Locally Differentially Private Frequency Estimation via Joint Randomized Response

Ye Zheng
Rochester Institute of Technology
ye.zheng@mail.rit.edu

Shafizur Rahman Seeam
Rochester Institute of Technology
ss6365@rit.edu

Yidan Hu
Rochester Institute of Technology
yidan.hu@rit.edu

Rui Zhang
University of Delaware
ruizhang@udel.edu

Yanchao Zhang
Arizona State University
yczhang@asu.edu

Abstract

Local Differential Privacy (LDP) has been widely recognized as a powerful tool for providing a strong theoretical guarantee of data privacy to data contributors against an untrusted data collector. Under a typical LDP scheme, each data contributor independently randomly perturbs their data before submitting them to the data collector, which in turn infers valuable statistics about the original data from received perturbed data. Common to existing LDP mechanisms is an inherent trade-off between the level of privacy protection and data utility in the sense that strong data privacy often comes at the cost of reduced data utility. Frequency estimation based on Randomized Response (RR) is a fundamental building block of many LDP mechanisms. In this paper, we propose a novel Joint Randomized Response (JRR) mechanism based on correlated data perturbations to achieve locally differentially private frequency estimation. JRR divides data contributors into disjoint groups of two members and lets those in the same group jointly perturb their binary data to improve frequency-estimation accuracy and achieve the same level of data privacy by hiding the group membership information in contrast to the classical RR mechanism. Theoretical analysis and detailed simulation studies using both real and synthetic datasets show that JRR achieves the same level of data privacy as the classical RR mechanism while improving the frequency-estimation accuracy in the overwhelming majority of the cases by up to two orders of magnitude.

Keywords

local differential privacy, randomized response, frequency estimation

1 Introduction

Differential privacy [18] is widely considered as the *de facto* framework for providing strong theoretical guarantee of data privacy. Recent years have also witnessed significant interests in developing data analysis techniques for ensuring differential privacy in the local setting, commonly referred to as Local Differential Privacy (LDP) [15]. A local differential privacy mechanism protects individual data contributors' data privacy against an untrusted data collector by having each data contributor randomly perturb their data value before submission while allowing the data collector to learn valuable statistics of the contributors' data. In addition to significant interests from academia, LDP techniques have seen growing adoption by industry for various data analysis applications. For example, Google has deployed RAPPOR [20] into Chrome to privately collect individual web browsing behavior. As another example, Apple adopts LDP algorithms [5] in Safari for privacy-preserving collection of users' typing history to better understand user behaviors.

Significant efforts have been made to achieve a good utility-privacy trade-off for various data analysis tasks. In particular, all existing LDP mechanisms exhibit a natural trade-off between data privacy and data utility at the data collector because strong data privacy for individual data contributors often comes at the cost of reduced data utility [15, 41]. Therefore, a major focus of current research is to design LDP mechanisms that achieve higher data utility without sacrificing privacy guarantees for individual contributors. For example, several recent proposals [6, 8, 12, 19, 33] show that it is possible to improve privacy protection while reducing the amount of noise needed by having an auxiliary server shuffle data contributors' perturbed data before sending them to the data collector. Other proposed approaches include parameter optimization [26, 49], developing advanced encoding schemes [7, 20, 25], random perturbation schemes and estimators [35, 49], interactive data collection schemes [60], cryptography-assisted solutions [41], post-processing techniques [21, 53], etc.

Frequency estimation is a classical data analysis problem in which the data collector aims to learn the number (or ratio) of data contributors with a certain attribute or possessing a particular data value. Randomized Response (RR) [55] is the first known and most classical LDP protocol for frequency estimation. Since frequency estimation is used in many other data analysis tasks such as heavy hitter estimation, mean value estimation, and range queries, RR is also widely used as a fundamental building block in many LDP mechanisms for these tasks [7, 13, 46, 50–52]. Common to these solutions is that every data contributor independently perturbs his/her data before submitting them to the data collector. *An open question is whether it is possible to improve data utility of RR-based LDP mechanisms without loss of LDP guarantees by introducing correlations among the random perturbations performed by different data contributors.*

In this paper, we make the first attempt to explore correlated random perturbations for frequency estimation to improve data

Proceedings on Privacy Enhancing Technologies 2025(3), 242–260

https://doi.org/10.56553/popets-2025-0097

utility without any sacrifice of LDP guarantees. We observe that it is possible to achieve much higher data utility in terms of estimation accuracy at the data collector by randomly dividing data contributors into disjoint groups of two and introducing carefully crafted correlations to each group's random perturbations. At the same time, no additional information can be inferred as long as the group membership is kept secret from the data collector. Based on these observations, we introduce a novel **Joint Randomized Response (JRR)** mechanism for locally differentially private frequency estimation. By carefully tuning the parameters, JRR can achieve significantly higher data utility in the overwhelming majority of cases while offering the same level of LDP protection as the classical RR mechanism.

Our contributions in this paper can be summarized as follows.

- We are the first to explore correlated random perturbations for frequency estimation to improve data utility at the data collector without sacrificing LDP guarantee for individual data contributors.
- We introduce a general Joint Randomized Response (JRR) mechanism that achieves the same level of LDP protection as the classical RR mechanism, while improving the data utility in an overwhelming majority of the cases, especially for a large number of data contributors.
- We present a practical instantiation of JRR by utilizing a non-colluding auxiliary server.
- We thoroughly evaluate JRR via a combination of theoretical analysis and detailed simulation studies using both real and synthetic datasets. Our results show that JRR outperforms the classical RR mechanism for over 97% of the possible frequencies and improve the estimation accuracy by as much as two orders of magnitude.

The rest of the paper is structured as follows. Section 2 presents the problem formulation and reviews LDP and the RR mechanism. Section 3 uses two examples to demonstrate the impact of correlated random perturbations. Section 4 introduces a general JRR mechanism, its performance analysis, and a practical instantiation. Section 5 evaluates the performance of JRR. Section 6 discusses related work. Section 7 concludes this paper and points out several future research directions.

2 Preliminaries

In this section, we formulate the problem and then reviews LDP and the RR mechanism.

2.1 Problem Formulation

We consider a system consisting of a data collector and a set of data contributors $\mathcal{U} = \{u_1, u_2, \cdots, u_n\}$. Each contributor has a binary value $x_i \in D = \{0, 1\}$, and the data collector wants to learn the number of data contributors having value v for each $v \in D$. Data contributors are concerned about their data privacy. As a result, instead of submitting the original value x_i, each contributor u_i randomly perturbs his/her value using a random perturbation mechanism $\mathcal{M}$ and submits the perturbed value $y_i = \mathcal{M}(x_i)$ to the data collector. After receiving the perturbed data from n contributors, the data collector estimates the number of data contributors having value v for each $v \in D$.

We assume the data collector is honest but curious, meaning it faithfully carries out system operations but is interested in inferring the original data values of the contributors. Specifically, we assume the data collector will not register or create fake contributor accounts to participate in data collection, as doing so would risk damaging its business reputation if detected. Moreover, we assume that normal data contributors are concerned about their data privacy and will not disclose their original data values to the data collector. Even if a few data contributors collude with the data collector, we assume that the number of such contributors does not exceed a predefined threshold M, e.g., a small fraction of all the contributors.

We seek to design a locally differentially private frequency estimation scheme that enables the data collector to estimate n_v with high accuracy while providing individual contributors with the same ε-LDP guarantee as the classical RR mechanism.

2.2 Local Differential Privacy

Local Differential Privacy (LDP) is considered a gold standard for privacy-preserving data collection against an untrusted data collector. It requires a perturbation mechanism that provides enough randomness to the private data.

Definition 1 (Local Differential Privacy). *A randomized mechanism $\mathcal{M} : \mathcal{X} \rightarrow \mathsf{Range}(\mathcal{M})$ satisfies ε-LDP if*

$$\frac{\Pr[\mathcal{M}(x) = y]}{\Pr[\mathcal{M}(x') = y]} \leq e^{\varepsilon}, \tag{1}$$

for any inputs $x, x' \in \mathcal{X}$ and any output $y \in \mathsf{Range}(\mathcal{M})$, where $\mathsf{Range}(\mathcal{M})$ is the output range of $\mathcal{M}$.

Here ε is a parameter controlling the level of privacy protection commonly referred to as *privacy budget*. The smaller the ε, the stronger the privacy protection, and vice versa. Intuitively, ε-LDP means that by observing the output y, the data collector cannot infer whether the input is x or x' with high confidence, which provides contributors submitting sensitive data with plausible deniability.

2.3 Review of Randomized Response

Randomized Response (RR) [55] was originally proposed to provide plausible deniability to interviewees answering a sensitive boolean question in a survey. Under RR, each interviewee reports the answer truthfully with probability p (the opposite answer with $q = 1 - p$). RR mechanism satisfies ε-LDP if $p \leq e^{\varepsilon}/(1 + e^{\varepsilon})$.

Assume that the total number of data contributors is n and that n_v contributors have value v for each $v \in D$. Suppose that the data collector receives I_v perturbed value v. The data collector estimates the number of data contributors having value v as

$$\hat{n}_v = \frac{I_v - nq}{p - q}\ , \tag{2}$$

which is an unbiased estimator of n_v [49, 55].

The data utility of RR is commonly measured by the variance of the unbiased estimator $\hat{n}_v$, which is given by

$$\mathrm{Var}[\hat{n}_v] = \frac{\mathrm{Var}[I_v]}{(p - q)^2} = \frac{npq}{(p - q)^2}. \tag{3}$$

Table 1: Joint probability distribution in Example 2.

	$T_1 = 1$	$T_1 = 0$
$T_2 = 1$	0.61	0.19
$T_2 = 0$	0.19	0.01

3 Impact of Correlation Among Data Contributors

In this section, we discuss the potential impact of introducing correlations among the random perturbations performed by different data contributors on data privacy and data utility through examples.

The data utility of LDP protocols such as [49, 55], is commonly measured by the variance of the estimator of the value of interest. A smaller variance indicates higher data utility. Traditional LDP protocols involve each contributor independently perturbing their data. Consequently, the estimator of an LDP protocol is reduced to the sum of the individual contributors' reported values, and its variance is proportional to the sum of the variances of individual reported values.

We find that if multiple contributors jointly perturb their data, the variance of the estimator also depends on the covariance of the jointly perturbed values. By carefully designing the joint perturbation to introduce a negative covariance, it is possible to achieve higher data utility. In what follows, we use two concrete examples to illustrate this finding.

Example 1. *Suppose that there are two data contributors, u_1 and u_2 with values $x_1 = 1$ and $x_2 = 1$, respectively. Each contributor independently perturbs their value using RR with $p = 0.8$. Let T_j be the indicator of reporting truthfulness of u_j, i.e., $T_j = 1$ if $y_j = x_j$ and 0 otherwise. We have*

$$T_j = \begin{cases} 1 & \textit{with probability } p = 0.8, \\ 0 & \textit{with probability } q = 0.2. \end{cases} \tag{4}$$

Estimation of n_1: Assume that the data collector has received I_1 perturbed values of 1. According to Eq. (2), the data collector can estimate n_1 as

$$\hat{n}_1 = \frac{(I_1 - 2 \times 0.2)}{0.6} \ . \tag{5}$$

Data privacy: Since $p/q = 0.8/0.2 = 4$, the RR mechanism in the above example satisfies ln 4-LDP.

Data utility: According to Eq. (3), the variance of $\hat{n}_1$ can be computed as

$$\mathrm{Var}[\hat{n}_1] = \frac{npq}{(2p-1)^2} = \frac{2 \cdot 0.8 \cdot 0.2}{(2 \cdot 0.8 - 1)^2} = 0.89 \ . \tag{6}$$

Example 2. *Consider the same two contributors in Example 1. Let T_j be a binary indicator of a random variable for whether a data contributor u_j reports truthfully, i.e., $T_j = 1$ if $y_j = x_j$ and 0 otherwise. The two contributors jointly perturb their data according to the joint probability distribution shown in Table 1.*

Estimation of n_1: It is easy to see that the marginal probability distribution of both T_1 and T_2 in Table 1 is the same as the probability distribution of T_j in Example 1. Define Y_j to be the indicator random variable for data contributor u_j reports a perturbed value $y_j = 1$ for all $1 \leq j \leq 2$. There are two cases. First, if $x_j = 0$, we have $\Pr[Y_j = 1 | x_j = 0] = \Pr[T_j = 0] = 0.2$. Second, if $x_j = 1$, we have $\Pr[Y_j = 1 | x_j = 1] = \Pr[T_j = 1] = 0.8$. Let I_1 be the random variable for the number of contributors reporting a perturbed value 1. We have $I_1 = Y_1 + Y_2$. Taking the expectation on both sides, we have

$$\begin{aligned} \mathrm{E}[I_1] &= \mathrm{E}[\sum_{j=1}^{2} Y_j] = \sum_{j=1}^{2} \mathrm{E}[Y_j] = \sum_{j=1}^{2} \Pr[Y_j = 1] \\ &= n_1 \cdot \Pr[T_j = 1] + (2 - n_1) \cdot 1 \cdot \Pr[T_j = 0] \\ &= 0.8 n_1 + 0.2 \cdot (2 - n_1) \\ &= 0.4 + 0.6 n_1 \ . \end{aligned} \tag{7}$$

The data collector can estimate n_1 as $\hat{n}_1 = (I_1 - 0.4)/0.6$, which is an unbiased estimator of n_1 and also identical to Eq. (5) in Example 1.

Data privacy: Since the marginal probability distribution of both T_1 and T_2 in Table 1 is the same as the one in Example 1, those marginal probability distributions also satisfy ln 4-LDP. However, it does not indicate that each contributor can enjoy the same level of ln 4-LDP as in Example 1. In fact, the introduction of correlation among different contributors will inevitably reduce privacy guarantee for individual contributors. We postpone the discussion of the potential privacy leakage from the correlation between two contributors in the same group to Section 4.2.

Data utility: The variance of the unbiased estimator is

$$\begin{aligned} \mathrm{Var}[\hat{n}_1] &= \frac{\mathrm{Var}[I_1]}{0.36} = \frac{25}{9} \mathrm{Var}[Y_1 + Y_2] \\ &= \frac{25}{9} (\mathrm{Var}[Y_1] + \mathrm{Var}[Y_2] + 2\mathrm{Cov}[Y_1, Y_2]) \ , \end{aligned} \tag{8}$$

where $\mathrm{Cov}[Y_1, Y_2]$ is the covariance between Y_1 and Y_2.

First, $\mathrm{Var}[Y_1]$ and $\mathrm{Var}[Y_2]$ are the same due to the same marginal distribution. Moreover, since both contributors have the same original value of 1, we have $\mathrm{E}[Y_j] = \Pr[Y_j = 1] = \Pr[T_j = 1] = 0.8$ and $\mathrm{E}[Y_j^2] = \Pr[Y_j = 1] = 0.8$. It follows that

$$\begin{aligned} \mathrm{Var}[Y_1] + \mathrm{Var}[Y_2] = 2\mathrm{Var}[Y_1] &= 2(\mathrm{E}[Y_1^2] - \mathrm{E}^2[Y_1]) \\ &= 2 \cdot (0.8 - 0.8^2) = 0.32 \ . \end{aligned} \tag{9}$$

We now compute $\mathrm{Cov}[Y_1, Y_2]$. Since $x_1 = x_2 = 1$, we have

$$\begin{aligned} \mathrm{E}[Y_1 Y_2] = &0 \times \Pr[T_1 = 0, T_2 = 0] + \\ &1 \times \Pr[T_1 = 1, T_2 = 1] = 0.61, \end{aligned}$$

and it follows that

$$\begin{aligned} \mathrm{Cov}[Y_1, Y_2] &= \mathrm{E}[Y_1 Y_2] - \mathrm{E}[Y_1]\mathrm{E}[Y_2] \\ &= 0.61 - 0.8 \cdot 0.8 = -0.03. \end{aligned} \tag{10}$$

Substitute Eqs. (9) and (10) into Eq. (8), we have

$$\mathrm{Var}[\hat{n}_1] = \frac{1}{0.36}(0.32 + 2 \times (-0.03)) \approx 0.72, \tag{11}$$

which is smaller than the $\mathrm{Var}[\hat{n}_1]$ of 0.89 in Example 1.

From the above two examples, we can see that it is possible to improve data utility, i.e., reduce the variance of the estimator, through the introduction of a negative correlation between Y_1 and Y_2 via joint perturbation of two contributors. Theoretical analysis of generalizing n and x_i in the above examples will be presented in the next section. Meanwhile, several key questions must be answered to fully exploit the potential of joint random perturbation.

Table 2: Joint reporting probability in JRR.

	$T_{2i-1} = 1$	$T_{2i-1} = 0$
$T_{2i} = 1$	$p^2 + \rho pq$	$(1-\rho)pq$
$T_{2i} = 0$	$(1-\rho)pq$	$q^2 + \rho pq$

(1) How can we generalize the above joint perturbation mechanism given in Table 1?
(2) Can the joint perturbation mechanism provide the same level of data privacy as RR? If so, under what condition? In particular, is it possible for the data collector to infer additional information about a target contributor's value by exploiting the correlations among different data contributors?
(3) How can we quantify the data utility of a joint perturbation mechanism?
(4) How can we optimize the joint perturbation mechanism to maximize the data utility while guaranteeing the same level of data privacy as RR?
(5) How can we design a practical joint perturbation mechanism?

We provide answers to these questions in the next section.

4 Joint Randomized Response

This section first introduces a general joint randomized response (JRR) mechanism as a generalization of the classical RR mechanism. We then generalize the definition of LDP to ensure that JRR can provide the same level of privacy protection as the classical RR. We quantify the data utility of JRR in Section 4.3. Section 4.4 presents a heuristic algorithm to choose its parameters for maximized data utility given the desirable level of privacy protection. Finally, we present two practical instantiations of the JRR mechanism.

4.1 A General JRR Mechanism

Assume there are n contributors, $\mathcal{U} = \{u_1, \dots, u_n\}$ each having a binary value and n is an even number. We first divide the n contributors into $n/2$ disjoint groups of two $G_1, \dots, G_{n/2}$ uniformly at random. Without loss of generality, assume that each group G_i consists of contributors u_{2i-1} and u_{2i} for all $1 \le i \le n/2$. Each group G_i of two contributors jointly perturb their values according to the joint probability distribution shown in Table 2, where $0.5 < p \le 1$, $q = 1 - p$, and $1 - 1/p \le \rho \le 1$ are system parameters. In particular, each contributor u_j, $1 \le j \le n$, reports $y_j = x_j$ if $T_j = 1$ and $1 - x_j$ if $T_j = 0$.

4.1.1 Properties of JRR. The general JRR mechanism has several key properties, which are summarized as follows.

First, the marginal probability distribution of every T_j $(1 \le j \le n)$ is identical. In particular, it is easy to verify that

$$\begin{aligned} \Pr[T_j = 1] &= p^2 + \rho pq + (1-\rho)pq = p \\ \Pr[T_j = 0] &= (1-\rho)pq + q^2 + \rho pq = q. \end{aligned} \tag{12}$$

This means that each data contributor reports their value truthfully with probability p and untruthfully with probability $q = 1 - p$, which aligns with the classical RR mechanism with parameter p.

Second, parameter ρ is the correlation coefficient between random variables T_{2i-1} and T_{2i}. Specifically, the correlation coefficient between random variables T_{2i-1} and T_{2i} is given by

$$\frac{\mathrm{Cov}[T_{2i-1}, T_{2i}]}{\sigma_1 \sigma_2} = \frac{\mathrm{E}[T_{2i-1}T_{2i}] - \mathrm{E}[T_{2i-1}]\mathrm{E}[T_{2i}]}{\sigma_1 \sigma_2} \tag{13}$$

where $\mathrm{Cov}[T_{2i-1}, T_{2i}]$ is the covariance of T_{2i-1} and T_{2i}, and σ_1 and σ_2 are the standard deviation of T_{2i-1} and T_{2i}, respectively. According to the joint probability distribution in Table. 2, we have

$$\begin{aligned} \mathrm{E}[T_{2i-1}T_{2i}] &= \Pr[T_{2i-1}T_{2i} = 1] \cdot 1 + \Pr[T_{2i-1}T_{2i} = 0] \cdot 0 \\ &= \Pr[T_{2i-1} = 1, T_{2i} = 1] \cdot 1 \\ &= \rho pq + p^2 \ . \end{aligned} \tag{14}$$

We can also compute

$$\mathrm{E}[T_j] = \Pr[T_j = 1] \cdot 1 + \Pr[T_j = 0] \cdot 0 = p \tag{15}$$

and

$$\begin{aligned} \sigma_j^2 &= \mathrm{E}[T_j^2] - \mathrm{E}^2[T_j] \\ &= \Pr[T_j^2 = 1] \cdot 1 + \Pr[T_j^2 = 0] \cdot 0 \\ &\quad - (\Pr[T_j = 1] \cdot 1 + \Pr[T_j = 0] \cdot 0)^2 \\ &= \Pr[T_j^2 = 1] - \Pr[T_j = 1]^2 \\ &= \Pr[T_j = 1] - \Pr[T_j = 1]^2 \\ &= p - p^2 \end{aligned} \tag{16}$$

for all $1 \le j \le n$.

Substituting Eqs. (14) to (16) into Eq. (13), we get the correlation coefficient between T_{2i-1} and T_{2i} as

$$\begin{aligned} \frac{\mathrm{Cov}[T_{2i-1}, T_{2i}]}{\sigma_1 \sigma_2} &= \frac{\rho pq + p^2 - p \cdot p}{p - p^2} \\ &= \rho \ . \end{aligned} \tag{17}$$

Note that ρ must not be smaller than $1 - 1/p$ to ensure that every probability value in Table 2 is non-negative.

Third, the classical RR mechanism is a special case of the general JRR mechanism. In particular, when $\rho = 0$, T_{2i-1} and T_{2i} are independent, and the JRR mechanism is equivalent to the case of each two in each perturbs his/her data value using the RR independently.

Last but not least, the data collector can estimate n_v using the same estimator as in the classical RR. Specifically, assume that the data collector receives I_v values of v for each $v \in D$, it can estimate the number of contributors having true value v as

$$\hat{n}_v = \frac{I_v - nq}{p - q} \ , \tag{18}$$

Theorem 1. *The estimator in Eq.(18) is unbiased.*

We give the proof in Appendix A.

4.2 Data Privacy Analysis

Assume that the data collector wants to infer a target contributor u_i's value x_i. Let C be the set of contributors who collude with the data collector so that the data collector knows whether each contributor in C reports truthfully. In other words, besides all perturbed values $y_1, \cdots, y_n$, the data collector also knows whether each colluding contributor reports truthfully or not, which we denote by $\mathcal{T}_c = \{T_j | j \in C\}$. Under the uniformly random grouping of JRR, the group peer of u_i, say u_j, may be a colluding contributor in C. If this happens, even if the data contributor does not know who the group peer of contributor u_i is, the correlation between the two contributors' perturbation would still inevitably reduce the data privacy of the targeted contributor u_i.

We therefore extend the definition of LDP to measure the individual privacy provided by the JRR scheme. Specifically, the following theorem shows that the JRR scheme can offer individual contributors a form of data privacy similar to ε-LDP.

Theorem 2. *Assume that there is a set of contributors C whose reporting truthfulness $\mathcal{T}_c$ is known to the adversary. For any contributor u_i, the JRR mechanism $\mathcal{M}$ satisfies*

$$\frac{\Pr[\mathcal{M}(x_i) = y_i|\mathcal{T}_c]}{\Pr[\mathcal{M}(x_i') = y_i|\mathcal{T}_c]} \le e^{\varepsilon} \tag{19}$$

for any pair of inputs $x_i, x_i' \in D$ and any output $y_i \in \mathrm{Range}(\mathcal{M})$, where

$$\varepsilon = \ln \frac{mp_{\max} + (n-m-1)p}{mp_{\min} + (n-m-1)q}, \tag{20}$$

with $p_{\max} = \max\{(1-\rho)p, p+\rho q\}$, $p_{\min} = \min\{(1-\rho)q, q+\rho p\}$, $m = |\mathcal{T}_c|$, and $0 \le m \le n-1$.

The detailed proof is provided in Appendix B.

4.3 Data Utility Analysis

Theorem 3. *Assume that n contributors are divided into $n/2$ groups uniformly at random. The variance of the estimator $\hat{n}_v$ given in Eq. (18) by JRR is*

$$\mathrm{Var}[\hat{n}_v] = \frac{pq}{(p-q)^2} \cdot \left(n + \frac{\rho\left((2n_1-n)^2 - n\right)}{n-1}\right), \tag{21}$$

where n_v is the number of contributors with a value of v.

The proof is given in Appendix C.

We can see that $\mathrm{Var}[\hat{n}_v] = npq/(p-q)^2$ when $\rho = 0$, which is the same as that of RR. Moreover, $\mathrm{Var}[\hat{n}_v] < npq/(p-q)^2$, i.e., smaller than that of RR, if $\rho((2n_1-n)^2-n) < 0$, which provides opportunities to achieve higher data utility than RR.

4.4 Choice of p and ρ

In this subsection, we show how to choose parameters p and ρ to achieve high data utility at the data collector under a given data privacy requirement.

4.4.1 An Optimization Problem Formulation. Assume that we want to provide the same level of privacy guarantee with an RR scheme that satisfies ε-LDP. We need to choose parameters p and ρ that satisfies Ineq. (19) in Theorem 2. One challenge is that Ineq. (19) involves the parameter m, i.e., the number of contributors colluding with the data collector, which is often unknown in practice. Fortunately, we find that for any pair of (p, ρ), given the privacy budget ε, if m_1 satisfies Ineq. (19), so does m_2 for all $m_2 \le m_1$, because $f(m) = \frac{mp_{\max}+(n-m-1)p}{mp_{\min}+(n-m-1)q}$ is monotonically increasing with respect to m. The detailed proof is given in Appendix H. Assume that there could be at most M data contributors colluding with the data collector. We can then replace m in Ineq. (19) by M when choosing p and ρ.

Let $h(p,\rho) = \mathrm{Var}[\hat{n}_1]$ as given in Eq. (21). We can formulate the choice of p and ρ as the following optimization problem, which seeks to minimize the objective function $h(p,\rho)$ while satisfying the privacy constraint and the domain constraint of ρ and p.

$$\begin{aligned} \min \quad & h(p,\rho) \\ \text{s.t.} \quad & \frac{Mp_{\max} + p(n-M-1)}{Mp_{\min} + q(n-M-1)} \le e^{\varepsilon}, \\ & 1 - 1/p \le \rho \le 1, \\ & 0.5 < p \le 1\ , \end{aligned} \tag{22}$$

where $p_{\max} = \max\{(1-\rho)p, p+\rho q\}$ and $p_{\min} = \min\{(1-\rho)q, q+\rho p\}$.

Unfortunately, since the objective function $h(p,\rho)$ involves n_1 that we want to estimate, the above optimization problem cannot be directly solved without knowing n_1 in advance. However, certain properties of the objective function $h(p,\rho)$ and the constraints make it possible to design an effective heuristic to choose p and ρ that can yield good performance in the majority of the cases.

Specifically, we have the following three lemmas, with proofs in Appendix D, Appendix E, and Appendix F, respectively.

Lemma 1.

$$(2n_1-n)^2 - n \begin{cases} \ge 0 & \text{\textit{if}}\ \frac{n_1}{n} \in [0, \frac{1}{2} - \frac{1}{2\sqrt{n}}] \bigcup [(\frac{1}{2} + \frac{1}{2\sqrt{n}}, 1), \\ < 0 & \text{\textit{if}}\ \frac{n_1}{n} \in [\frac{1}{2} - \frac{1}{2\sqrt{n}}, \frac{1}{2} + \frac{1}{2\sqrt{n}}]\ . \end{cases} \tag{23}$$

Lemma 2. *For any $n \ge 2, \rho \in [-1,1]$ and $0 \le n_1 \le n$,*

$$n + \frac{\rho((2n_1-n)^2-n)}{n-1} > 0\ . \tag{24}$$

Lemma 3. *$pq/(p-q)^2$ is monotonically decreasing with respect to $p \in (0.5, 1]$.*

We then have the following theorem regarding the monotonicity of the objective function $h(p,\rho)$.

Theorem 4. *The objective function $h(p,\rho)$ is*

- *monotonically increasing with respect to ρ if $n_1/n \le 1/2 - 1/2\sqrt{n}$ or $n_1/n \ge 1/2 + 1/2\sqrt{n}$ and monotonically decreasing with respect to ρ if $1/2 - 1/2\sqrt{n} < n_1/n < 1/2 + 1/2\sqrt{n}$,*
- *monotonically decreasing with respect to $p \in (0.5, 1]$.*

The proof uses the results from Lemmas 1 to 3, which is quite straightforward and given in Appendix G.

4.4.2 A Heuristic Algorithm for Selecting ρ and p. We now introduce a heuristic to choose ρ and p that can yield good performance in most cases by exploiting the monotonicity of $h(p,\rho)$. First, we assume that the data collector would collude with at most $m = M$ contributors, where M is a system parameter indicating the data collector's inference ability.

According to Theorem 4, $h(p,\rho)$ is monotonically increasing with respect to ρ if $n_1/n \in [0, 1/2 - 1/2\sqrt{n}] \bigcup [1/2 + 1/2\sqrt{n}, 1]$ and monotonically decreasing with respect to ρ if $n_1/n \in [1/2 - 1/2\sqrt{n}, 1/2 + 1/2\sqrt{n}]$. We notice that the size of the range $[1/2 - 1/2\sqrt{n}, 1/2 + 1/2\sqrt{n}]$ is $1/\sqrt{n}$, which is relatively small for large n and inversely proportional to $\sqrt{n}$. For example, when $n = 100$ and $10,000$, the ranges in which $h(p,\rho)$ is monotonically increasing with respect to ρ are $[0.45, 0.55]$ and $[0.495, 0.505]$, respectively. It is therefore reasonable to assume that $h(p,\rho)$ is monotonically increasing with respect to ρ when n is large in practice.

Algorithm 1: Search for (ρ, p)

```
   input  :$n$, $\varepsilon$, $M$, $\triangle\rho$, and $\triangle p$
   output :$\rho$ and $p$
1  $p \leftarrow \frac{e^{\varepsilon}}{1+e^{\varepsilon}} - \triangle p$;
2  while $p > 0.5$ do
3      $\rho \leftarrow 1 - \frac{1}{p}$;
4      while $\rho \leq 1$ do
5          if $(p, \rho)$ satisfies all constraints in Eq. (22) then
6              return $p$ and $\rho$;
7          end
8          $\rho \leftarrow \rho + \triangle\rho$;
9      end
10     $p \leftarrow p - \triangle p$ ;
11 end
12 return Null;
```

Assume that $n_1/n \notin [1/2 - 1/2\sqrt{n}, 1/2 + 1/2\sqrt{n}]$. The objective function $h(p, \rho)$ is then monotonically increasing with respect to ρ and monotonically decreasing with respect to p according to Theorem 4. Let $R \in \mathbb{R}^2$ be the feasible domain of (p, ρ). We can define a partial ordering relationship among different pairs of (p, ρ)s. In particular, for any two pairs of (p, ρ) and (p', ρ'), we have $h(p, \rho) \leq h(p', \rho')$ if $p \geq p'$ and $\rho \leq \rho'$. Moreover, let $\text{range}(p)$ be the feasible range of ρ for a given p. The optimal choice of the parameters must be $(p, \min(\text{range}(p)))$ for some p.

Based on the above observation, our key idea is to first find a maximal feasible p and then find the corresponding minimal feasible ρ. Let $\triangle p$ and $\triangle\rho$ be two small constants, e.g., 0.0001. Algorithm 1 provides the detailed procedure. Specifically, we initialize p to $\frac{e^{\varepsilon}}{1+e^{\varepsilon}} - \triangle p$ (Line 1). We intentionally skip $p = \frac{e^{\varepsilon}}{1+e^{\varepsilon}}$ because if $p = \frac{e^{\varepsilon}}{1+e^{\varepsilon}}$ then ρ would be zero and JRR would be equivalent to RR. We then search for a feasible pair of p and ρ using two nested loops. For each given p, we initialize ρ to $1 - \frac{1}{p}$ (Line 3) and then iteratively check whether the current (p, ρ) satisfy all the constraints in Eq. 22 (Lines 3-9). If so, we output (p, ρ). Otherwise, we increase ρ by $\triangle\rho$ until $\rho = 1$, in which case a new iteration starts. Algorithm 1 returns "Null" for completeness. In practice, Algorithm 1 always returns a feasible pair (ρ, p) in the worst case. The reason is that when $\rho = 0$, we have $p_{\max} = p$ and $p_{\min} = q$, so the first constraint is simplified to $\frac{p}{q} \leq e^{\varepsilon}$. In this case, any $p \leq \frac{e^{\varepsilon}}{1+e^{\varepsilon}}$ with $\rho = 0$ is always a feasible pair. The search space for p and ρ is $[0.5, \frac{e^{\epsilon}}{1+e^{\epsilon}})$ and $[1 - \frac{1}{p}, 1]$, respectively, and the computational complexity of Algorithm 1 is proportional to the area of search space $\frac{1}{p} \cdot (\frac{e^{\varepsilon}}{1+e^{\varepsilon}} - 0.5)$ and inversely proportional to both Δp and $\Delta\rho$.

4.5 Practical Instantiations of JRR

In this subsection, we present two practical instantiations of the JRR mechanism for completeness. The first instantiation is highly efficient but requires an auxiliary non-colluding server to facilitate random grouping and joint random perturbation, as described in Table 2. The second instantiation, in contrast, leverages multi-party computation techniques (MPC) to eliminate the need for a non-colluding server but comes at the cost of increased computational and communication overhead. While our proposed instantiations provide practical implementations, they are not necessarily optimal. There remains significant potential for further refinement and efficiency improvements, which we leave as future work.

4.5.1 A Non-colluding Server-Based Instantiation. Recall that the JRR mechanism relies on the two key assumptions: (1) all n contributors are divided into $n/2$ groups uniformly at random, and (2) the data collector is unaware of the group membership. Our first instantiation leverages an auxiliary non-colluding server to satisfy these requirements. Notably, similar non-colluding servers have been employed in recent works such as [6, 8, 12, 19]. Moreover, various approaches have been proposed to enforce non-collusion, as discussed in [9, 17, 28].

Random grouping. The auxiliary server divide n contributors into $n/2$ disjoint groups of two uniformly at random. Without loss of generality, assume that the i-th group G_i consists of contributors u_{2i-1} and u_{2i} for all $1 \leq i \leq \frac{n}{2}$. For each group G_i, the server generates $R_{2i-1} \in \{1, -1\}$ uniformly at random and computes $R_{2i} = -R_{2i-1}$. It then sends R_{2i-1} to u_{2i-1} and R_{2i} to u_{2i}.

The grouping information is kept by the auxiliary server, which is unknown to both the collector and individual contributors.

Correlated perturbation in each group. Each group G_i of contributors then perform correlated perturbation with the help of R_{2i-1} and R_{2i} received from the auxiliary server. Specifically, each contributor u_j generates a random variable C_j independently according to the following probability distribution

$$C_j = \begin{cases} 1.5 & \text{with probability } p - \sqrt{-\rho pq}, \\ 0.5 & \text{with probability } \sqrt{-\rho pq}, \\ -0.5 & \text{with probability } \sqrt{-\rho pq}, \\ -1.5 & \text{with probability } q - \sqrt{-\rho pq}, \end{cases} \tag{25}$$

for $j = 2i - 1$ and $2i$, where p, q and ρ are given in Table 2.

Finally, each contributor u_j determines whether to report truthfully according to the following rule

$$T_j = \begin{cases} 1 & \text{if } C_j + R_j > 0, \\ 0 & \text{if } C_j + R_j < 0, \end{cases} \tag{26}$$

for all $j = 1, \ldots, n$.

Theorem 5. *Under the practical mechanism, for any two contributors in the same group, the joint probability distribution of the truthfulness of the two contributors' reports is equivalent to the one described in Table. 2.*

We provide the detailed proof in Appendix I.

This practical scheme can guarantee the data privacy of each individual contributor against the auxiliary server, the data collector, and any other contributor in the system. First, while the auxiliary server knows the which contributor receives −1 or 1 and the membership information, it has no access to perturbed value submitted by individual contributors. Similarly, the data collector receives the perturbed values from contributors but is unaware of the group membership. Even in the worst case where all contributors but one victim contributor say u_j collude with the data collector, the data collector can only infer the random variable R_j that u_j received from the auxiliary server but cannot recover T_j because C_j is unknown to the data collector. While this can lead to reduced data privacy guarantee for u_j, such cases of collusion are extremely impractical in a system with a large number of contributors. Last but not the least, each contributor knows only whether the other group member receives -1 or 1 from the auxiliary server and is unaware of the identity of or the perturbed value submitted by the other group member.

4.5.2 An MPC-Based Instantiation. Now we introduce another practical instantiation that utilizes secure multi-party shuffling (SMPS) [34] to eliminate the need for a non-colluding server. SMPS is a secure multiparty computation protocol that allows a group of contributors to agree on a random permutation of their individual inputs while keeping the inputs private. After permutation, each contributor receives only one of the shuffled inputs, and no one can determine the complete mapping between the original and shuffled inputs. Our second instantiation uses SMPS to securely shuffle contributors IDs (i.e., original positions) at random and achieving random grouping and group membership private according to the shuffled IDs. The detailed steps are as follows.

1. Upon joining the system, each contributor is assigned a unique ID $j \in \{1, 2, \ldots, n\}$ from the data collector.
2. Assuming an even number n of participating contributors in a specific round, all contributors perform SMPS (Algorithm 1 in [34]) to securely shuffle their IDs, i.e., a random permutation $\pi : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$. After the permutation, each contributor j gets a shuffled ID $\pi(j)$ without learning any other contributor's shuffled ID.
3. Let $\pi^{-1}(\cdot)$ be the inverse permutation of $\pi(\cdot)$. Every two contributors with adjacent shuffled IDs form a group without knowing each other's original ID, i.e., the kth group consists of contributors $\pi^{-1}(2k-1)$ and $\pi^{-1}(2k)$ for each $k \in \{1, \ldots, n/2\}$
4. Each contributor with shuffle ID $\pi^{-1}(j)$ sets $R_j = 1$ if $\pi^{-1}(j)$ is even and -1 if $\pi^{-1}(j)$ is odd.
5. Each contributor with shuffle ID $\pi^{-1}(j)$ generates C_j according to Eq. (25), computes T_j according to Eq. (26). Finally, each contributor randomly perturbs their data according to T_j for submission.

This scheme ensures random grouping due to the randomness introduced by SMPS. Moreover, Step 3 guarantees that in every group k, consisting of data contributors $\pi^{-1}(2k-1)$ and $\pi^{-1}(2k)$, one contributor has $R_{2k-1} = 1$ while the other has $R_{2k} = -1$, effectively replicating the role of the trusted auxiliary server in the first scheme. Additionally, this approach ensures individual data privacy against both the data collector and other participants. Since each contributor j only knows his shuffled position $\pi(j)$ without knowing the other group member, group membership privacy is maintained. This setup allows each contributor to randomly decide whether to report truthfully, enhancing privacy while maintaining the functionality of the first scheme. On the other hand, this scheme requires all contributors to be online and participate in SMPS, leading to higher computation and communication overheads. In particular, SMPS has a computation and communication overheads of $O(n \log n)$, which is much higher than the first scheme.

It is also worth noting that SMPS and data submission can be conducted at different times. We anticipate that the JRR scheme will be implemented as a mobile app, while SMPS will run as a background service, executing periodically without contributor involvement. Upon successfully participating in SMPS, contributors can submit their data via JRR at any time as needed. Newly registered contributors who have not yet completed any SMPS procedure can still submit data using classical RR. Similarly, if a contributor's group peer fails to submit their value, it effectively reduces to that contributor submitting his value under RR.

4.6 Discussions

In this subsection, we examine several key issues related to JRR and its practical instantiations.

4.6.1 Extension to Non-Binary Data. It is possible to extend JRR to support non-binary data by redesigning the joint reporting probability in Table 2. In particular, for a group of two contributors with data $v_1, v_2 \in [k]$, where $k \geq 2$ is the domain size, they report their data according to the following joint probability distribution.

$$\Pr[(v_1', v_2')] = \begin{cases} p^2 + \rho pq & \text{if } v_1' = v_1, v_2' = v_2, \\ pq - \frac{1}{k-1}\rho pq & \text{if } v_1' = v_1, v_2' \neq v_2, \\ pq - \frac{1}{k-1}\rho pq & \text{if } v_1' \neq v_1, v_2' = v_2, \\ q^2 + \frac{1}{(k-1)^2}\rho pq & \text{if } v_1' \neq v_1, v_2' \neq v_2. \end{cases} \tag{27}$$

where $p + (k-1)q = 1$ and ρ is the correlation coefficient between T_1 and T_2. This is similar to the extension from RR to Generalized RR (GRR).

It is easy prove that the marginal probability distribution of each contributor reporting his value is the same as the binary case, i.e. k-JRR maps each value v to itself with probability p and to any other value with probability q. As a result, $\hat{n}_v = (I_v - nq)/(p-q)$ is also an unbiased estimator for n_v, where I_v is the number of contributors reporting v. We prove this property in Appendix K.

The data privacy and utility can be analyzed in the same way as the binary case. The remaining task is to choose p and ρ by solving an optimization problem similar to the one in Eq. (22), which we leave as our future work.

4.6.2 Integration with Advanced LDP Mechanisms. JRR can be integrated with advanced LDP mechanisms built upon RR to improve their utility and privacy tradeoff. Here we present its integration with Optimized Unary Encoding (OUE) [49] and Optimized Local Hashing (OLH) [49] as two examples.

Integration with OUE: OUE is an LDP mechanism that encodes a data value in a k-size domain into k-bit binary vector [49]. OUE follows a three-step encoding-perturbation-aggregation procedure: (i) Encoding: given an original value $x \in [k]$, OUE first encodes it into $\mathsf{Encode}(x) = \{0, \cdots, 0, 1, 0, \cdots, 0\}$, where the x-th bit is 1 and the rest are 0. (ii) Perturbation: given the k-bit vector $B = \mathsf{Encode}(x)$, OUE then generates a perturbed vector B' using RR according to the following probability distribution

$$\Pr[B'[j] = 1] = \begin{cases} p, & \text{if } B[j] = 1, \\ q, & \text{if } B[j] = 0, \end{cases}$$

where $B[j]$ denotes the j-th bit of B. (iii) Aggregation: the data collector counts the number of 1 in B' from all data contributors whereby to estimate the frequency of each value as in RR.

JRR can be seamlessly integrated into OUE by modifying the perturbation step. Specifically, instead of applying independent RR, two contributors jointly perturb each bit of their encoded vectors using JRR. This collaborative approach enhances the estimation accuracy of OUE while preserving the original privacy guarantees.

Integration with OLH: OLH follows a similar encoding-perturbation-aggregation procedure: (i) Each contributor randomly picks one hash function $H(\cdot)$ among a universal hash family to map his value v in a s-size domain into a much smaller domain of size k, i.e., $x = H(v)$. (ii) Each contributor then perturbs the encoded value x into y using GRR according to the following probability distribution

$$\Pr[y = i] = \begin{cases} p, & \text{if } i = x, \\ q, & \text{for each } i \in \{1, \ldots, k\} \setminus \{x\}, \end{cases}$$

(iii) The data collector then counts, for each $i \in [k]$, the number of reports equal to i, to estimate the frequency of i. This frequency also serves as an estimated frequency of original values that are mapped to i (i.e., for which $i = H(v)$).

We can easily integrate JRR with OLH by replacing the perturbation procedure in Step (ii) with the extended JRR for non-binary data introduced in the previous subsection, i.e., k-JRR given in Eq. 27, to enhance the estimation accuracy of OLH while preserving its original privacy guarantees.

4.6.3 Impact of Large Domain Size. Like other existing LDP mechanisms, the data utility of JRR inherently declines as the data domain size increases. This is because any LDP mechanism must allocate probability mass across all possible data values to satisfy ϵ-LDP, reducing the probability that each data contributor reports their true value as the domain expands. While the extended JRR method introduced in Section 4.6.1 i.e., k-JRR, consistently outperforms GRR (a special case of k-JRR with $\rho = 0$) by achieving a better utility-privacy trade-off through tuning the parameter ρ, its advantage over GRR diminishes as the domain size grows.

However, since JRR is designed to replace the RR or GRR component in other LDP mechanisms—many of which incorporate domain reduction techniques to mitigate the impact of large domain sizes—JRR can effectively handle large domains as long as the underlying LDP mechanisms can before being enhanced by JRR. This capability is demonstrated in the integration of JRR with OLH discussed earlier.

4.6.4 Extension to Larger-size Group. While this paper focuses on two-contributor groups, JRR can theoretically be extended to larger groups with more than two contributors. For a group of $k > 2$ contributors, their joint probability distribution of truthful reporting can be represented by a k-dimensional table. Intuitively, increasing the group size could enhance JRR's privacy-utility trade-off compared to two-contributor groups by allowing for greater correlation, which can help mitigate the added noise. As the group size continues to increase, the probability of a group including colluding contributors grows significantly, leading to additional privacy leakage and limiting further improvements in the utility-privacy trade-off.

However, designing such a scheme becomes increasingly complex as the group size grows. Consider a group of three contributors as an example. To ensure that the data collector cannot infer additional information beyond standard RR by analyzing each contributor's reported value in isolation, the marginal probability of truthful reporting for each contributor must remain consistent with RR and JRR. However, fully defining their joint probability distribution requires specifying three pairwise correlation coefficients ($\rho_{12}, \rho_{13}, \rho_{23}$) and a triple correlation coefficient (ρ_{123}) to capture higher-order dependencies. As the group size increases, the number of required correlation coefficients grows exponentially. Even if we leverage symmetry to reduce the number of independent correlation parameters to one less than the group size, selecting appropriate values while maintaining privacy guarantees remains a challenging problem. Therefore, we leave the extension of JRR to larger group sizes as future work.

Table 3: Summary of Datasets.

Dataset	Total (n)	# of "1" (n_1)	Pct. of "1" (n_1/n)
Kosarak	2×10^4	659	0.033
Amazon	1×10^4	762	0.076
E-commerce	23, 486	19, 314	0.822
Census	1×10^4	9, 528	0.953
Synthetic	$20 \sim 2 \times 10^6$	$0 \sim 2 \times 10^6$	$0 \sim 1.0$

4.6.5 Relationship with the Shuffle Model. We would like to clarify the relationship between JRR and the Shuffle Model [6, 8, 12, 19, 22, 31] which also uses a trusted server to improve the utility-privacy trade-off. In the shuffle model, data contributors perturb their data using an LDP mechanism and send the perturbed data to a trusted shuffler, which shuffles all the received data before forwarding them to the data collector. It has been shown that randomly shuffling the data can improve data privacy without sacrificing any data utility.

We stress that JRR is not intended to replace the shuffle model. Instead, they can be easily integrated to further improve data privacy. Specifically, each contributor can first perturbs their data via JRR and then send them to a shuffler, which in turn shuffles all the received data values before forwarding them to the data collector. The data collector can estimate n_v using the same estimator as JRR.

5 Performance Evaluation

This section thoroughly evaluates the performance of the proposed JRR mechanism using both real and synthetic datasets.

5.1 Datasets and Simulation Setting

We use four real-world datasets, Kosarak [1], Amazon Rating [4], E-commerce [2], Census [42], for performance evaluation. Detailed descriptions of them are shown in Appendix J. In addition to these four real datasets, we also generate synthetic datasets with n varying from 20 to 2×10^6 and n_1/n varying from 0 to 1. Table 3 summarizes these datasets.

We compare the proposed JRR mechanism with the RR mechanism because RR is not only the most classical LDP protocol for frequency estimation but also a special case of JRR. We do not compare JRR with the shuffle model because a fair comparison between them is challenging for two reasons. First, the privacy guarantee offered by the shuffle model is derived based on the (ε, δ)-DP definition with $\delta \neq 0$ [6, 12, 19, 31], whereas JRR provides ε-LDP, i.e., $\delta = 0$. It is therefore difficult to compare their estimation accuracy under the same data privacy guarantee. Second, the

Table 4: Default Simulation Setting

Parameter	Value	Description
n	10,000	# of participated contributors
n_1/n	0.1	Ratio of contributors with value 1
ε	0.1	Privacy budget
M	5	# of colluding contributors
$\triangle p$	0.0001	Search granularity
$\triangle \rho$	0.0001	Search granularity

shuffle model measures the lower bound of estimation error using the (α, β)-accuracy notion [22], which is weaker than the standard mean square error we use to measure the estimation error of JRR. Additionally, we do not compare JRR with other more advanced LDP mechanisms, as they rely on RR as a building block and target different data types, such as OLH [50] for categorical data and PCKV [23] for key-value data. As discussed in Section 4.6.2, JRR has the potential to replace RR in these schemes and improve their privacy-utility trade-off.

Data utility comparisons are performed at the same privacy level ε. For RR, utility is maximized by setting $p = e^{\varepsilon}/(1 + e^{\varepsilon})$. For JRR, we employ the heuristic solution of p and ρ outlined in Algorithm 1. Notably, to ensure a fair comparison at the same privacy level ε, the parameter p in JRR differs from that in RR. We use the following two metrics to evaluate the performance of JRR.

- *Mean squared error (MSE)* [24, 49]: it is the mean squared errors of the estimated $\hat{n}_v$ with respect to the real one n_v across all values, which is defined as

$$MSE = \frac{1}{|D|} \sum_{v \in D} (\hat{n}_v - n_v)^2 . \quad (28)$$

- *Averaged relative error (ARE)* [29, 60]: it is the mean relative error across all values that is defined as

$$ARE = \frac{1}{|D|} \sum_{v \in D} \frac{|\hat{n}_v - n_v|}{n_v} . \quad (29)$$

In the above formula, $|D| = 2$ is the size of $D = \{0, 1\}$. Note that as shown in Eq. (3) the MSE of RR is not affected by n_v, whereas its ARE is. Consequently, RR's ARE performance may exhibit multiple distinct lines for different n_v, whereas its MSE remains a single line.

Table 4 lists the default simulation settings. Using MATLAB, each point in the figures represents the average of 1000 runs with unique random seeds.

5.2 Results From Real-world Datasets

Fig. 1 presents the MSE of JRR and RR on the four real-world datasets when ε is 0.01, 0.1, and 1, respectively. We can see that the JRR achieves a lower MSE than RR for all four datasets under all three εs. This is expected because the negative correlation between two contributors' random perturbations under JRR can effectively reduce the expected MSE when the ratio n_1/n is not close to 0.5, which is true for all four real datasets with n_1/n being either smaller than 0.1 or larger than 0.8. Moreover, we can see that the larger the ε, the smaller the MSE under both JRR and RR. This is also anticipated as the larger the privacy budget ε, the more likely that each contributor reports truthfully, the smaller the MSE under both mechanisms and vice versa. In addition, we can see that JRR outperforms RR by a larger margin on Kosarak and Census datasets in comparison with the Amazon Rating and E-commerce datasets, especially when ε is small, e.g., $\varepsilon = 0.01$. This is because the ratio, n_1/n, in the Kosarak and Census datasets are farther away from 0.5 than those in the other two datasets. We will carefully evaluate the impact of ε, n, and n_1/n on the MSE using the synthetic datasets shortly.

Figs. 2 show the distributions of ARE over the 1,000 runs under JRR and RR on the four real datasets with $\varepsilon = 0.01$ and 1, respectively. A percentile indicates the percentage of error values that are lower than the corresponding ARE. We can see that for any specific percentile, JRR consistently outperforms RR with a lower ARE across all four datasets. For example, as shown in Fig. 2(b), when $\varepsilon = 0.1$, the 80th percentile under JRR on the Kosarak dataset is 0.7, i.e., 800 out of 1,000 runs have ARE lower than 0.7. In contrast, the 80th percentile under RR is 1.45. These results confirm that JRR consistently offers stable performance with a lower ARE.

5.3 Results From Synthetic Datasets

5.3.1 *Impact of ε.*

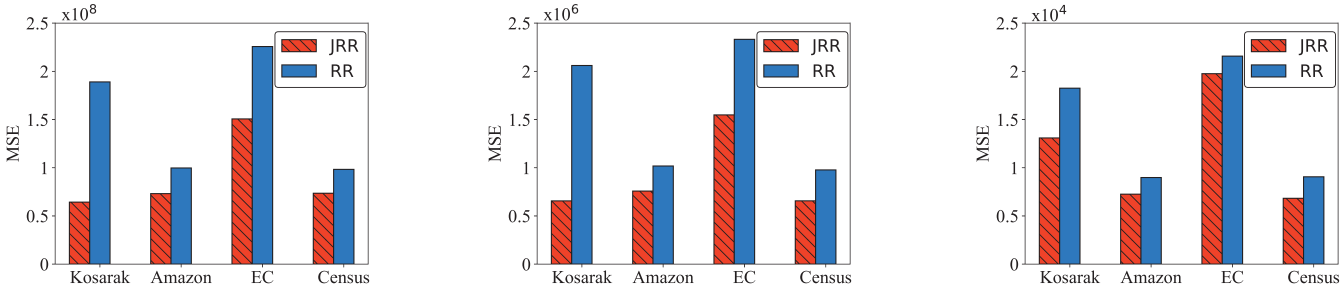
Figs. 3(a) to 3(c) illustrates the MSEs of RR and JRR with privacy budgets ε varying from 0.01 to 1 under different number of data contributors. As expected, MSE decreases for both RR and JRR as ε increases because higher ε increases the probability of reporting truthfully. Moreover, JRR consistently achieves a smaller MSE than RR, with the performance gap widening as n increases. For example, as shown in Fig. 3(a), when $\varepsilon = 0.01$, the MSE under JRR is 5.8%, 40.9%, and 1.0% of that under RR when n_1/n is 0.01, 0.1, and 1, respectively. In contrast, when $n = 8 \times 10^4$ (see Fig. 3(c)), the corresponding MSE is 4.5%, 37.2%, and 0.9% of that under RR. This trend occurs because a larger n reduces the likelihood of the data collector correctly identifying group members, limiting the additional information inferred from correlated reporting. The advantage of JRR over RR becomes even more pronounced as ε decreases. In Fig. 3(c), for $\varepsilon = 0.1$, the MSE under JRR is 86.6%, 55.8%, and 90.9% smaller than RR's when n_1/n is 0.01, 0.1, and 1, respectively. In contrast, when $\varepsilon = 0.01$, the corresponding MSE is 95.4%, 62.8%, and 99.0% smaller than that under RR, respectively. When $\varepsilon = 0.1$ and $n_1/n = 1$, the MSE under JRR is 7.3×10^5, which is about 74.1% of the one under RR. In contrast, when $\varepsilon = 0.01$, the MSE under JRR $n_1/n = 1$ is 8.6×10^6, which is only about 1.0% of the one under RR. These results demonstrate that JRR outperforms RR with a large margin, especially when n is large and ε is small.

Figs. 3(d) to 3(f) compare the AREs under RR and JRR with ε varying from 0.01 to 1. We can observe similar trends to Figs. 3(a) to 3(c) that the AREs under both RR and JRR decreases as ε increases. Moreover, we can see that a larger n_1/n leads to a smaller ARE. JRR achieves a smaller ARE than RR in all the cases, and the margin by which JRR outperforms RR increases as ε decreases due to the same reasons that we mentioned earlier.

5.3.2 *Impact of n.*

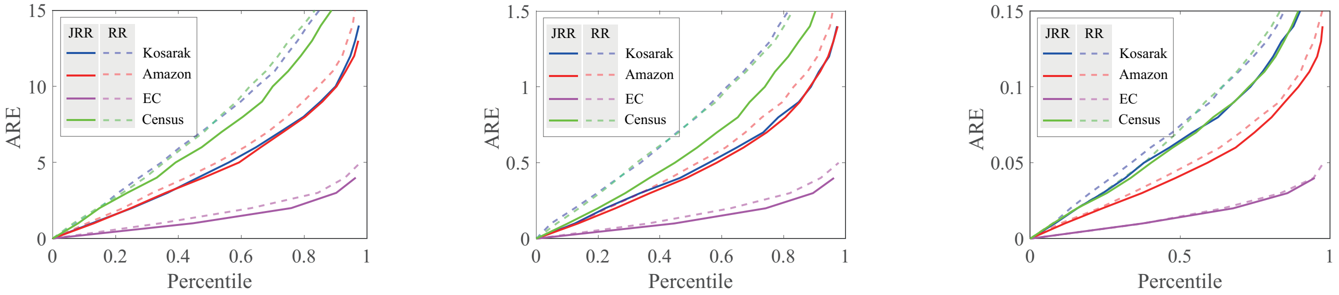
Figs. 4 compares the MSE and ARE under JRR and RR with the n varying from 20 to 2, 000, 000 for ε= 0.01, 0.1, and 1, respectively.

Figs. 4(a) to 4(c) show that the MSE under RR increases linearly as n increases, which is expected. In contrast, the MSE under JRR initially increases linearly as n increases from 20 to 2,000, then increases at a slower rate or even decreases as n increases from



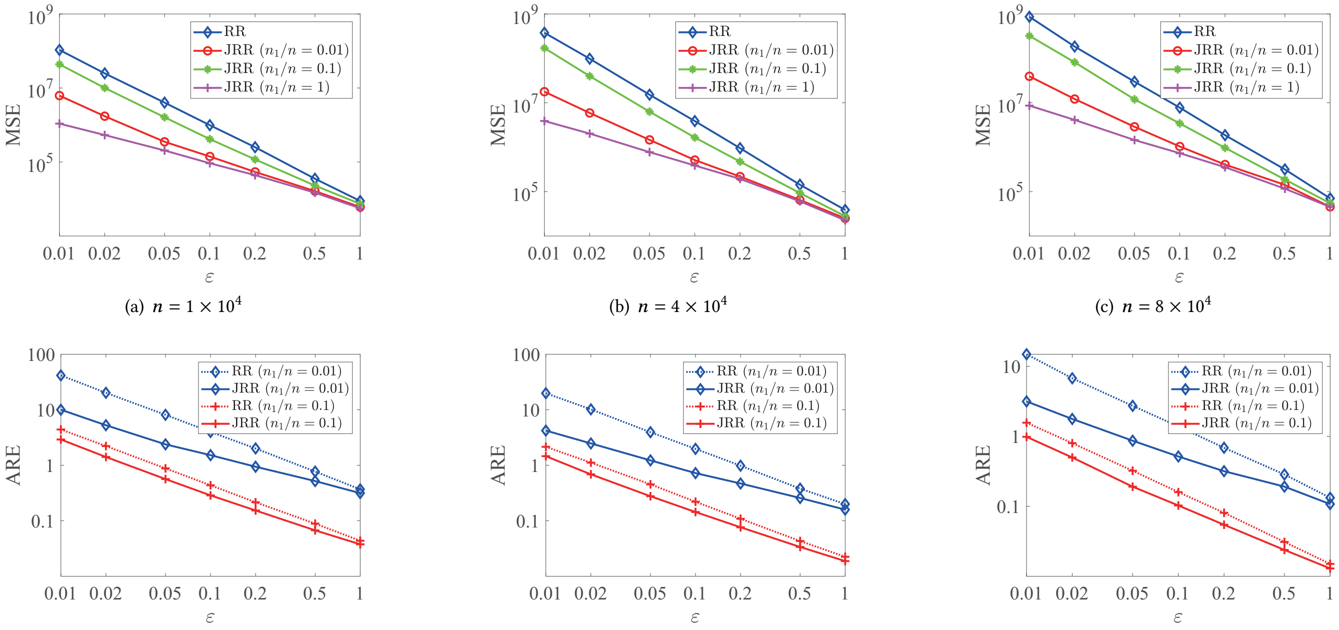
(a) $\varepsilon = 0.01$ (b) $\varepsilon = 0.1$ (c) $\varepsilon = 1$

Figure 1: Comparison of MSE under RR and JRR on four real datasets when the privacy budget $\varepsilon = 0.01, 0.1$ and 1.



(a) $\varepsilon = 0.01$ (b) $\varepsilon = 0.1$ (c) $\varepsilon = 1$

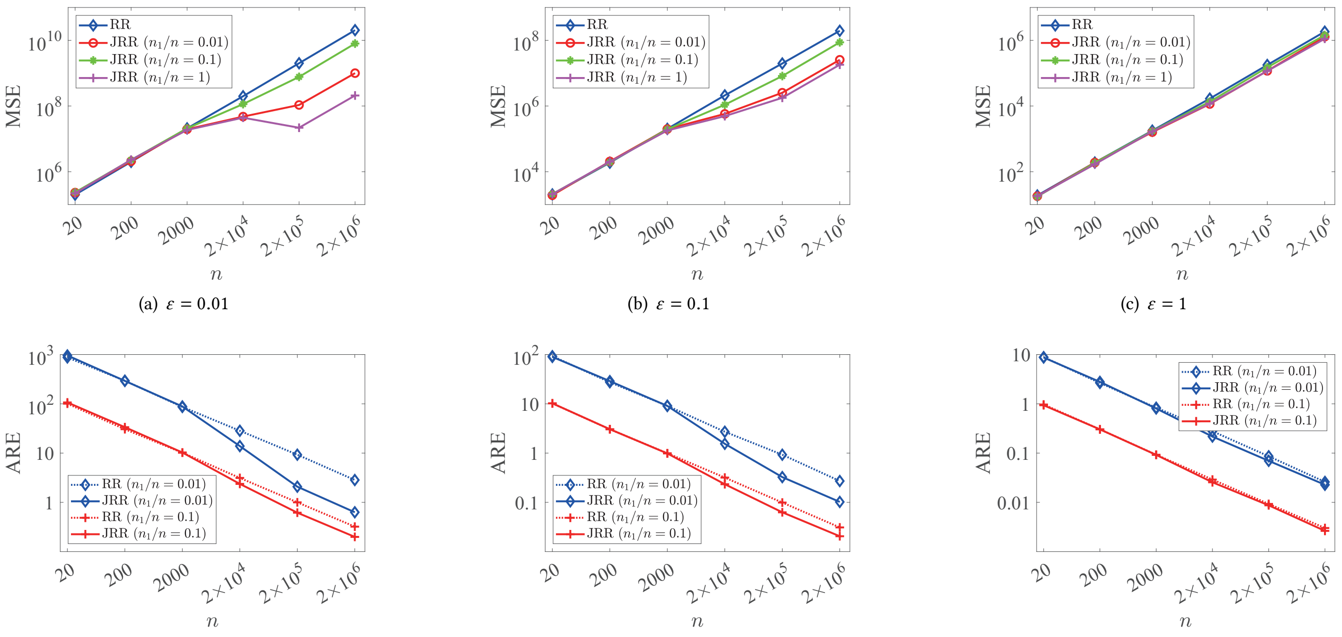
Figure 2: Percentiles of ARE under RR and JRR on four real datasets when the privacy budget $\varepsilon = 0.01, 0.1$ and 1.



(a) $n = 1 \times 10^4$ (b) $n = 4 \times 10^4$ (c) $n = 8 \times 10^4$

(d) $n = 1 \times 10^4$ (e) $n = 4 \times 10^4$ (f) $n = 8 \times 10^4$

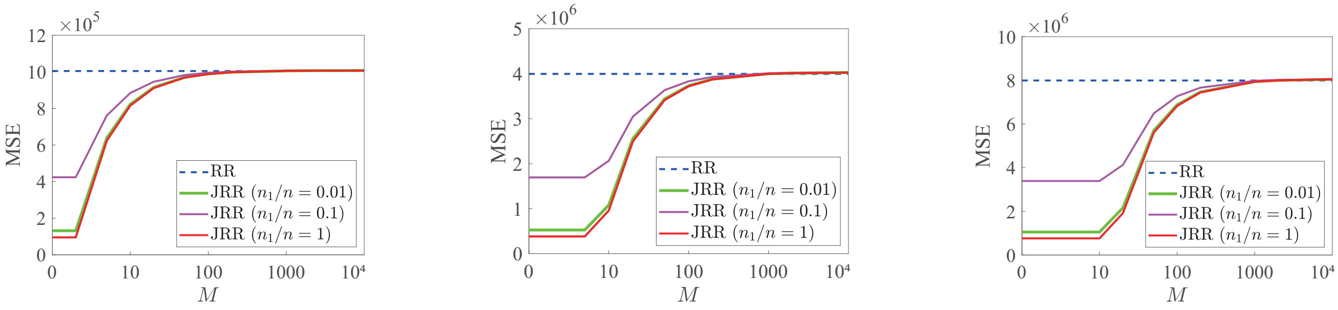
Figure 3: Comparison of MSE (top row) and ARE (bottom row) under RR and JRR with privacy budget $\varepsilon = 0.01$ to 1.



(a) $\varepsilon = 0.01$ (b) $\varepsilon = 0.1$ (c) $\varepsilon = 1$

(d) $\varepsilon = 0.01$ (e) $\varepsilon = 0.1$ (f) $\varepsilon = 1$

Figure 4: Comparison of MSE under RR and JRR with $n = 20$ to 2×10^6.



(a) $n = 1 \times 10^4$ (b) $n = 4 \times 10^4$ (c) $n = 8 \times 10^4$

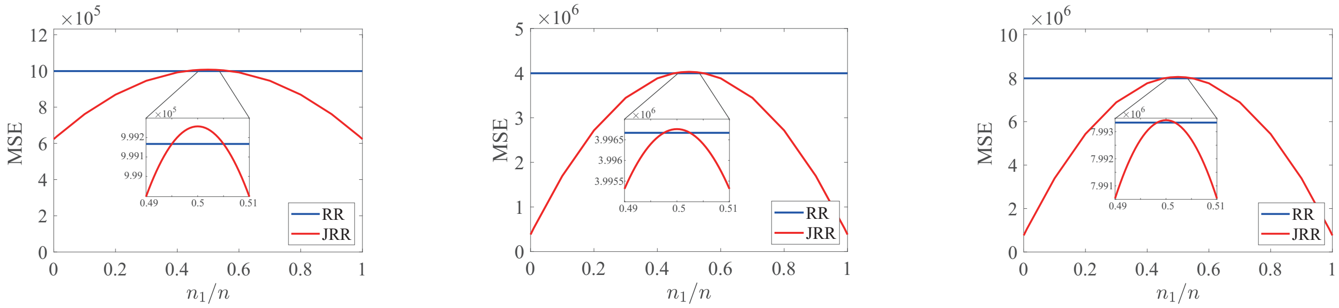
Figure 5: Comparison of MSE of JRR's heuristic solution with RR.

2, 000 to 2, 000, 000. This is because when there are relatively few contributors, ρ needs to be close to zero to guarantee sufficient data privacy, and JRR and RR have similar MSE. As n increases, ρ output by Algorithm 1 decreases, and a smaller negative correlation is introduced between the two contributors in each group, resulting in a smaller MSE than RR. Since the MSE under JRR is the sum of the variance of all $n/2$ groups, the change in the MSE under JRR is the joint effect of the decreased variance in each group and the increased number of groups. As n increases, the MSE under JRR inevitably increases but is still lower than that under RR. Most notably, in Fig. 4(a) when $n = 200, 000$ and $n_1/n = 1$, JRR outperforms RR by two orders of magnitude. Additionally, although the margin JRR outperforms RR decreases as ε increases, the improvement remains significant even when ε is large. Taking Fig. 4(c) as an example, the MSE under JRR is still 7.2×10^5 lower than that of RR when $n = 2 \times 10^6$ and $n_1/n = 1$.

Figs. 4(d) to 4(f) show a similar trend of ARE under RR and JRR. Specifically, the ARE under JRR and RR both decrease as n increases. As n increases from 2, 000 to 2, 000, 000, the ARE under JRR decreases much faster than that under RR due to the joint effect of increasing n and decreasing variance in each group.

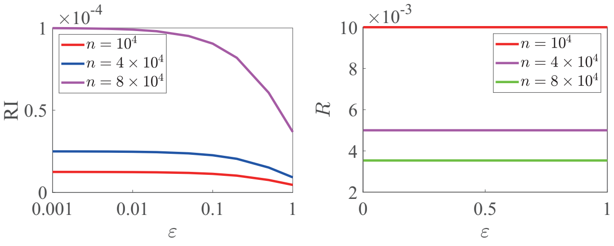
These results show that JRR is particularly favorable for the cases of small ε, large n, and n_1/n close to 0 or 1, reducing RR's MSE by up to two orders of magnitude and ARE by over 70%.

5.3.3 *Impact of M.* Figs. 5(a) to 5(c) show the MSE under JRR and RR as M increases from 0 to $n - 1$, where the MSE under RR is not affected by M and is plotted for reference. We can see that the MSE of JRR initially almost stays stable and then increases until it



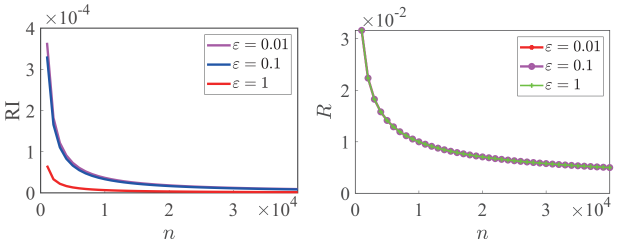
(a) $n = 1 \times 10^4$ (b) $n = 4 \times 10^4$ (c) $n = 8 \times 10^4$

Figure 6: Comparison of MSE under RR and JRR with $n_1/n = 0$ to 1.



(a) RI vs. ε (b) R vs. ε

Figure 7: The impact of ε on RI and R.



(a) RI vs. n (b) R vs. n

Figure 8: The impact of n on RI and R.

reaches that of RR. The reason is that the p selected by Algorithm 1 is always close to the one achieved under RR, but the corresponding ρ increases as M increases. Specifically, when M is small, ρ is always very close to the minimal value $1 - 1/p$, resulting in a relatively stable MSE that is much smaller than the one under RR. As M increases, a small ρ no longer satisfies the privacy constraint, and an increased negative ρ leads to an increased MSE. When M is very large, e.g., $M = 90\%n$, ρ is close to 0, and JRR degrades to RR.

These results indicate that JRR consistently outperforms RR in terms of MSE for any M from 0 to $n-1$, and is particularly favorable when M is small.

5.3.4 *Impact of n_1/n.*

Fig. 6 compares the MSE of JRR and RR on synthetic datasets, with $\varepsilon = 1$ and n_1/n ranging from 0 to 1. We can see that the MSE under JRR initially increases and then decreases as n_1/n increases from 0 to 1 and is symmetric with respect to n_1/n. The reason for the initial increase is that the MSE under JRR has term $\rho((2n_1 - n)^2 - n) = \rho n^2((2n_1/n - 1)^2 - 1/n)$, which is monotonically increasing with respect to $n_1/n \in [0, 0.5]$ when $\rho < 0$. In addition, the symmetry comes from the fact that $MSE = (\hat{n}_1 - n_1)^2 = (\hat{n}_0 - n_0)^2$, so MSE does not change if every contributor's original value is flipped. Moreover, we can see that the MSE of JRR exceeds that under RR when n_1/n is close to 0.5. There are two reasons. First, we choose p and ρ by Algorithm 1 under the assumption that $n_1/n \notin [1/2 - 1/2\sqrt{n}, 1/2 + 1/2\sqrt{n}]$. When n_1/n is close to 0.5, this assumption does not hold, and the choice of p and ρ results in higher MSE than RR. Second, the p and ρ chosen by Algorithm 1 are not the optimal solution for the optimization problem given in Eq. (22), which may further increase the MSE.

5.3.5 *The Cases of JRR Underperforming RR.*

We further evaluate the conditions under which JRR underperforms RR using the following two metrics:

- *Relative increases (RI):* It is defined as the ratio of the difference between the MSE of JRR and that of RR to the MSE of RR when $\frac{n_1}{n} = 0.5$ (i.e., the worst case for JRR):

$$\mathrm{RI} = \frac{MSE_{\mathrm{JRR}} - MSE_{\mathrm{RR}}}{MSE_{\mathrm{RR}}}, \tag{30}$$

  where MSE_{JRR} and MSE_{RR} are the MSE of JRR and RR, respectively.
- *Ratio of underperforming range (R):* Since the MSE of JRR and RR are roughly symmetric to $\frac{n_1}{n} = 0.5$, the range of $\frac{n_1}{n}$ in which the MSE of JRR exceeds that of RR is $[0.5 - R/2, 0.5 + R/2]$ for some $R \in [0, 0.5]$. Therefore, we define R as the ratio of the underperforming range.

Figs. 7(a) and 7(b) show RI and R with ε varying from 0.001 to 1. We can see from Fig. 7(a) that RI decreases as ε, which is anticipated because a larger ε means a larger p under both JRR and RR. Notably, RI is always less than 10^{-4} even for an extremely small $\varepsilon = 0.001$. From Fig. 7(b), we can see that R remain stable as ε increases, but a larger n (e.g., $n = 80,000$) results in a smaller R, coinciding with theoretical analysis in Section 4.4 that the range of underperforming is $1/\sqrt{n}$, i.e., independent of ε but decreases as n increases.

Figs. 8(a) and 8(b) show RI and R with n varying from $1,000$ to $40,000$. We can see from Fig. 8(a) that RI initially decreases sharply and then gradually decreases as n increases. In particular, even when n is small, e.g., $n = 5,000$, RI is 7.3×10^{-5}, which is negligible. Moreover, as we can see in Fig. 8(b), R decreases as n increases, but it is not affected by ε, which is consistent with Fig. 7(b). In addition, even when $n = 1,000$, R is less than 3%, indicating JRR outperforms RR in terms of the MSE for more than 97% of value n_1/n.

These results indicate that JRR outperforms RR for an overwhelming majority of n_1/n with only a negligible relative increase in the worst case.

5.4 Summary of Simulation Results

We summarize the simulation results as follows.

- JRR achieves smaller MSE and ARE than RR as long as the numbers of contributors having value 1 and 0 are not very close, i.e., $n_1/n \notin [1/2 - 1/2\sqrt{n}, 1/2 + 1/2\sqrt{n}]$.
- JRR significantly outperforms RR when the numbers of contributors with values 1 and 0 are not close, the total number of contributors is large, and the number of colluding contributors is small.
- When n_1 and n_0 are not very close, the margin by which JRR outperforms RR is inversely proportional to the privacy budget ε and the maximum number of colluding contributors M but proportional to the number of contributors n. As M increases, the MSE of JRR approaches that of RR.
- JRR underperforms RR if the numbers of contributors having value 1 and 0 are very close, i.e., when $n_1/n \in [1/2 - 1/2\sqrt{n}, 1/2 + 1/2\sqrt{n}]$. However, the margin by which RR outperforms JRR is very small or negligible.

6 Related Work

Privacy-preserving frequency estimation dates back to Warner [55], who introduced the RR mechanism for collecting sensitive data in social science research. RAPPOR [20] extends RR to non-binary data by encoding values as d-bit vectors and applying RR to each bit. OLH [49] refines this by introducing a *local hash* to compress the d-bit vector, reducing communication cost. A comparative analysis of these mechanisms and their variants is in [14].

Significant efforts have improved the privacy-utility tradeoff in LDP. A variance analysis framework was proposed in [49] to optimize the parameters of RR-based mechanisms, thereby enhancing data utility. Post-processing techniques can also improve the utility. For example, the non-negative and sum-to-one constraints were applied in [53], in which they referred to as *consistency*. As another example, the convolution framework in [21] added Wiener filter-based deconvolution to existing LDP protocols for improved data utility. Interactive protocols such as PrivKV [60] can iteratively improve estimation accuracy. Estimation of the most frequent items, or *heavy hitters*, can be accomplished through random projection, as shown in [7]. Cryptographic methods enhance privacy without sacrificing utility, as seen in Cryptε [41]. However, none of them address correlated perturbation among contributors. Some techniques, including post-processing, can be integrated with JRR for further utility gains.

Privacy leakage due to data correlation has long been a concern. Prior research [10, 27, 37, 43, 59] explores this from both theoretical and practical perspectives. The Pufferfish framework [27] enables customized privacy definitions for correlated data, later adapted in [43]. Bayesian differential privacy [59] analyzes correlated data privacy from a Bayesian perspective, with [10] using Bayesian networks to determine the minimum required noise. A game-theoretic model [57] examines the privacy-utility tradeoff in data sharing. Applications such as graph data publication [30], trajectory and network data release [10, 38], and trading statistics aggregation [37] have also been studied under differential privacy. However, these works focus on protecting correlated data, not the correlation among different contributors' perturbations.

The shuffle model [6, 8, 12, 19, 33] enhances privacy by having a trusted auxiliary server shuffle perturbed data before forwarding it to the data collector, breaking the linkage between contributors and their data. Originally proposed in Prochlo [8], the model's theoretical privacy guarantees have since been extensively studied. The first instantiation of JRR also utilizes a non-colluding auxiliary server, but unlike the shuffle model, this server never accesses contributor-submitted data. Moreover, shuffling is complementary to JRR and can be integrated to further strengthen privacy.

A separate line of research focused on designing LDP mechanisms for various types of data, including real-valued data [16, 32, 36, 46], multi-dimensional data [11, 40, 46, 58], set-valued data [39, 47, 48], time-series data [54], social graph data [44], key-value pairs [23, 45, 60], sparse vector [61], and directional data [56]. However, similar to existing LDP frequency estimation techniques, these works do not consider correlated perturbation.

7 Conclusions and Future Works

In this paper, we explored correlated random data perturbations for locally differentially private frequency estimation to achieve a better utility-privacy tradeoff. We have presented a general Joint Randomized Response (JRR) mechanism, along with two practical instantiations, which can provide the same level of data privacy as the classical RR mechanism while improving the data utility in an overwhelming majority of the cases. We have confirmed the advantages of JRR over RR through theoretical analysis and detailed simulation studies using both real and synthetic datasets.

There are several directions to extend this work. First, since JRR may underperform RR if the ratio n_1/n is very close to 0.5, it is possible to avoid this situation via a two-phase frequency estimation. In the first phase, we use the standard RR to obtain a rough estimate of n_1 using a portion of privacy budget whereby to choose optimal p and ρ for JRR. In the second phase, we use JRR with these parameters to refine the estimation of n_1 using the remaining privacy budget. Additionally, we plan to extend JRR for groups with more than two contributors. Moreover, we will seek to extend JRR to support other data types such as non-binary data and explore its integration with advanced LDP mechanisms for other data analysis problems such as mean value estimation.

Acknowledgments

We would like to thank the anonymous reviewers for their insightful and constructive feedback, which has greatly helped us enhance the quality of our work. This work was supported in part by the US National Science Foundation under grants CNS-2245689 (CRII), CNS-2325564 and CNS-2325563, as well as the 2022 Meta Research Award for Privacy-Enhancing Technologies. This work was also partially sponsored by the Army Research Laboratory and was accomplished under Cooperative Agreement Number W911NF-23-2-0225. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

References

[1] 2003. Kosarak. Downloaded from http://fimi.uantwerpen.be/data/.

[2] 2018. Women's E-Commerce Clothing Reviews. Downloaded from https://www.kaggle.com/nicapotato/womens-ecommerce-clothing-reviews.

[3] 2024. Law of total variance. https://en.wikipedia.org/w/index.php?title=Law_of_total_variance&oldid=1226284431#cite_note-1 Page Version ID: 1226284431.

[4] Amazon. 2013. Ratings (Beauty Products). Downloaded from https://www.kaggle.com/skillsmuggler/amazon-ratings.

[5] Apple 2017. Learning with Privacy at Scale. https://machinelearning.apple.com/research/learning-with-privacy-at-scale

[6] Borja Balle, James Bell, Adrià Gascón, and Kobbi Nissim. 2019. The Privacy Blanket of the Shuffle Model. In *CRYPTO'19*. 638–667.

[7] Raef Bassily and Adam Smith. 2015. Local, Private, Efficient Protocols for Succinct Histograms. In *Proceedings of the Forty-Seventh Annual ACM Symposium on Theory of Computing (STOC'15)* (Portland, Oregon, USA). 127–135.

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

[9] David Chaum. 1981. Untraceable Electronic Mail, Return Addresses, and Digital Pseudonyms. *Commun. ACM* 24, 2 (1981), 84–88. https://doi.org/10.1145/358549.358563

[10] J. Chen, H. Ma, D. Zhao, and L. Liu. 2017. Correlated Differential Privacy Protection for Mobile Crowdsensing. *IEEE Transactions on Big Data* (2017), 1–12.

[11] Jiangnan Cheng, Ao Tang, and Sandeep Chinchali. 2021. Task-aware Privacy Preservation for Multi-dimensional Data. *arXiv preprint arXiv:2110.02329* (2021).

[12] Albert Cheu, Adam Smith, Jonathan Ullman, David Zeber, and Maxim Zhilyaev. 2019. Distributed Differential Privacy via Shuffling. In *EUROCRYPT'19*. 375–403.

[13] Graham Cormode, Tejas Kulkarni, and Divesh Srivastava. 2019. Answering Range Queries under Local Differential Privacy. *Proc. VLDB Endow.* 12, 10 (June 2019), 1126–1138.

[14] Graham Cormode, Samuel Maddock, and Carsten Maple. 2021. Frequency estimation under local differential privacy. *Proceedings of the VLDB Endowment* 14, 11 (2021), 2046–2058.

[15] J. C. Duchi, M. I. Jordan, and M. J. Wainwright. 2013. Local Privacy and Statistical Minimax Rates. In *2013 IEEE 54th Annual Symposium on Foundations of Computer Science (FOCS'13)* (Berkeley, CA, USA). 429–438.

[16] John C Duchi, Michael I Jordan, and Martin J Wainwright. 2018. Minimax optimal procedures for locally private estimation. *J. Amer. Statist. Assoc.* 113, 521 (2018), 182–201.

[17] Cynthia Dwork, Krishnaram Kenthapadi, Frank McSherry, Ilya Mironov, and Moni Naor. 2006. Our Data, Ourselves: Privacy Via Distributed Noise Generation. In *Advances in Cryptology - EUROCRYPT 2006, 25th Annual International Conference on the Theory and Applications of Cryptographic Techniques, St. Petersburg, Russia, May 28 - June 1, 2006, Proceedings (Lecture Notes in Computer Science, Vol. 4004)*, Serge Vaudenay (Ed.). Springer, 486–503. https://doi.org/10.1007/11761679_29

[18] Cynthia Dwork and Aaron Roth. 2014. The Algorithmic Foundations of Differential Privacy. *Foundations and Trends in Theoretical Computer Science* 9, 3–4 (2014), 211–407.

[19] Úlfar Erlingsson, Vitaly Feldman, Ilya Mironov, Ananth Raghunathan, Kunal Talwar, and Abhradeep Thakurta. 2019. Amplification by Shuffling: From Local to Central Differential Privacy via Anonymity. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA '19)* (San Diego, California). 2468–2479.

[20] Úlfar Erlingsson, Vasyl Pihur, and Aleksandra Korolova. 2014. RAPPOR: Randomized Aggregatable Privacy-Preserving Ordinal Response. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security (CCS '14)* (Scottsdale, Arizona, USA). 1054–1067.

[21] Huiyu Fang, Liquan Chen, Yali Liu, and Yuan Gao. 2023. Locally Differentially Private Frequency Estimation Based on Convolution Framework. In *2023 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2208–2222.

[22] Badih Ghazi, Noah Golowich, Ravi Kumar, Rasmus Pagh, and Ameya Velingker. 2019. On the Power of Multiple Anonymous Messages. *IACR Cryptol. ePrint Arch.* (2019), 1382. https://eprint.iacr.org/2019/1382

[23] Xiaolan Gu, Ming Li, Yueqiang Cheng, Li Xiong, and Yang Cao. 2020. PCKV: Locally Differentially Private Correlated Key-Value Data Collection with Optimized Utility. In *29th USENIX Security Symposium (USENIX Security 20)*.

[24] X. Gu, M. Li, L. Xiong, and Y.Cao. 2020. Providing Input-Discriminative Protection for Local Differential Privacy. In *2020 IEEE 36th International Conference on Data Engineering (ICDE'20)* (Dallas, Texas, USA). 13 pages.

[25] Peter Kairouz, Keith Bonawitz, and Daniel Ramage. 2016. Discrete Distribution Estimation under Local Privacy. In *Proceedings of the 33rd International Conference on International Conference on Machine Learning* (New York, NY, USA) *(ICML'16)*. 2436–2444.

[26] Peter Kairouz, Sewoong Oh, and Pramod Viswanath. 2016. Extremal Mechanisms for Local Differential Privacy. 17, 1 (Jan. 2016), 492–542.

[27] Daniel Kifer and Ashwin Machanavajjhala. 2014. Pufferfish: A Framework for Mathematical Privacy Definitions. *ACM Trans. Database Syst.* 39, 1 (Jan. 2014), 1–36.

[28] Albert Kwon, David Lazar, Srinivas Devadas, and Bryan Ford. 2016. Riffle: An Efficient Communication System With Strong Anonymity. *Proc. Priv. Enhancing Technol.* 2016, 2 (2016), 115–134. https://doi.org/10.1515/POPETS-2016-0008

[29] Ninghui Li, Wahbeh Qardaji, Dong Su, and Jianneng Cao. 2012. PrivBasis: Frequent Itemset Mining with Differential Privacy. *Proc. VLDB Endow.* 5, 11 (July 2012), 1340–1351.

[30] X. Li, C. Zhang, T. Jung, J. Qian, and L. Chen. 2016. Graph-based privacy-preserving data publication. In *Proceedings of the 35th Annual IEEE International Conference on Computer Communications (INFOCOM 2016)*. 1–9.

[31] Qiyao Luo, Yilei Wang, and Ke Yi. 2022. Frequency Estimation in the Shuffle Model with Almost a Single Message. In *Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security, CCS 2022, Los Angeles, CA, USA, November 7-11, 2022*, Heng Yin, Angelos Stavrou, Cas Cremers, and Elaine Shi (Eds.). ACM, 2219–2232. https://doi.org/10.1145/3548606.3560608

[32] Fei Ma, Renbo Zhu, and Ping Wang. 2021. OPTT: Optimal Piecewise Transformation Technique for Analyzing Numerical Data under Local Differential Privacy. *arXiv preprint arXiv:2112.04745* (2021).

[33] Casey Meehan, Amrita Roy Chowdhury, Kamalika Chaudhuri, and Somesh Jha. 2022. Privacy Implications of Shuffling. In *International Conference on Learning Representations*.

[34] Mahnush Movahedi, Jared Saia, and Mahdi Zamani. 2015. Secure Multi-Party Shuffling. *IACR Cryptol. ePrint Arch.* (2015), 664. http://eprint.iacr.org/2015/664

[35] Takao Murakami and Yusuke Kawamoto. 2019. Utility-Optimized local differential privacy mechanisms for distribution estimation. In *28th USENIX Security Symposium (USENIX Security 19)*. 1877–1894.

[36] Thông T Nguyên, Xiaokui Xiao, Yin Yang, Siu Cheung Hui, Hyejin Shin, and Junbum Shin. 2016. Collecting and analyzing data from smart device users with local differential privacy. *arXiv preprint arXiv:1606.05053* (2016).

[37] Chaoyue Niu, Zhenzhe Zheng, Fan Wu, Shaojie Tang, Xiaofeng Gao, and Guihai Chen. 2018. Unlocking the Value of Privacy: Trading Aggregate Statistics over Private Correlated Data. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD 2018, London, UK, August 19-23, 2018*, Yike Guo and Faisal Farooq (Eds.). ACM, 2031–2040. https://doi.org/10.1145/3219819.3220013

[38] L. Ou, Z. Qin, S. Liao, Y. Hong, and X. Jia. 2018. Releasing Correlated Trajectories: Towards High Utility and Optimal Differential Privacy. *IEEE Transactions on Dependable and Secure Computing* (2018), 1–13.

[39] Zhan Qin, Yin Yang, Ting Yu, Issa Khalil, Xiaokui Xiao, and Kui Ren. 2016. Heavy Hitter Estimation over Set-Valued Data with Local Differential Privacy. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security* (Vienna, Austria) *(CCS '16)*. 192–203.

[40] Xuebin Ren, Chia-Mu Yu, Weiren Yu, Shusen Yang, Xinyu Yang, Julie McCann, and Philip Yu. 2018. LoPub: High-Dimensional Crowdsourced Data Publication with Local Differential Privacy. *IEEE Transactions on Information Forensics and Security* 13, 9 (Sep. 2018), 2151–2166.

[41] Amrita Roy Chowdhury, Chenghong Wang, Xi He, Ashwin Machanavajjhala, and Somesh Jha. 2020. CryptEpsilon: Crypto-Assisted Differential Privacy on Untrusted Servers. In *Proceedings of the 2020 ACM SIGMOD International Conference on Management of Data* (Portland, OR, USA) *(SIGMOD '20)*. 603–619.

[42] Steven Ruggles, J Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B Schroeder, Matthew Sobek, et al. 2010. Integrated public use microdata series: Version 5.0 [Machine-readable database]. *Minneapolis: University of Minnesota* 42 (2010).

[43] Shuang Song, Yizhen Wang, and Kamalika Chaudhuri. 2017. Pufferfish Privacy Mechanisms for Correlated Data. In *Proceedings of the 2017 ACM International Conference on Management of Data (SIGMOD '17)* (Chicago, Illinois, USA). 1291–1306.

[44] Haipei Sun, Xiaokui Xiao, Issa Khalil, Yin Yang, Zhan Qin, Hui (Wendy) Wang, and Ting Yu. 2019. Analyzing Subgraph Statistics from Extended Local Views with Decentralized Differential Privacy. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security (CCS'19)* (London, United Kingdom). 703–717.

[45] Lin Sun, Jun Zhao, Xiaojun Ye, Shuo Feng, Teng Wang, and Tao Bai. 2019. Conditional analysis for key-value data with local differential privacy. *arXiv preprint arXiv:1907.05014* (2019).

[46] Ning Wang, Xiaokui Xiao, Yin Yang, Jun Zhao, Siu Cheung Hui, Hyejin Shin, Junbum Shin, and Ge Yu. 2019. Collecting and Analyzing Multidimensional Data with Local Differential Privacy. In *2019 IEEE 35th International Conference on Data Engineering (ICDE)*. 638–649.

[47] Shaowei Wang, Liusheng Huang, Yiwen Nie, Pengzhan Wang, Hongli Xu, and Wei Yang. 2018. PrivSet: Set-Valued Data Analyses with Locale Differential Privacy. In *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications*. 1088–1096.

[48] Shaowei Wang, Yuqiu Qian, Jiachun Du, Wei Yang, Liusheng Huang, and Hongli Xu. 2020. Set-Valued Data Publication with Local Privacy: Tight Error Bounds and Efficient Mechanisms. *Proc. VLDB Endow.* 13, 8 (April 2020), 1234–1247.

[49] Tianhao Wang, Jeremiah Blocki, Ninghui Li, and Somesh Jha. 2017. Locally Differentially Private Protocols for Frequency Estimation. In *26th USENIX Security Symposium (USENIX Security 17)* (Vancouver, BC, Canada). 729–745.

[50] Tianhao Wang, Bolin Ding, Jingren Zhou, Cheng Hong, Zhicong Huang, Ninghui Li, and Somesh Jha. 2019. Answering multi-dimensional analytical queries under local differential privacy. In *Proceedings of the 2019 International Conference on Management of Data*. 159–176.

[51] T. Wang, N. Li, and S. Jha. 2018. Locally Differentially Private Frequent Itemset Mining. In *2018 IEEE Symposium on Security and Privacy (SP'18)*. 127–143.

[52] Tianhao Wang, Ninghui Li, and Somesh Jha. 2021. Locally Differentially Private Heavy Hitter Identification. *IEEE Transactions on Dependable and Secure Computing* 18, 2 (2021), 982–993.

[53] Tianhao Wang, Milan Lopuhaä-Zwakenberg, Zitao Li, Boris Skoric, and Ninghui Li. 2020. Locally Differentially Private Frequency Estimation with Consistency. In *27th Annual Network and Distributed System Security Symposium (NDSS)* (San Diego, California, USA).

[54] Zhibo Wang, Wenxin Liu, Xiaoyi Pang, Ju Ren, Zhe Liu, and Yongle Chen. 2020. Towards Pattern-aware Privacy-preserving Real-time Data Collection. In *IEEE INFOCOM 2020 - IEEE Conference on Computer Communications*. 109–118.

[55] Stanley L. Warner. 1965. Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias. *J. Amer. Statist. Assoc.* 60, 309 (1965), 63–69.

[56] Benjamin Weggenmann and Florian Kerschbaum. 2021. Differential Privacy for Directional Data. In *Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security* (Virtual Event, Republic of Korea) *(CCS '21)*. 1205–1222.

[57] X. Wu, T. Wu, M. Khan, Q. Ni, and W. Dou. 2017. Game Theory Based Correlated Privacy Preserving Analysis in Big Data. *IEEE Transactions on Big Data* (2017), 1–14.

[58] Min Xu, Bolin Ding, Tianhao Wang, and Jingren Zhou. 2020. Collecting and Analyzing Data Jointly from Multiple Services under Local Differential Privacy. *Proc. VLDB Endow.* 13, 12 (July 2020), 2760–2772.

[59] Bin Yang, Issei Sato, and Hiroshi Nakagawa. 2015. Bayesian Differential Privacy on Correlated Data. In *Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data (SIGMOD'15)* (Melbourne, Victoria, Australia). 747–762.

[60] Q. Ye, H. Hu, X. Meng, and H. Zheng. 2019. PrivKV: Key-Value Data Collection with Local Differential Privacy. In *2019 IEEE Symposium on Security and Privacy (SP'19)*. 317–331.

[61] Mingxun Zhou, Tianhao Wang, TH Hubert Chan, Giulia Fanti, and Elaine Shi. 2022. Locally differentially private sparse vector aggregation. In *2022 IEEE Symposium on Security and Privacy (SP)*. IEEE, 422–439.

A Proof of Theorem 1

Proof. It is easy to see that the marginal probability distribution of T_{2i-1} and T_{2i} in Table. 2 is the same. More specifically, for any data contributor u_j,

$$T_j = \begin{cases} 1 & \text{with probability } p, \\ 0 & \text{with probability } q. \end{cases}$$

Define Y_j as the random indicator variable for data contributor u_j reporting a perturbed value $y_j = 1$. There are two cases:

- $x_j = 0$, then $\Pr[Y_j = 1 | x_j = 0] = \Pr[T_j = 0] = q$;
- $x_j = 1$, then $\Pr[Y_j = 1 | x_j = 1] = \Pr[T_j = 1] = p$.

Denote I_v as the random variable for the number of contributors reporting a perturbed value of v, where $v \in \{0, 1\}$. When $v = 1$, we have

$$I_1 = \sum_{j=1}^{n} Y_j.$$

Taking the expectation on both sides follows

$$\begin{aligned} \mathrm{E}[I_1] = \mathrm{E}[\sum_{j=1}^{n} Y_j] &= \sum_{j=1}^{n} \mathrm{E}[Y_j] = \sum_{j=1}^{n} \Pr[Y_j = 1] \\ &= n_1 \cdot \Pr[T_j = 1] + (n - n_1) \cdot \Pr[T_j = 0] \\ &= n_1 \cdot p + (n - n_1) \cdot q \\ &= (p - q)n_1 + nq. \end{aligned} \tag{31}$$

Therefore, the data collector can estimate the number of contributors having value 1 as

$$\hat{n}_1 = \frac{I_1 - nq}{p - q}, \tag{32}$$

which is an unbiased estimator of n_1.

Similar to the proof of $\hat{n}_1$, we have $E[I_0] = (p-q)n_0 + nq$ followed by $I_0 = n - I_1$, which leads to the same unbiased estimator in the theorem. □

B Proof of Theorem 2

Proof. Denote by C the set of contributors who collude with the data collector and $\mathcal{T}_c = \{T_j | j \in C\}$. We prove the theorem by showing

$$\begin{aligned} \frac{\Pr[\mathcal{M}(x_i) = y_i \mid \mathcal{T}_c]}{\Pr[\mathcal{M}(x'_i) = y_i \mid \mathcal{T}_c]} &\leq \frac{\max \Pr[\mathcal{M}(x_i) = y_i \mid \mathcal{T}_c]}{\min \Pr[\mathcal{M}(x'_i) = y_i \mid \mathcal{T}_c]} \\ &= \frac{mp_{\max} + (n - m - 1)p}{mp_{\min} + (n - m - 1)q}, \end{aligned} \tag{33}$$

for all $x_i, x'_i, y_i \in D$.

We start by analyzing $\Pr[\mathcal{M}(x_i) = y_i \mid \mathcal{T}_c]$. Specifically, denote by u_j the contributor that is assigned to the same group as contributor u_i. We first have

$$\begin{aligned} &\Pr[\mathcal{M}(x_i) = y_i \mid \mathcal{T}_c] \\ &= \Pr[\mathcal{M}(x_i) = y_i \mid \mathcal{T}_c, j \notin C] \cdot \Pr[j \notin C] \\ &\quad + \Pr[\mathcal{M}(x_i) = y_i \mid \mathcal{T}_c, j \in C] \cdot \Pr[j \in C] \end{aligned} \tag{34}$$

Under uniform random grouping, we have

$$\Pr[j \notin C] = \frac{n - m - 1}{n - 1}, \tag{35}$$

and

$$\Pr[j \in C] = \frac{m}{n - 1}, \tag{36}$$

where m is the number of contributors who collude with the data collector.

We now analyze the conditional probabilities of contributor u_i reporting y_i under these two cases.

Case 1: $j \notin C$. If u_j is not a colluder, the probabilities of contributor u_i reporting y_i is independent of $\mathcal{T}_c$, and we have

$$\Pr[\mathcal{M}(x_i) = y_i \mid \mathcal{T}_c, j \notin C] = \Pr[\mathcal{M}(x_i) = y_i]. \tag{37}$$

It follows that

$$\begin{aligned}&\max_{x_i,y_i\in D}\Pr[\mathcal{M}(x_i)=y_i \mid \mathcal{T}_c, j\notin C]\\ =&\Pr[\mathcal{M}(x_i)=x_i]=p\ ,\end{aligned}\tag{38}$$

where the first equality means that the maximum is achieved when reporting truthfully (i.e., $y_i = x_i$). Similarly, we have

$$\begin{aligned}&\min_{x_i,y_i\in D}\Pr[\mathcal{M}(x_i)=y_i \mid \mathcal{T}_c, j\notin C]\\ =&\Pr[\mathcal{M}(x_i)=1-x_i]=q\ .\end{aligned}\tag{39}$$

Case 2: $j \in C$. If u_j colludes with the data collector, the conditional probabilities of contributor u_i reporting y_i only depend on T_j. We then have

$$\Pr[\mathcal{M}(x_i)=y_i \mid \mathcal{T}_c, j\in C]=\Pr[\mathcal{M}(x_i)=y_i \mid T_j].\tag{40}$$

There are four cases:

- Case 2.1: If $T_i = 1, T_j = 1$, then we have

$$\begin{aligned}\Pr[\mathcal{M}(x_i)=y_i \mid T_j]&=\frac{\Pr[y_i=x_i,T_j=1]}{\Pr[T_j=1]}\\ =&\frac{p^2+\rho pq}{p}=p+\rho q.\end{aligned}\tag{41}$$

- Case 2.2: If $T_i = 0, T_j = 1$, then we have

$$\begin{aligned}\Pr[\mathcal{M}(x_i)=y_i \mid T_j]&=\frac{\Pr[y_i=1-x_i,T_j=1]}{\Pr[T_j=1]}\\ =&\frac{(1-\rho)pq}{p}=(1-\rho)q.\end{aligned}\tag{42}$$

- Case 2.3: If $T_i = 1, T_j = 0$, then we have

$$\begin{aligned}\Pr[\mathcal{M}(x_i)=y_i \mid T_j]&=\frac{\Pr[y_i=x_i,T_j=0]}{\Pr[T_j=0]}\\ =&\frac{(1-\rho)pq}{q}=(1-\rho)p.\end{aligned}\tag{43}$$

- Case 2.4: If $T_i = 0, T_j = 0$, then we have

$$\begin{aligned}\Pr[\mathcal{M}(x_i)=y_i \mid T_j]&=\frac{\Pr[y_i=1-x_i,T_j=0]}{\Pr[T_j=0]}\\ =&\frac{q^2+\rho pq}{q}=q+\rho p.\end{aligned}\tag{44}$$

The maximum and the minimum of the above four cases are given by

$$\begin{aligned}p_{\max}&=\max\{p+\rho q,(1-\rho)q,(1-\rho)p,q+\rho p\}\\ &=\max\{p+\rho q,(1-\rho)p\},\\ p_{\min}&=\min\{p+\rho q,(1-\rho)q,(1-\rho)p,q+\rho p\}\\ &=\min\{q+\rho p,(1-\rho)q\}.\end{aligned}\tag{45}$$

It follows that

$$\max_{x_i,y_i\in D}\Pr[\mathcal{M}(x_i)=y_i|\mathcal{T}_c, j\in C]=p_{\max}\tag{46}$$

and

$$\min_{x_i,y_i\in D}\Pr[\mathcal{M}(x_i)=y_i|\mathcal{T}_c, j\in C]=p_{\min}\tag{47}$$

Substituting Eqs. (35), (36),(38) and (46) into Eq. (34), we have

$$\max_{x_i,y_i}\Pr[\mathcal{M}(x_i)=y_i|\mathcal{T}_c]=\frac{mp_{\max}}{n-1}+\frac{n-1-m}{n-1}\cdot p,\tag{48}$$

Similarly, substituting Eqs. (35), (36), (39) and (47) into Eq. (34), we have

$$\min_{x_i,y_i}\Pr[\mathcal{M}(x_i)=y_i|\mathcal{T}_c]=\frac{mp_{\min}}{n-1}+\frac{n-1-m}{n-1}\cdot q,\tag{49}$$

It follows that

$$\frac{\max\Pr[\mathcal{M}(x_i)=y_i \mid \mathcal{T}_c]}{\min\Pr[\mathcal{M}(x_i')=y_i \mid \mathcal{T}_c]}=\frac{mp_{\max}+(n-m-1)p}{mp_{\min}+(n-m-1)q}\ ,\tag{50}$$

for all $x_i, x_i', y_i \in D$. The theorem is thus proved. □

C Proof of Theorem 3

Proof. First, the variance of the estimator $\hat{n}_v$ is given by

$$\mathrm{Var}[\hat{n}_v]=\frac{\mathrm{Var}[I_v-nq]}{(p-q)^2}=\frac{\mathrm{Var}[I_v]}{(p-q)^2}\ ,\tag{51}$$

where the second equality holds because both n and q are constant. Since $n = n_0 + n_1$, we have $\mathrm{Var}[\hat{n}_0] = \mathrm{Var}[n - \hat{n}_1] = \mathrm{Var}[\hat{n}_1]$. In what follows, we focus on the analysis of $\mathrm{Var}[\hat{n}_1]$.

Again define Y_j to be the indicator random variable such that $Y_j = 1$ if contributor u_j reports a perturbed value of "1" and 0 otherwise for all $1 \le j \le n$. Without loss of generality, assume that group G_i consists of contributors u_{2i-1} and u_{2i} for all $1 \le i \le n/2$. Since the perturbation of different groups is independent of each other, we have

$$\mathrm{Var}[I_1]=\mathrm{Var}[\sum_{j=1}^{n}Y_j]=\sum_{i=1}^{n/2}\mathrm{Var}[Y_{2i-1}+Y_{2i}].\tag{52}$$

The $n/2$ groups can be classified into three categories: Type-1 groups with both contributors having value 1, Type-2 group with one contributor having value 1 and the other having value 0, and Type-3 groups with both contributors having value 0. The variance of each group's variance $\mathrm{Var}[Y_{2i-1} + Y_{2i}]$ depends on its type, and groups of the same type have the same variance. Define $V_z = \mathrm{Var}[Y_{2i-1} + Y_{2i}]$ if group G_i is a type-z group for all $1 \le z \le 3$ and $1 \le i \le \frac{n}{2}$. Let m_1, m_2, and m_3 be the numbers of Type-1, Type-2, and Type-3 groups, respectively, which are themselves random variables due to uniform random grouping. For any given n_0 and n_1, we have $2m_1 + m_2 = n_1$ and $m_1 + m_2 + m_3 = n/2$. It follows that $m_2 = n_1 - 2m_1$ and $m_3 = m_1 + n/2 - n_1$, which indicates that the random grouping only produces one independent random variable m_1.

For any given m_1, the conditional variance of I_1 is given by

$$\begin{aligned}\mathrm{Var}[I_1|m_1]&=\sum_{i=1}^{n/2}\mathrm{Var}[Y_{2i-1}+Y_{2i}]\\ =&m_1V_1+m_2V_2+m_3V_3\\ =&m_1V_1+(n_1-2m_1)V_2+(m_1+\frac{n}{2}-n_1)V_3\ .\end{aligned}\tag{53}$$

According to the law of total variance [3], the (unconditional) variance of I_1 is given by

$$\mathrm{Var}[I_1]=\mathrm{E}[\mathrm{Var}[I_1|m_1]]+\mathrm{Var}[\mathrm{E}[I_1|m_1]]\ .\tag{54}$$

Next, we calculate the two terms in Eq. (54) one by one.

The first term $\mathrm{E}[\mathrm{Var}[I_1|m_1]]$. We first calculate V_1, V_2 and V_3. For any group G_i, we have

$$\begin{aligned}&\mathrm{Var}[Y_{2i-1}+Y_{2i}]\\ =&\mathrm{Var}[Y_{2i-1}]+\mathrm{Var}[Y_{2i}]+2\mathrm{Cov}[Y_{2i-1},Y_{2i}]\\ =&\mathrm{Var}[Y_{2i-1}]+\mathrm{Var}[Y_{2i}]\\ &+2(\mathrm{E}[Y_{2i-1}Y_{2i}]-\mathrm{E}[Y_{2i-1}]\mathrm{E}[Y_{2i}]).\end{aligned} \tag{55}$$

There are three cases.

- Case 1: If G_i is of Type-1, then we have

$$\mathrm{E}[Y_{2i-1}Y_{2i}]=\Pr[T_{2i-1}=1,T_{2i}=1]=\rho pq+p^2\ , \tag{56}$$

  and

$$\mathrm{E}[Y_{2i-1}]\mathrm{E}[Y_{2i}]=\Pr[T_{2i-1}=1]\cdot\Pr[T_{2i}=1]=p^2\ . \tag{57}$$

- Case 2: If G_i is of Type-2, then we have

$$\mathrm{E}[Y_{2i-1}Y_{2i}]=(1-\rho)pq \tag{58}$$

  and

$$\mathrm{E}[Y_{2i-1}]\mathrm{E}[Y_{2i}]=pq\ . \tag{59}$$

- Case 3: If G_i is of Type-3, then we have

$$\mathrm{E}[Y_{2i-1}Y_{2i}]=q^2+\rho pq\ , \tag{60}$$

  and

$$\mathrm{E}[Y_{2i-1}]E[Y_{2i}]=q^2\ . \tag{61}$$

Substituting Eqs. (56) to (61) into Eq. (55), we get

$$\begin{aligned}V_1&=2pq(1+\rho)\ ,\\ V_2&=2pq(1-\rho)\ ,\\ V_3&=2pq(1+\rho)\ .\end{aligned} \tag{62}$$

Substituting Eq. (62) into Eq. (53), we have

$$\mathrm{Var}[I_1|m_1]=npq+(8m_1+n-4n_1)\rho pq, \tag{63}$$

Taking the expectation on both sides, we have

$$\begin{aligned}\mathrm{E}[\mathrm{Var}[I_1|m_1]]=&\mathrm{E}[npq+(8m_1+n-4n_1)\rho pq]\\ =&npq+(8\mathrm{E}[m_1]+n-4n_1)\rho pq.\end{aligned} \tag{64}$$

Since the expectation of the number of Type-1 groups is

$$\mathrm{E}[m_1]=\frac{n}{2}\cdot\frac{n_1(n_1-1)}{n(n-1)}=\frac{n_1(n_1-1)}{2(n-1)}, \tag{65}$$

Substituting Eq. (65) into Eq. (64), we have

$$\mathrm{E}[\mathrm{Var}[I_1|m_1]]=npq+\frac{(2n_1-n)^2-n}{n-1}\rho pq. \tag{66}$$

The second term $\mathrm{Var}[\mathrm{E}[I_1|m_1]]$. According to the definition of conditional expectation, we have

$$\begin{aligned}\mathrm{E}[I_1|m_1]&=\mathrm{E}[\sum_{j=1}^{n}Y_j|m_1]=\sum_{j=1}^{n}\mathrm{E}[Y_j|m_1]\\ &=n\cdot 1\cdot\Pr(Y_j=1|m_1)\ .\end{aligned} \tag{67}$$

Under JRR, whether an arbitrary contributor u_j reports 1 or 0 only depends on the contributor's original value x_j and the identical marginal probability distribution $\Pr[T_j]$. Since the numbers of contributors with the original value 1 and 0, n_1 and n_0, are predetermined. Thus, we have

$$\begin{aligned}\mathrm{E}[I_1|m_1]&=n\cdot\Pr[Y_j=1|m_1]\\ &=n_1\cdot\Pr[T_j=1]+(n-n_1)\cdot\Pr[T_j=0]\\ &=n_1p+(n-n_1)(1-p)\\ &=(2n_1-n)p+n-n_1\end{aligned} \tag{68}$$

which is a constant independent with m_1. It follows that

$$\mathrm{Var}[\mathrm{E}[I_1|m_1]]=0\ . \tag{69}$$

Substituting Eqs. (69) and (66) into Eq. (54), we have

$$\mathrm{Var}[I_1]=npq+\frac{(2n_1-n)^2-n}{n-1}\rho pq. \tag{70}$$

Finally, substituting Eq. (70) into Eq. (51), we have

$$\mathrm{Var}[\hat{n}_v]=\frac{pq}{(p-q)^2}\cdot(n+\frac{\rho((2n_1-n)^2-n)}{n-1})\ .$$

The theorem is thus proved. □

D Proof of the Lemma 1

Proof. Let $(2n_1-n)^2-n<0$. Solving the inequality, we have

$$-\sqrt{n}<2n_1-n<\sqrt{n}\ . \tag{71}$$

By simple algebraic manipulation, we get

$$\frac{n-\sqrt{n}}{2}<n_1<\frac{n-\sqrt{n}}{2}\ . \tag{72}$$

Dividing all three sides by n, we can obtain

$$\frac{n-\sqrt{n}}{2n}<\frac{n_1}{n}<\frac{n-\sqrt{n}}{2n}\ .$$

We therefore have $(2n_1-n)^2-n<0$ if $\frac{1}{2}-\frac{1}{2\sqrt{n}}<\frac{n_1}{n}<\frac{1}{2}+\frac{1}{2\sqrt{n}}$ and $(2n_1-n)^2-n\geq 0$ if $\frac{n_1}{n}\in[0,\frac{1}{2}-\frac{1}{2\sqrt{n}}]\bigcup[\frac{1}{2}+\frac{1}{2\sqrt{n}}]$ The lemma is thus proved. □

E Proof of Lemma 2

Proof. Since $0\leq n_1\leq n$, we have $0\leq(2n_1-n)^2\leq n^2$. Subtracting n from all three sides and then dividing them by $n-1$, we get

$$\frac{0-n}{n-1}\leq\frac{(2n_1-n)^2-n}{n-1}\leq\frac{n^2-n}{n-1}\ . \tag{73}$$

It follows that

$$\frac{1}{n-1}-1\leq\frac{(2n_1-n)^2-n}{n-1}\leq n\ . \tag{74}$$

Since $\rho\in[-1,1]$ and $n\geq 2>|\frac{1}{n-1}-1|$, multiplying ρ by all three sides of Inequality (74), we get

$$-n<\rho(\frac{1}{n-1}-1)\leq\rho\cdot\frac{(2n_1-n)^2-n}{n-1}\leq\rho n\leq n\ . \tag{75}$$

It follows that

$$-n<\rho\cdot\frac{(2n_1-n)^2-n}{n-1}\leq n\ . \tag{76}$$

Adding n to all three sides of Inequality (76), we get

$$0<n+\rho\cdot\frac{(2n_1-n)^2-n}{n-1}\leq 2n\ . \tag{77}$$

It follows that

$$n+\rho\cdot\frac{(2n_1-n)^2-n}{n-1}>0\ . \tag{78}$$

The lemma is thus proved.

□

F Proof of Lemma 3

Proof. Since $q = 1 - p$, we have

$$\frac{pq}{(p-q)^2} = \frac{p(1-p)}{(2p-1)^2} = \frac{1}{4}(\frac{1}{(2p-1)^2} - 1).$$

It is easy to see that $\frac{1}{(2p-1)^2}$ is monotonically decreasing with respect to $p \in (0.5, 1]$. Therefore, $\frac{pq}{(p-q)^2}$ is also monotonically decreasing with respect to $p \in (0.5, 1]$. The lemma is therefore proved. □

G Proof of Theorem 4

Proof. Since $h(p, \rho)$ is the product of $\frac{pq}{(p-q)^2}$ and $n + \frac{\rho((2n_1-n)^2-n)}{n-1}$ according to Eq. (21), we can analyze its monotonicity with respect to p and ρ based on the monotonicity of $\frac{pq}{(p-q)^2}$ and $n + \frac{\rho((2n_1-n)^2-n)}{n-1}$.

First, since $\frac{pq}{(p-q)^2}$ is monotonically decreasing with respect to $p \in (0.5, 1]$ according to Lemma 3, and $n + \frac{\rho((2n_1-n)^2-n)}{n-1} > 0$ according to Lemma 1 and is independent of p, $h(p, \rho)$ is monotonically decreasing with respect to $p \in (0.5, 1]$.

Second, since $\frac{pq}{(p-q)^2} > 0$ and is independent of ρ, the monotonicity of $h(p, \rho)$ with respect to ρ is the same as that of $n + \frac{\rho((2n_1-n)^2-n)}{n-1}$. Since $n \geq 2$ and $((2n_1 - n)^2 - n) < 0$ if $n_1/n \in (\frac{1}{2} - \frac{1}{2\sqrt{n}}, \frac{1}{2} + \frac{1}{2\sqrt{n}})$ according to Lemma 2, $h(p, \rho)$ is also monotonically decreasing with respect to ρ if $\frac{n_1}{n} \in (\frac{1}{2} - \frac{1}{2\sqrt{n}}, \frac{1}{2} + \frac{1}{2\sqrt{n}})$. By similar deduction, it is also easy to prove that $h(p, \rho)$ is also monotonically increasing with respect to ρ if $\frac{n_1}{n} \in [0, \frac{1}{2} - \frac{1}{2\sqrt{n}}] \bigcup [\frac{1}{2} + \frac{1}{2\sqrt{n}}, 1]$.

The theorem is therefore proved. □

H Monotonicity of $f(m)$

Denote by $g_1(m) = mp_{\max} + (n - m - 1)p$ and $g_2(m) = mp_{\min} + (n - m - 1)q$. We have $f(m) = \frac{g_1(m)}{g_2(m)}$, and its derivative is

$$\begin{aligned} f'(m) &= \frac{g_1'(m) \cdot g_2(m) - g_1(m) \cdot g_2'(m)}{g_2^2(m)} \\ &= \frac{(p_{\max} \cdot q - p_{\min} \cdot p) \cdot (n-1)}{g_2^2(m)} \end{aligned} \tag{79}$$

We now consider the following two cases.

- Case 1: if $\rho \leq 0$, we have $p_{\max} = (1-\rho)p$ and $p_{\min} = q + \rho p$. It follows that $p_{\max} \cdot q - p_{\min} \cdot p = -2\rho pq > 0$,
- Case 2: if $\rho > 0$, we have $p_{\max} = p + \rho q$ and $p_{\min} = (1-\rho)q$. It follows that $p_{\max} \cdot q - p_{\min} \cdot p = 2\rho pq > 0$.

Notice that $g_2^2(m) > 0$ and $n - 1 > 0$. We then have $f'(m) > 0$, and $f(m)$ is monotonically increasing with respect to m.

I Proof of Theorem 5

Proof. We prove the reporting trustfulness in Section 4.5 is the same as in Table 2.

For any group with two contributors u_{2i-1} and u_{2i}, let T_{2i-1} and T_{2i} be the truthfulness of the two contributors' reports.

First, for the case $T_{2i-1} = 1, T_{2i} = 1$, we have:

$$\begin{aligned} &\Pr[T_{2i-1} = 1, T_{2i} = 1] \\ =&\Pr[T_{2i-1} = 1, T_{2i} = 1 | R_1 = 1] \\ &+ \Pr[T_{2i-1} = 1, T_{2i} = 1 | R_1 = -1] \\ =&\frac{1}{2}(p^2 + \rho pq) + \frac{1}{2}(p^2 + \rho pq) = p^2 + \rho pq. \end{aligned} \tag{80}$$

Second, for the case $T_{2i-1} = 1, T_{2i} = 0$, we have:

$$\begin{aligned} &\Pr[T_{2i-1} = 1, T_{2i} = 0] \\ =&\Pr[T_{2i-1} = 1, T_{2i} = 0 | R_1 = 1] \\ &+ \Pr[T_{2i-1} = 1, T_{2i} = 0 | R_1 = -1] \\ =&\frac{1}{2}((1-\rho)pq - \sqrt{-\rho pq}) \\ &+ \frac{1}{2}((1-\rho)pq + \sqrt{-\rho pq}) \\ =&(1-\rho)pq. \end{aligned} \tag{81}$$

$T_{2i-1} = 0, T_{2i} = 1$ is symmetric to the case of $T_{2i-1} = 1, T_{2i} = 0$, so we have $\Pr[T_{2i-1} = 0, T_{2i} = 1] = (1-\rho)pq$.

For the case of $T_{2i-1} = 0, T_{2i} = 0$, we have

$$\begin{aligned} &\Pr[T_{2i-1} = 0, T_{2i} = 0] \\ =&\Pr[T_{2i-1} = 0, T_{2i} = 0 | R_1 = 1] \\ &+ \Pr[T_{2i-1} = 0, T_{2i} = 0 | R_1 = -1] \\ =&\frac{1}{2}(q^2 - \rho pq) \\ &+ \frac{1}{2}(q^2 - \rho pq) \\ =&q^2 - \rho pq. \end{aligned} \tag{82}$$

These results are the same as in Table 2. □

J Details of Real-world Datasets

We use the following four real-world datasets to evaluate the performance of JRR:

- **Kosarak [1]**: a dataset containing the click stream of a Hungarian news website that records about 8 million click events for 41,270 different pages. For our purpose, we randomly select 100 pages as the target pages and 20,000 click events as contributors. If a click event's visited page belongs to the target pages, that contributor's true value is *"1: visited"* and *"0: not"* otherwise. The frequency of the clicks on the target pages is deemed as the ground truth.
- **Amazon Rating Dataset [4]:** a dataset that contains over 2 million customer ratings of beauty-related products sold on Amazon. We randomly select 10,000 customers as contributors and set each contributor's true value to *1* if his/her rating is "1 star" and *0* otherwise.
- **E-commerce [2]:** a women's clothing E-Commerce dataset consisting of 23,486 records and 10 features variables. We select the binary variable "Recommended IND" as each contributor's true data.
- **Census[42]:** a dataset of the United States census in 2010 from the Integrated Public Use Microdata Series (IPUMS). We randomly select 10,000 records and set each contributor's true value to *1* if the code of group quarter (GQ) is 1 and *0* otherwise.

K Marginal distribution and estimator of k-JRR

We first prove that the marginal distribution of k-JRR (Section 4.6.1) is identical, with each contributor reporting their true value with probability p and any other value with probability q.

Proof. Let v_1 and v_1' be a contributor's true value and reported value. We have

$$\begin{aligned}\Pr[v_1' = v_1] = &\sum_{v_2' \in [k]} \Pr[v_1' = v_1, v_2']\\ =&p^2 + \rho pq + (k-1)(pq - \frac{1}{k-1}\rho pq)\\ =&p.\end{aligned} \tag{83}$$

Similarly, for each $v_1' \neq v_1$ in the data domain, we have

$$\begin{aligned}\Pr[v_1' \neq v_1] = &\sum_{v_2' \in [k]} \Pr[v_1' \neq v_1, v_2']\\ =&(pq - \frac{1}{k-1}\rho pq) + (k-1)(q^2 + \frac{1}{(k-1)^2}\rho pq)\\ =&q.\end{aligned} \tag{84}$$

We now prove the estimator $\hat{n}_v = (I_v - nq)/(p-q)$ is unbiased. First, we have

$$\begin{aligned}\mathrm{E}[I_v] = &n_v \cdot \Pr[v' = v] + (n - n_v) \cdot \Pr[v' \neq v]\\ =&n_v p + (n - n_v)q.\end{aligned} \tag{85}$$

Plugging $\mathrm{E}[I_v]$ into $\hat{n}_v$ gives

$$\begin{aligned}\mathrm{E}[\hat{n}_v] = &\frac{\mathrm{E}[I_v] - nq}{p-q}\\ =&\frac{n_v p + (n-n_v)q - nq}{p-q}\\ =&n_v,\end{aligned} \tag{86}$$

i.e. $\hat{n}_v$ is an unbiased estimator for n_v. □