Journal of Privacy and Confidentiality
Vol. 14 (2) 2024

Submitted August 2023
Published June 2024

GENERALIZED RAINBOW DIFFERENTIAL PRIVACY

YUZHOU GU[1], ZIQI ZHOU[2], ONUR GÜNLÜ[3], RAFAEL G. L. D'OLIVEIRA[4], PARASTOO SADEGHI [5], MURIEL MÉDARD[6], AND RAFAEL F. SCHAEFER[7]

[1] School of Mathematics, Institute for Advanced Study, Princeton, NJ 08540
e-mail address: yuzhougu@ias.edu

[2] Telecommunication Networks Group (TKN), Department of Telecommunication Systems, Technische Universität Berlin, 10587 Berlin, Germany
e-mail address: zhou@ccs-labs.org

[3] Information Coding Division, Department of Electrical Engineering, Linköping University, 58183 Linköping, Sweden
e-mail address: onur.gunlu@liu.se

[4] School of Mathematical and Statistical Sciences, Clemson University, Clemson, SC 29634
e-mail address: rdolive@clemson.edu

[5] School of Engineering and Technology, The University of New South Wales, Canberra, Australia
e-mail address: p.sadeghi@unsw.edu.au

[6] Research Laboratory of Electronics (RLE), Massachusetts Institute of Technology, Cambridge, MA 02139
e-mail address: medard@mit.edu

[7] Chair of Information Theory and Machine Learning, the BMBF Research Hub 6G-life, the Cluster of Excellence "Centre for Tactile Internet with Human-in-the-Loop (CeTI)", and the 5G Lab Germany, Technische Universität Dresden, 01062 Dresden, Germany
e-mail address: rafael.schaefer@tu-dresden.de

Abstract. We study a new framework for designing differentially private (DP) mechanisms via randomized graph colorings, called rainbow differential privacy. In this framework, datasets are nodes in a graph, and two neighboring datasets are connected by an edge. Each dataset in the graph has a preferential ordering for the possible outputs of the mechanism, and these orderings are called rainbows. Different rainbows partition the graph of connected datasets into different regions. We show that if a DP mechanism at the boundary of such regions is fixed and it behaves identically for all same-rainbow boundary datasets, then a unique optimal (ϵ, δ)-DP mechanism exists (as long as the boundary condition is valid) and can be expressed in closed-form. Our proof technique is based on an interesting relationship between dominance ordering and DP, which applies to any finite number of colors and for (ϵ, δ)-DP, improving upon previous results that only apply to at most three colors and for ϵ-DP. We justify the homogeneous boundary condition assumption by giving an example with non-homogeneous boundary condition, for which there exists no optimal DP mechanism.

Key words and phrases: differential privacy, optimal mechanism, dominance ordering.

www.journalprivacyconfidentiality.org
DOI:10.29012/jpc.896
© Gu, Zhou, Günlü, D'Oliveira, Sadeghi, Médard, and Schaefer
Creative Commons (CC BY-NC-ND 4.0)

1. Introduction

Differential privacy (DP) is a general framework that aims to limit the statistical capability of a curious analyst, irrespective of its computational power, in determining whether or not the data of a specific participant was used in response to its query[1] [2, 3]; see [4] for a treatment of the subject and [5] for a survey. Recently, DP is applied in the 2020 US Census [6], as well as by Apple, Google, and Microsoft [7–9].

DP imposes constraints on all neighboring datasets, which traditionally differ only in data from one participant. These constraints are *relative* (specified as ratios of mechanism probability distributions), *local* (specified for neighboring datasets), and *dataset-independent* (agnostic to the underlying data distribution or structure) [4], contributing to the success of DP. However, many DP implementations are agnostic to the actual dataset at hand, which is nonadaptive and undesirable [10]. Majority of output perturbation DP mechanisms consider the worst-case query sensitivity between any two neighboring datasets to determine the scale of noise [4]. This approach is pessimistic and can negatively affect the query utility [10].

Several solutions are available that improve the query utility. For instance, noise calibration was proposed in [11] to smooth the sensitivity, but a chosen utility level is then not guaranteed and the mechanism suffers from a heavy tail that leads to outliers. Another direction is to relax the DP constraints [10, 12, 13]. For example, [10] proposed *individual*-DP that defines DP constraints only for *given* datasets and their neighbors. The individual-DP framework destroys the group DP, i.e., DP constraints for non-neighboring datasets are no longer valid.

Recently, [14] proposed a method to design dataset-dependent DP mechanisms for binary-valued queries that guarantee optimal utility without weakening the original DP constraints; see also [15]. In the model in [14], each dataset has a true query value (e.g., `blue` or `red`) and is represented as a node on a graph with edges, representing neighboring datasets. Moreover, they consider DP mechanisms which act homogeneously at the boundary datasets.[2] They then show how these initial constraints can be optimally extended in closed-form for all other datasets, where the probability of giving the truthful query response is maximized by taking into account the distance to the boundary.

The framework in [14] was generalized in [16] by increasing the number of possible query outputs to three (e.g., `blue`, `red`, and `green` that represent majority votes among three choices). This extension is challenging in several ways. In the binary case, the optimal probability assignment for one color (e.g., `blue`) automatically determines the whole mechanism. In the multi-color case, this is not possible. Thus, a preferential order of colors at each dataset is assumed to solve this problem, in which a mechanism is defined to be better than another if the preferred colors are output with larger probabilities. When there are at most three colors, it is shown that for a DP mechanism that is homogeneous at the boundary, at most one optimal ϵ-DP mechanism exists, for which a closed-form expression is also given. This result recovers the binary case of [14] as a special case.

1.1. Main Contributions.

In this work, we significantly improve [16] by providing a new proof technique that allows us to extend the results to any number of colors and any (ϵ, δ)-DP requirements. We show that given a valid boundary homogeneous DP mechanism, at most

[1] DP variants assuming a finite computational power for the adversary have been studied in works including [1] but are not within the scope of our work.

[2] Boundary datasets are neighbors whose true query value is different from each other.

one optimal (ϵ, δ)-DP mechanism exists, for which we provide a closed-form expression. Our results recover the result of [16], in which there are three colors and $\delta = 0$. We note that our definition of optimality of a mechanism is through dominance ordering (see Definition 3 below), while in [16], optimality is defined through lexicographic ordering. Because dominance ordering is stronger than lexicographic ordering (i.e., $x \preceq y$ in dominance ordering implies $x \preceq y$ in lexicographic ordering, but two elements comparable in lexicographic ordering are not necessarily comparable in dominance ordering), our optimality result is strictly stronger than [16] even in the ternary and $\delta = 0$ case.

At its core, our proof uses an interesting relationship between DP and dominance ordering. Namely, for any $\epsilon, \delta \geq 0$ and any distribution P on an ordered set $\mathcal{V}$, there exists a unique distribution Q that is (ϵ, δ)-close to P and that dominates any other distribution Q' that is (ϵ, δ)-close to P; see Section 4 below for more details. Finally, we justify the homogeneous boundary condition assumption by presenting an example with a non-homogeneous boundary condition, such that valid DP mechanisms exist, but no DP mechanisms are optimal; see Example 1 below.

1.2. **Organization of the Paper.** In Section 2, we introduce the setting for rainbow DP. In Section 3, we show that to construct optimal DP mechanisms for general graphs under homogeneous boundary conditions, it suffices to do so for a special class of graphs called line graphs. Furthermore, we show that an optimal DP mechanism may not exist for non-homogeneous boundary conditions, and discuss the relationship between our optimality condition and that of previous work [16]. In Section 4, we construct optimal DP mechanisms for line graphs. In Section 5, we give explicit formulas for the optimal rainbow DP mechanism and present several examples. In Section 6, we summarize our results, discuss related approaches and possible further directions.

1.3. **Notation.** All logarithms in this paper are natural logarithms unless otherwise noted. For a non-negative integer n, we use $[n]$ to denote the set $\{1, \ldots, n\}$. For two integers $n \leq m$, we use $[n : m]$ to denote the set $\{n, \ldots, m\}$. For two distributions P, Q on a measurable space $\mathcal{X}$, we define their total variation (TV) distance as

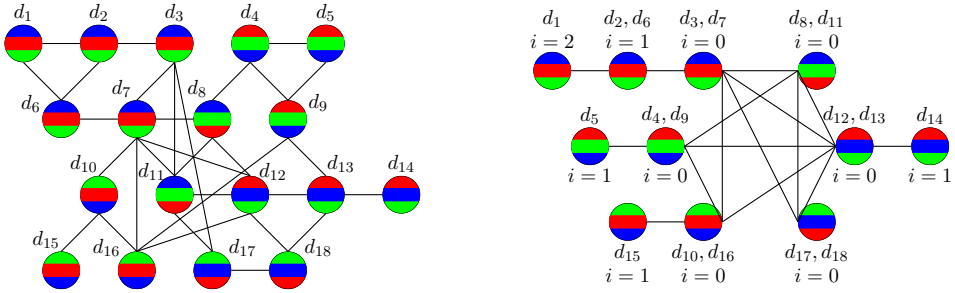
$$\mathrm{TV}(P, Q) = \sup_{\mathcal{S}} |P(\mathcal{S}) - Q(\mathcal{S})|, \tag{1.1}$$

where $\mathcal{S}$ goes over measurable subsets of $\mathcal{X}$.

2. Rainbow Differential Privacy

We denote by $(\mathcal{D}, \sim)$ a family of datasets together with a symmetric neighborhood relationship, where $d, d' \in \mathcal{D}$ are neighbors if $d \sim d'$. We consider a finite output space $\mathcal{V}$. Each dataset $d \in \mathcal{D}$ has an ordered preference for the elements of $\mathcal{V}$, captured by what we call a *rainbow* that represents each preference order.

Definition 1. Let $\mathcal{V}$ be a finite output space. A rainbow on $\mathcal{V}$ is a total ordering of $\mathcal{V}$. We denote a rainbow as a permutation vector $c \in \mathrm{Sym}(\mathcal{V})$, where $\mathrm{Sym}(\mathcal{V})$ is the set of all permutations of $\mathcal{V}$.



(A) A rainbow graph. (B) The boundary rainbow graph.

FIG. 1. A rainbow graph and its corresponding boundary graph. A vertex represents a dataset and its neighboring datasets are connected by an edge. The function output space is represented by three colors `blue`, `red`, and `green`. Each dataset has a color preference, represented by the ordering inside the vertex. For instance, vertex d_1 prefers `blue` to `red` and `red` to `green`. We call each such color ordering a rainbow. A DP mechanism is then a probability distribution over colors for every vertex. In (B), we show the boundary rainbow graph of the rainbow graph shown in (A), as described in Definition 8 below. In Theorem 3 we show how, for homogeneous boundary conditions (defined in Definition 6 below), optimal (ϵ, δ)-DP mechanisms on (A) can be retrieved from optimal ones on (B). For example, for rainbow $c = (\texttt{red}, \texttt{green}, \texttt{blue})$, the vertex $(c, 0)$ in the boundary rainbow graph corresponds to datasets d_4, d_9 in the original rainbow graph because they are on the boundary of B^c in the original graph. There is an edge between $(c = (\texttt{red}, \texttt{green}, \texttt{blue}), 0)$ and $(c' = (\texttt{red}, \texttt{blue}, \texttt{green}), 0)$ in the boundary rainbow graph, because there is an edge (d_9, d_{13}) in the original rainbow graph, with $i \in B^c$, $m \in B^{c'}$.

The preference of a dataset is captured by the *preference function* $f : \mathcal{D} \to \mathrm{Sym}(\mathcal{V})$ that assigns a rainbow to each dataset $d \in \mathcal{D}$. Thus, if $f(d) = (\texttt{blue, red, green})$, then it means that the dataset $d \in \mathcal{D}$ prefers `blue` to `red` and `red` to `green`. Moreover, the goal is to construct a random function $\mathcal{M} : \mathcal{D} \to \mathcal{V}$ that, for each dataset $d \in \mathcal{D}$, randomly puts out an element of $\mathcal{V}$ such that for a given DP constraint a pre-specified utility function is maximized. As commonly done in the DP literature, we refer to the random function as a *mechanism*. A mechanism is DP if the distribution of its output on neighboring datasets are approximately indistinguishable, as we formalize next.

Definition 2 ([4])**.** Let ϵ, δ be non-negative real numbers with $\delta \leq 1$. For two distributions P and Q on $\mathcal{V}$, we say P and Q are (ϵ, δ)-close if for any $\mathcal{S} \subseteq \mathcal{V}$, we have $P(\mathcal{S}) \leq e^{\epsilon} Q(\mathcal{S}) + \delta$ and $Q(\mathcal{S}) \leq e^{\epsilon} P(\mathcal{S}) + \delta$. A mechanism $\mathcal{M} : \mathcal{D} \to \mathcal{V}$ is called (ϵ, δ)-DP if for any $d \sim d'$, the distributions of $\mathcal{M}(d)$ and $\mathcal{M}(d')$ are (ϵ, δ)-close. If $\mathcal{M}$ is $(\epsilon, 0)$-DP, then we also say it is ϵ-DP. We denote the set of all (ϵ, δ)-DP mechanisms by $\mathfrak{M}$.

The performance of a mechanism is measured via a utility function $U : \mathfrak{M} \to \mathbb{R}$, where $U[\mathcal{M}] \geq U[\mathcal{M}']$ means that the mechanism $\mathcal{M}$ outperforms $\mathcal{M}'$. In this work, we consider utility functions that agree with the preference function $f : \mathcal{D} \to \mathrm{Sym}(\mathcal{V})$, i.e., it is preferable for a dataset $d \in \mathcal{D}$ to output a color it prefers according to its rainbow $f(d) \in \mathrm{Sym}(\mathcal{V})$.

Definition 3. Let $\preceq$ be the dominance ordering on the probability simplex

$$\Delta(\mathcal{V}) = \{x \in [0,1]^{|\mathcal{V}|} : x_1 + \cdots + x_{|\mathcal{V}|} = 1\}, \tag{2.1}$$

i.e., for $x, y \in \Delta(\mathcal{V})$, $x \preceq y$ if and only if $x_1 + \cdots + x_k \leq y_1 + \cdots + y_k$ for all $1 \leq k \leq |\mathcal{V}|$. For every mechanism $\mathcal{M} \in \mathfrak{M}$ and dataset $d \in \mathcal{D}$, let $\vec{\mathcal{M}}(d) \in \Delta(\mathcal{V})$ be the vector with coordinates $\vec{\mathcal{M}}(d)_k = \mathbb{P}[\mathcal{M}(d) = f(d)_k]$. Then, a mechanism $\mathcal{M} \in \mathfrak{M}$ dominates another mechanism $\mathcal{M}' \in \mathfrak{M}$ (denoted by $\mathcal{M} \succeq \mathcal{M}'$) if for every dataset $d \in \mathcal{D}$, $\vec{\mathcal{M}}(d) \succeq \vec{\mathcal{M}}'(d)$. Moreover, we say a utility function $U : \mathfrak{M} \to \mathbb{R}$ is *order reasonable* if whenever a mechanism $\mathcal{M} \in \mathfrak{M}$ dominates another mechanism $\mathcal{M}' \in \mathfrak{M}$, we have $U[\mathcal{M}] \geq U[\mathcal{M}']$.

The notion of domination in Definition 3 induces a partial order on the set $\mathfrak{M}$ of all (ϵ, δ)-DP mechanisms. When a mechanism $\mathcal{M}$ dominates $\mathcal{M}'$, it means that $\mathcal{M}$ outperforms $\mathcal{M}'$ for any order reasonable utility. In this setting, we say that a mechanism is *optimal* if no other mechanism dominates it.

An interesting subclass of order reasonable utility functions is the set of functions U of the form $U[\mathcal{M}] = \mathbb{E}\left[\sum_{d \in \mathcal{D}} u_d(\mathcal{M}(d))\right]$, where the expectation is taken over the randomness of the output of the DP mechanism and where $u_d(\cdot)$ is a monotone function for all $d \in \mathcal{D}$ in the sense that $u_d(f(d)_i) \geq u_d(f(d)_{i+1})$ for $1 \leq i \leq |\mathcal{V}| - 1$.

As in [14], we represent a family of datasets together with their neighboring relation $(\mathcal{D}, \sim)$ by a simple graph, where the vertices are the datasets in $\mathcal{D}$ and there is an edge between $d, d' \in \mathcal{D}$ if and only if they are neighbors, i.e., $d \sim d'$.

Definition 4 ([14])**.** A morphism between $(\mathcal{D}_1, \overset{1}{\sim})$ and $(\mathcal{D}_2, \overset{2}{\sim})$ is a function

$$g : (\mathcal{D}_1, \overset{1}{\sim}) \to (\mathcal{D}_2, \overset{2}{\sim}) \tag{2.2}$$

such that $d \overset{1}{\sim} d'$ implies in either $g(d) \overset{2}{\sim} g(d')$ or $g(d) = g(d')$ for every $d, d' \in \mathcal{D}_1$.

An example of a morphism is shown in Fig. 1 above. A morphism $g : (\mathcal{D}_1, \overset{1}{\sim}) \to (\mathcal{D}_2, \overset{2}{\sim})$ allows to transport (ϵ, δ)-DP mechanisms from its codomain to its domain.

Theorem 1 ([14])**.** Let $g : (\mathcal{D}_1, \overset{1}{\sim}) \to (\mathcal{D}_2, \overset{2}{\sim})$ be a morphism and $\mathcal{M}_2 : \mathcal{D}_2 \to \mathcal{V}$ be an (ϵ, δ)-DP mechanism on $(\mathcal{D}_2, \overset{2}{\sim})$. Then, the mechanism $\mathcal{M}_1 : \mathcal{D}_1 \to \mathcal{V}$ given by the pullback operation $\mathcal{M}_1 = \mathcal{M}_2 \circ g$ is an (ϵ, δ)-DP mechanism on $\mathcal{D}_1$.

3. Optimal Rainbow Differential Privacy Mechanisms

In [14], DP schemes were interpreted as randomized graph colorings. In that setting, each dataset's preference was characterized by a single color. In general, for larger output spaces, each dataset has a corresponding rainbow according to its ordering preference. Thus, we call the triple $(\mathcal{D}, \sim, f)$ a *rainbow graph*, where $\mathcal{D}$ is the family of datasets, $\sim$ is the neighborhood relationship, and $f : \mathcal{D} \to \mathrm{Sym}(\mathcal{V})$ is the preference function. We say a morphism $g : (\mathcal{D}_1, \overset{1}{\sim}, f_1) \to (\mathcal{D}_2, \overset{2}{\sim}, f_2)$ is *rainbow-preserving* if $f_1 = f_2 \circ g$. Indeed, the morphism in Fig. 1 above is rainbow-preserving. We consider the following topological notions.

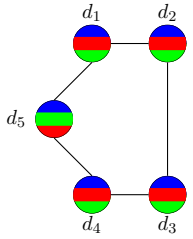


FIG. 2. Illustration of the rainbow graph in Example 1. We use `blue`, `red`, `green` to represent the choices $1, 2, 3$ respectively.

Definition 5. Let $(\mathcal{D}, \sim, f)$ be a rainbow graph. Then, for every $c \in \mathrm{Sym}(\mathcal{V})$, we denote $B^c = \{d \in \mathcal{D} : f(d) = c\}$. The interior of B^c is the set

$$(B^c)^\circ = \left\{d \in B^c : d \sim d' \Rightarrow d' \in B^c\right\} \tag{3.1}$$

and its boundary is the set

$$\partial B^c = B^c - (B^c)^\circ. \tag{3.2}$$

We next study optimal DP mechanisms given a rainbow boundary condition. Since dominance (between DP mechanisms) is a partial order, there exists at most one optimal DP mechanism. In the binary case, it is known that when there exists a valid DP mechanism on the boundary, then there exists an optimal DP mechanism on the whole graph [14]. Surprisingly, this is no longer the case when there are more than two colors, as shown in the following example.

Example 1. Let $\mathcal{D} = \{d_1, d_2, d_3, d_4, d_5\}$, with neighboring relations $d_1 \sim d_2 \sim d_3 \sim d_4 \sim d_5 \sim d_1$. Let $e^\epsilon = 2$, $\delta = 0$. Suppose the output space $\mathcal{V} = \{1, 2, 3\}$, and the preference function $f(d_i) = (1, 2, 3)$ for $1 \le i \le 4$ and $f(d_5) = (1, 3, 2)$. Consider the rainbow $c = (1, 2, 3)$. Then $B^c = \{d_1, d_2, d_3, d_4\}$ with boundary $\partial B^c = \{d_1, d_4\}$ and interior vertices $(B^c)^\circ = \{d_2, d_3\}$. Suppose the boundary condition is such that $\mathcal{M}(d_1) = (0.2, 0.1, 0.7)$ and $\mathcal{M}(d_4) = (0.4, 0.1, 0.5)$. We claim that under this setting, there exist valid DP mechanisms, but there is no optimal DP mechanism.

First, we define two valid DP mechanisms $\mathcal{M}_1$ and $\mathcal{M}_2$. Let $\mathcal{M}_1(d_2) = \mathcal{M}_1(d_3) = (0.4, 0.2, 0.4)$, $\mathcal{M}_2(d_2) = (0.4, 0.1, 0.5)$, and $\mathcal{M}_2(d_3) = (0.7, 0.05, 0.25)$. It is straightforward to verify that $\mathcal{M}_1$ and $\mathcal{M}_2$ are both valid DP mechanisms. Suppose, for the sake of illustrating a contradiction, that $\mathcal{M}_3$ is an optimal DP mechanism. Because $\mathcal{M}_3 \succeq \mathcal{M}_1$, we have $\mathcal{M}_3(d_2) \succeq \mathcal{M}_1(d_2)$. Because of the boundary condition $\mathcal{M}(d_1) = (0.2, 0.1, 0.7)$, we must have $\mathcal{M}_3(d_2) = \mathcal{M}_1(d_2)$. Because $\mathcal{M}_3 \succeq \mathcal{M}_2$, we have $\mathcal{M}_3(d_3) \succeq \mathcal{M}_2(d_3)$. Because of the boundary condition $\mathcal{M}(d_4) = (0.4, 0.1, 0.5)$, we must have $\mathcal{M}_3(d_3) = \mathcal{M}_2(d_3)$. So we have fully determined $\mathcal{M}_3$. However, $\mathcal{M}_3$ is not a valid DP mechanism because $\mathcal{M}_3(d_2)_2 > e^\epsilon \mathcal{M}_3(d_3)_2 + \delta$. Thus, there exists no optimal DP mechanism.

The main issue in the above example is that the boundary condition is not homogeneous. That is, the conditions on datasets d_1 and d_4 are different. We next define a homogeneity condition for DP mechanisms, generalizing [14] to non-binary functions.

Definition 6. A mechanism $\mathcal{M} : \mathcal{D} \to \mathcal{V}$ is *boundary homogeneous* if, for every rainbow $c \in \mathrm{Sym}(\mathcal{V})$, it holds that any two boundary datasets $d, d' \in \partial B^c$ satisfy the condition $\mathbb{P}[\mathcal{M}(d) = v] = \mathbb{P}[\mathcal{M}(d') = v]$ for every $v \in \mathcal{V}$.

We next provide our main result, which shows that under a valid homogeneous boundary condition, there exists a unique optimal DP mechanism and this optimal DP mechanism can be expressed in closed form.

Theorem 2. Let $(\mathcal{D}, \sim, f)$ be a rainbow graph and, for every rainbow $c \in \mathrm{Sym}(\mathcal{V})$ and $d \in \partial B^c$, let $\vec{m}^c \in \Delta(\mathcal{V})$ be a fixed homogeneous boundary condition. Suppose that the boundary condition is valid, i.e., for every pair $d \sim d' \in \bigcup_{c \in \mathrm{Sym}(\mathcal{V})} \partial B^c$, we have that $\vec{m}^{f(d)}$ and $\vec{m}^{f(d')}$ are (ϵ, δ)-close to each other. Then, there exists a unique optimal (ϵ, δ)-DP mechanism satisfying the boundary condition.

Proof of Theorem 2 and description of the optimal DP mechanism is delayed to Section 4 below. We next define the notion of a line graph, which is used in the proof.

Definition 7. Let $c \in \mathrm{Sym}(\mathcal{V})$ be a rainbow and $n \in \mathbb{N}$. The (c, n)-line is the rainbow graph $(\mathcal{D}, \sim, f)$ with datasets $\mathcal{D} = [0 : n]$, neighboring relation $i \sim j$ if $|i - j| = 1$, and preference function $f(d) = c$ for every $d \in \mathcal{D}$. Dataset $0 \in \mathcal{D}$ is considered the boundary of the line.

The other new notion we need for the proof of Theorem 2 is that of the boundary rainbow graph of a rainbow graph, defined next.

Definition 8. Let $(\mathcal{D}, \sim, f)$ be a rainbow graph. We define its boundary rainbow graph $(\mathcal{D}_\partial, \overset{\partial}{\sim}, f_\partial)$ as follows. For $c \in \mathrm{Sym}(\mathcal{V})$, let

$$d_c = \max_{d \in \mathcal{D} : f(d) = c} \mathrm{dist}(d, \partial B^c), \tag{3.3}$$

where we define $\mathrm{dist}(x, S) = \min_{y \in S} \mathrm{dist}(x, y)$. That is, d_c is the maximum distance of a dataset with preference c to the boundary ∂B^c. Let $\mathcal{D}_\partial = \{(c, i) : c \in \mathrm{Sym}(\mathcal{V}), i \in [0 : \ d_c]\}$ and $f_\partial((c, i)) = c$. Define $(c, i) \overset{\partial}{\sim} (c, i + 1)$ for $i \in [0 : d_c - 1]$, and $(c, 0) \overset{\partial}{\sim} (c', 0)$ (for $c, c' \in \mathrm{Sym}(V)$ and $c \neq c'$) if there exists $d \in B^c$, $d' \in B^{c'}$ such that $d \sim d'$. In other words, for each preference c there is a chain with $d_c + 1$ vertices, with $(c, 0)$ being the head and (c, d_c) being the tail. There is an edge between two heads $(c, 0)$, $(c', 0)$ if and only if there are two adjacent datasets, one with preference c, the other with preference c'.

Note that the boundary rainbow graph is a union of (c, d_c)-lines for $c \in \mathrm{Sym}(V)$ with a few possible additional edges between the endpoints $\{(c, 0) : c \in \mathrm{Sym}(\mathcal{V})\}$; see Fig. 1 above. Therefore, the boundary rainbow graph consists of a series of line graphs, each for a different rainbow occurring in the original graph. We define the boundary morphism $g_\partial : \mathcal{D} \to \mathcal{D}_\partial$ by sending $d \in \mathcal{D}$ to $(f(d), \mathrm{dist}(d, \partial B^{f(d)})) \in \mathcal{D}_\partial$. We now show that optimal mechanisms for boundary-homogeneous rainbow graphs can be obtained by pulling them back via Theorem 1 from their boundary rainbow graphs.

Theorem 3. Let $(\mathcal{D}, \sim, f)$ be a rainbow graph and $\mathcal{M}_\partial : \mathcal{D}_\partial \to \mathcal{V}$ be the optimal (ϵ, δ)-DP mechanism on its boundary rainbow graph subject to fixed boundary probabilities. Then, the pullback $\mathcal{M} = \mathcal{M}_\partial \circ g_\partial$, where g_∂ is the boundary morphism, is the optimal boundary homogeneous (ϵ, δ)-DP mechanism subject to the same boundary probabilities.

Proof. From Theorem 1, it follows that the morphism $g_\partial : \mathcal{D} \to \mathcal{D}_\partial$ induces an (ϵ, δ)-DP mechanism on $\mathcal{D}$ defined by $\mathcal{M}_\partial \circ \ g_\partial$. This mechanism is clearly boundary homogeneous. Let $\mathcal{M}$ be a valid (ϵ, δ)-DP mechanism satisfying the boundary condition. We prove that $\mathcal{M} \preceq \mathcal{M}_\partial \circ g_\partial$.

Let $B^c \subseteq \mathcal{D}$ be the subset of datasets with the same preference function. Let d_0 be the closest dataset in ∂B^c to $d \in B^c$. Let $G = \{d, d_{\mathrm{dist}(d,\partial B^c)-1}, \ldots, d_0\}$ be a set of datasets which forms a shortest path from d to d_0. Since $g_\partial|_G$ is injective, it has a left inverse, which we denote as $h : g_\partial(G) \to G$. However, h is a morphism and, therefore, from Theorem 1, $\mathcal{M} \circ h$ is an (ϵ, δ)-DP mechanism on $\mathcal{D}_\partial$. Then, since $\mathcal{M}_\partial$ is the optimal mechanism on $\mathcal{D}_\partial$, it follows that $\mathcal{M}_\partial|_{g_\partial(G)}$ is the optimal mechanism on $g_\partial(G)$. Thus, $\overrightarrow{\mathcal{M}_\partial \circ g_\partial}(d) \succeq \vec{\mathcal{M}}(d)$. Because the choice of d is arbitrary, we obtain $\mathcal{M} \preceq \mathcal{M}_\partial \circ g_\partial$, as desired. □

Since a boundary rainbow graph consists of a series of line graphs, the problem of finding optimal mechanisms can be reduced to finding them for line graphs, which is discussed in the next section after comparing the two orderings considered.

4. Optimal Differentially Private Mechanisms for Line Graphs

In this section, we derive optimal DP mechanisms for line graphs and use this to prove Theorem 2, and describe optimal DP mechanisms for general graphs. In this section, we use the shorthand $q := |\mathcal{V}|$.

Theorem 4. For any (c, n)-line graph $(\mathcal{D}, \sim, f)$ with boundary condition $\vec{m}$, i.e.,

$$\mathbb{P}[\mathcal{M}(0) = k] = m_k \quad \text{for all} \quad 1 \le k \le q, \tag{4.1}$$

there exists a unique optimal DP mechanism. Furthermore, under the unique optimal DP mechanism, $\mathcal{M}(d)$, $1 \le d \le n$, has distribution $T^d_{\epsilon,\delta}(\vec{m})$, where $T_{\epsilon,\delta} : \Delta(\mathcal{V}) \to \Delta(\mathcal{V})$ maps $p \in \Delta(\mathcal{V})$ to $p' \in \Delta(\mathcal{V})$ defined as follows:

$$p'_k = s'_k - s'_{k-1}, \tag{4.2}$$

where we have

$$s'_k = \min\{1, \min\{e^\epsilon s_k, 1 - e^{-\epsilon}(1 - s_k)\} + \delta\} \quad \text{for} \quad 1 \le k \le q. \tag{4.3}$$

In the above, $s_0 = s'_0 = 0$ and $s_k = \sum_{1 \le i \le k} p_i$ for $1 \le k \le q$.

We next describe the main results we require to prove Theorem 4, whose proof is given at the end of this section. Theorem 4 is used to prove Theorem 2, whose proof is also given at the end of this section. Now, without loss of generality, we assume that $c = (1, \ldots, |\mathcal{V}|)$, and $\mathcal{V} = \{1, \ldots, q\}$. Our key lemma is the following.

Lemma 1. Given a distribution $p \in \Delta(\mathcal{V})$, $p' = T_{\epsilon,\delta}(p)$ is the unique distribution such that
(1) p' is (ϵ, δ)-close to p;
(2) for any p'' that is (ϵ, δ)-close to p, we have $p' \succeq p''$.

Proof. We claim that p' defined above satisfies both conditions of Lemma 1.

Step 0. We verify that p' is a valid distribution. Because $s_q = 1$, we have $s'_q = 1$. Because s_k is monotone increasing in k and both functions $e^\epsilon x$ and $1 - e^{-\epsilon}(1 - x)$ are monotone increasing in x, s'_k is monotone increasing in k. So p' is a valid distribution.

Step 1. We verify that p' is (ϵ, δ)-close to p. We define another distribution $\widetilde{p}$. Let

$$\widetilde{s}_k = \min\{e^\epsilon s_k, 1 - e^{-\epsilon}(1 - s_k)\} \quad \text{for} \quad 1 \le k \le q \tag{4.4}$$

and

$$\widetilde{p}_k = \widetilde{s}_k - \widetilde{s}_{k-1}. \tag{4.5}$$

Note that $\widetilde{p}$ is a valid distribution and $\mathrm{TV}(\widetilde{p}, p') \le \delta$. It suffices to prove that $\widetilde{p}$ is $(\epsilon, 0)$-close to p.

We first prove that for any $1 \le k \le q$, we have

$$e^{-\epsilon} p_k \le \widetilde{p}_k \le e^{\epsilon} p_k. \tag{4.6}$$

Note that $(e^{\epsilon} s_k - (1 - e^{-\epsilon}(1 - s_k)))$ is monotone increasing in k. So there exists $k_0 \in \{0, \dots, q\}$ such that

$$\widetilde{s}_k = \begin{cases} e^{\epsilon} s_k & \text{if } k \le k_0, \\ 1 - e^{-\epsilon}(1 - s_k) & \text{if } k \ge k_0 + 1. \end{cases} \tag{4.7}$$

Then,

(1) for all $k \le k_0$, we have
$$\widetilde{p}_k = e^{\epsilon}(s_k - s_{k-1}) = e^{\epsilon} p_k;$$
(2) for all $k \ge k_0 + 2$, we have
$$\widetilde{p}_k = (1 - e^{-\epsilon}(1 - s_k)) - (1 - e^{-\epsilon}(1 - s_{k-1})) = e^{-\epsilon} p_k;$$
(3) for $k = k_0 + 1$, we have
$$\widetilde{p}_k = (1 - e^{-\epsilon}(1 - s_k)) - e^{\epsilon} s_{k-1} \le e^{\epsilon} s_k - e^{\epsilon} s_{k-1} = e^{\epsilon} p_k$$
and
$$\begin{aligned} \widetilde{p}_k &= (1 - e^{-\epsilon}(1 - s_k)) - e^{\epsilon} s_{k-1} \\ &\ge (1 - e^{-\epsilon}(1 - s_k)) - (1 - e^{-\epsilon}(1 - s_{k-1})) = e^{-\epsilon} p_k. \end{aligned}$$

This proves that $\widetilde{p}$ is $(\epsilon, 0)$-close to p.

Step 2. We next prove that for any p'' that is (ϵ, δ)-close to p, we have $p' \succeq p''$. Because p'' is (ϵ, δ)-close to p, there exists a distribution $p^\sharp$ such that $\mathrm{TV}(p'', p^\sharp) \le \delta$ and that $p^\sharp$ is $(\epsilon, 0)$-close to p. Let $s^\sharp$ be the prefix sum of $p^\sharp$ and s'' be the prefix sum of p''.

For any $1 \le k \le q$, taking $\mathcal{S} = \{1, \dots, k\}$ in the DP condition (Definition 2) on $(p, p^\sharp)$, we obtain

$$s^\sharp_k \le e^{\epsilon} s_k. \tag{4.8}$$

Taking $\mathcal{S} = \{k + 1, \dots, q\}$, we have

$$s^\sharp_k \le 1 - e^{-\epsilon}(1 - s_k). \tag{4.9}$$

Thus, we have $s^\sharp_k \le \widetilde{s}_k$. Furthermore, because $\mathrm{TV}(p'', p^\sharp) \le \delta$, we have

$$s''_k \le \min\{1, s^\sharp_k + \delta\} \le \min\{1, \widetilde{s}_k + \delta\} = s'_k. \tag{4.10}$$

Because k can be any integer between 1 and q, this proves that $p' \succeq p''$. □

Lemma 2. $T_{\epsilon,\delta}$ preserves dominance ordering, i.e., $p \succeq p' \implies T_{\epsilon,\delta}(p) \succeq T_{\epsilon,\delta}(p')$.

Proof. This holds because $\min\{1, \min\{e^{\epsilon} x, 1 - e^{-\epsilon}(1 - x)\} + \delta\}$ is a monotone increasing function in x. □

Now, we are ready to prove Theorem 4.

Proof of Theorem 4. Let $\mathcal{M}$ be the DP mechanism defined as in the theorem statement. By Lemma 1 part (1), $\mathcal{M}$ is an (ϵ, δ)-DP mechanism. Now let us prove its optimality. Let $\mathcal{M}'$ be another (ϵ, δ)-DP mechanism with the same boundary condition.

Let us use induction on d to prove that $\mathcal{M}(d) \succeq \mathcal{M}'(d)$ for all $d \in [0:n]$. Because of the boundary condition, we have $\mathcal{M}(0) = \mathcal{M}'(0)$. In particular, $\mathcal{M}(0) \succeq \mathcal{M}'(0)$. This is the base case of our induction. Now suppose that we have proved $\mathcal{M}(d) \succeq \mathcal{M}'(d)$ for some $d \in [0:n-1]$. Then

$$\mathcal{M}(d+1) = T_{\epsilon,\delta}(\mathcal{M}(d)) \succeq T_{\epsilon,\delta}(\mathcal{M}'(d)) \succeq \mathcal{M}'(d+1), \tag{4.11}$$

where the first step is by definition of $\mathcal{M}$, the second step is by Lemma 2 and induction hypothesis, the third step is by Lemma 1. This completes the induction. Therefore $\mathcal{M}(d) \succeq \mathcal{M}'(d)$ for all $d \in [1:n]$. □

Using Theorem 4, we next prove Theorem 2.

Proof of Theorem 2. The proof is straightforwardly obtained by combining the results of Theorems 3 and 4. Theorems 3 shows that to construct an optimal DP mechanism for a general graph under a homogeneous boundary condition, it suffices to construct an optimal DP mechanism for the boundary rainbow graph, which is achieved by Theorem 4. □

By expanding the proof of Theorem 2, we can describe the optimal DP mechanism for general graphs as follows. For any $c \in \mathrm{Sym}(\mathcal{V})$, $d \in B^c$, let $\vec{m}^c$ be the boundary condition. Define $\widetilde{m}^c \in \Delta(\mathcal{V})$ as $\widetilde{m}^c_i = \vec{m}^c_{c(i)}$ for $1 \le i \le q$. This is a permuted version of $\vec{m}^c$ such that the preference order is $(1, \dots, |\mathcal{V}|)$. Then the optimal DP mechanism is given by $\mathbb{P}[\mathcal{M}(d) = c(k)] = \left(T_{\epsilon,\delta}^{\mathrm{dist}(d,\partial B^c)}(\widetilde{m}^c)\right)_k$, for $1 \le k \le q$.

5. Designing the Optimal Differentially Private Mechanism

In this section, we discuss how to design optimal DP mechanisms. Throughout, we consider a (c, n)-line with $c = (1, \dots, |\mathcal{V}|)$ and boundary condition $\vec{m} \in \Delta(\mathcal{V})$, i.e., $\mathbb{P}[\mathcal{M}(0) = k] = m_k$. The case $c \neq (1, \dots, |\mathcal{V}|)$ can be handled by performing a permutation, as in the end of Section 4. For $t \in [n]$, the distribution of $\mathcal{M}(t)$ can be straightforwardly computed using the definition of $T_{\epsilon,\delta}$ (see Theorem 4 above). In fact, it can be expressed in an even more explicit form, as we explain below.

5.1. Special Case $\delta = 0$.

As a warm up, let us first consider the case $\delta = 0$. By Theorem 4, for the optimal DP mechanism, the distribution of $\mathcal{M}(t)$ is equal to $T^t_{\epsilon,\delta}(\vec{m})$. For $t \in \mathbb{Z}_{\ge 0}$, let s^t denote the prefix sum of $T^t_{\epsilon,\delta}(\vec{m})$. By construction of the operator $T_{\epsilon,\delta}$, we have

(1) if $s^t_k \le 1/(e^\epsilon + 1)$, then
$$s^{t+1}_k = e^\epsilon s^t_k;$$
(2) if $s^t_k \ge 1/(e^\epsilon + 1)$, then
$$s^{t+1}_k = 1 - e^{-\epsilon}(1 - s^t_k).$$

For $1 \le k \le q$, define

$$\tau_k = \left\lfloor \max\left\{ -\frac{\log(s_k^0 \cdot (e^\epsilon + 1))}{\epsilon} + 1, 0 \right\} \right\rfloor. \tag{5.1}$$

Then we have $s_k^t \le 1/(e^\epsilon + 1)$ for $t \le \tau_k - 1$ and $s_k^t \ge 1/(e^\epsilon + 1)$ for $t \ge \tau_k$. Thus, we obtain as the solution

$$s_k^t = \begin{cases} e^{t\epsilon} s_k^0 & \text{if } t \le \tau_k, \\ 1 - e^{-\epsilon(t-\tau_k)}(1 - s_k^{\tau_k}) & \text{if } t \ge \tau_k + 1. \end{cases} \tag{5.2}$$

Then, we obtain $\mathbb{P}[\mathcal{M}(t) = k] = s_k^t - s_{k-1}^t$.

Note that in the case of three colors ($q = 3$), our expression recovers the one given in [16].

5.2. **General Case $\delta > 0$.** Now, we consider the case $\delta > 0$. The formula is slightly more complicated, but the derivation method is similar. Recall that $\vec{m}$ is the boundary mechanism. Let s^t denote the prefix sum of $T_{\epsilon,\delta}^t(\vec{m})$. Define

$$\rho = \frac{\delta}{e^\epsilon - 1}. \tag{5.3}$$

By construction of the operator $T_{\epsilon,\delta}$, we have

(1) if $s_k^t \le 1/(e^\epsilon + 1)$, then

$$s_k^{t+1} = \min\{1, e^\epsilon s_k^t + \delta\}$$

which is equal to

$$s_k^{t+1} + \rho = \min\{e^\epsilon(s_k^t + \rho), 1 + \rho\};$$

(2) if $s_k^t \ge 1/(e^\epsilon + 1)$, then

$$s_k^{t+1} = \min\{1, 1 - e^{-\epsilon}(1 - s_k^t) + \delta\}$$

which is equal to

$$1 - s_k^{t+1} + e^\epsilon \rho = \max\{e^{-\epsilon}(1 - s_k^t + e^\epsilon \rho), e^\epsilon \rho\}.$$

Define

$$\tau_k = \left\lfloor \max\left\{ \frac{\log\left(\left(\frac{1}{e^\epsilon + 1} + \rho\right) / \left(s_k^0 + \rho\right)\right)}{\epsilon} + 1, 0 \right\} \right\rfloor. \tag{5.4}$$

Then we have $s_k^t \le 1/(e^\epsilon + 1)$ for $t \le \tau_k$ and $s_k^t \ge 1/(e^\epsilon + 1)$ for $t \ge \tau_k + 1$. Solving this, we obtain

$$s_k^t = \begin{cases} \min\{1, e^{t\epsilon}(s_k^0 + \rho) - \rho\} & \text{if } t \le \tau_k, \\ \min\{1, 1 + e^\epsilon \rho - e^{-\epsilon(t-\tau_k)}(1 - s_k^{\tau_k} + e^\epsilon \rho)\} & \text{if } t \ge \tau_k + 1. \end{cases} \tag{5.5}$$

Finally, we have $\mathbb{P}[\mathcal{M}(t) = k] = s_k^t - s_{k-1}^t$.

When $\delta = 0$, our expressions reduce to the one we derived in Section 5.1.

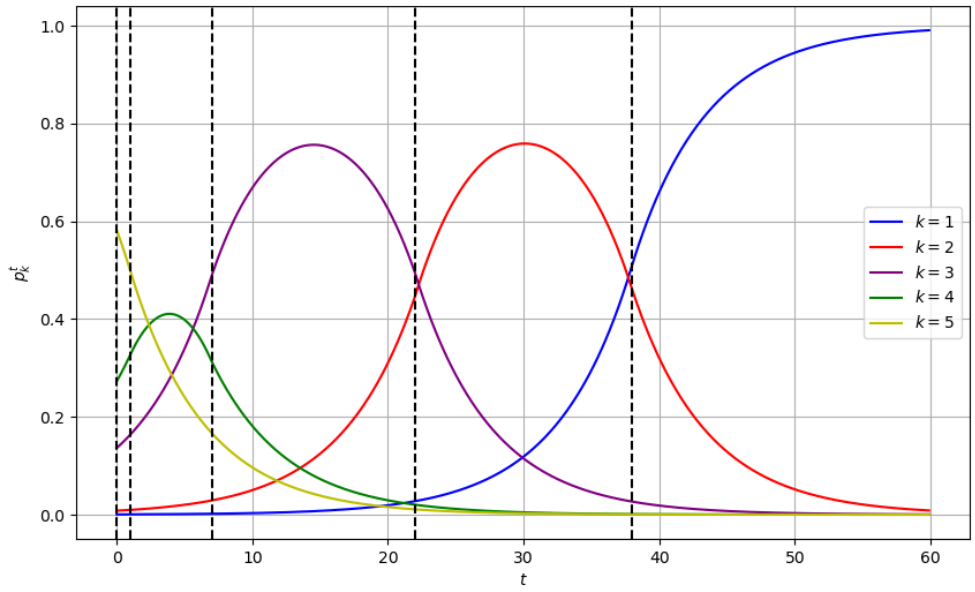


FIG. 3. The optimal $(\log(1.2), 0)$-DP mechanism with homogeneous boundary condition $\vec{m} = (0.0005, 0.0081, 0.1364, 0.2727, 0.5822)$. We have $\tau_1 = 38, \tau_2 = 22, \tau_3 = 7, \tau_4 = 1$, and $\tau_5 = 0$.

5.3. **Numerical Results.** We depict several examples for optimal DP mechanism designs in Fig. 3, 4, and 5. In these examples, we consider $|\mathcal{V}| = 5$, $c = (1, \dots, 5)$, and $\epsilon = \log 1.2$. We choose the boundary condition to be

$$\vec{m} = (0.0005, 0.0081, 0.1364, 0.2727, 0.5822)$$

which corresponds to $0.0005 = 0.001 \times \frac{e^\epsilon}{1+e^\epsilon}$, $0.0081 = 0.015 \times \frac{e^\epsilon}{1+e^\epsilon}$, $0.1364 = 0.25 \times \frac{e^\epsilon}{1+e^\epsilon}$, $0.2727 = 0.5 \times \frac{e^\epsilon}{1+e^\epsilon}$, and 0.5822 is 1 minus all the other values.

Fig. 3 illustrates the optimal (ϵ, δ)-DP mechanism for $\delta = 0$. At integer t, the figure shows $T^t_{\epsilon,\delta}\vec{m}$. At non-integer t, the values are interpolated using (5.2). By (5.2), we observe that for fixed k, s_k^t goes through a phase transition at τ_k. For $t \le \tau_k$, s_k^t increases exponentially and for $t \ge \tau_k$, $1 - s_k^t$ decreases exponentially, respectively. By (5.1), τ_k is non-increasing in k. Therefore, the probabilities $\mathbb{P}[\mathcal{M}(t) = k] = s_k^t - s_{k-1}^t$ goes through at most three phases as t increases. We have the following results:

- in the first phase with $t \le \tau_k$, s_k^t and s_{k-1}^t both increase exponentially;
- in the second phase with $\tau_k \le t \le \tau_{k-1}$, s_{k-1}^t increases exponentially, while $1 - s_k^t$ decreases exponentially;
- in the third phase with $t \ge \tau_{k-1}$, $1 - s_k^t$ and $1 - s_{k-1}^t$ both decrease exponentially.

When $\tau_k = 0$, the first phase is degenerate (i.e., has length 0); when $\tau_k = \tau_{k-1}$, the second phase is degenerate; and when $\tau_{k-1} = \infty$ (i.e., $s_{k-1}^0 = 0$), the third phase is degenerate. In Fig. 3, the first phase for $k = 5$ is degenerate (no phase transition), and the third phase for $k = 1$ is degenerate.

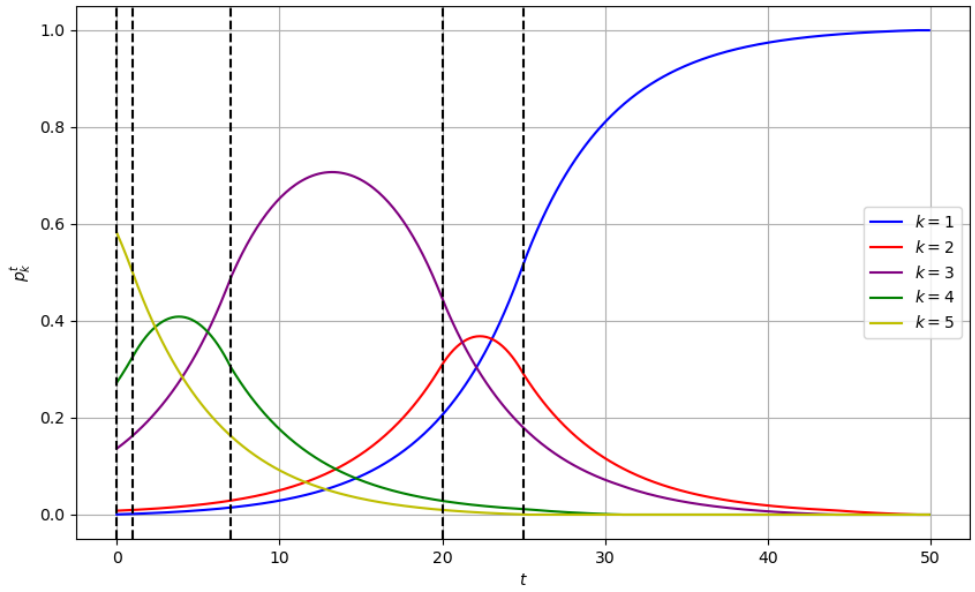


FIG. 4. The optimal $(\log(1.2), 10^{-3})$-DP mechanism with homogeneous boundary condition $\vec{m} = (0.0005, 0.0081, 0.1364, 0.2727, 0.5822)$. We have $\tau_1 = 25, \tau_2 = 20, \tau_3 = 7, \tau_4 = 1$, and $\tau_5 = 0$.

Secondly, Fig. 4 shows the optimal DP mechanisms with the same parameters as Fig. 3, but with $\delta = 10^{-3}$. At integer t, the figure shows $T^t_{\epsilon,\delta}\vec{m}$. At non-integer t, the values are interpolated using (5.5). From (5.5) we see that for fixed k, s^t_k goes through a phase transition at τ_k. For $t \le \tau_k$, s^t_k increases exponentially (with a drift); and for $t \ge \tau_k$, $1 - s^t_k$ decreases exponentially (with a drift). The phase transition behavior is similar to the $\delta = 0$ case.

Finally, Fig. 5 shows the optimal DP mechanisms with the same parameters as Fig. 3 and Fig. 4, but with $\delta = 0.01$. In this case, the phase transitions happen earlier (i.e., the τ's are smaller than previous examples).

6. Discussions

6.1. **Contributions.** In this paper, we presented optimal rainbow DP mechanisms given valid homogeneous boundary conditions for any finite query output sizes by using a new proof technique.

We remark that it is a priori not clear whether an optimal mechanism exists for either homogeneous or non-homogeneous boundary conditions. In fact, our Example 1 shows that there exist examples with an inhomogeneous boundary condition where there are valid mechanisms but there is no optimal mechanism. Therefore, we believe it is an interesting and non-trivial result to show that under homogeneous boundary condition, an optimal mechanism exists, and it is optimal under a wide range of utility functions. Furthermore, we give an explicit formula for the optimal mechanism. As shown in previous works [14, 16], as

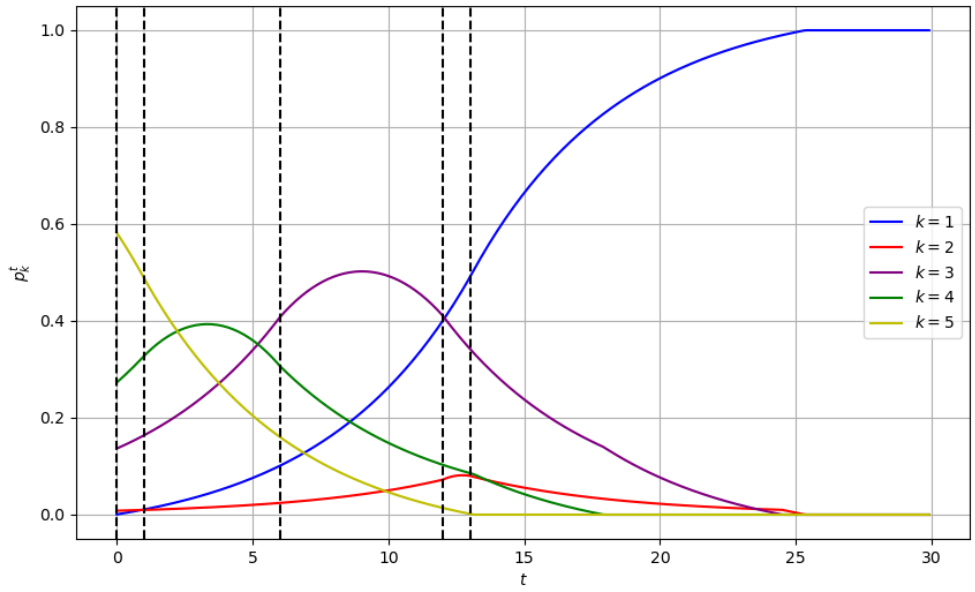


FIG. 5. The optimal $(\log(1.2), 0.01)$-DP mechanism with the homogeneous boundary condition $\vec{m} = (0.0005, 0.0081, 0.1364, 0.2727, 0.5822)$. We have $\tau_1 = 13, \tau_2 = 12, \tau_3 = 6, \tau_4 = 1$, and $\tau_5 = 0$.

the number of colors grows, the complexity of the optimal DP mechanism also increases, and it is a priori not clear whether the complexity will be out of reach as the number of colors become larger. Our result shows that the answer is no, as we give a uniform treatment for any number of colors.

6.2. **Dataset dependency.** Our mechanism is a dataset-dependent mechanism. In general, dataset dependency is desirable from the perspective of DP mechanism design because it allows more room for mechanism design and may lead to improved utility. In our rainbow DP setting, a dataset-independent mechanism would be suboptimal because such a mechanism will not be aware of whether a dataset is close to or far away from the boundary. A dataset-dependent mechanism such as our optimal mechanism can be more aggressive when choosing the output distribution for datasets far away from the boundary. Our result advocates using data-dependent mechanisms in DP mechanism design.

6.3. **Lexicographic vs. Dominance Ordering.** In [16], optimal rainbow DP mechanisms under lexicographic ordering are studied, while we focus on dominance ordering. We briefly discuss the relationship between the two orderings.

On the probability simplex $\Delta(\mathcal{V})$, dominance ordering is strictly stronger than lexicographic ordering. Therefore, a rainbow DP mechanism that is optimal under dominance ordering is also optimal under lexicographic ordering. As we show in Theorem 2 above, under homogeneous boundary conditions, there exists a unique optimal rainbow DP mechanism

under dominance ordering, which implies the same result for lexicographic ordering. However, under non-homogeneous boundary conditions, as we illustrate in Example 1 above, there exist scenarios for which there is no optimal rainbow DP mechanism, under either lexicographic ordering or dominance ordering. An interesting question is whether scenarios exist for which there is an optimal rainbow DP mechanism under lexicographic ordering, but no optimal mechanism under dominance ordering. Our efforts in this direction have not resulted in any such scenario. We know that such an example, if it exists, should have a non-homogeneous boundary condition. However, it seems difficult to control a DP mechanism to be optimal under lexicographic ordering, unlike under dominance ordering.

6.4. **Comparison with exponential mechanism.** We compare the exponential mechanism with our mechanism on the rainbow DP problem. First of all, because our mechanism is optimal, any mechanism produced using the exponential mechanism cannot be better than ours. So the crux of the question is whether our mechanism can be produced using the exponential mechanism. We note that a dataset-independent version cannot produce our mechanism, because our mechanism is dataset-dependent. If we consider dataset-dependent versions, then the question is how to design the score function used in the exponential mechanism. A clever design of the score function could give the same mechanism as ours, but it seems that to find such a clever design, it is necessary to use proofs similar to ours, rather than using the usual proof of utility for exponential mechanisms. Therefore, using exponential mechanism in this setting seems to have little benefit.

It is an interesting question whether our mechanism can be applied to other problems such as median queries. It might be related to a continuous generalization we will discuss below. We would leave this direction for future work.

6.5. **Further directions.** A possible extension to our result is the case of continuous alphabets (e.g., an interval $[0, 1]$). For example, when we have $\mathcal{V} = [0, 1]$ and smaller numbers are preferred over larger numbers, we could replace dominance ordering with a suitable generalization (e.g., first-order stochastic dominance [17]), for which case generalizations of our main results (Theorems 2 and 4) are expected to hold. We leave this direction for further research.

Acknowledgment

This work has been supported in part by NSF DMS-1926686, the German Federal Ministry of Education and Research (BMBF) within the National Initiative on 6G Communication Systems through the Research Hub *6G-life* under Grant 16KISK001K, the German Research Foundation (DFG) as part of Germany's Excellence Strategy – EXC 2050/1 – Project ID 390696704 – Cluster of Excellence *"Centre for Tactile Internet with Human-in-the-Loop" (CeTI)* of Technische Universität Dresden, the ARC Future Fellowship FT190100429, by the ELLIIT funding endowed by the Swedish government, by the ZENITH Research and Leadership Career Development Award fund, by NSF RINGS-2148132, and NSF CNS-2008624. We thank anonymous reviewers for helpful suggestions and comments. We thank Ying Zhao for pointing out a numerical issue in a figure in a previous version of the paper.

References

[1] I. Mironov, O. Pandey, O. Reingold, and S. Vadhan, "Computational differential privacy," in *Proc. Int. Cryptology Conf.*, (Santa Barbara, CA), pp. 126–142, Aug. 2009. doi: 10.1007/978-3-642-03356-8_8.

[2] C. Dwork, F. McSherry, K. Nissim, and A. Smith, "Calibrating noise to sensitivity in private data analysis," *Journal of Privacy and Confidentiality*, vol. 7, no. 3, pp. 17–51, 2016. doi: 10.29012/jpc.v7i3.405.

[3] C. Dwork, "Differential privacy," in *Proc. Int. Colloq. Automata Lang. Program.*, (Venice, Italy), pp. 1–12, July 2006. doi: 10.1007/11787006_1.

[4] C. Dwork, A. Roth, *et al.*, "The algorithmic foundations of differential privacy," *Foundations and Trends® in Theoretical Computer Science*, vol. 9, no. 3–4, pp. 211–407, 2014. doi: 10.1561/0400000042.

[5] T. Zhu, G. Li, W. Zhou, and P. S. Yu, "Differentially private data publishing and analysis: A survey," *IEEE Trans. Knowl. Data Eng.*, vol. 29, pp. 1619–1638, Aug. 2017. doi: 10.1109/TKDE.2017.2697856.

[6] *Disclosure Avoidance and the 2020 Census*, 2020. www.census.gov/about/policies/privacy/statistical_safeguards/disclosure-avoidance-2020-census.html.

[7] Ú. Erlingsson, V. Pihur, and A. Korolova, "Rappor: Randomized aggregatable privacy-preserving ordinal response," in *Proc. ACM SIGSAC Conf. Comput. Commun. Security*, (New York, NY), pp. 1054–1067, Nov. 2014. doi: 10.1145/2660267.2660348.

[8] D. P. Team, *Learning with Privacy at Scale*. machinelearning.apple.com/2017/12/06/learning-with-privacy-at-scale.html, year=2017 (last accessed May 2021).

[9] B. Ding, J. Kulkarni, and S. Yekhanin, "Collecting telemetry data privately," in *Advances in Neural Information Processing Systems* (I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, eds.), vol. 30, Curran Associates, Inc., 2017. https://proceedings.neurips.cc/paper_files/paper/2017/file/253614bbac999b38b5b60cae531c4969-Paper.pdf.

[10] J. Soria-Comas, J. Domingo-Ferrer, D. Sánchez, and D. Megías, "Individual differential privacy: A utility-preserving formulation of differential privacy guarantees," *IEEE Trans. Inf. Forensics Security*, vol. 12, pp. 1418–1429, June 2017. doi: 10.1109/TIFS.2017.2663337.

[11] K. Nissim, S. Raskhodnikova, and A. Smith, "Smooth sensitivity and sampling in private data analysis," in *Proc. ACM Symp. Theory Comput.*, (San Diego, CA), pp. 75–84, June 2007. doi: 10.1145/1250790.1250803.

[12] X. He, A. Machanavajjhala, and B. Ding, "Blowfish privacy: Tuning privacy-utility trade-offs using policies," in *Proc. ACM SIGMOD Int. Conf. Management Data*, (Snowbird, UT), pp. 1447–1458, 2014. doi: 10.1145/2588555.2588581.

[13] J. Geumlek and K. Chaudhuri, "Profile-based privacy for locally private computations," in *Proc. IEEE Int. Symp. Inf. Theory*, (Paris, France), pp. 537–541, July 2019. doi: 10.1109/ISIT.2019.8849549.

[14] R. G. L. D'Oliveira, M. Médard, and P. Sadeghi, "Differential privacy for binary functions via randomized graph colorings," in *Proc. IEEE Int. Symp. Inf. Theory*, (Melbourne, Victoria, Australia), pp. 473–478, July 2021. doi: 10.1109/ISIT45174.2021.9517935.

[15] N. Holohan, D. J. Leith, and O. Mason, "Optimal differentially private mechanisms for randomised response," *IEEE Trans. Inf. Forensics Security*, vol. 12, pp. 2726–2735, Nov. 2017. doi: 10.1109/TIFS.2017.2718487.

[16] Z. Zhou, O. Günlü, R. G. L. D'Oliveira, M. Médard, P. Sadeghi, and R. F. Schaefer, "Rainbow differential privacy," in *IEEE Int. Symp. Inf. Theory*, (Espoo, Finland), pp. 614–619, June/July 2022. doi: 10.1109/ISIT50566.2022.9834887.

[17] J. Hadar and W. R. Russell, "Rules for ordering uncertain prospects," *The American Economic Review*, vol. 59, no. 1, pp. 25–34, 1969. http://www.jstor.org/stable/1811090.

This work is licensed under the Creative Commons License Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0). To view a copy of this license, visit https://creativecommons.org/licenses/by-nc-nd/4.0/ or send a letter to Creative Commons, 171 Second St, Suite 300, San Francisco, CA 94105, USA, or Eisenacher Strasse 2, 10777 Berlin, Germany