma_x^2}}\frac{1}{\pi}\int_{-\tan^{-1}(\frac{r+\rho}{\sqrt{1-\rho^2}})}^{\tan^{-1}(\frac{r-\rho}{\sqrt{1-\rho^2}})}d\theta,\end{aligned}$$

since $\cos^2(\theta) \in [0, 1]$. Notice that P in (20) can be written as $P = \frac{1}{\pi}\int_{-\tan^{-1}(\frac{r+\rho}{\sqrt{1-\rho^2}})}^{\tan^{-1}(\frac{r-\rho}{\sqrt{1-\rho^2}})} d\theta$. Hence, we know that the conditional tail probability is

$$Pr(|X| > t \big| |X| > |Y|) = \frac{Pr(|X| > t, |X| > |Y|)}{Pr(|X| > |Y|)} \leq \exp\left(-\frac{t^2}{2\sigma_x^2}\right), \quad \forall r > 0, \rho \in (-1, 1).$$

At the boundaries $\rho = 1, \rho = -1$, one can verify $Pr(|X| > |Y|) = 0$. This concludes the proof. $\square$

E.1 Proof of Lemma 4.2

Proof. Since X_i's are independent, we know that

$$Pr(\max_{i=1,...,p} |X| < |Y|) = \prod_{i=1}^{p} Pr(|X_i| < |Y|).$$

By Lemma E.1, among all the possible dependency structures, the above probability reaches its minimum when every X_i is independent of Y. Therefore, $Pr(\max_{i=1,...,p} |X| > |Y|) = 1 - Pr(\max_{i=1,...,p} |X| < |Y|)$ achieves maximum when $\rho(X_i, Y) = 0$, $\forall i = 1, ..., p$. Since $|X_i|$ follows a half-normal distribution with cdf being $\mathrm{erf}(\frac{x}{\sqrt{2}\sigma_x})$, we have

$$Pr(\max_{i=1,...,p} |X_i| \leq t) = \mathrm{erf}\left(\frac{t}{\sqrt{2}\sigma_x}\right)^p = \left[2\Phi\left(\frac{x}{\sigma_x}\right) - 1\right]^p,$$

and probability density function $g(x) = 2p[\Phi(\frac{x}{\sigma_x}) - 1]^{p-1}\frac{1}{\sqrt{2\pi}\sigma_x}e^{-\frac{x^2}{2\sigma_x^2}}$. When Y is independent of all X_i's (which gives the upper bound), we have

$$\begin{aligned}
Pr(\max_{i=1,...,p} |X| > |Y|) &= \int_0^\infty 2p\left[\Phi\left(\frac{x}{\sigma_x}\right) - 1\right]^{p-1} \frac{1}{\sqrt{2\pi}\sigma_x} e^{-\frac{x^2}{2\sigma_x^2}} Pr(|Y| < x) dx \\
&= \int_0^\infty 2p\left[\Phi\left(\frac{x}{\sigma_x}\right) - 1\right]^{p-1} \frac{1}{\sqrt{2\pi}\sigma_x} e^{-\frac{x^2}{2\sigma_x^2}} \mathrm{erf}\left(\frac{x}{\sqrt{2}\sigma_y}\right) dx \\
&= \int_0^\infty 2p[2\Phi(t) - 1]^{p-1}[2\Phi(rt) - 1]\phi(t)dt,
\end{aligned}$$

with a proper change of variables. This gives an upper bound as shown above. $\square$

E.2 Proof of Proposition 4.4

Proof. Consider a single Gaussian projection vector w with iid $N(0, 1)$ entries. Since $w^T u = \sum_{i=1}^p u_i w_i$ and each $w_i \sim N(0, 1)$, we know that $\begin{pmatrix} \beta w_i \\ x \end{pmatrix} \sim N\begin{pmatrix} \beta^2 & \rho_i\beta\|u\| \\ \rho_i\beta\|u\| & \|u\|^2 \end{pmatrix}$ where $\rho_i = \frac{u_i}{\|u\|}$ is the correlation coefficient. Since $|w^T(u - u')| \leq \beta \max_{i=1,...,p} |w_i|$ by Definition 2.2 of β-neighboring (and more generally, when $\|u - u'\|_1 \leq \beta$), we have

$$Pr(\max_{u' \in Nb(u)} |w^T(u - u')| \geq |w^T u|) = Pr(\beta \max_{i=1,...,p} |w_i| \geq |w^T u|).$$

Note that, $\beta \max_{i=1,...,p} |w_i| \geq |w^T u|$ is a necessary condition for the event that there exists a neighbor such that $sign(w^T u) \neq sign(w^T u')$. Denote $I = \mathbb{1}\{\beta \max_{i=1,...,p} |w_i| \geq |w^T u|\}$. Applying Lemma 4.2 with $r = \beta/\|u\| \leq 1$ yields

$$\mathbb{E}[I] = Pr(\beta \max_{i=1,...,p} |w_i| \geq |w^T u|) \leq F_{\|u\|,p} = \int_0^\infty 2p[2\Phi(t) - 1]^{p-1}[2\Phi(rt) - 1]\phi(t)dt \quad (21)$$

as given by (5). Let I_j be the corresponding indicator function w.r.t. each column in the projection matrix W. Denote $N_+ = \sum_{j=1}^k I_j$, and by the above reasoning, we know that $|S| \leq N_+$ where S is defined in the theorem. Since the columns of W are independent, N_+ follows a $Binomial(k, \mathbb{E}[I])$

distribution with k trials and success probability $\mathbb{E}[I]$ bounded as above. Applying Chernoff's bound on binomial variable (Lemma 4.3), we obtain

$$Pr(N_+ \geq (1+\eta)F_{\|u\|,p}k) \leq \exp(-\frac{\eta^2 F_{\|u\|,p}k}{\eta+2}).$$

Setting the RHS to δ gives $\eta = \frac{\log(1/\delta)+\sqrt{(\log(1/\delta))^2+8F_{\|u\|,p}k\log(1/\delta)}}{2F_{\|u\|,p}k}$. Therefore, with probability $1-\delta$,

$$N_+(\|u\|,\delta,k,p) \leq F_{\|u\|,p}k + \frac{1}{2}\Big[\log(1/\delta) + \sqrt{(\log(1/\delta))^2 + 8F_{\|u\|,p}k\log(1/\delta)}\Big].$$

In addition, $N_+ \leq k$ trivially. The proof is complete. □

E.3 Proof of Theorem 4.5

Proof. Let $s = sign(W^T u) \in \{-1,+1\}^k$, $s' = sign(W^T u') \in \{-1,+1\}^k$. We denote the collision probability of non-private SignRP as

$$P_{SRP} = Pr(s_{1j} = s_{2j}) = 1 - \frac{\cos^{-1}(\rho)}{\pi} = 1 - \frac{\theta}{\pi}.$$

Hence, the collision probability of DP-SignRP-RR can be computed as

$$\begin{aligned}
\tilde{P} := Pr(\tilde{s}_{1j} = \tilde{s}_{2j}) &= Pr(s_{1j} = s_{2j}, \text{both change sign or not change sign}) \\
&\qquad\qquad + Pr(s_{1j} \neq s_{2j}, \text{one sign changes}) \\
&= P_{SRP}[(\frac{e^{\epsilon'}}{e^{\epsilon'}+1})^2 + (\frac{1}{e^{\epsilon'}+1})^2] + 2(1-P_{SRP})\frac{e^{\epsilon'}}{(e^{\epsilon'}+1)^2} \\
&= P_{SRP}\frac{(e^{\epsilon'}-1)^2}{(e^{\epsilon'}+1)^2} + \frac{2e^{\epsilon'}}{(e^{\epsilon'}+1)^2},
\end{aligned}$$

which increases linearly in P_{SRP}. Thus, it holds that

$$\mathbb{E}[\hat{P}_{RR}] = \frac{(e^{\epsilon'}+1)^2}{(e^{\epsilon'}-1)^2}\tilde{P} - \frac{2e^{\epsilon'}}{(e^{\epsilon'}-1)^2} = P_{SRP} = 1 - \frac{\theta}{\pi},$$

which implies $\mathbb{E}[\hat{\theta}_{RR}] = \pi\left(1-(1-\frac{\theta}{\pi})\right) = \theta$. To compute the variance, we first estimate $\theta = \cos^{-1}(\rho)$ by

$$\hat{\theta} = \pi(1-\hat{P}_{RR}).$$

Then according to the Central Limit Theorem (CLT), for the sample mean of iid Bernoulli's, as $k\to\infty$, we have

$$\frac{1}{k}\sum_{j=1}^k \mathbb{1}\{\tilde{s}_{1j} = \tilde{s}_{2j}\} \to N(\tilde{P}, \frac{\tilde{P}(1-\tilde{P})}{k}).$$

As a result, we have $\hat{\theta} \to N(\theta, \frac{V_{RR}}{k})$, where

$$\begin{aligned}
V_{RR} &= \frac{\pi^2(e^{\epsilon'}+1)^4}{(e^{\epsilon'}-1)^4}\Big[(1-\frac{\theta}{\pi})\frac{(e^{\epsilon'}-1)^2}{(e^{\epsilon'}+1)^2} + \frac{2e^{\epsilon'}}{(e^{\epsilon'}+1)^2}\Big]\Big[\frac{e^{2\epsilon'}+1}{(e^{\epsilon'}+1)^2} - (1-\frac{\theta}{\pi})\frac{(e^{\epsilon'}-1)^2}{(e^{\epsilon'}+1)^2}\Big] \\
&= \frac{\pi^2(e^{\epsilon'}+1)^4}{(e^{\epsilon'}-1)^4}\Big[(1-\frac{\theta}{\pi})\frac{(e^{\epsilon'}-1)^2}{(e^{\epsilon'}+1)^2} + \frac{2e^{\epsilon'}}{(e^{\epsilon'}+1)^2}\Big]\Big[\frac{\theta}{\pi}\frac{(e^{\epsilon'}-1)^2}{(e^{\epsilon'}+1)^2} + \frac{2e^{\epsilon'}}{(e^{\epsilon'}+1)^2}\Big] \\
&= \frac{\pi^2\theta}{\pi}(1-\frac{\theta}{\pi}) + (1-\frac{\theta}{\pi})\frac{2e^{\epsilon'}}{(e^{\epsilon'}-1)^2} + \frac{\theta}{\pi}\frac{2e^{\epsilon'}}{(e^{\epsilon'}-1)^2} + \frac{4e^{2\epsilon'}}{(e^{\epsilon'}-1)^4} \\
&= \theta(\pi-\theta) + \frac{2\pi^2 e^{\epsilon'}}{(e^{\epsilon'}-1)^2} + \frac{4\pi^2 e^{2\epsilon'}}{(e^{\epsilon'}-1)^4}.
\end{aligned}$$

We conclude the proof by replacing $\epsilon' = \epsilon/N_+$. □

E.4 Proof of Theorem 4.6

Proof. Let us consider a single projection vector $w_j = W_{[:,j]}$. Denote $x_j = w_j^T u$ and $x'_j = w_j^T u'$ for a neighboring data u' of u, and $s_j = sign(x_j)$, $s'_j = sign(x'_j)$. Also, let $L_j = \lceil \frac{|x_j|}{\beta \max_{i=1,...,p} |W_{ij}|} \rceil$ and $L'_j = \lceil \frac{|x'_j|}{\beta \max_{i=1,...,p} |W_{ij}|} \rceil$. W.l.o.g., we can assume $s_j = 1$ by the symmetry of random projection and the symmetry of DP. Consider two cases:

- Case I: $L_j \geq 2$. In this case, we know that $s'_j = s_j$, i.e., the change from u to u' will not change the sign of the projection. Thus, in Algorithm 4, we have

$$\frac{Pr(\tilde{s}_j = 1)}{Pr(\tilde{s}'_j = 1)} = \exp(\frac{L_j - L'_j}{k}\epsilon)\frac{\exp(\frac{L'_j}{k}\epsilon) + 1}{\exp(\frac{L_j}{k}\epsilon) + 1}.$$

  By the definition of β-adjacency, $|L_j - L'_j|$ equals either 0 or 1. When $L_j = L'_j$, $\frac{Pr(\tilde{s}_j=1)}{Pr(\tilde{s}'_j=1)} = 1$. When $L_j - L'_j = 1$, we have

$$\frac{Pr(\tilde{s}_j = 1)}{Pr(\tilde{s}'_j = 1)} = \frac{\exp(\frac{L_j}{k}\epsilon) + \exp(\frac{1}{k}\epsilon)}{\exp(\frac{L_j}{k}\epsilon) + 1}.$$

  Hence, we have $1 \leq \frac{Pr(\tilde{s}_j=1)}{Pr(\tilde{s}'_j=1)} \leq e^{\frac{\epsilon}{k}}$ by the numeric identity $1 \leq \frac{a+c}{b+c} \leq \frac{a}{b}$ for $a \geq b > 0$ and $c > 0$. Thus, by symmetry, $e^{-\frac{\epsilon}{k}} \leq \frac{Pr(\tilde{s}_j=1)}{Pr(\tilde{s}'_j=1)} \leq e^{\frac{\epsilon}{k}}$. On the other hand,

$$\frac{Pr(\tilde{s}_j = -1)}{Pr(\tilde{s}'_j = -1)} = \frac{\exp(\frac{L'_j}{k}\epsilon) + 1}{\exp(\frac{L_j}{k}\epsilon) + 1}.$$

  Similarly, when $L_j = L'_j$, the ratio equals 1. When $L_j = L'_j - 1$, we have $\frac{Pr(\tilde{s}_j=-1)}{Pr(\tilde{s}'_j=-1)} \leq \exp(\frac{L'_j}{k}\epsilon - \frac{L_j}{k}\epsilon) = e^{\frac{\epsilon}{k}}$. By symmetry we obtain $e^{-\frac{\epsilon}{k}} \leq \frac{Pr(\tilde{s}_j=-1)}{Pr(\tilde{s}'_j=-1)} \leq e^{\frac{\epsilon}{k}}$.

- Case II: $L_j = 1$. In this case, s_j might be different from s'_j. First, if $L'_j = 2$, then the above analysis also applies that $\frac{Pr(\tilde{s}_j=1)}{Pr(\tilde{s}'_j=1)}$ and $\frac{Pr(\tilde{s}_j=-1)}{Pr(\tilde{s}'_j=-1)}$ are both lower and upper bounded by $e^{-\frac{\epsilon}{k}}$ and $e^{\frac{\epsilon}{k}}$, respectively. It suffices to examine the case when $L'_j = 1$. In this case, if $s'_j = s_j = 1$ then the probability ratios simply equal 1. If $s'_j = -1$, we have

$$\frac{Pr(\tilde{s}_j = 1)}{Pr(\tilde{s}'_j = 1)} = \frac{\frac{\exp(\frac{\epsilon}{k})}{\exp(\frac{\epsilon}{k})+1}}{\frac{1}{\exp(\frac{\epsilon}{k})+1}} = e^{\frac{\epsilon}{k}}, \quad \frac{Pr(\tilde{s}_j = -1)}{Pr(\tilde{s}'_j = -1)} = \frac{\frac{1}{\exp(\frac{\epsilon}{k})+1}}{\frac{\exp(\frac{\epsilon}{k})}{\exp(\frac{\epsilon}{k})+1}} = e^{-\frac{\epsilon}{k}}.$$

Combining two cases, we have that $\log \frac{Pr(\tilde{s}_j=t)}{Pr(\tilde{s}'_j=t)} \leq \frac{\epsilon}{k}$, for $t = -1, 1$, and for all $j = 1, ..., k$. That is, each single perturbed sign achieves $\frac{\epsilon}{k}$-DP. Since the k projections are independent, by Theorem 2.1, we know that the output bit vector $\tilde{s} = [\tilde{s}_1, ..., \tilde{s}_k]$ is ϵ-DP as claimed. □

References

[1] Dimitris Achlioptas. Database-friendly random projections: Johnson-lindenstrauss with binary coins. *J. Comput. Syst. Sci.*, 66(4):671–687, 2003.

[2] Damien Desfontaines and Balázs Pejó. Sok: Differential privacies. *Proc. Priv. Enhancing Technol.*, 2020(2):288–313, 2020.

[3] Jinshuo Dong, Aaron Roth, and Weijie J Su. Gaussian differential privacy. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 84(1):3–37, 2022.

[4] Cynthia Dwork and Guy N Rothblum. Concentrated differential privacy. *arXiv preprint arXiv:1603.01887*, 2016.

[5] Justin Hsu Marco Gaboardi Andreas Haeberlen and Sanjeev Khanna. Differential privacy: An economic method for choosing epsilon. *arXiv preprint arXiv:1402.3329*, 2014.

[6] Bargav Jayaraman and David Evans. Evaluating differentially private machine learning in practice. In *Proceedings of the 28th USENIX Security Symposium (USENIX Security)*, pages 1895–1912, Santa Clara, CA, 2019.

[7] Christopher T Kenny, Shiro Kuriwaki, Cory McCartan, Evan TR Rosenman, Tyler Simko, and Kosuke Imai. The use of differential privacy for census data and its impact on redistricting: The case of the 2020 us census. *Science advances*, 7(41):eabk3283, 2021.

[8] Ping Li, Trevor J Hastie, and Kenneth W Church. Very sparse random projections. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD)*, pages 287–296, Philadelphia, PA, 2006.

[9] Ping Li and Xiaoyun Li. OPORP: One permutation + one random projection. *arXiv preprint arXiv:2302.03505*, 2023.

[10] Ilya Mironov. Rényi differential privacy. In *Proceedings of the 30th IEEE Computer Security Foundations Symposium (CSF)*, pages 263–275, Santa Barbara, CA, 2017.

[11] Donald Bruce Owen. A table of normal integrals: A table. *Communications in Statistics-Simulation and Computation*, 9(4):389–419, 1980.

[12] Jordi Soria-Comas, Josep Domingo-Ferrer, David Sánchez, and David Megías. Individual differential privacy: A utility-preserving formulation of differential privacy guarantees. *IEEE Trans. Inf. Forensics Secur.*, 12(6):1418–1429, 2017.