On the Discrete Geometry of Differential Privacy via Ehrhart Theory*

November 3, 2017

Arun Padakandla
Dept. of Computer Science
Purdue University
W. Lafayette, IN 47907
U.S.A.
arunpr@purdue.edu

P.R. Kumar
Dept. of Elec. & Comp. Eng.
Texas A&M
College Station, TX 77843
U.S.A.
prk.tamu@gmail.com

Wojciech Szpankowski
Dept. of Computer Science
Purdue University
W. Lafayette, IN 47907
U.S.A.
szpan@purdue.edu

Abstract

We study a linear program (LP) endowed with rich combinatorial structure, the solution to which will provide characterization of a fundamental utility-privacy tradeoff in Differential Privacy (DP). This LP is of central interest and has been recognized to be notoriously hard, leading previous works to consider some relaxations. While these relaxations have yielded near-order-optimal results, the original problem has remained open. In this work, we solve this problem directly using tools from discrete geometry, combinatorics and analytic methods. For large database we provide a complete characterization of this tradeoff for the expected $L_1$-distortion via a simple closed form computable expression. In particular, we prove that the minimum expected $L_1$-distortion is a simple functional of the Ehrhart series of a suitable defined cross polytope - thereby establishing a connection between DP and Ehrhart theory. We leverage weak duality theorem to derive the lower bound. Our findings lead us to characterizing dual variables corresponding to the truncated geometric mechanism whose corresponding (dual) objective value is proven to yield an optimal lower bound. Going beyond order optimal characterization, our solution is of value to designing practical DP mechanisms. For practically viable values of the DP parameter, higher order terms cannot be ignored. Designing practical DP mechanisms involves fine tuning the parameters governing the DP constraint and our precise characterization of the utility as a function of this parameter trade-off aids such efforts.

Index Terms: Linear programming, discrete geometry, analytic combinatorics, Ehrhart theory, dual LP, stability, robustness, differential privacy.

*This work was supported by NSF Center for Science of Information (CSoI) Grant CCF-0939370, NSF Grants CCF-1524312, NSF ECCS-1646449, NIH Grant 1U01CA198941-01 and USARO under Contract W911NF-15-1-0279.
1 Problem Description, Contribution and Significance

Accurate statistical information extracted from databases yield useful information for informed decision making in terms of placement of optimal resources. However, extraction of such information has to respect privacy of subjects participating in the database. Differential Privacy [1, 2] has emerged as a popular mathematical framework to quantify the vulnerability of algorithms extracting statistical information to privacy breaches. In this work, we address the problem of characterizing a fundamental utility-privacy tradeoff arising in differentially private histogram sanitization mechanisms - a generic multi-dimensional query.

Problem Description: Our study is focused on a linear program (LP) endowed with rich mathematical structure, the solution to which will provide a fundamental utility-privacy tradeoff in Differential Privacy (DP), whose precise characterization has remained open. To fix ideas, and present our precise result, we begin with a description of the LP. Let

\[ H^n := \{(h_1, \ldots, h_K) \in \mathbb{Z}^K : h_k \geq 0, \sum_{k=1}^K h_k = n\} \tag{1} \]

be the set of possible histograms corresponding to databases storing records of \( n \) subjects, where each record can be chosen from one among \( K \) possible choices. A histogram sanitization mechanism (HSM) \( W : H^n \Rightarrow H^n \) is a randomized algorithm that takes a histogram in \( H^n \) as input, say \( h \in H^n \), and outputs a random histogram \( G \in H^n \) sampled with probability distribution (probability mass function (pmf)) \( (W(g|h) : g \in H^n) \). Our design variables are the pmfs \( (W(g|h) : g \in H^n) : h \in H^n \) of the HSM. The goal is to design these pmfs to meet pure DP constraints, while maximizing expected accuracy between the input and output histograms. We model a prior pmf for the input histogram. Each of the \( n \) subject’s records are chosen independently and identically with a generic pmf \( p = (p_1, \ldots, p_K) \) on the set of \( K \) possible records. The histogram corresponding to this random database has a multinomial pmf. The random input histogram \( H \in H^n \) is therefore sampled with a pmf \( (\frac{1}{n} h^K)^2 : (h_1, \ldots, h_K) \in \mathbb{N}^K \) : \( h \in H^n \). When input \( H \), the HSM produces output \( G \in H^n \), and the accuracy between \( H, G \) is quantified through a distortion measure \( F : H^n \times H^n \rightarrow [0, \infty) \). Consider the LP

\[ D^n_(\theta, p, F) := \min_{W(\cdot)} \mathbb{E}_p \left\{ \mathbb{E}_W \{ F(H, G) \} \right\} = \sum_{g \in H^n} \sum_{h \in H^n} \left( n \right) \frac{h_{\lambda}}{h} \mathbb{W}(g|h)F(g, h) \text{ subject to} \tag{2} \]

\[ \sum_{g \in H^n} \mathbb{W}(g|h) = 1 \text{ for all } h \in H^n, \text{ (ii) } \mathbb{W}(g|h) - \theta \mathbb{W}(g|h) \geq 0 \text{ for all } (h, h) \in H^n \times H^n \text{ satisfying } |h - h| = 2 \text{ and all } g \in H^n, \text{ and (iii) } \mathbb{W}(g|h) \geq 0 \text{ for all } (g, h) \in H^n \times H^n. \tag{3} \]

The objective function (2) quantifies utility of the HSM, in terms of its accuracy and the constraint (3b) quantifies privacy in the framework of pure DP. Note that histograms \( h, h \in H^n \) satisfy \( |h - h| = 2 \) if and only if the corresponding databases differ in a single record, i.e. neighboring databases [3]. The constraints (3b) play a central role in DP [3] and a HSM whose pmfs \( \mathbb{W}(\cdot|\cdot) \) satisfy (3) is said to be a \( \theta \)-DP HSM. Essentially, (3b) which is equivalent to the constraint \( \theta \mathbb{W}(g|h) \leq \mathbb{W}(g|h) \leq \theta^{-1} \mathbb{W}(g|h) \), constrains the point-wise ratio of pmfs corresponding to databases differing in a single record, thereby limiting how much one can glean about any individual record, by observing the output. Solution to LP (2-3) therefore characterizes the minimum expected distortion of a \( \theta \)-DP HSM operating over databases with \( n \) records. Throughout, we assume \( \theta \in (0, 1) \) and \( p_k > 0 \) for \( k \in [K] \).

The discrete optimization problem (2-3), and in particular the limit \( D^n_K(\theta, p, F) = \lim_{n \to \infty} D^n(\theta, p, F) \) are the objects of central interest in this article. Our study will harness the structure of LP (2-3) using tools from discrete geometry, in particular Ehrhart theory and we provide a complete characterization for \( D^n_K(\theta, p, F) \) and its dependence on all orders of \( \theta \) via a power series - Ehrhart series of the convex polytope (5).

LP (2-3) is of central interest in DP. This is evidenced by several prior works that have studied similar instances. Hardt and Talwar [4, Sec 2.4] formulate the minimum expected error
as an LP equivalent to (2-3), recognize its complexity and state “· · · these linear programs can be prohibitive in size. Moreover, it is not apriori clear how one can use this formulation to understand the asymptotic behavior of the error of the optimum mechanism. · · ·”, thus leaving the solution to (2-3) open. [4] focuses on the utility maximization problem in the min-max setting and bring to light novel techniques based on Markov inequality-type bounds to derive near-order-optimal lower bounds for the continuous version of the corresponding min-max optimization problem. Our work may be viewed as providing an answer to the question left in [4]. Ghosh, Roughgarden and Sundararajan [5] focus on LP (2-3) for the particular \( K = 2 \) case and study the structure of an optimal mechanism for a general class of symmetric loss measures. Our work puts forth a notion of asymptotic generalization of universal optimality [5]. Though a weakening, it provides an approach to get beyond impossibility results [6]. Furthermore, [7] studies LP (2-3) in the context of identifiability. The importance of LP (2-3) stems from the fact that a histogram is equivalent to a database (modulo permutations) and hence the expected accuracy of a mechanism responding to any statistical query can be modeled as an instantiation of LP (2-3) by an appropriate choice of distortion measure. Moreover the output of a HSM can be employed as an impenetrable first-step sanitization mechanism for further querying.

Results at a Glance: In spite of its central nature, there is no precise characterization of \( D^*_K(\theta, p, F) \) even for specific distortion measures and/or specific values of \( K \). Subject to \( \mathcal{O}(K^3(n + 1)^{2(K - 1)}) \) constraints imposed on \( \mathcal{O}((n + 1)^{2K}) \) decision variables, LP (2-3) is indeed notoriously hard, justifying the order optimal results [4] obtained via skillful relaxations. Here, we focus on the original discrete setting and prove for the popular \( \mathbb{L}_1 \)-metric that

\[
D^*_K(\theta, p, | \cdot |_1) = \lim_{n \to \infty} D^*(\theta, p, | \cdot |_1) = \frac{2\theta}{\mathcal{E}(\theta)} \frac{\mathcal{E}(\theta)}{d\theta} = \frac{2\theta}{1 - \theta},
\]

where \( \mathcal{E}(\theta) \) is the Ehrhart series of the cross-polytope whose \( d \)-th dilation is given by

\[
P_d = \{(x_1, \cdots, x_K) \in \mathbb{R}^K : \sum_{k=1}^K x_k = 0, \sum_{k=1}^K |x_k| \leq 2d\}.
\]

Simple combinatorial arguments enable us prove (Appendix 10) that

\[
\mathcal{E}(\theta) = \frac{1}{1 - \theta} + \sum_{d=1}^{\infty} \left\{ \sum_{r=1}^{K-1} \binom{K}{r} \binom{d + r - 1}{r - 1} \binom{d - 1}{K - r - 1} \right\} \frac{\theta^d}{1 - \theta}.
\]

Observe that (4) in conjunction with (6) provides us with a characterization for the minimum expected \( \mathbb{L}_1 \)-fidelity in terms of a power series in \( \theta \) thereby providing us with its dependence on every order of \( \theta \), and hence answer the question left open in [4]. We go further, and by using results from analytic combinatorics, we recognize that

\[
\mathcal{E}(\theta) = \frac{S_K-1(\theta)}{(1-\theta)^K},
\]

and hence

\[
D^*_K(\theta, p, | \cdot |_1) = 2\theta \left( \frac{K - 1}{1 - \theta} + \frac{S_{K-1}(\theta)}{S_{K-1}(\theta)} \right)^2, \quad \text{where } S_{K-1}(\theta) = \sum_{j=0}^{K-1} \theta^j \left( \binom{K - 1}{j} \right)^2.
\]

We have thus expressed the minimum \( \mathbb{L}_1 \)-distortion of an optimal \( \theta \)-DP HSM in the asymptotic regime of large databases to infinite precision via a simple closed form computable expression. Moreover, \( D^*_K(\theta, p, | \cdot |_1) \) has connections (Corollary 1) to the Legendre polynomial [8].

Why is \( D^*_K(\theta, p, | \cdot |_1) \) invariant wrt pmf \( p \)? For large \( n \), \( \binom{n}{k} \frac{n!}{k!(n-k)!} \) approximates a near-uniform pmf [9] on the set of histograms within an \( \mathbb{L}_1 \)-ball of radius \( \mathcal{O}(\sqrt{n}) \) centered at \( (np_1, \cdots, np_K) \). This radius being sub-linear, for any \( p \) with positive entries, the \( \mathbb{L}_1 \)-ball that contains most of the mass is eventually supported on the set of histograms. Since we are concerned only in the eventual limit, the effect of \( p \) is only a shift of the center of this \( \mathbb{L}_1 \)-ball containing a near-uniform pmf. As we will see, the optimal mechanism in the limit of large databases are identical for different \( p \) except for ‘a shift of its center’, with its distortion invariant wrt \( p \).
We are thus led to an interesting property of a sequence of optimal mechanisms that achieve $D^*_K(\theta, p, |·|_1)$. We prove the existence of a sequence $\mathbb{W}^n : \mathcal{H}^n \Rightarrow \mathcal{H}^n : n \in \mathbb{N}$ of HSMs that satisfy constraints (3) and can be realized as a cascade $\mathbb{U}^n : \mathcal{H}^n \Rightarrow \mathcal{H}^*_{ext}$, $\mathbb{V}^n : \mathcal{H}^*_{ext} \Rightarrow \mathcal{H}^n$ i.e., $\mathbb{W}^n(\cdot|h) = \sum_{h \in \mathcal{H}^*_{ext}} \mathbb{U}^n(\cdot|h)\mathbb{V}^n(\cdot|h)$, where $\mathcal{H}^*_{ext}$ is an extended set of histograms, $\mathbb{W}^n(·|·)$ is geometric, $\mathbb{V}^n(·|·)$ is a truncation, such that that (i) the $L_1$-distortion of $\mathbb{W}^n(·|·)$ approaches $D^*_K(\theta, p, |·|_1)$ in the limit of large databases, and (ii) $\mathbb{V}^n : \mathcal{H}^n \Rightarrow \mathcal{H}^*_{ext}$ is $θ$-DP and (iii) $\mathbb{W}^n(·|·)$ is invariant with $p$. Indeed, $\mathbb{W}^n$ is essentially universally optimal, where the notion of universal optimality was defined in [5]. While universal optimality considered the entire class of pmfs, we restrict our pmfs to multinomial pmfs. Since modelling records of a database to be independent and identically distributed is generic and appropriate, it suffices to consider multinomial pmfs on $\mathcal{H}^n$. Our findings give rise to a new definition - an asymptotic generalization of universal optimality [5]. As we will explain (Rem. 1) our findings do not contradict those of [6].

**Significance**: We highlight significance of our methods and results. In regard to the former, we bring tools from different disciplines - discrete geometry, combinatorics, analytic methods and LP - to harness the discrete combinatorial structure of the underlying domain. By expressing the minimum $L_1$-distortion as a functional of the Ehrhart series - a fundamental object in Ehrhart theory - we have established a connection between two disciplines. Analytic methods provide a simple closed form computable expression for the resulting power series. Generalization to convex distortion measures such as $L_2$ will involve integer point enumeration on the intersection of different dilates of two convex polytopes (Fig. 1). The nature of the resulting power series and whether they admit closed form computable expressions are questions of interest.

In regards to our results, a precise characterization of accuracy, in contrast to order approximations, provides a practitioner useful information in choosing the DP parameter $\theta$. Outputs are blatantly non-private or too noisy inaccurate for very small or very large values of $\theta$ respectively. For practically viable intermediate values $0 < \theta < 1$, higher order terms cannot be ignored. Moreover, accuracy has been observed to be very sensitive to $\theta$. DP is at a crucial juncture wherein ideas proposed in theory are being tested in practice. Practitioners will benefit from a precise characterization of accuracy. In our work, we derive lower bounds by identifying dual variables and evaluating the corresponding dual objective. Each dual variable (shadow price) quantifies sensitivity of utility to the corresponding constraint. Our assignment of dual variables corresponding to the popular geometric mechanism provides a practitioner information on how and which constraints to tighten or relax to finely tradeoff accuracy and privacy. In essence, it provides more control for a partitioner designing DP mechanisms.

**Prior Work**: DP [3] has been a subject of intense research. Following early works [2, 10], analysis of noise adding mechanisms for important classes of queries - releasing $k$-way marginals [11, 12, 13], counting queries [14] - were undertaken. Geng and Viswanath [15, 16] proved that ‘staircase mechanisms’ [17] are optimal for a general class of convex utility functions. These, and most other works restrict attention to noise-adding mechanisms. Our study makes no (such) assumptions/restrictions. With regard to lower bounds, Hardt and Talwar [4], followed by De [18], developed novel geometry-based techniques, while [19] develops lower bounds based on non-existence of certain fingerprinting codes.

## 2 Statement of the Main Results

We begin with remarks on notation. Calligraphic letters such as $\mathcal{R}$ represent sets. Random variables, and (generic) parameters that remain fixed, are represented via upper case letters. For $a, b \in \mathbb{Z}$, $n \in \mathbb{N}$, we let $[a, b] := \{a, a+1, \cdots, b\}$ whenever $a \leq b$, and $[n] := [1, n]$. $\cdot$ denotes the ‘proportional to’ symbol. A mechanism (randomized algorithm) $M : A \Rightarrow B$ with set $A$ of inputs and set $B$ of outputs is a map $\mathbb{W}_M : A \to \mathbb{P}(B)$ where $\mathbb{P}(B)$ is the set of pmfs on $B$. When input $a \in A$, the mechanism $M$ produces the output $b \in B$ with probability $\mathbb{W}_M(b|a)$.

---

\footnote{In the computer science literature, such as in [1, 20], $\mathbb{W}(b|a)$ is denoted $M(b|a)$.}
Since $M : A \Rightarrow B$ is uniquely characterized by the corresponding collection $(\mathbb{W}_M(\cdot | a) : a \in A)$ of pmfs, we refer to it either through $\mathbb{W}_M : A \Rightarrow \mathbb{P}(B)$ or $\mathbb{W}_M : A \Rightarrow B$. A mechanism $\mathbb{W} : \mathcal{H}^n \Rightarrow \mathcal{H}^n$ whose input and output are histograms is a HSM. HSM $\mathbb{W} : \mathcal{H}^n \Rightarrow \mathcal{H}^n$ that satisfy constraints (3) is said to be a $\theta$–DP HSM. If $G = (V, E)$ is a graph, $d_G(v_1, v_2)$ denotes the length of a shortest path from $v_1 \in V$ to $v_2 \in V$. When $h = (h_1, \ldots, h_K)$, $\sum_{k=1}^{K} h_k = n$, we let $(\frac{n}{h}) = (h_1, n-h_K)$ denote a multinomial coefficient, and $p^{h} : = p_1^{h_1}p_2^{h_2} \cdots p_K^{h_K}$. $(\frac{n}{h})^p$ denotes a generic multinomial pmf on $\mathcal{H}^n$. For any HSM $\mathbb{W} : \mathcal{H}^n \Rightarrow \mathcal{H}^n$, we let

$$D^n(p, \mathbb{W}) : = \mathbb{E}_p \{ \mathbb{E}_{\mathbb{W}} \{ |H - G|_1 \} \} = \sum_{g \in \mathcal{H}^n} \sum_{h \in \mathcal{H}^n} \left( \frac{n}{h} \right)^p \mathbb{W}(g|h) |g - h|_1. \quad (8)$$

denote the expected $L_1$–distortion of HSM $\mathbb{W}(\cdot | \cdot)$.

Our goal is to characterize $D^n_K(\theta, p_1 | \cdot | 1) = \lim_{n \to \infty} D^n_K(\theta, p_1 | \cdot | 1)$ as defined in (2). Since we restrict attention to the $L_1$–distortion measure for the rest of the article, we employ the following simplified notation. We let $D^n_K(\theta, p) : = D^n_K(\theta, p_1 | \cdot | 1)$ and $D^n_K(\theta, p) : = D^n_K(\theta, p_1 | \cdot | 1) = \lim_{n \to \infty} D^n_K(\theta, p)$. With this, we restate our problem for ease of reference. We aim to characterize $D^n_K(\theta, p) : = \lim_{n \to \infty} D^n(\theta, p)$, where

$$D^n(\theta, p) : = \min_{\mathbb{W}(\cdot | \cdot)} D^n(\theta, p, \mathbb{W}) = \sum_{g \in \mathcal{H}^n} \sum_{h \in \mathcal{H}^n} \left( \frac{n}{h} \right)^p \mathbb{W}(g|h) |g - h|_1 \text{ with } \mathbb{W} : \mathcal{H}^n \Rightarrow \mathcal{H}^n \text{ subject to the constraints in (3)}. \quad (9)$$

$D^n(\theta, p)$ is the solution of an LP with $O((n + 1)^2 K)$ variables subject to $O(K^2(n + 1)^{2(K-1)})$ constraints,\footnote{For every $h \in \mathcal{H}^n$ except for which one or more of the co-ordinates are 0, we have $|\{h \in \mathcal{H}^n : |h - 1|_1 = 2\}| = K(K - 1)$. Also, $|\mathcal{H}^n| = (n+K-1) \sim (n+1)^{K-1}$ [21, Lemma II.1], [22, Chap 2, Lemma1].} lending computation infeasible for practical values of $K$ and $n$. In fact, we are unaware of a solution of this LP even for $K = 2$. One of our main contributions is a simple computable analytic expression for $D^n_K(\theta, p)$. Thm. 1 provides one such expression - a hyper-geometric series. In Sec. 3.1, we identify this power series and elaborate on how and why it is related to $D^n_K(\theta, p)$.

**Theorem 1.** (i) When $K = 2$, the limit $D^2(\theta, p) = \lim_{n \to \infty} D^2(\theta, p)$ is $\frac{4d}{1-\theta^2}$. (ii) In general, the limit $D^n_K(\theta, p) = \lim_{n \to \infty} D^n_K(\theta, p)$ of the solutions to the LP (9) is

$$D^n_K(\theta, p) = 20 \left\{ \frac{K-1}{1-\theta} + \frac{S_{K-1}(\theta)}{S_K(\theta)} \right\}, \text{ where } S_K(\theta) = \sum_{j=0}^{K-1} \theta^j \left( \binom{K-1}{j} \right)^2 \quad (10)$$

with $S_{K-1}(\theta) : = \frac{d}{d\theta}S_{K-1}(\theta)$. An optimal HSM is obtained as a truncation of a geometric mechanism $\mathbb{W}^*(g|h) = (1 - \theta)^{-1} \mathbb{E}_p(\theta)^{-1} \mathbb{E}_p(\theta^{|g-h|_1})$, where $\mathbb{E}_p(\theta)$ is defined in (6).

Below, we express $D^n_K(\theta)$ in terms of another important construct in analysis - the Legendre polynomial. We note that $S_{K-1}(\theta) = (1 - \theta)^{-1} L_{K-1}(\frac{1}{1-\theta})$ [23, Prob. 85], where $L_n(x) : = \frac{1}{2n! \pi} \frac{d^n}{dx^n} (x^2 - 1)^n$ is the Legendre polynomial of degree $n$ defined in [8, Prob. 219].

**Corollary 1.** The limit $D^n_K(\theta) = \lim_{n \to \infty} D^n_K(\theta)$ of the sequence of solutions to the LP (9) is

$$D^n_K(\theta, p) = K \left\{ \frac{1 + 4\theta - \theta^2}{1 - \theta} + \frac{L_K(y)}{L_{K-1}(y)} \right\}, \text{ where } y = \frac{1 + \theta}{1 - \theta}. \quad (11)$$

While $D^n_K(\theta, p)$ is invariant with $p$, a sequence $\mathbb{W}^*(\cdot | \cdot) : \mathcal{H}^n \Rightarrow \mathcal{H}^n$ of HSMs that is optimal in the limit is not necessarily invariant with $p$. In this regard, the existence of essentially universally optimal mechanisms for the multinomial family of pmfs is noteworthy. Towards that end, we say a sequence $\mathbb{W}^n : \mathcal{H}^n \Rightarrow \mathcal{H}^n$ of pmfs is essentially universally optimal (Ess-Univ-Opt) if for each $n \in \mathbb{N}$, $\mathbb{W}^n$ can be realized as a cascade $\mathbb{U}^n : \mathcal{H}^n \Rightarrow \mathcal{H}_{ext}^n$, $\mathbb{V}^n : \mathcal{H}_{ext}^n \Rightarrow \mathcal{H}^n$, i.e., $\mathbb{W}^n(g|h) = \sum_{h \in \mathcal{H}_{ext}^n} \mathbb{U}^n(b|h) \mathbb{V}^n(g|b)$ for every $g, h \in \mathcal{H}^n$, where $\mathcal{H}_{ext}^n$ is any (not necessarily finite) set, such that (i) $\lim_{n \to \infty} D^n(p, \mathbb{W}^n) = D^n_K(\theta, p)$ for every pmf $p$ on a set of $K$ elements, and (ii) $\mathbb{U}^n : \mathcal{H}^n \Rightarrow \mathcal{H}_{ext}^n$ is invariant with $p$.
Remark 1. Ess-Univ-Opt is a relaxed/weaker form of universal optimality [5] in two respects. Firstly, we restrict the class of pmfs on histograms to multinomial pmfs. Indeed, our definition of $D^n_k(\theta, p)$ in (9) is wrt $\binom{\hat{h}}{k} L^k_h$. Secondly, we only ask for asymptotic optimality of the sequence of mechanisms. This relaxed notion is of interest for the following reasons. Firstly, we operate with large databases. For sufficiently large $n$ the distortion of an Asym-Univ-Opt sequence of mechanisms might be sufficiently close to the true optimum for that $n$. Secondly, as the reader will note, it suffices to consider multinomial pmfs on $H^n$. In the light of non-existence of ‘strict’ universally optimal mechanisms [6], it is worth pursuing this relaxed notion.

As mentioned in [5], the existence of Ess-Univ-Opt is noteworthy. The proof of our main result will bring to light a sequence of Ess-Univ-Opt mechanisms

**Theorem 2.** Ess-Univ-Opt mechanisms for histogram sanitization wrt $L_1$-distortion exist.

The proof of Thm. 2 follows from proof of Thm. 1 wherein a sequence of truncated geometric mechanisms are proven to be Ess-Univ-Opt.

3 Analysis and Proofs

In Sec. 3.1, we derive an upper bound on $D^n_K(\theta, p)$ using tools from Ehrhart theory and combinatorics. We then provide a lower bound for $K = 2$ in Section 3.2 via complementary slackness conditions. For general $K$, we provide an assignment for dual variables in Sec. 3.3 whose (dual) objective value coincides with the upper bound, and we conclude via weak duality thm. (WDT).

3.1 Upper bound via Ehrhart Theory

To prove an upper bound on $D^n_K(\theta, p) = \lim_{n \to \infty} D^n_k(\theta, p)$, we identify a sequence $W^n : H^n \Rightarrow H^n : n \in \mathbb{N}$ of mechanisms and evaluate their $L_1$-distortions $D^n(p, W^n)$. Note that we are interested only in the limit $\lim_{n \to \infty} D^n(p, W^n)$ of one such (optimal) sequence, it suffices to define $W^n : H^n \Rightarrow H^n : n \in \mathbb{N}$ only for $n$ sufficiently large and evaluate their $L_1$-distortions. Towards that end, it is natural to consider geometric mechanisms $W^n_{\theta}(g|h) \propto \theta^{-\frac{|g-h|}{2}}$. Specifically, suppose we define

$$W_{\theta}(g|h) = \frac{\theta^{-\frac{|g-h|}{2}}}{E_{\theta}(\theta)}$$

where $E_{\theta}(\theta) = 1 + \sum_{d=1}^{n} N_d(\hat{h})\theta^d$, $N_d(\hat{h}) = \{ g \in H^n : |g - h|_1 = 2d \}$, (12)

then it can be verified that $W_{\theta}(\cdot\cdot)$ satisfy (3a) and non-negativity constraints. Does $W_{\theta}(\cdot\cdot)$ satisfy (3b)? Observe that for neighboring histograms $\hat{h}, \hat{h} \in H^n$ i.e., $|\hat{h} - \hat{h}|_1 = 2$, we have

$$\frac{W_{\theta}(g|h)}{W_{\theta}(\hat{g}|\hat{h})} = \frac{E_{\theta}(\theta)\theta^{-\frac{|g-h|}{2}}}{E_{\theta}(\theta)\theta^{-\frac{|\hat{g}-\hat{h}|}{2}}} = \frac{E_{\theta}(\theta)\theta^{-\frac{|\hat{g}-\hat{h}|}{2}-\frac{|g-h|}{2}}}{E_{\theta}(\theta)\theta^{-\frac{|\hat{g}-\hat{h}|}{2}}} \in \left[ \frac{E_{\theta}(\theta)}{E_{\theta}(\theta)}, \frac{E_{\theta}(\theta)}{E_{\theta}(\theta)} \cdot \frac{1}{\theta} \right]$$

by the triangular-inequality. $W_{\theta}(\cdot\cdot)$ is $\theta-\text{DP}$ if and only if $E_{\theta}(\theta) = E_{\theta}(\theta)$. It can be easily verified this is not true by considering simple values of $K$ and $n$. For example, consider $K = 3$, $n = 5$ and $\hat{h} = (1, 3, 1)$ and $\hat{h} = (1, 2, 2)$. Furthermore, $N_2(\hat{h}), N_2(\hat{h}) : 1 \leq d \leq 5$ can be computed through the distance distribution of vertices $\hat{h}, \hat{h}$ in the graph Fig. 2 and verified that $E_{\theta}(\theta) \neq E_{\theta}(\theta)$ for $\theta \in (0, 1)$. We therefore note that the finite $(K - 1)$-dimensional lattice $H^n$ throws challenges to defining a simple geometric mechanism.

We overcome this impediment by considering a discrete enlargement of $H^n$ and designing a cascade mechanism. In particular, we let $H^n_{\text{ext}}$ be a discrete set and define mechanisms $U^n : H^n \Rightarrow H^n_{\text{ext}}$ and $V^n : H^n_{\text{ext}} \Rightarrow H^n$ and let $W^n_{\theta}(g|h) = \sum_{\hat{h} \in H^n_{\text{ext}}} U^n(h|\hat{h})V^n(\hat{g}|\hat{h})$.

**Remark 2.** Enlargement of $H^n$ into $H^n_{\text{ext}}$ markedly differs from earlier works that considered enlargement of discrete sets into continuous domains. This key step leads us into discrete tools such as Ehrhart theory.
Let $\mathcal{H}_{ext}^n := \{(h_1, \cdots, h_K) \in \mathbb{Z}^K : \sum_{k=1}^K h_k = n\}$. The only difference between $\mathcal{H}^n$ defined in (1), and $\mathcal{H}_{ext}^n$ is that the latter permits co-ordinates to be negative resulting in it being (countably) infinite. We now define a pair of mechanisms $\mathbb{U}^n : \mathcal{H}^n \Rightarrow \mathcal{H}_{ext}^n$ and $\mathbb{V}^n : \mathcal{H}_{ext}^n \Rightarrow \mathcal{H}^n$. For every $h \in \mathcal{H}^n$ and every $g \in \mathcal{H}_{ext}^n$, let

$$\mathbb{U}^n(g|h) = \left(1 + \sum_{d=1}^{\infty} N_d \theta^d\right)^{-1} \theta^{|g-h|}, \quad \text{where} \quad N_d := |\mathcal{N}_d(h)|.$$  

(13)

Why is $N_d$, as defined above, invariant with $h$? Suppose one considers $a, b \in \mathcal{H}^n$. For any $b_2 \in \mathcal{H}_{ext}^n$, such that $|b_2 - a| = 2d$, consider the element $b_2 = b + b_2 - a$. It can be verified that $|b_2 - b| = 2d$ and there is a one-to-one correspondence between such elements $h_2$ and $h_2$. This proves invariance of $N_d$ with respect to $h$. We now prove $\mathbb{U}^n(\cdot|\cdot)$ satisfies (3a). Suppose $h \in \mathcal{H}^n$, note that $h \in \mathcal{H}_{ext}^n$,

$$\sum_{g \in \mathcal{H}_{ext}^n} \mathbb{U}^n(g|h) = \sum_{d=0}^{\infty} \sum_{g \in \mathcal{N}_d(h)} \theta^{|g-h|} = \sum_{d=0}^{\infty} \sum_{g \in \mathcal{N}_d(h)} \theta^d = \sum_{d=0}^{\infty} N_d \theta^d = 1 + \sum_{d=1}^{\infty} N_d \theta^d,$$

where $N_d(h)$ is as defined in (13), and hence $\sum_{g \in \mathcal{H}_{ext}^n} \mathbb{U}^n(g|h) = 1$ for all $h \in \mathcal{H}^n$. If $h, n \in \mathcal{H}^n$ and $|h - n|_1 = 2$, then $\mathbb{U}^n(g|h) = \theta^{|g-h|-1} \in [\theta, \theta^{-1}]$ by triangular-inequality which states that $-2 = -|h - n|_1 \leq |g - h|_1 - |g - n|_1$ for any $h, n \in \mathcal{H}^n$. We conclude $\mathbb{U}^n(\cdot|\cdot)$ is $\theta$–DP.

**Remark 3.** Note that $\mathbb{U}^n(\cdot|\cdot)$ is $\theta$–DP and invariant with $p$. This will be used in proving Thm. 2.

We now define $\mathbb{V}^n : \mathcal{H}_{ext}^n \Rightarrow \mathcal{H}^n$. The choice of $\mathbb{V}^n$ is based on the fact that the DBs whose histograms differ widely from the mean histogram $n \mu$ contribute an exponentially (in $n$) small amount to the expected value. Towards that end, choose any $R > 0$ invariant with $n$. We let $\mathbb{V}^n$ map the histogram outside the $L_1$–ball of radius $Rn^2$ centered at $np$ to the histogram $\hat{n}$. The histograms within radius $Rn^2$ of $np$ remain unchanged. Formally, let

$$\mathbb{V}^n(g|h) = 1 \text{ if } g = h, |h - np|_1 \leq Rn^2, \quad \mathbb{V}^n(g|h) = 1 \text{ if } g = np, |h - np|_1 > Rn^2,$$

(14)

and $\mathbb{V}^n(g|h) = 0$ otherwise, where $R > 0$ is invariant with $n$.

Does $\mathbb{V}^n$ output a histogram? The output of $\mathbb{V}^n$ is contained within a $L_1$–ball of radius $\alpha_n = Rn^2$ centered at $np \in \mathcal{H}^n$. The boundary of $\mathcal{H}^n$ is at a $L_1$–distance of at least $\beta_n = \min_{k=1,\cdots,K} np_k$ from $np \in \mathcal{H}^n$. Since $p_k > 0$ for all $k \in [K]$, for every $\alpha_n$ sufficiently large, $\alpha_n \leq \beta_n$, and the range of $\mathbb{V}^n$ is contained within $\mathcal{H}^n$, for all those $n \in \mathbb{N}$. The output of mechanism $\mathbb{V}^n$ is indeed a histogram for every $n \in \mathbb{N}$ sufficiently large. We provide a formal proof in Appendix 8.

We now define $\mathcal{W}^n(g|h) := \sum_{b \in \mathcal{H}_{ext}^n} \mathbb{V}^n(g|b) \mathbb{U}^n(b|h)$. Note that $\mathcal{W}^n : \mathcal{H}_{ext}^n \Rightarrow \mathcal{H}^n$ takes only the output of $\mathbb{U}^n$, as input. By the post-processing theorem of DP, the cascade mechanism $\mathcal{W}^n$ is $\theta$-DP. In other words, $\mathcal{W}^n(\cdot|\cdot)$ satisfies constraints in (3). We now prove that $\lim_{n \to \infty} D^n(p, \mathcal{W}^n) \preceq \lim_{n \to \infty} D(\mathbb{U}^n)$, where $D(\mathbb{U}^n) := \sum_{h \in \mathcal{H}^n} \sum_{g \in \mathcal{H}_{ext}^n} \left(\frac{n}{n} \right)^2 \mathbb{U}^n(g|h) |g - h|_1$ is the $L_1$–distortion of $\mathbb{U}^n(\cdot|\cdot)$ that outputs a histogram in the enlarged set $\mathcal{H}_{ext}^n$. We describe the argument here and provide a formal proof in Appendix 9. Let $d_\mu(\mathcal{W}^n) := \sum_{g \in \mathcal{H}^n} \mathcal{W}^n(g|h) |g - h|_1, \ d_\mu(\mathbb{U}^n) := \sum_{g \in \mathcal{H}_{ext}^n} \mathbb{U}^n(g|h) |g - h|_1$ denote (unweighted) contributions of $h$ to $D^n(p, \mathcal{W}^n)$ and $D(\mathbb{U}^n)$ respectively. Refer to Fig. 6. Let $B(\frac{1}{2})$ and $B(1)$ be the $L_1$–balls centered at $np$ of radii $\frac{R}{2}n^2$ and $Rn^2$ respectively. Let $B^*(1) := \mathcal{H}_{ext}^n \setminus B(1)$. For each $h \in B(\frac{1}{2})$, the mechanism $\mathcal{W}^n$ has the effect of decreasing $h$’s contribution. In other words, for any $h \in B(\frac{1}{2}), d_\mu(\mathcal{W}^n) \leq d_\mu(\mathbb{U}^n)$. This is because (i) $\mathcal{W}^n$ transfers mass placed on $\hat{g} \in B^*(1)$ - an element farther from $np$ to $np$, and (ii) $\mathcal{W}^n$ does not alter the mass placed on elements $\hat{g} \in B(1)$ (other than $np$). This is made precise in the sequence of
steps (35) - (37) in Appendix 9 (and can be followed ignoring the earlier steps therein). What about for $h \in B^{n}(\frac{1}{2})$? The weights $\binom{n}{h} \rho_{h}$ associated with these elements, when summed up, contribute an exponentially small amount. Formally, $\sum_{h \in B^{n}(\frac{1}{2})} \binom{n}{h} \rho_{h} \leq \exp(-n\gamma)$ for some $\gamma > 0$. Since $|g - h|_{1} \leq 2n$ whenever $h, g \in H^{n}$, we have $d_{h}(\mathbb{W}^{n}) \leq 2n \exp(-n\gamma)$ and hence $\sum_{h \in B^{n}(\frac{1}{2})} \binom{n}{h} \exp(d_{h}(\mathbb{W}^{n})) \to 0$ as $n \to 0$. These details are fleshed out in Appendix 9. Therein we have proved that for any $\epsilon > 0$, there exists $N_{\epsilon} \in \mathbb{N}$ such that $D^{n}(p, \mathbb{W}^{n}) \leq D(\mathbb{U}) + \epsilon$ for every $n \geq N_{\epsilon}$. We conclude $\lim_{n \to \infty} D^{n}(p, \mathbb{W}^{n}) \leq \lim_{n \to \infty} D(\mathbb{U})$.

Our next step is to characterize $\lim_{n \to \infty} D(\mathbb{U})$. A first simple question is: Does $D(\mathbb{U})$ depend on $\mathbb{U}$? We know, $H^{n}$, the domain of $\mathbb{U}^{n}(\cdot)$, depends on $\mathbb{U}$. The reader will recognize that the pmfs $(\mathbb{U}^{n}(g) : g \in H^{n}_{\text{ext}})$ and $(\mathbb{U}^{n}(g) : g \in H^{n}_{\text{vat}})$ are just shifts of each other, and hence $d_{n}(\mathbb{U}^{n}) = \sum_{g \in H^{n}_{\text{rat}}} \mathbb{U}^{n}(g)\|g - q\|_{1}$, as defined earlier, does not depend on $g \in H^{n}$, and in fact is invariant with $\mathbb{U}$ too. We will make this precise now. By the definition of $\mathbb{U}^{n}(\cdot)$ we will identify $\mathbb{U}^{n}(g) = \mathbb{U}(g)$ for $g \in H^{n}_{\text{rat}}$, precisely in our proof. As $\mathbb{U}^{n}(\cdot)$ is an integral convex polytope (Fig. 1). Our pursuit of $\mathbb{U}^{n}(\cdot)$ is the number of integral points in the $\delta$-dilation of the integral convex polytope $\mathbb{P} = |\mathbb{P}|$. Ehrhart theory concerns the enumeration of integer points in an integral convex polytope and the objects associated with these counts. We briefly present the foundational results in Ehrhart theory that we will have opportunity to use in our study. A convex $l$-polytope is a convex polytope of dimension $l$. A convex $l$-polytope whose vertices have integral co-ordinates is an integral convex $l$-polytope. $L_{P}(d)$ is the number of integral points in the $d$-th dilation of the integral convex polytope $P = \{x \in \mathbb{R}^{l} : \sum_{k=1}^{K} x_{k} = 0, \sum_{k=1}^{K} |x_{k}| \leq 2d\}$.

Indeed, if $L_{P}(d) = \|\mathbb{Z}^{K} \cap \mathbb{P}\|$, then $N_{d} = L_{P}(d) - L_{P}(d - 1)$. Notice that $L_{P}(d)$ is the number of integral points in the $d$-th dilation of the integer convex polytope $P = \{1\}$. Our pursuit of $L_{P}(d)$ and the associated objects is aided by the following fundamental fact due to Eugène Ehrhart. Ehrhart’s theorem states that if $P$ is an integral convex $l$-polytope, then $L_{P}(d)$ is a polynomial in $d$ of degree $l$. We refer to $L_{P}(d)$ as Ehrhart’s polynomial. We will identify $N_{d}$, and hence $L_{P}(d)$, precisely in our proof. As evidenced by (15), we need to study the generating function of the counts $L_{P}(d) : d \in \mathbb{N}$. We refer to the formal power series.
\( \mathcal{E}_P(z) = 1 + \sum_{d=1}^{\infty} \lambda \mathcal{L}(d) z^d \) as the Ehrhart series of \( P \), and let \( \mathcal{F}_P(z) : = (1 - z) \mathcal{E}_P(z) \). (18)

Since \( N_d = \mathcal{L}(d) - \mathcal{L}(d - 1) \), the definitions for \( \mathcal{F}_P(z) \) in (15) and (18) are consistent and this explains the occurrence of \( P \) in \( \mathcal{F}_P(z) \). By substituting \( \mathcal{F}_P(\theta) = (1 - \theta) \mathcal{E}_P(\theta) \), it can be verified that the RHS of (4) is equal to \( \mathcal{D}_K(\theta) = \frac{d}{d\theta} \mathcal{F}_P(\theta) \). In Appendix 10, we characterize \( N_d \) explicitly using simple combinatorial arguments. This proves (6). Our final step in the upper bound analysis is to prove \( \mathcal{D}_K(\theta) \) is given by the expression in (7). It suffices to characterize either the Ehrhart series \( \mathcal{E}_P(\theta) \) or \( \mathcal{F}_P(\theta) \), of \( P = P_1 \), where \( P_1 \) is the polytope characterized in (17). In the interest of brevity, we refer to [24], wherein \( \mathcal{F}_P(\theta) \) is characterized as

\[
\mathcal{F}_P(\theta) \bigg|_{\theta=1} = S_{K-1}(\theta) \quad (19)
\]

Substituting this in (15), one can verify that

\[
\mathcal{D}(\mathcal{U}^n) = \mathcal{D}_K(\theta) = 20 \frac{K - 1}{1 - \theta} + S_{K-1}(\theta)
\]

[25, Thm 3.1, Eqn (3.7)] provides a proof of (19(a)) via contour integration. We shall provide, in a subsequent version of this article, an alternate proof via a simple counting principle, Ehrhart theory and analytic methods.

### 3.2 Lower bound for \( K = 2 \) via Complementary Slackness

We identify a feasible solution to the dual of the LP in (9). The corresponding objective value, say \( C^m \), by the weak duality thm. (Sec. 6) evaluates to a lower bound on \( D^k(\theta, p) \). By proving \( \lim_{m \to \infty} C^m = \mathcal{D}_K(\theta) \), we conclude \( D^k(\theta, p) = \mathcal{D}_K(\theta) \). Towards that end, let us identify the dual of the LP in (9). With each DP constraint (2b), we have a non-negative dual variable \( \lambda_{g(h,h)} \). Note that \( \lambda_{g(h,h)} \) and \( \lambda_{g(h,h)} \) are distinct dual variables. With each sum constraint (2a) we have a free dual variable \( \mu_h \). It can be verified that the dual of (9) is

\[
S^n(\theta) : = \max \sum_{h \in \mathcal{H}^n} \mu_h \text{ subject to (i) } \mu_h \leq \left( \begin{array}{c} n \\ h \end{array} \right) |g| |h - g|_1 + \theta \sum_{h \in \mathcal{N}(h)} \lambda_{g(h,h)} - \sum_{h \in \mathcal{N}(h)} \lambda_{g(h,h)} \text{ for } h, g \in \mathcal{H}^n \times \mathcal{H}^n \text{ and (ii) } \lambda_{g(h,h)} \geq 0 \text{ for } g \in \mathcal{H}^n \text{ and } h, g \in \mathcal{H}^n \times \mathcal{H}^n \text{ satisfying } |h - g|_1 = 2, (20)
\]

where \( \mathcal{N}(h) : = \{ h \in \mathcal{H}^n : |h - h|_1 = 2 \} \) is the set of neighbors of \( h \in \mathcal{H}^n \). We let \( C^n(\lambda, \mu) = \sum_{h \in \mathcal{H}^n} \mu_h \) denote the objective value corresponding to a feasible solution \( \lambda, \mu \), where \( \lambda \) and \( \mu \) represent the aggregate of \( \lambda_{g(h,h)} \) and \( \mu_h \) variables respectively.

We first focus on \( K = 2 \) case and provide an assignment and prove the lower bound. The assignment is then interpreted via the shadow prices interpretation (Appendix 13). This interpretation leads us to a natural assignment for the general \( K \) case presented next.

We identify the histogram \((i, n - i) \in \mathcal{H}^n\) with just its first co-ordinate. We also have \( \mathbb{W}(n-j|i) \) denote \( \mathbb{W}(n-j|i, n-i) \), \( \lambda_{j(i-1,i)} \) denote \( \lambda_{j(i-j)|(i-1,n-i),(i,n-i)} \), and so on. With this notational simplification, we state below the primal and dual LP for \( K = 2 \) case.
where $\mathcal{E}_i^n = (n) p_1^n (1 - p_1)^{n-i}$. We have suppressed dependence of $\mathcal{E}_i^n$ on $p_1$. Furthermore, we let $p = p_1$ and $p_2 = 1 - p$. We provide a complete solution i.e., primal and dual feasible solutions that satisfy complementary slackness conditions. Recall from complementary slackness, we are required to prove that (i) either the primal constraint is tight or the corresponding dual variable is 0, and (ii) either the primal variable is 0 or the dual constraint is tight. For ease of verification, we have stated variables and constraints that are duals of each other on the same row of (21).

Let us begin with a primal feasible solution. Let $f_i = \sum_{j=0}^{n} 2\mathcal{E}_i^n \theta^{-j}$, $b_i = \sum_{k=0}^{n} 2\mathcal{E}_i^n \theta^{-k}$ and

$$A_n := \min \left\{ i \in [0, n] : f_{k-1} - \theta b_k \geq 0 \right\} \quad B_n := \max \left\{ i \in [0, n] : b_{k+1} - \theta f_k \geq 0 \right\}.$$ (22)

Using [9, Theorem 4], it can be proved that $A_n < np_1 < B_n$ and $B_n - np_1 = \Omega(\sqrt{n})$ and $B_n - np_1 = \Omega(\sqrt{n})$. We will use $A_n < np_1 < B_n$ in the assignment of primal and feasible variables, and proving that they satisfy complementary conditions. The latter $(np_1 - A_n = \Omega(\sqrt{n})$ and $B_n - np_1 = \Omega(\sqrt{n}))$ is used to prove the objective value of these sequence of primal feasible values tend to $\frac{4\theta}{1+\theta} = D_K(\theta)$. Consider the truncated geometric mechanism that are folded at $A_n - 1$ on the left and $B_n + 1$ on the right. Specifically, consider the $\theta-$DP HSM

$$\mathcal{W}(j|i) = \begin{cases} \theta^{j-i} \frac{1-\theta}{1+\theta} & i \in [0, n], j \in [A_n, B_n] \\ \frac{1}{1+\theta} & j = B_n + 1, i \leq j \\ 1 & i > B_n + 1, j = B_n + 1 \\ 0 & j \notin [A_n - 1, B_n + 1] \end{cases} \quad \mathcal{W}(j|i) = \begin{cases} 1 - \frac{\theta^{A_n-i}}{1+\theta} & i < A_n - 1, j = A_n - 1 \\ 1 - \frac{\theta^{B_n-i}}{1+\theta} & i > B_n + 1, j = B_n + 1 \\ 0 & \text{otherwise} \end{cases} \quad (23)$$

It can be easily verified that the above assignment satisfies (3). Indeed, this is just the truncated geometric. What are the complementary slackness conditions with regard to the above primal feasible assignment? Firstly, our assignment must guarantee

$$\mu_i = 2\mathcal{E}_i^n \theta^{-j} = \theta \lambda_{j(i-1)} + \theta \lambda_{j(i+1)} - \lambda_{j(i,i-1)} - \lambda_{j(i,i+1)} \quad \text{for } j \in [A_n - 1, B_n + 1] \quad (24)$$

since the corresponding primal variables are set to strictly positive values. We also observe

$$\theta < \frac{\mathcal{W}(A_n - 1|i)}{\mathcal{W}(A_n - 1|i - 1)} < \frac{1}{\theta} \quad \text{if } i < A_n - 1 \quad \text{and similarly } \theta < \frac{\mathcal{W}(B_n + 1|i)}{\mathcal{W}(B_n + 1|i - 1)} < \frac{1}{\theta} \quad \text{if } i > B_n + 1$$

From complementary slackness, we have $\lambda_{A_n - 1(i-1)} = 0$ for $i \leq A_n - 1$ and $\lambda_{B_n - 1}(i+1)$ is the dual variable associated with the constraint $\mathcal{W}(j|i) - \theta \mathcal{W}(j|i) \geq 0$. For $j \in \{A_n - 1, \ldots, B_n + 1\}$, and $i$ such that $|j - (i-1)| = |j - i - 1|$, verify that the constraint $\mathcal{W}(j|i - 1) - \theta \mathcal{W}(j|i) > 0$ is not tight. From complementary slackness, $\lambda_{j(i-1)} = 0$ if $j \leq i - 1$ and $j \in \{A_n - 1, \ldots, B_n + 1\}$. Similarly, if $j \geq i + 1$ and $j \in \{A_n - 1, \ldots, B_n + 1\}$, we set $\lambda_{j(i+1)} = 0$.

\footnote{We assume, without loss of generality $p_1 \leq \frac{1}{2}$}

\footnote{For the general $K$, we will assign $\lambda_{g|ar{k} \notin k} = 0$ if $|g - \bar{k}| \leq |g - l|$. Note that this simple observation halves the number of decision variables.}
(21). We are only left to provide an assignment for the rest of the dual variables that satisfy (24). For \( i \in \{1, \ldots, A_n - 1\} \) and \( j \in \{i, \ldots, A_n - 1\} \), set \( \lambda_{j(i-1,i)} = 0 \).

For \( i \in \{1, \ldots, A_n - 1\} \) and \( j \in \{A_n - 1, \ldots, n\} \), set \( \lambda_{j(i-1,i)} = [j-(A_n - 1)]f_{i-1} \). (25)

For \( i \in [A_n, n] \) and \( j \in [i, n] \), set \( \lambda_{j(i-1,i)} = \frac{f_{i-1} - \theta_i}{1-\theta_i} + (j-i)f_{i-1} \). (26)

For \( i \in [B_n + 1, n-1] \) and \( j \in [B_n + 1, i] \), set \( \lambda_{j(i-1,i)} = 0 \).

For \( i \in [B_n + 1, n-1] \) and \( j \in [0, B_n + 1] \), set \( \lambda_{j(i-1,i)} = [(B_n + 1) - j]b_{i+1} \).

For \( i \in [0, B_n] \) and \( j \in [0, i] \), set \( \lambda_{j(i-1,i)} = \frac{b_{i+1} - \theta_i}{1-\theta_i} + (i-j)b_{i+1} \). For \( i < A_n - 1 \), set (27)

\[
\mu_i = 2|G^*|[(A_n - 1) - i], \quad i > B_n + 1, \quad \mu_i = 2|G^*|i - (B_n + 1)]
\]

For \( i \in [A_n - 1, B_n + 1] \), set \( \mu_i = \theta_i(f_{i-1} + b_{i+1}) - \frac{4\theta_i}{(1-\theta_i)}|G^*| \). (29)

For \( i \in [A_n - 1, B_n + 1] \) verify \( \mu_i = \theta(f_{i-1} + b_{i+1}) - \frac{4\theta_i}{(1-\theta_i)}|G^*| \). (30)

Refer to Appendix 13 for an interpretation of the above assignment via shadow prices. This interpretation will prove very valuable in arriving at the dual variable assignment for the general \( K \) case in (32). We will now use the above assignment to verify (24). This is provided in Appendix 11.

### 3.3 Lower bound for General \( K \) via Weak Duality Theorem of LP

We provide an assignment for dual variables. The proof of feasibility of the following assignment follows from arguments analogous to those presented in Eqns. (38) - (42) for \( K = 2 \) case.

Recall the PC graph for general \( K \). For \( a \in \mathcal{H}^n \), let \( \mathcal{N}(a) : = \{a \in \mathcal{H}^n : |a - \tilde{a}|_1 = 2\} \) be the set of neighbors of \( a \). For \( a, b \in \mathcal{H}^n \), let

\[
\mathcal{F}(b, a) := \{a \in \mathcal{N}(a) : |b - a|_1 > |b - \tilde{a}|_1\}, \mathcal{C}(b, a) := \{a \in \mathcal{N}(a) : |b - a|_1 < |b - \tilde{a}|_1\}
\]

be the set of histograms farther to, closer to, and at equidistant from \( b \) than \( a \) respectively. Recall that \( 2d_G(a, b) = |a - b|_1 \) (Lemma 1). Complementary slackness conditions imply

\[
\lambda_{g(b, h)} = 0 \text{ whenever } |g - \tilde{h}|_1 > |g - h|_1.
\]

When \( |g - \tilde{h}|_1 < |g - h|_1 \), let

\[
\lambda_{g(b, h)} = \frac{\sum_{a \in \mathcal{C}(b, a)} \left( \frac{n}{a} \right)^2 2 \theta^{d_G(a, b)} - \theta \sum_{b \in \mathcal{C}(b, h)} \left( \frac{n}{b} \right)^2 2 \theta^{d_G(b, h)}}{1 + |\mathcal{C}(h, \tilde{h})| \theta^2 - (K(K - 1)) \theta^2 + \theta |\mathcal{E}(h, \tilde{h})|} + |g - \tilde{h}|_1 \sum_{a \in \mathcal{C}(b, a)} \left( \frac{n}{a} \right)^2 2 \theta^{d_G(a, b)} (32)
\]

\[
\mu_g = -\frac{\theta \sum_{b \in \mathcal{C}(g, b)} \left( \frac{n}{b} \right)^2 2 \theta^{d_G(a, b)} - \theta \sum_{b \in \mathcal{C}(g, h)} \left( \frac{n}{b} \right)^2 2 \theta^{d_G(b, h)}}{1 + |\mathcal{C}(h, \tilde{h})| \theta^2 - (K(K - 1)) \theta^2 + \theta |\mathcal{E}(h, \tilde{h})|}. (33)
\]

Having provided the above assignments, the natural question that arises is whether these are feasible for (20), and if yes, what do they evaluate to? A couple of remarks are in order. The first term in (32) is negative if \( g = \tilde{h}, |h - np| > |h - np| + 2 \) and \( |h - np| > \Theta(\sqrt{n}) \). This is the case analogous to (25). Therein, note that when \( i \in [A, B] \), the assignment is (26). In fact, the fraction in (32) is analogous to the fraction in (26). The reader will recognize \( \mathcal{E}(h, \tilde{h}) = 0 \) and \( \mathcal{C}(h, \tilde{h}) = \mathcal{F}(h, \tilde{h}) = 1 \). The first term in the numerator of the fraction in (32) is analogous to \( f_{i-1} \) in (26). The rest of the terms can also be related to the assignment in (25) - (29). The above assignment is a slightly simplified version, in the sense that the variables corresponding to non-active constraints have been ignored. Our thorough description for the \( K = 2 \) case, its feasibility and evaluation of its objective value will complete our proof. Appendix 13 provides a clear interpretation for the above assignment for \( K = 2 \).
References


4 Figures

Figure 3: Privacy-constraint graph for $k = 2$ and general $n$. The vertices are labeled by the corresponding histogram. Two vertices are connected by an edge if their corresponding histograms are at an $L_1$-distance 2.

Figure 4: The PC graphs for $K = 3, N = 2$ are depicted. The decision variables ($W(g | (1, 0, 1)) : g \in H_3^2$) are associated with the nodes of the graph on the left. On the right, the decision variables ($W(g | (1, 1, 0)) : g \in H_3^2$) are associated with the nodes of the graph. Since $(1, 1, 0)$ and $(1, 0, 1)$ are neighbors, at every node, the two values have to be within $\theta$ and $\frac{1}{\theta}$ of each other.
Figure 5: Consider distance distribution of nodes within the dotted circle.
Figure 6:
Figure 7: $B(R, h)$ denotes the ball of radius $R$ centered at $h$. $C_h$ denotes the collection of centers of the smaller disjoint balls of radius $r$ contained within $B(R, h)$. $g$ denotes a generic element in $C_h$. By optimal packing, we know that there exists a $C_h$ that contains $\Theta((K - 1)^{\frac{R}{r}})$ centers. An upper bound on the minimax expected fidelity implies, by Markov inequality, a lower bound on $\mathbb{W}(B(r, g) | g)$. The latter translates to a lower bound on $\mathbb{W}(B(r, g) | h)$ via the DP constraints and an upper bound of $R$ on $|g - h|_1$. 


Figure 8: The shells depicted above are of width $n^{\frac{1}{3}}$. The idea is to pack each of these shells with disjoint balls $B(R, h)$. We can bound the multinomial probabilities in each of these shells using bounds found in [9, Proof of Thm. 4].
5 Properties of Privacy-Constraint Graph and $\mathcal{H}^n$

We list and prove some simple properties of the set of histograms $\mathcal{H}^n$ and the PC graph involved in our study.

**Lemma 1.** Consider the set $\mathcal{H}^n_K$ of histograms defined in (1) and the PC graph $G = (V, E)$, where $V = \mathcal{H}^n_K$ and $E = \{(h, \tilde{h}) \in \mathcal{H}^n \times \mathcal{H}^n : |h - \tilde{h}|_1 = 2\}$. The following are true (i) For any $g, h \in \mathcal{H}^n$, $|g - h|_1$ is an even integer and at most $2n$. (ii) $d_G(g, h) = 2|g - h|_1$.

**Proof.** (i) For any $g, h \in \mathcal{H}^n$, we have $\sum_{k=1}^{K} g_k = \sum_{k=1}^{K} h_k = n$, and hence for any subset $S \subseteq [K]$, we have $\sum_{i \in S} (g_i - h_i) = \sum_{j \in [K] \setminus S} (h_j - g_j)$. Note that

$$|g - h|_1 = \sum_{i=1}^{n} |g_i - h_i| = \sum_{i; g_i \geq h_i} (g_i - h_i) + \sum_{j; h_j > g_j} (h_j - g_j) = 2 \sum_{i; g_i \geq h_i} (g_i - h_i),$$

which is an even integer. Moreover $\sum_{i; g_i \geq h_i} (g_i - h_i) \leq \sum_{i=1}^{n} g_i = n$, and hence $|g - h|_1 \leq 2n$.

(ii) We prove this by induction on $K$. When $K = 1$, we have $\mathcal{H}^1_1 = \{(n)\}$ and the statement is true. When $K = 2$, we note that $|(n-i, j)-(n-j, i)|_1 = 2|i-j|$ and the nodes $(n-i, i), (n-j, j)$ are indeed $|i-j|$ hops apart (Fig. 3). Hence $|i-j| = d_G((n-i, i), (n-j, j))$ and the statement is true. We assume the truth of this statement for $K = 1, \ldots, L-1$ and any $n$. Suppose $K = L$ and let $g, h \in \mathcal{H}^n_L$. If for some co-ordinate $i$, we have $g_i = h_i$. Then, let $\tilde{g} : = (g_j : j \neq i)$ and $\tilde{h} : = (h_j : j \neq i)$. We have $\tilde{g}, \tilde{h} \in \mathcal{H}^n_{L-1}$. By our induction hypothesis, we have $2d_G(\tilde{g}, \tilde{h}) = |\tilde{g} - \tilde{h}| = |g - h|_1$, where $\tilde{G}$ is the PC graph corresponding to $\mathcal{H}^{n-\text{g}}_{L-1}$. It can now be verified that a shortest path from $\tilde{g}$ to $\tilde{h}$ on $\tilde{G}$ corresponds to a shortest path between $\text{ulinetilde}g$ to $\text{ulinetilde}h$ and hence $d_G(g, h) = d_G(\tilde{g}, \tilde{h})$. In fact, observe that the graph induced on the set of vertices on a horizontal line in Fig. 2 is isomorphic to the graph in Fig. 3 for an appropriate choice of $n$. Let us now consider the alternate case where $g, h \in \mathcal{H}^n_L$ are such that for no co-ordinate $i$ do we have $g_i = h_i$. Without loss of generality, assume $a = g_1 - h_1 > 0$. Let $i_1, \ldots, i_R \in [2, L]$ be coordinates such that $h_{i_r} > g_{i_r} \text{ for } r \in [1, R]$ and $\sum_{r=1}^{R} (h_{i_r} - g_{i_r}) \geq a$. (The existence of coordinates $i_1, \ldots, i_R$ can be easily proved by using the fact that $g, h \in \mathcal{H}^n_L$.) Now, let $b_1, \ldots, b_R > 0$ be integers such that $h_{i_r} - g_{i_r} > b_r > 0 \text{ for } r \in [R]$ and $\sum_{r=1}^{R} b_r = a$. Now consider $f \in \mathcal{H}^n_L$ such that $f_1 = g_1 - a, f_{i_r} = g_{i_r} + b_r$ and $f_j = g_j$ if $j \notin \{1, i_1, \ldots, i_R\}$. It can now be verified, by using the induction hypothesis on $f, h$ that $d_G(f, h) = d_G(g, f) + d_G(f, h)$ and $2d_G(f, h) = |f - h|_1 + |f - h|_1 = |\tilde{g} - \tilde{h}|_1$, thereby proving the statement for $K = L$.

6 The weak duality theorem of LP

We refer the reader to [26] for a lucid description of the dual linear program. Following the same notation, we state WDT below.

**Weak Duality Theorem:** Consider the following primal and dual LP problems. Let $\mathbf{A}$ be a matrix with rows $\mathbf{a}_i^j$ and columns $\mathbf{A}_j$.

**Minimize** $c^j \mathbf{x}$ **subject to** $\mathbf{a}_i^j \geq \mathbf{b}_i$, $i \in M_1$; $\mathbf{a}_i^j = \mathbf{b}_i$, $i \in M_2$; $x_j \geq 0$, $j \in N_1$

**Maximize** $\mathbf{p}^i \mathbf{b}$ **subject to** $p_i \geq 0$, $i \in M_1$; $p_i \text{ free}$, $i \in M_2$; $\mathbf{p}^i \mathbf{A}_j \leq c_j$, $j \in N_1$

If $\mathbf{x}$ and $\mathbf{p}$ are feasible solutions to the primal and dual problems respectively, then $\mathbf{p^i b} \leq c^j \mathbf{x}$.
7 Mechanism $\mathcal{U} : \mathcal{H}^n \Rightarrow \mathcal{H}^n_{\text{ext}}$ is a $\theta$–DP mechanism

Recall, $\mathcal{U} : \mathcal{H}^n \Rightarrow \mathcal{H}^n_{\text{ext}}$ is specified in (13), and we let

$$\mathcal{F}_P(\theta) = (1 - \theta)\mathcal{E}_P(\theta) = 1 + \sum_{d=1}^{\infty} N_d \theta^d$$  \hspace{1cm} (34)

Clearly, $\mathcal{U}^n(g|h) \geq 0$. We note that

$$\sum_{g \in \mathcal{H}^n_{\text{ext}}} \mathcal{U}^n(g|h) = \frac{1}{\mathcal{F}_P(\theta)} \sum_{g \in \mathcal{H}^n_{\text{ext}}} g^{(\frac{|g-h|}{2})} = \frac{1}{\mathcal{F}_P(\theta)} \sum_{g \in \mathcal{H}^n_{\text{ext}}} \sum_{|g-h| = 2d} \theta^{\frac{|g-h|}{2}} = \frac{1}{\mathcal{F}_P(\theta)} \sum_{d=0}^{\infty} \sum_{|g-h| = 2d} \theta^d$$

Lastly, suppose $\mathbf{h} \in \mathcal{H}^n$ and $\tilde{\mathbf{h}} \in \mathcal{H}^n$ are a pair of neighboring histograms,

$$\mathcal{U}^n(g|h)/\mathcal{U}^n(g|\tilde{h}) = \theta^{\frac{|g-h|}{2}}/\theta^{\frac{|g-\tilde{h}|}{2}} = \theta^{\frac{|g-h| - |g-\tilde{h}|}{2}}.$$ 

By the triangular inequality, $-2 = -|g - \tilde{\mathbf{h}}| \leq |g - \mathbf{h}| - |g - \tilde{\mathbf{h}}| \leq |\mathbf{h} - \tilde{\mathbf{h}}| = 2$, and we therefore have the above ratio in $[\theta, 1]$. $\mathcal{U}^n$ is therefore a $\theta$–DP mechanism.

8 The output of mechanism $\mathcal{V}^n$ is a histogram

We recall $\mathcal{V}^n : \mathcal{H}^n_{\text{ext}} \rightarrow \mathcal{H}^n$ is defined as

$$\mathcal{V}^n(g|h) = \begin{cases} 1 & \text{if } g = \mathbf{h}, |\mathbf{h} - np| \leq Rn^2 \\ 1 & \text{if } g = np, |\mathbf{h} - np| > Rn^2 \\ 0 & \text{otherwise,} \end{cases}$$

where $R > 0$ is any constant invariant with $n$. Since $\mathcal{V}^n$ is a deterministic map, it can also be defined through the map $f_{\mathcal{V}^n} : \mathcal{H}^n_{\text{ext}} \Rightarrow \mathcal{H}^n$ where

$$f_{\mathcal{V}^n}(\mathbf{h}) = \begin{cases} \mathbf{h} & \text{if } |\mathbf{h} - np| \leq Rn^2 \\ np & \text{otherwise, i.e., } |\mathbf{h} - np| > Rn^2, \end{cases}$$

and $\mathcal{V}^n(g|h) = \mathbb{1}_{\{g = f_{\mathcal{V}^n}(\mathbf{h})\}}$, where $R > 0$ is a constant, invariant with $n$. Let us analyze what ‘extended histograms’ are within the range of $f_{\mathcal{V}^n}$. $\mathbf{h} \in \mathcal{H}^n_{\text{ext}}$ falls in the range of $f_{\mathcal{V}^n}$, or in other words, is output by mechanism $\mathcal{V}^n$ only if $|\mathbf{h} - np| \leq Rn^2$, which is true, only if $|h_k - np_k| \leq Rn^2$. The latter is equivalent to $np_k - Rn^2 < h_k \leq np_k + Rn^2$ for every $k \in [K]$. Observe that, since we assumed $p_k > 0$ for all $k \in [K]$, the lower bound $np_k - Rn^2 > 0$ for any $R > 0$ and sufficiently large $n$. For sufficiently large $n$, $\mathcal{V}^n$ outputs an extended histogram whose co-ordinates are non-negative. From (1), and the definition $\mathcal{H}^n_{\text{ext}}$, the output of $\mathcal{V}^n$ is indeed a histogram from $\mathcal{H}^n$. The output of mechanism $\mathcal{V}^n$ is indeed a histogram.

9 For $n \in \mathbb{N}$ sufficiently large, $D^n(p, \mathcal{V}^n) \leq D(\mathcal{U}^n)$

We now prove that the expected distortion of $\mathcal{W}^n(\cdot|\cdot)$ is in the limit, at most that of $\mathcal{U}^n$. Recall, $\mathcal{W}^n(\cdot|\cdot)$ is defined as $\mathcal{W}^n(g|h) = \sum_{b \in \mathcal{H}^n_{\text{ext}}} \mathcal{U}^n(b|h) \mathcal{V}^n(g|b)$, where $\mathcal{W}^n(\cdot|\cdot), \mathcal{V}^n(\cdot|\cdot)$ are as defined in (13), (14) respectively. We let $B(\delta, \mathbf{h}) := \{g \in \mathcal{H}^n : |g - \mathbf{h}| \leq \delta\}$ and $B^*(\delta, \mathbf{h}) := \{g \in \mathcal{H}^n : |g - \mathbf{h}| \leq \delta\}$. 

19
Characterizing $N_d$, its complement. We abbreviate $B(\frac{1}{2}) = B(\frac{R}{2} n^\frac{3}{2}, np)$, $B'(\frac{1}{2}) = B'(\frac{R}{2} n^\frac{3}{2}, np)$, $B(1) = B(Rn^\frac{3}{2}, np)$, $B'(1) = B'(Rn^\frac{3}{2}, np)$, where $R > 0$ is any positive radius that is invariant with $n$. Observe that

$$D^n(p, \mathcal{W}^n) = \sum_{h \in \mathcal{H}^n} \sum_{g \in \mathcal{H}^n} \left( \frac{n}{h} \right) p \mathcal{W}^n(g|h) |g - h|_1$$

$$= \sum_{h \in B(\frac{1}{2})} \sum_{g \in \mathcal{H}^n} \left( \frac{n}{h} \right) p \mathcal{W}^n(g|h) |g - h|_1 + \sum_{h \in B'(\frac{1}{2})} \sum_{g \in \mathcal{H}^n} \left( \frac{n}{h} \right) p \mathcal{W}^n(g|h) |g - h|_1$$

$$\leq \sum_{h \in B(\frac{1}{2})} \sum_{g \in \mathcal{H}^n} \left( \frac{n}{h} \right) p \mathcal{W}^n(g|h) |g - h|_1 + 2n \sum_{h \in B'(\frac{1}{2})} \left( \frac{n}{h} \right) p \mathcal{W}^n(g|h)$$

It can be easily shown that $\sum_{h \in B'(\frac{1}{2})} \left( \frac{n}{h} \right) p \mathcal{W}^n(g|h) |g - h|_1 \leq \exp \{-n\gamma\}$ for some $\gamma > 0$ using Hoeffding’s inequality or [9, Theorem 4] and hence the second term above can be made arbitrarily small by choosing $n$ large enough. We will be using the fact that $p_k > 0$ for all $k \in [K]$. We henceforth focus on the first term above which is given by

$$\sum_{h \in B(\frac{1}{2})} \sum_{g \in \mathcal{H}^n} \left( \frac{n}{h} \right) p \mathcal{W}^n(g|h) |g - h|_1 + \sum_{h \in B'(\frac{1}{2})} \sum_{g \in \mathcal{H}^n} \left( \frac{n}{h} \right) p \mathcal{W}^n(g|h) |g - h|_1 = \sum_{h \in B(\frac{1}{2})} \left( \frac{n}{h} \right) p \mathcal{W}^n(g|h) |g - h|_1 + \sum_{h \in B'(\frac{1}{2})} \left( \frac{n}{h} \right) p \mathcal{W}^n(g|h) |g - h|_1$$

where (i) (35) follows from $\mathcal{W}^n(\tilde{g}|h) = 0$ for $\tilde{g} \in B'(1)$ implying$^5$ that the second term is zero, (ii) (36) follows from the definition of $\mathcal{W}^n$ in terms of $\mathcal{U}^n$, (iii) (37) is true since, for every $h \in B(\frac{1}{2})$ and every $\tilde{g} \in B'(1)$, $|np - h|_1 \leq \frac{R}{2} n^\frac{3}{2} \leq Rn^\frac{3}{2} \leq |\tilde{g} - h|_1$.

10 Characterizing $N_d$

Recall $N_d = |S_d|$, where

$$S_d := \mathbb{Z}^K \cap \mathcal{P}_d = \left\{ (x_1, \ldots, x_K) \in \mathbb{Z}^K : \sum_{k=1}^K x_k = 0, \sum_{k=1}^K |x_k| = 2d \right\}$$

$^5$Note that the range of $\mathcal{U}^n(\cdot|\cdot)$ is $B(1)$.
$S_d$ can be partitioned into *disjoint* sets based on the co-ordinates (in set $A_{|P|}$ below) corresponding to its non-negative indices. Let

$$A_n := \left\{ (a_1, \cdots, a_n) \in \mathbb{Z}^n : a_i \geq 0, \sum_{i=1}^{n} a_i = d \right\},$$

$$B_m = \left\{ (b_1, \cdots, b_m) \in \mathbb{Z}^m : b_j < 0, -\sum_{j=1}^{m} b_i = d \right\}$$

$$= \left\{ (b_1, \cdots, b_m) \in \mathbb{Z}^m : b_j > 0, \sum_{j=1}^{m} b_i = d \right\}.$$

It can be verified that,

$$S_d = \bigcup_{P \subseteq [K]} A_{|P|} \times B_{K-|P|} = \bigcup_{P \subseteq [K]} A_{K-|P|} \times B_{|P|}.$$ 

We can now compute $|A_{|P|}|$ and $|B_{|P|}|$. Since

$$|A_n| = \binom{d + n - 1}{n - 1}, |B_m| = \binom{d - 1}{m - 1},$$

we have

$$N_d = \sum_{r=1}^{K-1} \binom{K}{r} \binom{d + r - 1}{r - 1} \binom{d - 1}{K - r - 1} = \sum_{r=1}^{K-1} \binom{K}{r} \binom{d + K - r - 1}{K - r - 1} \binom{d - 1}{r - 1},$$

where the running variable $r$ denotes the cardinality of the (running set) $P \subseteq [K]$. An alternate count can be obtained by explicitly considering the set of zero co-ordinates. Suppose $0 \leq z \leq K - 1$ denotes the number of 0–co-ordinates and $p$, the number of positive co-ordinates, then, for $d \geq 1$, it can be verified that

$$S_d = \bigcup_{Z \subseteq [K]|Z|} \bigcup_{P \subseteq [K] |Z| : 1 \leq |P| \leq K - 2} A_{|P|} \times B_{K-|P|-|Z|},$$

and hence

$$N_d = \sum_{z=0}^{K-2} \sum_{p=1}^{K-z-1} \binom{K}{z} \binom{K - z}{p} \binom{d - 1}{p - 1} \binom{d - 1}{K - z - p - 1}.$$ 

Here, we conclude

$$\mathcal{F}_p(\theta) = 1 + \sum_{d=1}^{\infty} \left\{ \sum_{r=1}^{K-1} \binom{K}{r} \binom{d + r - 1}{r - 1} \binom{d - 1}{K - r - 1} \right\} \theta^d$$

$$= 1 + \sum_{d=1}^{\infty} \left\{ \sum_{r=1}^{K-1} \binom{K}{r} \binom{d + K - r - 1}{K - r - 1} \binom{d - 1}{r - 1} \right\} \theta^d$$

$$= 1 + \sum_{d=1}^{\infty} \left\{ \sum_{z=0}^{K-2} \sum_{p=1}^{K-z-1} \binom{K}{z} \binom{K - z}{p} \binom{d - 1}{p - 1} \binom{d - 1}{K - z - p - 1} \right\} \theta^d.$$ 

In a subsequent version of this article, we prove (19(a)), using above counts.
11 Verification of Complementary Slackness conditions for $K = 2$

We proceed to verify (24) for the above assignment.\(^6\) Recall $\ell_i^n = \binom{n}{i} p^i (1 - p)^{n-i}$.

We first prove that for any $i < A_n - 1$, $j \in [A_n - 1, B_n + 1]$, 

\[
\binom{n}{i} p^i (1 - p)^{n-i} 2|j - i| + \theta \lambda_j(j-1,i) + \theta \lambda_j(j,i-1) = \binom{n}{i} p^i (1 - p)^{n-i} 2|(A_n - 1) - i| \tag{38}
\]

Towards that end, note that $\lambda_j(j+1,i) = \lambda_j(j,i-1) = 0$ for the considered values for $i,j$. Substituting $\lambda_j(j-1,i) = [j - (A_n - 1)] f_{i-1}$ from (25), we have $\theta \lambda_j(j-1,i) - \lambda_j(j,i+1) = [j - (A_n - 1)](\theta f_{i-1} - f_i) = -[j - (A_n - 1)](\binom{n}{i} p^i (1 - p)^{n-i})$, and we therefore have (38). From the assignment (27), (28), we conclude validity of (24) for $i < A_n - 1$. Before we continue, we note that

\[
f_i = \theta f_{i-1} + \binom{n}{i} 2 p^i (1 - p)^{n-i}, \quad \text{and} \quad b_i = \theta b_{i+1} + \binom{n}{i} 2 p^i (1 - p)^{n-i} \tag{39}
\]

We now consider upper bounds on $\mu_i$ for the range $i \in [A - 1, B + 1], j \in [i + 1, n]$. Substituting (26), (27) and using (39), verify that

\[
\binom{n}{i} p^i (1 - p)^{n-i} 2|j - i| = \binom{n}{i} p^i (1 - p)^{n-i} 2|j - i| + \theta f_{i-1} - f_i - \theta^2 b_i + \theta b_{i+1} + (j - i)(\theta f_{i-1} - f_i) + f_i = f_i + b_i - \frac{4}{(1 - \theta^2)} \binom{n}{i} p^i (1 - p)^{n-i} \tag{40}
\]

Similarly, for $i \in [A - 1, B + 1], j \in [0, i - 1]$, substitute (26), (27) and use (39), to establish

\[
\binom{n}{i} p^i (1 - p)^{n-i} 2|j - i| = \binom{n}{i} p^i (1 - p)^{n-i} 2|j - i| + \theta b_{i+1} - \theta^2 f_i - b_i + \theta f_{i-1} + (i - j)(\theta b_{i+1} - b_i) + b_i = f_i + b_i - \frac{4}{(1 - \theta^2)} \binom{n}{i} p^i (1 - p)^{n-i} \tag{41}
\]

Suppose $i \in [A - 1, B + 1]$ and $j = i$, the upper bound on $\mu_i$ is

\[
\theta \lambda_i(i-1,i) + \theta \lambda_i(i+1,i) = \frac{\theta f_{i-1} - \theta^2 b_i + \theta b_{i+1} - \theta^2 f_i}{1 - \theta^2} = \frac{\theta f_{i-1} + \theta b_{i+1} - \frac{4 \theta^2}{(1 - \theta^2)} \binom{n}{i} p^i (1 - p)^{n-i}}{1 - \theta^2} \tag{42}
\]

The expressions in (40), (41) and (42) being equal to the assignment (29) for $\mu_i$ in the range $i \in [A - 1, B + 1]$, we conclude validity of (24). We are left to prove validity of (24) for $i \geq B_n + 1$. This is similar to (38). Substituting (28), verify that

\[
\binom{n}{i} p^i (1 - p)^{n-i} 2|j - i| + \theta \lambda_j(j, i+1, i) = \binom{n}{i} p^i (1 - p)^{n-i} 2|i - (B_n + 1)|. \tag{43}
\]

\(^6\)Note that we have not assigned values to every dual variable. For example, we have not assigned values to $\lambda_j(j+1,i)$ when $j \geq i+1$ and $j \in \{0, \ldots, A-1\}$. We can set this to any positive value, since the corresponding primal constraints are tight.
From the assignment for $\mu_i$ in (28) for $i > B_n + 1$, we have validity of (24) for $i > B_n + 1$. We have thus proved validity of (24) for all values of $i$ and $j \in [A_n - 1, B_n + 1]$. The non-negativity of $\lambda_j(i, i, i)$ and $\lambda_j((i + 1, i))$ follows from (i) definition of $A_n, B_n$, and (ii) non-negativity of $f_i, b_i$. We have thus proved that the above assignment assignment are valid primal and feasible assignments and satisfy complementary slackness conditions. We only need to evaluate the objective of one of these values and prove that it tends to $\frac{4\theta}{1 - \theta}$ in the limit $n \to \infty$. It is easier to evaluate the primal. This being a simple step of evaluating the $L_1$-distance of the geometric truncated within $[A_n - 1, B_n + 1]$, we leave this to reader. In evaluating this, we will utilize the fact that $np_1 - A_n = \mathcal{O}(\sqrt{n})$ and $B_n - np_1 = \mathcal{O}(\sqrt{n})$.

A generalization of our assignment if worthy of mention. Suppose $C^n = \frac{1}{n + 1}$, then it can be recognized that the above assignment for the primal and dual satisfies complementary slackness conditions. We note that an important element in these feasible solutions being valid is the fact that $A_n < np < B_n$. Binomial and Uniform distributions do not have ‘side lobes’, and hence $A_n < np < B_n$. In Section 12, we study $A_n, B_n$.

12 Characterization of $A_n, B_n$ defined in (22)

$A_n$ on the left and $B_n$ on the right constitute the boundaries of the support of the truncated geometric mechanism. It is instructive to study $A_n, B_n$ for different distributions $\mathcal{C}^n_i$. Suppose one replaces $\mathcal{C}^n_i$ by $\frac{1}{n + 1}$ - the uniform pmf on the set of histograms $\mathcal{H}_2^n$, then it can be shown, by simple hand calculation that $A_n \leq N_0 := \min\{|i \in \mathbb{N} : \theta^i < 1 - \theta\}$ and $B_n \geq n - N_0$.

Since this will provide us with important intuition, we first proceed with these steps. We recall definitions for ease of reference.

\[
\begin{align*}
  f_i & := 2 \sum_{j=0}^{i} \mathcal{C}^n_j \theta^{j-i}, \quad b_i := 2 \sum_{k=i}^{n} \mathcal{C}^n_k \theta^{k-i} \\
  A_n & := \min \left\{ i \in [0, n] : \; f_{k-1} - \theta b_k \geq 0 \right\}, \quad B_n := \max \left\{ i \in [0, n] : \; b_{k+1} - \theta f_k \geq 0 \right\} \\
  & \text{for every } k \geq i.
\end{align*}
\]

Since we are interested in $f_i - \theta b_i$ and $b_{i+1} - \theta f_i$, we will ignore the multiplier 2 in the definitions of $f_i$ and $b_i$. We work out a simple case to understand the core problem. Let us begin with the case $C^n_i = \frac{1}{n + 1}$ for $i \in [0, n]$.

It can be verified that

\[
\begin{align*}
  f_{i-1} - \theta b_i & = \frac{1}{n + 1} \left[ \theta^{i-1} + \theta^{i-2} + \cdots + \theta + 1 - \theta \left( 1 + \theta + \theta^2 + \cdots + \theta^{n-i} \right) \right] \\
  & = \frac{1}{n + 1} \left[ \frac{1 - \theta^i}{1 - \theta} - \theta \left( \frac{1 - \theta^{n-i+1}}{1 - \theta} \right) \right] \\
  & = \frac{1}{n + 1} \left[ \frac{1 - \theta^i}{1 - \theta} \right] \\
  & \geq \frac{1}{n + 1} \left[ \frac{1 - \theta^i}{1 - \theta} \right]
\end{align*}
\]

Clearly, $A_n < \min\{i : \theta^i < 1 - \theta\}$. A similar sequence of steps leads one to conclude that $B_n > \max\{i : \theta^{n-i} < 1 - \theta\}$. We observe $A_n = \mathcal{O}(1)$ and $n - B_n = \mathcal{O}(1)$. Our characterization for $A_n$ and $B_n$ for $C^n_i = \binom{n}{i} p^i (1 - p)^{n-i}$

13 Shadow price interpretation of dual variable assignments

We provide interpretation to assignments to dual variables in Eq. (25)-(29) via shadow prices. Assignment (26) for $j = i$ can be interpreted via mechanism $\hat{\mathbb{W}}(\cdot|\cdot)$ defined as $\hat{\mathbb{W}}(k|j) = \mathbb{W}(k|j) + \ldots + \mathbb{W}(k|n)$.
\[ d\mathcal{W}(k|j), \text{ where } \mathcal{W}(|\cdot|) \text{ is the truncated geometric mechanism defined in (23) and} \]
\[ d\mathcal{W}(k|j) = \begin{cases} 0 & \text{if } k \neq (i - 1), \text{ and } k \neq i, \\ -\epsilon \theta^{j-(i-1)} & \text{if } k = (i - 1), \\ +\epsilon \theta^{j-(i-1)} & \text{if } k = i. \end{cases} \tag{45} \]

It is straightforward to verify that \( \mathcal{W} \) satisfies all the constraints of a \( \theta \)-DP mechanism (just as \( \mathcal{W} \)), and more importantly, \( \mathcal{W}(i|i - 1) - \theta \mathcal{W}(i|i) = \epsilon(1 - \theta^2) \). In fact, except for this constraint, \( \mathcal{W} \) and \( \mathcal{W} \) are identical wrt all other constraints. \( \mathcal{W} \) and \( \mathcal{W} \) are identical vertices in their corresponding feasible region, with the only difference being \( \mathcal{W} \) satisfies the constraint \( \mathcal{W}(i|i - 1) - \theta \mathcal{W}(i|i) \geq \epsilon(1 - \theta^2) \). Moreover, it can be verified that \( D^n_{\mathcal{H}}(\mathcal{W}) - D^n_{\mathcal{H}}(\mathcal{W}) = \epsilon(f_{i-1} - \theta b_i) \).

Recognize that
\[ \lim_{\epsilon \to 0} \frac{D^n_{\mathcal{H}}(\mathcal{W}) - D^n_{\mathcal{H}}(\mathcal{W})}{\mathcal{W}(i|i - 1) - \theta \mathcal{W}(i|i)} = \lim_{\epsilon \to 0} \frac{D^n(\mathcal{W})}{\mathcal{W}(i|i - 1) - \theta \mathcal{W}(i|i)} = \frac{\epsilon(f_{i-1} - \theta b_i)}{\epsilon(1 - \theta^2)} = \lambda|_{i(i-1,i)} \]

This is indeed the shadow prices interpretation that we alluded to. We continue and discuss the interpretation for the rest of the variables. Consider assignment (26) for \( j > i \). Consider \( \mathcal{W}(\cdot|\cdot) \) defined as \( \mathcal{W}(a|b) = \mathcal{W}(a|b) + d\mathcal{W}(a|b) \), where \( \mathcal{W}(\cdot|\cdot) \) is the truncated geometric mechanism defined in (23) and \( d\mathcal{W} \) is now defined as

\[ d\mathcal{W}(a|b) = \begin{cases} 0 & \text{if } a \neq (i - 1), \text{ and } a \neq i, \text{ and } a \neq j \\ -\epsilon \theta^{b-(i-1)} & \text{if } a = (i - 1), \\ +\epsilon \theta^{b-(i-1)} & \text{if } a = i, b \geq i \\ +\epsilon \theta^{b-(i-1)+2} & \text{if } a = i, b \leq i - 1 \\ +\epsilon \theta^{b-(i-1)} - \epsilon \theta^{b-(i-1)+2} & \text{if } a = j, b \leq i - 1 \\ 0 & \text{if } a = j, b \geq i. \end{cases} \tag{46} \]

As earlier, it is straightforward to verify that \( \mathcal{W} \) satisfies all the constraints of a \( \theta \)-DP mechanism (just as \( \mathcal{W} \)), and more importantly, \( \mathcal{W}(j|i - 1) - \theta \mathcal{W}(j|i) = \epsilon(1 - \theta^2) \). In fact, except for this constraint, \( \mathcal{W} \) and \( \mathcal{W} \) are identical wrt all other constraints. Moreover, it can be verified that \( D^n_{\mathcal{H}}(\mathcal{W}) - D^n_{\mathcal{H}}(\mathcal{W}) = \epsilon(f_{i-1} - \theta b_i) \).

Recognize that
\[ \lim_{\epsilon \to 0} \frac{D^n_{\mathcal{H}}(\mathcal{W}) - D^n_{\mathcal{H}}(\mathcal{W})}{\mathcal{W}(j|i - 1) - \theta \mathcal{W}(j|i)} = \lim_{\epsilon \to 0} \frac{D^n(\mathcal{W})}{\mathcal{W}(j|i - 1) - \theta \mathcal{W}(j|i)} = \frac{\epsilon(\theta^2 f_{i-1} + (j - i + 1)(1 - \theta^2)f_{i-1} - \theta b_i)}{\epsilon(1 - \theta^2)} = \lambda|_{j(i-1,i)} \]

Now consider (27) with \( j = i \). Analogous to (45), consider

\[ d\mathcal{W}(k|j) = \begin{cases} 0 & \text{if } k \neq (i + 1), \text{ and } k \neq i, \\ -\epsilon \theta^{j-(i+1)} & \text{if } k = (i + 1), \\ +\epsilon \theta^{j-(i+1)} & \text{if } k = i. \end{cases} \tag{47} \]

Following the same arguments as above, it can be verified by straightforward substitutions that
\[ \lim_{\epsilon \to 0} \frac{D^n_{\mathcal{H}}(\mathcal{W}) - D^n_{\mathcal{H}}(\mathcal{W})}{\mathcal{W}(i|i + 1) - \theta \mathcal{W}(i|i)} = \lim_{\epsilon \to 0} \frac{D^n(\mathcal{W})}{\mathcal{W}(i|i + 1) - \theta \mathcal{W}(i|i)} = \frac{\epsilon(b_{i+1} - \theta f_i)}{\epsilon(1 - \theta^2)} = \lambda|_{i(i+1,i)} \]

where, as before, \( \mathcal{W}(\cdot|\cdot) \) defined as \( \mathcal{W}(k|j) = \mathcal{W}(k|j) + d\mathcal{W}(k|j), \mathcal{W}(\cdot|\cdot) \) is the truncated geometric mechanism. Similarly, for \( j < i \) we can verify the assignment in (27) through the following. Define mechanism \( \mathcal{W}(\cdot|\cdot) = \mathcal{W}(k|j) + d\mathcal{W}(k|j), \mathcal{W}(\cdot|\cdot) \) is the truncated geometric mechanism
The following can be verified easily.

\[
dW(a|b) = \begin{cases} 
0 & \text{if } a \neq (i+1), \text{ and } a \neq i, \text{ and } a \neq j \\
-\epsilon \theta^{(i+1)j} & \text{if } a = (i+1), \\
+\epsilon \theta^{i(i+1)} & \text{if } a = i, b > i \\
+\epsilon \theta^{i(i+1)+2} & \text{if } a = i, b \geq i + 1 \\
+\epsilon (\theta^{i(i+1)} - \theta^{(i+1)i+2}) & \text{if } a = j, b \geq i + 1 \\
0 & \text{if } a = j, b \leq i.
\end{cases}
\]

Following the same arguments as above, it can be verified by straightforward substitutions that

\[
\lim_{\epsilon \to 0} \frac{D^n_H(\hat{W}) - D^n_H(W)}{\hat{W}(j|i+1) - \theta W(j|i)} = \lim_{\epsilon \to 0} \frac{D^n(dW)}{\hat{W}(j|i+1) - \theta W(j|i)} = \frac{\epsilon(\theta^2 b_{i+1} + (i+1-j)(1-\theta^2)b_{i+1} - \theta f_i)}{\epsilon(1-\theta^2)} = \lambda_{j(i+1,i)}.
\]

where, as before, \(\hat{W}(\cdot|\cdot)\) defined as \(\hat{W}(k|j) = W(k|j) + dW(k|j), \ W(\cdot|\cdot)\) is the truncated geometric mechanism. Finally, we explain the assignment for \(\mu_i\) in the range \([A-1, B+1]\). Consider \(W(b|a) = W(b|a) + dW(b|a)\) where \(W\) is the truncated Geometric mechanism as before and

\[
dW(a|b) = \begin{cases} 
0 & \text{if } a \neq (i-1), \text{ and } a \neq i, \text{ and } a \neq (i+1) \\
-\epsilon \theta^{(i-1)i} & \text{if } a = (i-1), \\
-\epsilon \theta^{i(i+1)} & \text{if } a = i + 1, \\
+\epsilon \theta^{(i-1)i+1} + \epsilon \theta^{i(i+1)+1} & \text{if } a = i, b \neq i \\
+\epsilon(1 + \theta^2) & \text{if } a = i, b = i
\end{cases}
\]

The following can be verified easily. \(\sum_{j=0}^n \hat{W}(j|i) = 1 + \epsilon - \epsilon \theta^2, \ W\) and \(W\) are identical in which DP constraints they satisfy and

\[
\lim_{\epsilon \to 0} \frac{D^n_H(\hat{W}) - D^n_H(W)}{\sum_{j=0}^n \hat{W}(j|i) - 1} = \lim_{\epsilon \to 0} \frac{D^n(dW)}{\sum_{j=0}^n W(j|i) - 1} = \frac{\epsilon[\theta(1 - \theta^2)(f_{i-1} + b_{i+1}) - 4\theta^2\binom{n}{i}p^i(1-p)^{n-i}]}{\epsilon(1-\theta^2)} = \mu_i.
\]

The key import of the above interpretation is the relationship between the assignments (45)-(49). (46) can be got from (45) by just shifting mass from \(i\) to \(j\). Similarly, (48) can be got from (47) by just shifting mass from \(i\) to \(j\). This provides an alternate proof of feasibility of this dual variable assignment. Also note that the assignment (49) is obtained as \(\theta\) times the assignment (45) summed to \(\theta\) times the assignment (47). The feasibility of this assignment is now an immediate consequence of these relationships. This shadow price interpretation is the basis for (32), whose feasibility follows immediately from the geometry of the constraints.