Finding ϵ and δ of Statistical Disclosure Control Systems

Saswat Das[1], Keyu Zhu[2], Christine Task[3], Pascal Van Hentenryck[2], Ferdinando Fioretto[1]

[1]University of Virginia
[2]Georgia Institute of Technology
[3]Knexus Research Corporation
saswatdas@email.virginia.edu, keyu.zhu@gatech.edu, pvh@isye.gatech.edu,
christine.task@knexusresearch.com, fioretto@virginia.edu

Abstract

This paper analyzes the privacy of traditional Statistical Disclosure Control (SDC) systems under a differential privacy interpretation. SDCs, such as cell suppression and swapping, promise to safeguard the confidentiality of data and are routinely adopted in data analyses with profound societal and economic impacts. Through a formal analysis and empirical evaluation on demographic data from real household in the U.S., the paper shows that widely adopted SDC systems not only induce vastly larger privacy losses than classical differential privacy mechanisms, but, they may also come at a cost of larger accuracy and fairness.

1 Introduction

Statistical Disclosure Control (SDC) techniques are used to protect confidentiality while still enabling data analyses and dissemination in various fields with profound societal impacts, such as economics, public health, social science, and data science. These techniques have a long history, with their use dating back to the 1930 decennial release by the US Census Bureau, which leveraged traditional SDC techniques such as suppressing certain tables based on the number of people or households in a given area and swapping data in records with similar characteristics (Kelly, Golden, and Assad 1992; Dalenius and Reiss 1982).

While SDC techniques have traditionally been important for protecting against accidental or intentional disclosure, they lack formal guarantees that quantify the privacy risks that individuals incur upon data releases. This limitation restricts the ability of participants to assess the impact of these protections on published data, leading to potential vulnerabilities and privacy leaks. In contrast, differential privacy (DP) (Dwork et al. 2006) offers a rigorous definition of privacy and provides quantifiable privacy guarantees. Its deployments are increasing at a fast rate, with the US Census Bureau adopting DP for their 2020 release (Abowd et al. 2022), marking a significant shift towards more rigorous privacy protections. However, this adoption has also created controversy among data users, citing errors introduced by the noisy process adopted to ensure differential privacy, leading to skepticism and even legal action to block the bureau from using DP (US Court 2021). Despite the debate, many data agencies and organizations around the world continue to rely on traditional SDC techniques to protect their data confidentiality. While these approaches can be effective at protecting against accidental or intentional disclosures, it is unclear what privacy guarantees they provide when compared to differential privacy. On the other hand, while differential privacy can provide formal privacy guarantees, it may come with a cost in terms of accuracy, fairness, and equity (Kuppam et al. 2019; Tran et al. 2021; Fioretto et al. 2022; Tran, Dinh, and Fioretto 2021; Zhu, Hentenryck, and Fioretto 2021), a topic of considerable debate recently.

Considering their significant societal and economic consequences, a rigorous comparison of traditional SDC and DP is essential. Conducting this comparison, however, is challenged by the absence of a standardized framework for evaluating privacy protections: While DP provides quantifiable privacy guarantees, traditional SDC techniques often lack a distinct set of privacy metrics, complicating direct comparisons of the privacy protection levels they offer.

Contributions. This paper aims to address this challenge: it proposes a framework to compare traditional SDC to differential privacy for the first time and makes four distinct contributions. **(1)** It first proposes *carefully randomized* versions of two widely adopted SDC: suppression and swapping. The resulting randomized mechanisms are designed to closely resemble their original counterparts to preserve fidelity while allowing us to derive (ϵ, δ)-DP bounds. *Given these bounds, a key takeaway of the paper is the recognition that SDC mechanisms often fail to provide meaningful privacy guarantees.* **(2)** The paper then derives bounds for the bias and variance of the SDC mechanisms, allowing for a direct comparison with classical DP techniques for which such bounds exist. **(3)** Next, we analyze the fairness impact induced by the considered SDC systems and show that the fairness violations incurred by the randomized SDC algorithms are close to those of their traditional counterparts. *A second takeaway of the paper is the recognition that traditional SDC algorithms can induce much higher fairness violations than those reported by classical DP mechanisms.* (4) Finally, the paper provides an extensive empirical analysis of the performance of the SDC mechanisms and a comparison with two classical DP algorithms on data release and classification tasks.

From a broader perspective, the paper demonstrates that,

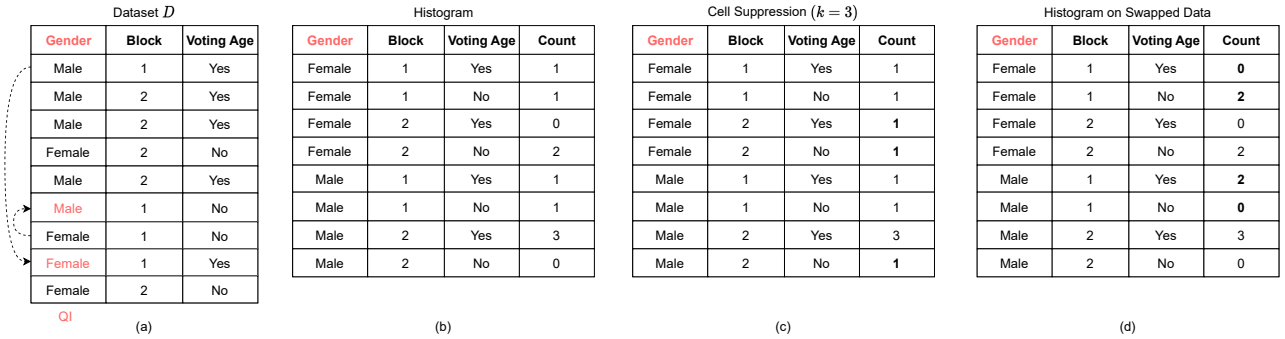

Figure 1: Illustration of the various traditional SDC mechanisms. Gender is taken as a quasi-identifier (QI). The dotted arrows represent swapping operations from donor to target records. Counts changed by SDC mechanisms are highlighted in bold.

contrary to popular belief, classical differential privacy mechanisms may be superior to traditional SDC systems in important data release and learning tasks in terms of accuracy and fairness for the same privacy levels. As a consequence, the results of this study have the potential to impact the way in which data agencies and organizations approach disclosure avoidance.

2 Problem Setting

The paper considers datasets $D = \{r_i\}_{i=1^n}$ of m records. Each record is a d-dimensional tuple of attributes associated with a unique individual from a data universe $\mathcal{X}$. A *histogram* $\boldsymbol{x}(D)$ of dataset D is an n-dimensional vector whose i^{th} entry, written $x_i(D)$, represents the count of the individual records with the i^{th} combination of attributes in $\mathcal{X}$. When there is no ambiguity, the dataset D is omitted in the expression $\boldsymbol{x}(D)$ for simplicity. Additionally, and without loss of generality, the histogram $\boldsymbol{x}$ is assumed to be sorted in some increasing order, i.e., $x_i \leq x_j$, for any $i < j$. Finally, each entry of the histogram $\boldsymbol{x}$ is assumed to be bounded by a value $B > 0$, i.e., $\boldsymbol{x} \in [B]^n$.

Consider, for example, the illustration in Figure 1(a); The dataset D contains records with three attributes: (geographic) "Block", "Gender", and "Voting Age". The associated histogram is illustrated in Figure 1(b). In this instance, the attribute "Gender", when combined with external information like "Zip code", can become personally identifying information and thus is known as a *quasi-identifier* (QI) while the remaining attributes are referred to as *non-quasi-identifiers*. Throughout the paper, the sets of quasi-identifiers and non-quasi-identifiers are denoted by Q and N, respectively. Given a record r and a set S of attributes, $r[S]$ is the vector of values for attributes S in r.

The goal of the paper is to analyze the privacy, utility, and fairness properties of traditional statistical disclosure control systems (reviewed next) on the task of releasing a privacy-preserving version $\tilde{\boldsymbol{x}}(D)$ of the histogram $\boldsymbol{x}(D)$. The notion of privacy considered in this paper is that of differential privacy, which is reviewed in the next section. The notions of utility and fairness central to the analysis rely on the concept of (statistical) *bias*. For any entry $i \in [n]$, the bias associated with a mechanism $\mathcal{M}$ is

$$\mathcal{B}(\mathcal{M})_i = \mathbb{E}\left[\mathcal{M}(D)_i\right] - x_i(D)\,,$$

where the expectation is over the randomness of the mechanism. Fairness is defined as the maximal difference in biases across the histogram entries.

Definition 1 (α-fairness (Zhu, Fioretto, and Van Hentenryck 2022))**.** *A mechanism $\mathcal{M}$ is said to be α-fair if the maximum difference among the biases is bounded by α, i.e.,*

$$\|\boldsymbol{\mathcal{B}}\left(\mathcal{M}\right)\|_{\rightleftharpoons} = \max_{i\in[n]} \mathcal{B}\left(\mathcal{M}\right)_i - \min_{i\in[n]} \mathcal{B}\left(\mathcal{M}\right)_i \leq \alpha,$$

where $\boldsymbol{\mathcal{B}}\left(\mathcal{M}\right) = [\mathcal{B}\left(\mathcal{M}\right)_1 \ \ldots \ \mathcal{B}\left(\mathcal{M}\right)_n]$.

3 SDC for Private Data Release

This section provides an overview of Differential Privacy and the prevalent SDC methods utilized by data agencies to safeguard sensitive information within datasets.

■ **Differential Privacy** (DP) (Dwork et al. 2006) is a privacy notion which quantifies and bounds the privacy loss of an individual participation to a computation. Changing a record from a dataset D, resulting in a new dataset D', defines the notion of *adjacency*, denoted $D \sim D'$.

Definition 2. *A mechanism $\mathcal{M}: \mathcal{D} \to \mathcal{R}$ with domain $\mathcal{D}$ and range $\mathcal{R}$ is (ϵ, δ)-differentially private, if, for any two inputs $D \sim D' \in \mathcal{D}$, and any subset of output responses $R \subseteq \mathcal{R}$:*

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

Parameter $\epsilon > 0$ describes the algorithm's *privacy loss* while parameter $\delta \in [0, 1)$ captures the probability of failure of the algorithm to satisfy ϵ-DP. In particular, the *Laplace mechanism* for histogram data release, defined by $\mathcal{M}_{\text{Lap}}(\boldsymbol{x}) = \boldsymbol{x} + \text{Lap}(2/\epsilon)$, where $\text{Lap}(\eta)$ is the Laplace distribution centered at 0 and with scaling factor η, satisfies $(\epsilon, 0)$-DP. Additionally, the *discrete Gaussian mechanism* (Canonne, Kamath, and Steinke 2020), defined by $\mathcal{M}_{\text{Gaus}}(\boldsymbol{x}) = \boldsymbol{x} + \mathcal{N}_{\mathbb{Z}}(0, 4/\epsilon^2)$, where $\mathcal{N}_{\mathbb{Z}}(0, \sigma)$ is the discrete Gaussian distribution with 0 mean and standard deviation σ, satisfies $(\frac{1}{2}\epsilon^2 + \epsilon\sqrt{2\log(1/\delta)}, \delta)$-DP.

We next discuss two predominant SDC systems which, in contrast to differential privacy, do not provide formal bounds on privacy leakage.

■ **Cell suppression** (Kelly, Golden, and Assad 1992), a technique frequently employed by statistical agencies (e.g., (Tatauranga Aotearoa 2020)), aims at concealing the low-frequency counts in histograms before data dissemination.

Definition 3. *Given a histogram $\boldsymbol{x}$ and a threshold value $k > 0$, cell suppression returns a private histogram $\tilde{\boldsymbol{x}}$ with entries $\tilde{x}_i = x_i$, for all $x_i \geq k$, and $\lfloor k/2 \rfloor$ otherwise.*

Figure 1(c) illustrates the application of cell suppression with threshold $k = 3$ to the histogram of Figure 1(b). The affected row counts are highlighted in red. Notice, how, for this choice of k, cell suppression returns a histogram where a large number of counts are revealed even after suppression, thus highlighting its privacy risks. A significant limitation of this approach is that it only protects sensitive attributes with a low number of records while neglecting others.

■ **Swapping** (Dalenius and Reiss 1982) replaces the values of sensitive attributes (the quasi-identifiers) in a record with those of another record. The version considered in this paper (Christ, Radway, and Bellovin 2022) consists of the following basic steps:

1. Select a target record r and then choose a donor record $r_s \neq r$ from the dataset with minimal discrepancy with r;
2. Replace r quasi-identifiers' values with those of r_s.

In the above, the discrepancy of a record from another is defined with respect to some metric. Like cell suppression, swapping aims to produce a privacy-preserving dataset, which can produce a private histogram $\tilde{\boldsymbol{x}}$. However, contrary to cell suppression (and DP mechanisms), swapping requires knowledge of the quasi-identifier attributes of the dataset. Figure 1(a) and (d) illustrate the application of swapping using "Gender" as the quasi-identifier attribute. The affected quasi-identifiers are highlighted in red. While swapping has been commonly used, for example by the US Census Bureau, to swap individuals with similar characteristics within close geographies, it is susceptible to reconstruction attacks (Garfinkel, Abowd, and Martindale 2019).

4 SDC Analysis Roadmap

This section outlines the methods used in the paper. Section 5 presents DP analogs of cell suppression and swapping. Our objective is to illustrate that these DP equivalents retain the primary attributes of the original methods while offering a comprehensive analysis of their privacy and error rates under a unified privacy framework. *Importantly, these DP analogs are designed to closely resemble their original SDC counterparts, even if it means not necessarily achieving the best possible δ bounds for a fixed ϵ. Indeed, a key takeaway of this paper is to demonstrate that existing SDC mechanisms fail to provide robust privacy guarantees.*

Note also that classical DP algorithms (e.g., Laplace mechanism) and cell suppression, make no assumptions about data attributes. In contrast, swapping relies on the use of quasi-identifiers. Finally, we emphasize that while the DP SDC mechanisms share many characteristics with their traditional SDC counterparts, *they should not be considered as "noisy" versions of them*. As a result, the presented results may not necessarily show a decrease in error as the privacy budget increases. In fact, they may even be more precise than the traditional mechanisms for some privacy budgets.

Next, we present the DP versions of the SDC algorithms and their privacy analyses. These analyses specify the value of the δ parameter for a given value of ϵ. Section 6 analyzes the fairness results. Finally, Section 7 presents an experimental evaluation on an extract of the 2019 American Community Survey (ACS) data (NIST 2021).

5 Privacy and Errors Analysis

This section presents the first main contribution of the paper. The approach involves introducing minimal modifications to the SDC presented earlier to derive their ϵ and δ parameters while maintaining their inherent characteristics as closely as possible. The section starts with a technical lemma that specifies a sufficient condition for (ϵ, δ)-DP, providing a crucial tool to derive the privacy guarantees of the randomized versions of the SDC discussed next.

Lemma 1. *Let $D \sim D'$ be adjacent datasets, and let $S := \left\{ \boldsymbol{o} \;\middle|\; \frac{\Pr(\mathcal{M}(D) = \boldsymbol{o})}{\Pr(\mathcal{M}(D') = \boldsymbol{o})} \leq \exp(\epsilon) \right\}$ and $S^{\complement}$ be its complement set. If $\Pr\left(\mathcal{M}(D) \in S^{\complement}\right) \leq \delta$, then $\mathcal{M}$ is (ϵ, δ)-DP.*

Differentially Private Cell Suppression

While cell suppression protects the privacy of minorities in the dataset, it does so deterministically and it neglects the privacy protection of the majority, thus, it does not satisfy the requirements of differential privacy. Indeed, the deterministic nature of this mechanism prevents it from generating different outputs for two neighboring datasets. Finding ϵ and δ values for cell suppression requires the introduction of a small amount of randomization. This randomized version, referred to as *DP cell suppression*, and denoted by $\mathcal{M}_{\text{CS}}$, releases a private count for every $i \in [n]$ as follows:

$$\mathcal{M}_{\text{CS}}(D)_i = \hat{x}_i = \begin{cases} k/2 & \text{if } x_i + \eta_i < k \\ x_i & \text{otherwise} \end{cases}, \qquad (1)$$

where $\eta_i \sim \text{Lap}(2/\epsilon)$ is an additive noise variable drawn from a 0-centered Laplace distribution with factor $2/\epsilon$ and k is the cell suppression threshold. DP-cell suppression is designed to avoid a significant alteration in the behavior of its deterministic counterparts while simultaneously, enabling us to quantify the worst-case privacy loss.

Figure 2 (left) illustrates the empirical errors of $\mathcal{M}_{\text{CS}}$ for several threshold values k (x-axis) and ϵ parameters. The errors are given for the ACS Massachusetts dataset (NIST 2021): they report the ℓ_1 distances $\|\tilde{\boldsymbol{x}} - \boldsymbol{x}\|_1$ between the histograms of the cell suppression and its DP counterpart. Notice how close the errors incurred by $\mathcal{M}_{\text{CS}}$ are to the original mechanism. This is important as it enables a meaningful comparison of $\mathcal{M}_{\text{CS}}$ and other DP mechanisms.

Privacy and Error Analysis. The next theorem reports the privacy guarantee provided by $\mathcal{M}_{\text{CS}}$.

Theorem 1. *Given $\epsilon > 0$ and a threshold $k < B$, mechanism $\mathcal{M}_{CS}$ is (ϵ, δ)-DP with $\delta = 1 - \frac{1}{4}\exp\left(-\epsilon(B - k)\right)$, where B is a bound on the histogram entries.*

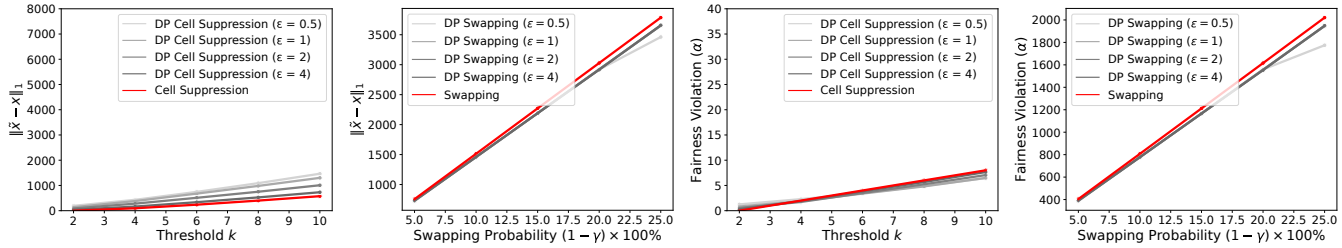

Figure 2: MA ACS dataset: Errors $\|\tilde{\boldsymbol{x}} - \boldsymbol{x}\|_1$ (left) and fairness values α (right) for cell suppression and swapping (right) in comparison to their differentially private counterparts (average of 200 repetitions). Note how the DP version resembles the original counterpart of the algorithms for the metrics of interest.

Notice that the δ-values can be large: the randomized mechanisms introduced in this paper do not aim to establish strong DP versions of their original counterparts. Their goal is to derive DP privacy bounds while closely resembling the behavior of their original counterparts, as shown in Figure 2.

Having examined privacy, we next show how close the histograms $\tilde{\boldsymbol{x}}$ returned by $\mathcal{M}_{\text{CS}}$ are to the original histogram $\boldsymbol{x}$, focusing on the statistical bias which, for each entry $i \in [n]$, is expressed as

$$\begin{aligned}\mathcal{B}\left(\mathcal{M}_{\text{CS}}\right)_i &= \mathbb{E}\left[\mathcal{M}_{\text{CS}}(D)_i\right] - x_i \\ &= \left(k/2 - x_i\right) \cdot \Pr\left(x_i + \eta_i < k\right).\end{aligned}$$

Observe that the error merely takes place when the noisy count is below the threshold k and is quantified as the difference between half of the threshold and the true count. Therefore, the following theorem relates the errors associated with $\mathcal{M}_{\text{CS}}$ with the probabilities of noisy counts being below the threshold, and the differences between half of the threshold and the counts of the original histogram.

Theorem 2. *The statistical bias of DP cell suppression $\mathcal{M}_{CS}$ is bounded as follows,*

$$\|\boldsymbol{\mathcal{B}}\left(\mathcal{M}_{CS}\right)\|_1 \leq \|k/2 \cdot \mathbf{1}_n - \boldsymbol{x}\|_2 \cdot \|\boldsymbol{p}\|_2 \ ,$$

where $\boldsymbol{p}$ is a shorthand for the vector $\left[\Pr\left(x_i + \eta_i < k\right)\right]_{i=1}^n$.

Differentially Private Swapping

Despite the randomized nature of swapping, the choice of swapping partners for a record is deterministic, and thus the swapping mechanism fails to meet the requirements of differential privacy. To illustrate its failure, let us take a look at an instance of two neighboring datasets D and D' with differing last records r and r', without loss of generality.

For these datasets to be the same after swapping, the QIs of r and r' need to be the same after swapping. However, it may be the case that the donor records (chosen as being the closest records to r and r' without the same QIs) in their respective datasets may not be the same and thus r and r' may receive different sets of QIs, leading to different outputs.

To obtain a DP counterpart to swapping, record swapping can be performed by choosing donor records with some randomness. In particular, we swap a record to its donor with inverse proportional probability to their discrepancy. Differentially private selection mechanisms such as the exponential mechanism (McSherry and Talwar 2007) and permute-and-flip (McKenna and Sheldon 2020) provide ways to select the best object with respect to a score function (here, discrepancy scores) in a differentially private manner. This paper adopts a permute-and-flip strategy to choose donor records in the DP analog of swapping.

The mechanism, referred to as DP swapping and denoted by $\mathcal{M}_{\text{SW}}$, works as follows: Given $\epsilon > 0$ and $\gamma \in [0, 1]$, $\mathcal{M}_{\text{SW}}$ chooses a record r_i and outputs its swapped version $\tilde{r}_i$, where $\tilde{r}_i[N] = r_i[N]$ and

$$\tilde{r}_i[Q] = \begin{cases} r_i[Q] & \text{w.p. } \gamma \\ r_i^d[Q] & \text{w.p. } 1-\gamma \end{cases} \tag{2}$$

where r_i^d is another record of the dataset chosen using the permute-and-flip mechanism using record distance/discrepancy scores as the scores. Notice that $\mathcal{M}_{\text{SW}}$ only modifies quasi-identifiers and produces a private dataset $\tilde{D}$, similar to what is done by the original swapping algorithm.

Figure 2 (center left) compares the ℓ_1 distances $\|\tilde{\boldsymbol{x}} - \boldsymbol{x}\|_1$ between the histograms generated by $\mathcal{M}_{\text{SW}}$ and its traditional counterpart for various amounts of rows swapped (in %) and parameters ϵ. Once again, observe how close the errors of the two mechanisms are; the lines for swapping and its DP variant are very close, and while almost indistinguishable, with increasing ϵ, DP swapping's errors approach those of swapping (refer to table 1). As with cell suppression, the goal here is to be able to modify the original swapping algorithm so that its performance is comparable to its deterministic counterpart while allowing us to derive DP bounds.

Privacy and Error analysis. With the definition of $\mathcal{M}_{\text{SW}}$ in hand, the privacy offered by (DP) swapping can now be quantified in terms of differential privacy. The privacy analysis of $\mathcal{M}_{\text{SW}}$ is reported in the following theorem.

Theorem 3. *For any $\epsilon > 0$, DP swapping is (ϵ, δ)-DP with*

$$\delta = 1 - 2\gamma(1-\gamma)\mathcal{L} - 3(1-\gamma)^2\mathcal{L}^2,$$

where $\mathcal{L} \triangleq \frac{\exp(-\epsilon/2)}{m}$, γ defined as in Equation (2), and m is the number of records in the dataset D.

Notice that δ is controlled by the privacy budget ϵ and the size of the dataset, but once again we stress that such values can be large, which is indicative of the poor privacy protection offered by SDC methods.

Next, we discuss bounds on how close histograms on the deidentified data returned by the swapping and its DP counterpart are. These bounds serve as useful worst-case quantifications of the utility we may expect from this mechanism and are useful to provide fairness guarantees, discussed below.

Proposition 1. *The bias of each element $i \in [n]$ of the histogram formed using DP swapped data is bounded as*

$$(1-\gamma)(\mathcal{L}-m^2) \leq \mathcal{B}(\mathcal{M})_i \leq m^2(1-\gamma).$$

Theorem 4. *The statistical bias of the DP swapping mechanism $\mathcal{M}_{SW}$ can be bounded as follows,*

$$0 \leq \|\boldsymbol{\mathcal{B}}(\mathcal{M}_{SW})\|_1 \leq \alpha_{SW},$$

where $\alpha_{SW} \triangleq (1-\gamma)m^3$.

6 Fairness analysis

The second main contribution of this paper is an analysis of the fairness of various SDC algorithms. Fairness (Definition 1) is expressed as the maximum difference in biases in the privacy-preserving histograms and the next result quantifies the unfairness of the DP cell suppression and swapping, along with the Laplace mechanism.

Theorem 5 (α-fairness for $\mathcal{M}_{\text{CS}}$)**.** *DP cell suppression is α_{CS}-fair with α_{CS} given by*

$$(x_n - x_1)\, p_1 + \max\left\{\left|k/2 - x_1\right|, \left|k/2 - x_n\right|\right\}(p_1 - p_n),$$

where p_1 and p_n are the first and last entries of $\boldsymbol{p}$ defined in Theorem 2 respectively.

Theorem 6 (α-fairness for $\mathcal{M}_{\text{SW}}$)**.** *DP swapping is α_{SW}-fair with α_{SW} given by α_{SW} as defined in Theorem 4.*

Theorem 7 (α-fairness for $\mathcal{M}_{\text{Lap}}$)**.** *The Laplace mechanism $\mathcal{M}_{Lap}$ is α_{Lap}-fair with α_{Lap} given by*

$$\frac{\exp(-\epsilon x_1/2)}{2}\|\boldsymbol{x}\|_{\rightleftharpoons} = \frac{\exp(-\epsilon x_1/2)}{2}(x_n - x_1).$$

Figure 2 (right) illustrates the fairness violations, represented by the value α, for cell suppression (third subplot) and swapping (fourth subplot), as well as their DP counterparts, for various privacy parameters ϵ and values of k (for cell suppression) or percentage of rows swapped (for swapping). It can be observed that the fairness violations of the differentially private mechanisms are comparable (or better) to those of their traditional counterparts. This is particularly noteworthy as the privacy parameter ϵ increases.

The following theorem is the third key result of this paper. *It establishes a relation between the Laplace mechanism and DP cell suppression and swapping with respect to unfairness.*

Theorem 8. *Suppose that the minimum count of the original histogram $\boldsymbol{x}(D)$ is between 2 and the threshold k, i.e., $2 \leq x_1 \leq k$ and that $\gamma \leq 1 - (B/2m^3)$.*

Then, the fairness error associated with the Laplace mechanism is not greater than that of the DP cell suppression or DP swapping mechanism, namely,

$$\alpha_{Lap} \leq \alpha_{CS} \quad \textit{and} \quad \alpha_{Lap} \leq \alpha_{SW}.$$

ϵ	$\mathcal{M}$	δ	Error	Fairness	Variance
0.5	Lap	**0**	763.77	3.65	39.50
	DGauss	0.36*	980.81	4.94	45.06
	DP-Sup	0.99*	**745.39**	**3.47**	**6.37**
	DP-Swap	0.99*	3462.16	1774.72	64920.40
1	Lap	**0**	**342.88**	**1.84**	9.97
	DGauss	0.13	659.21	3.06	22.94
	DP-Sup	0.99*	676.37	3.38	**2.94**
	DP-Swap	0.99*	3652.08	1945.24	113.87
2	Lap	**0**	**154.78**	**0.90**	**2.06**
	DGauss	0.01	436.92	2.20	12.03
	DP-Sup	0.99*	510.18	3.49	2.22
	DP-Swap	0.99*	3653.04	1946.84	112.22
4	Lap	**0**	**67.34**	**0.46**	**0.68**
	DGauss	3E-4	290.71	1.49	5.74
	DP-Sup	0.99*	336.54	3.83	1.37
	DP-Swap	0.99*	3658.16	1949.00	117.29

Table 1: MA dataset: Comparison of DP and SDC mechanisms on privacy violations δ, errors $\|\boldsymbol{\mathcal{B}}(\mathcal{M})\|_1$, fairness w.r.t. bias ($\|\boldsymbol{\mathcal{B}}(\mathcal{M})\|_{\rightleftharpoons}$), and ℓ_∞ norm of the empirical variance ($\|\boldsymbol{\mathcal{V}}(\mathcal{M})\|_\infty$).

Note that the upper bound on γ should be very close to 1 for real datasets as B is upper bounded by (and usually much smaller than) m.

The paper next presents empirical evidence that, under common regimes, the Laplace mechanism may have a significant advantage over the DP SDC mechanisms as well.

7 Experimental Evaluation

This section assesses the performance of the DP variants of cell suppression and swapping and compares them with two key DP mechanisms, the Laplace and the Discrete Gaussian Mechanisms, reviewed in Section 3. The experiments use the 2019 Diverse Community Data Excerpts datasets for Massachusetts, Texas, and National PUMAs (Task et al. 2022). All the experiments report averages of 200 repetitions. When not otherwise stated, we use $k = 6$ as the threshold for cell suppression, a swap rate of 25% for swapping, and a feature set comprising of PUMA, race, sex, house ownership status, and whether an individual earns $\geq$ \$50,000 per year for histograms. This section focuses on evaluating the mechanisms in three settings: data release, classification, and subgroup distribution. An empirical privacy assessment is also provided.

Data Release

The first task compares datasets reconstructed from histograms generated by the various DP mechanisms studied. Table 1 assesses the performance of the DP variants of the traditional SDC mechanisms and the Laplace and discrete Gaussian mechanisms in terms of errors and fairness violations. When negative counts are produced, a simple post-processing projection into the non-negative orthant is applied. In addition, while performing suppression the cell suppression algorithm and its DP variants, zero counts are

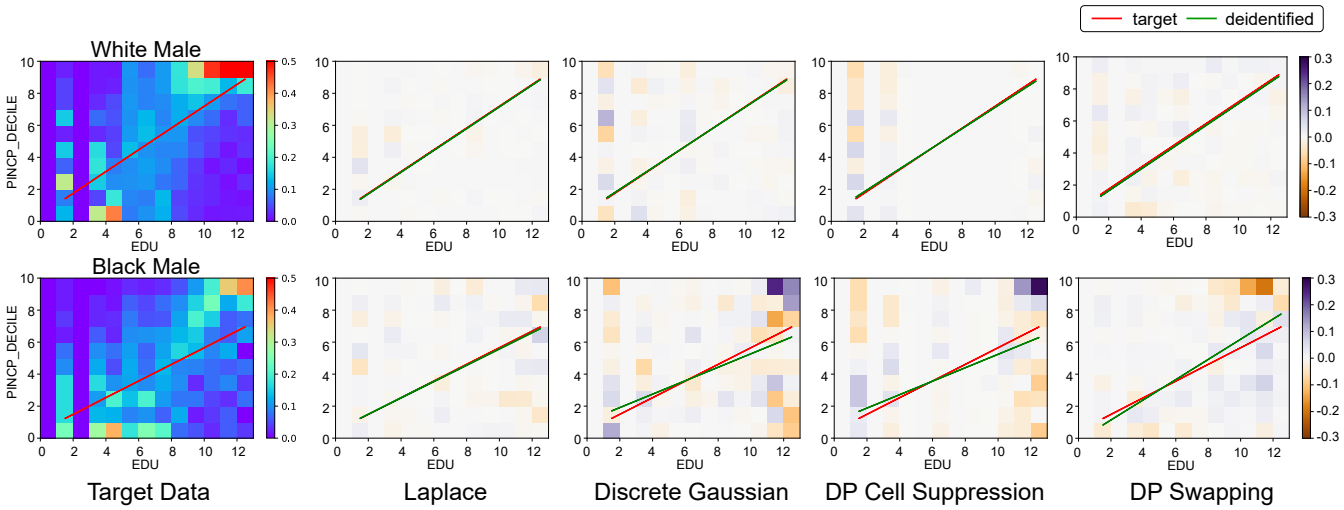

Figure 3: National ACS Dataset: Heatplots for two-way marginals on *income decile* and *education* for the Laplace mechanism, Discrete Gaussian mechanism, DP Cell Suppression, and DP Swapping (left to right) comparing White vs Black Men subgroups.

left untouched to reflect what is done in practice (Templ, Kowarik, and Meindl 2015). These results are particularly significant: contrary to commonly held beliefs, in these datasets, classical DP algorithms not only provide strong privacy guarantees (see the unreasonably large δ values, highlighted with *, for other mechanisms), but also produce histograms with better accuracy, fairness, and variance, in most cases.

While the above relies on worst-case analysis, we also report an empirical privacy assessment conducted on the datasets generated by the various privacy-preserving methods analyzed in a later subsection.

Classification

The next task compares the performance of the various SDC mechanisms studied in this paper in a classification task. Emulating classical studies performed by data agencies, this setting employs the private datasets obtained through a data-release query in order to train a logistic regression classifier.

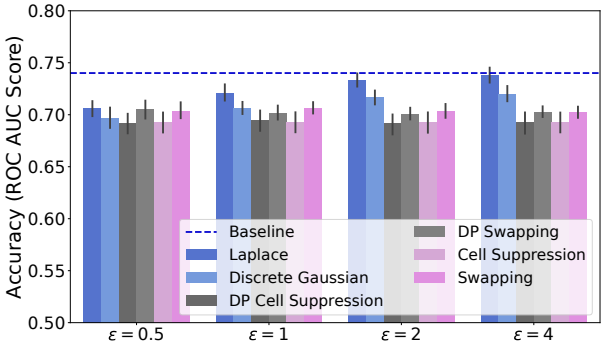

Figure 4: MA dataset: Logistic Regression on datasets generated by various privacy-preserving mechanisms.

The task is to predict whether an individual earns at least $50,000 per year, given their race, age (discretized into 5 bins of equal size), sex, and house ownership status, and the results in Figure 4 are presented in terms of accuracy on the original, non-private dataset.

Observe how the Laplace and discrete Gaussian mechanisms lead to classifiers with much higher accuracies than classifiers trained over data produced by other traditional SDC mechanisms in all instances but for $\epsilon = 0.5$. Also, notice that the accuracies for the DP variants of the SDC mechanisms are largely dependent on the values of the parameters (k and swap rate) than on ϵ. Notably, the classification accuracy of Laplace and discrete Gaussian is much closer to that of the baseline method (trained on non-private datasets) than any other method, for these parameter choices. Again, this is significant: *despite their simplicity, these tasks are the basis for numerous statistical analyses performed routinely by data agencies and organizations.*

Subgroup Distributions

The next task describes the errors attained by a basic statistical analysis: a two-way-marginal query and simple linear regression. The attributes adopted are *income decile*—the person's total income rank, with respect to their state, discretized into 10% bins—and *educational status*—which describes twelve educational attainment levels from no schooling to doctorate degree. A two-way marginal query reports the frequencies of individuals distributed across the combined categories of the two target variables. This is computed on both the original (target) data and the deidentified data, and the resulting errors are visually represented through a heatplot: redder hues indicate a deficit of individuals, while bluer hues signal a surplus. Linear regression is performed to fit a line summarizing the relationship between education and income—the red line indicates the true relationship, the green line is fit to the deidentified data.

Figure 3 presents the results for the National ACS dataset, focusing on a subpopulation of white men (top) and black men (bottom). The results were obtained using $\epsilon = 2$, with a swap rate of 25% for DP swapping, and a threshold $k = 10$ for DP cell suppression. Two main observations can be made: **(1)** The plots reveal relatively small errors for the Laplace, Discrete Gaussian, and DP cell suppression mechanisms. In contrast, DP Swapping exhibits significantly higher errors in its heatplot for black males for the same value of ϵ. This difference can be attributed to how DP mechanisms like Laplace and Discrete Gaussian add precisely calibrated noise to bin counts, while DP cell suppression only suppresses counts below a certain threshold, leaving the rest intact. Swapping, by exchanging quasi-identifiers (in this case, *race*) among rows, can cause more pronounced changes. Of all the methods analyzed, the Laplace mechanism induces the least amount of errors, consistent with the results observed in the previous section. **(2)** All methods display disparate impacts when comparing the errors induced on the white group versus those on the black group. Notably, the Laplace mechanism shows the least impact in such error disparity, while the Discrete Gaussian mechanism and DP cell suppression behave similarly.

It is worth noting that while DP algorithms offer distinct advantages in terms of worst-case privacy guarantees compared to cell suppression and swapping, they tend to perform worse than the DP versions of traditional SDC mechanisms as the sparsity of the analyzed subpopulation increases. This observation underscores the natural tradeoff between accuracy, privacy, and data complexity.

Empirical Privacy Assessment

The theoretical results presented in the paper and the empirical assessment reported in Table 1 focus on worst-case privacy leakage. Next, the paper considers an empirical assessment that offers a nuanced understanding of the potential risk of the various de-identification processes. Since the objective of the paper is to analyze the privacy of different mechanisms in data release tasks, we utilize a *unique exact match (UEM)* metric (NIST 2021) as an empirical measurement of privacy. This method counts the number of unique rows or singleton bins in the intersection of the original data and its deidentified counterpart. These are records that are uniquely identifiable in the target data, and are also easily reidentifiable as they appear unmodified in the de-identified data. Privacy risks are subsequently quantified as the proportion of unique individuals whose data was presumably deidentified but disclosed with their sensitive attributes unaltered. Thus, a higher score indicates a reduced privacy risk.

Figure 5 shows the UEM values (as the percentage of identified unique individuals over the set of all individuals) for the Massachusetts and Texas states of the ACS datasets obtained after using Laplace, Discrete Gaussian, and DP Swapping mechanisms for various ϵ values. DP cell suppression is excluded as, by removing individuals that fall below the threshold, it attains a (near) perfect score; however, such observation is not a cell suppression property but an artifact of the UEM metric, which fails to account for that data deidentified using cell suppression leaves many rows

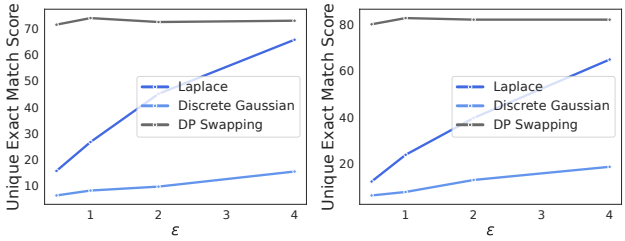

Figure 5: MA (left) and TX (right): Empirical privacy assessment with the Unique Exact Match metric.

untouched, posing reidentification risks. This phenomenon occurs because all singleton bins are imputed with the value $\lfloor k/2 \rfloor$ after being suppressed, unless their noised counts happen to exceed k in the case of DP cell suppression.

Not surprisingly, the results in Figure 5 reveal that DP data release methods become more susceptible to UEM as ϵ increases, indicating a greater risk of reidentification for individuals. However, notice that while DP swapping, by virtue of its closeness to swapping, is more robust to increase in ϵ, it also has a significantly higher risk of reidentification vis-à-vis DP methods, even for higher values of ϵ, for the swapping rate adopted.

Finally, we notice that the Discrete Gaussian mechanism appears more resilient to this privacy risk as ϵ increases. This behavior aligns with the observation made in the previous section, where this mechanism was shown to induce larger errors than the Laplace mechanism. As a result, it is more likely to produce fewer unique exact match values, reflecting a more robust stance against reidentification risks.

8 Conclusion

This paper presented a framework for comparing traditional statistical disclosure control (SDC) to differential privacy. It proposed carefully randomized versions of two widely adopted traditional SDC methods, i.e., suppression and swapping, and derived (ϵ, δ)-DP bounds for these mechanisms. The paper also analyzed these DP algorithms empirically and showed that they are close to their traditional counterparts both in terms of accuracy and fairness. The DP SDC mechanisms were then compared experimentally with traditional DP mechanisms (i.e., the Laplace or the discrete Gaussian mechanisms) on widely adopted data release and classification tasks. Importantly, and contrary to popular belief, the experimental evaluation showed that classical DP mechanisms not only achieve much higher privacy protections than SDC methods but they may also be superior in terms of accuracy and fairness for the same privacy levels.

We hope this work will stimulate additional research in the important direction of rigorously comparing distinct privacy-preserving techniques. Besides additional empirical assessments on different data formats, additional exploration venues may consider more elaborate settings where SDC and DP mechanisms may compose multiple operations (e.g., they may operate on different data universes and adopt post-processing steps), which is often the case in practice.

Acknowledgements

This research is partially supported by NSF grant 2133169 and NSF CAREER award 2143706. Fioretto is also supported by a Google Scholar Research Award and an Amazon Research Award. The views and conclusions of this work are those of the authors only.

References

Abowd, J. M.; Ashmead, R.; Cumings-Menon, R.; Garfinkel, S.; Heineck, M.; Heiss, C.; Johns, R.; Kifer, D.; Leclerc, P.; Machanavajjhala, A.; et al. 2022. The 2020 census disclosure avoidance system topdown algorithm. *Harvard Data Science Review*, (Special Issue 2).

Canonne, C. L.; Kamath, G.; and Steinke, T. 2020. The discrete gaussian for differential privacy. *Advances in Neural Information Processing Systems*, 33: 15676–15688.

Christ, M.; Radway, S.; and Bellovin, S. M. 2022. Differential Privacy and Swapping: Examining De-Identification's Impact on Minority Representation and Privacy Preservation in the U.S. Census. In *2022 IEEE Symposium on Security and Privacy (SP)*, 457–472.

Dalenius, T.; and Reiss, S. P. 1982. Data-swapping: A technique for disclosure control. *Journal of statistical planning and inference*, 6(1): 73–85.

Dwork, C.; McSherry, F.; Nissim, K.; and Smith, A. 2006. Calibrating noise to sensitivity in private data analysis. In *Theory of cryptography conference*, 265–284. Springer.

Fioretto, F.; Tran, C.; Van Hentenryck, P.; and Zhu, K. 2022. Differential Privacy and Fairness in Decisions and Learning Tasks: A Survey. In *In Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, 5470–5477.

Garfinkel, S.; Abowd, J. M.; and Martindale, C. 2019. Understanding database reconstruction attacks on public data. *Communications of the ACM*, 62(3): 46–53.

Kelly, J. P.; Golden, B. L.; and Assad, A. A. 1992. Cell suppression: Disclosure protection for sensitive tabular data. *Networks*, 22(4): 397–417.

Kuppam, S.; McKenna, R.; Pujol, D.; Hay, M.; Machanavajjhala, A.; and Miklau, G. 2019. Fair decision making using privacy-protected data. *arXiv preprint arXiv:1905.12744*.

McKenna, R.; and Sheldon, D. 2020. Permute-and-Flip: A New Mechanism for Differentially Private Selection. In *Proceedings of the 34th International Conference on Neural Information Processing Systems*, NIPS'20. Red Hook, NY, USA: Curran Associates Inc. ISBN 9781713829546.

McSherry, F.; and Talwar, K. 2007. Mechanism Design via Differential Privacy. In *48th Annual IEEE Symposium on Foundations of Computer Science (FOCS'07)*, 94–103.

NIST. 2021. SDNist v1.4 beta: Synthetic Data Report Tool. National Institute of Standards and Technology.

Task, C.; Bhagat, K.; Damon, S.; and Howarth, G. 2022. NIST Diverse Community Excerpts Data.

Tatauranga Aotearoa. 2020. Microdata output guide. https://www.stats.govt.nz/assets/Methods/Microdata-Output-Guide-2020-v5-Sept22update.pdf.

Templ, M.; Kowarik, A.; and Meindl, B. 2015. Statistical Disclosure Control for Micro-Data Using the R Package sdcMicro. *Journal of Statistical Software*, 67(4): 1–36.

Tran, C.; Dinh, M.; and Fioretto, F. 2021. Differentially Private Empirical Risk Minimization under the Fairness Lens. In *Advances in Neural Information Processing Systems*, volume 34, 27555–27565. Curran Associates, Inc.

Tran, C.; Fioretto, F.; Van Hentenryck, P.; and Yao, Z. 2021. Decision Making with Differential Privacy under the Fairness Lens. In *International Joinmogt Conference on Artificial Intelligence (IJCAI)*, 560–566.

US Court. 2021. State of Alabama Differential Privacy June 29, 2021 Memorandum Opinion and Order. DocumentCloud. Accessed: Apr. 27, 2023.

Zhu, K.; Fioretto, F.; and Van Hentenryck, P. 2022. Post-processing of Differentially Private Data: A Fairness Perspective. In Raedt, L. D., ed., *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI-22*, 4029–4035. International Joint Conferences on Artificial Intelligence Organization. Main Track.

Zhu, K.; Hentenryck, P. V.; and Fioretto, F. 2021. Bias and Variance of Post-processing in Differential Privacy. In *AAAI Conference on Artificial Intelligence*, 11177–11184. AAAI Press.