Differential Privacy on the Unit Simplex via the Dirichlet Mechanism

Parham Gohari, Bo Wu, Calvin Hawkins, Matthew Hale*, Ufuk Topcu

Abstract

As members of network systems share more information among agents and with network providers, sensitive data leakage raises privacy concerns. Motivated by such concerns, we introduce a novel mechanism that privatizes vectors belonging to the unit simplex. Such vectors can be found in many applications, such as privatizing a decision-making policy in a Markov decision process. We use differential privacy as the underlying mathematical framework for this work. The introduced mechanism is a probabilistic mapping that maps a vector within the unit simplex to the same domain using a Dirichlet distribution. We find the mechanism well-suited for inputs within the unit simplex because it always returns a privatized output that is also in the unit simplex. Therefore, no further projection back onto the unit simplex is required. We verify and quantify the privacy guarantees of the mechanism for three cases: identity queries, average queries, and general linear queries. We establish a trade-off between the level of privacy and the accuracy of the mechanism output, and we introduce a parameter to balance the trade-off between them. Numerical results illustrate the proposed mechanism.

I. Introduction

In many decision-making problems, a policy-maker forms a control policy based on data collected from individuals in a network. The gathered data often contains sensitive information, which raises privacy concerns, e.g., for smart appliances [1]. In some applications, privatizing sensitive data has been achieved by adding carefully calibrated noise to sensitive data and functions thereof [2], [3], [4]. These noise-additive approaches are well-suited to some classes of numerical data, though sensitive data may take a form ill-suited to them. For example, developments in [5] explored symbolic control systems in which additive noise cannot be meaningfully implemented.

In this work, we privatize data inputs that belong to the unit simplex, *i.e.*, the set of vectors with non-negative entries that sum to one. Such vectors are seen in many decision-making problems. For example, in Markov decision processes (MDPs), the goal is to find a total-reward-maximizing decision policy [6], [7]. In some cases, it is shown that the optimal policy is a randomized function that maps each of an MDP's states to a probability distribution

Parham Gohari is with the Department of Electrical and Computer Engineering, University of Texas at Austin, Austin, TX. Bo Wu and Ufuk Topcu are with the Department of Aerospace Engineering and Engineering Mechanics, and the Oden Institute for Computational Engineering and Sciences, University of Texas at Austin, Austin, TX. Email: {pgohari, bwu3, utopcu}@utexas.edu. PG, BW, and UT were supported by grants no. AFRL FA9550-19-1-0169, DARPA D19AP00004, and ARL ACC-APG-RTP W911NF.

Calvin Hawkins and Matthew Hale are with the Department of Mechanical and Aerospace Engineering at the University of Florida, Gainesville, FL. Email: {calvin.hawkins,matthewhale}@ufl.edu. CH and MH were supported by the AFOSR Center of Excellence on Assured Autonomy in Contested Environments (grant no. FA9550-19-1-0169) and by NSF CAREER Grant #1943275.

*Corresponding author.

on the set of actions available at that state, see, *e.g.*, [8], [9], [10]. Finite action sets give rise to discrete, finitely supported probability distributions, which can be represented as vectors with non-negative entries summing to one, i.e., elements of the unit simplex. Policies of this kind arise in applications such as autonomous driving [11] and the smart power grid [12], and revealing them can therefore reveal individuals' behaviors. Thus, there is a need to privatize such policies, and this represents one application of privatizing sensitive data in the unit simplex. Existing noise-additive approaches will not, in general, produce a privatized vector in the unit simplex. Projecting these privatized vectors back onto the simplex leads to poor accuracy of privatized data (which we illustrate in Section II-C). We therefore propose a new approach to privatization for this context.

In this paper, we use differential privacy as the underlying mathematical framework for privacy. Differential privacy, first introduced in [13], is designed to protect the exact values of sensitive pieces of data, while preserving their usefulness in statistical analyses. Two desirable properties of differential privacy are (i) that it is immune to post-processing [14], in the sense that arbitrary post-hoc transformations of privatized data do not weaken its privacy guarantees, and (ii) that it is robust to side information, in that gaining additional information about data-producing entities does not weaken its privacy guarantees by much [15]. As a result, differential privacy has been frequently used as the mathematical formulation of privacy in both computer science and, more recently, in control theory [16], [17], [18], [19], [20].

As the main contribution of this paper, we introduce a novel mechanism that privatizes a vector within the unit simplex. A mechanism is a probabilistic mapping from some pre-defined domain to a pre-defined range, and a mechanism is used to privatize sensitive data. This paper develops a novel mechanism using the Dirichlet distribution, and we therefore call it the Dirichlet mechanism. The Dirichlet distribution is a multivariate distribution supported on the unit simplex, which makes it a natural choice for this setting because its outputs are always elements of the unit simplex.

In our developments, we use probabilistic differential privacy, which is known to imply that the conventional form of differential privacy also holds [21]. Then, we show that the Dirichlet mechanism satisfies probabilistic differential privacy for identity queries. By an identity query, we mean privatizing a single vector within the unit simplex. In the course of proving these privacy guarantees, based on the assumptions we provide, we prove the log-concavity of the cumulative distribution function of a Dirichlet distribution. The proof that we present may be of independent interest in ongoing research on convexity analysis of special functions such as [22].

Beyond identity queries, we further show that the Dirichlet mechanism is differentially private for both average queries and general linear queries, in which we privatize operations over collections of vectors, each of which is contained in the unit simplex. We derive analytic expressions for privacy levels of both cases.

We also analyze the accuracy of the output of the Dirichlet mechanism. In particular, we evaluate the accuracy of the Dirichlet mechanism in terms of the expected value and the variance of its outputs. Similar to additive noise methods, the Dirichlet mechanism output has the same expected value as its input, which implies that its privatized outputs obey a distribution centered on the underlying sensitive data. We show that there exists a trade-off between privacy levels and the extent to which privatized data is concentrated around the underlying sensitive data.

We emphasize that additive noise privacy mechanisms are ill-suited to privacy on the unit simplex because they add noise of infinite support. As a result, such mechanisms will output vectors that do not belong to the unit simplex; attempting to normalize the noise would result in its distribution not being one known to provide differential privacy. Furthermore, the complexity of the projection onto the unit simplex makes it difficult to provide meaningful accuracy guarantees of privatized data, which are essential in differential privacy. It is for these reasons that we develop the Dirichlet mechanism. Although its form appears quite different from existing mechanisms, they are related through membership in a broad class of probability distributions. In particular, the Laplacian, Gaussian, and exponential mechanisms all use distributions belonging to a parameterized family of exponential distributions. The outputs of the Dirichlet distribution are equivalent to a normalized vector of i.i.d. exponential random variables, which means their distribution also belongs to the exponential family. This connection reveals why we should expect the Dirichlet mechanism to be well-suited to differential privacy, and the developments of this paper formalize and confirm this intuition.

We also point out here that the exponential mechanism is another widely used differentilly private mechanism which can be used for sensitive data ill-suited to additive approaches [14]. However, the exponential mechanism can be computationally demanding to implement for privacy applications with many possible outputs. The output space here is the unit simplex, which contains uncountably many elements. The resulting complexity of such an implementation therefore makes it infeasible [23], especially in large dimensions, and we avoid it here.

A preliminary version of this work appeared in [24]. The current paper develops additional privatization techniques for general linear queries, provides concentration bounds to assess the accuracy of the Dirichlet mechanism, and provides full proofs of all results.

The rest of the paper is organized as follows. Section II establishes the privacy preliminaries needed in the rest of the paper. Then, Section III establishes privacy guarantees for identity queries, Section IV establishes privacy guarantees for averaging queries, and Section V establishes privacy guarantees for general linear queries. Section VI provides accuracy bounds on the Dirichlet mechanism's outputs, and Section VII provides simulations to illustrate our results. Finally, Section VIII concludes the paper.

II. Preliminaries

A. Notation

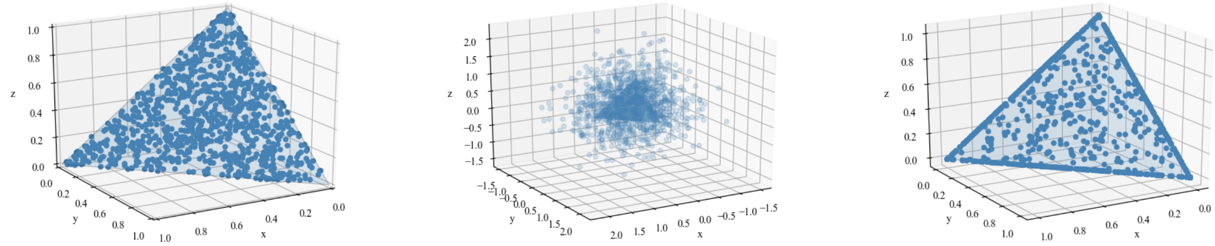
In this section we establish notation used throughout the paper. We represent the real numbers by $\mathbb{R}$ and the positive reals by $\mathbb{R}_+$. For a positive integer n, let $[n] := \{1, \ldots, n\}$. We denote the unit simplex in $\mathbb{R}^n$ by Δ_n where

$$\Delta_n := \left\{ x \in \mathbb{R}^n \mid \sum_{i=1}^{n} x_i = 1, x_i \geq 0 \text{ for all } i \in [n] \right\}.$$

We use Δ_n° to represent the interior of Δ_n. Letting $W \subseteq [n-1]$ with $|W| \geq 2$ we then define the set

$$\Delta_{n,W}^{(\eta,\bar{\eta})} := \left\{ p \in \Delta_n^\circ \mid \sum_{i \in W} p_i \leq 1 - \bar{\eta}, p_i \geq \eta \text{ for all } i \in W \right\}.$$

We impose the following assumption on η and $\bar{\eta}$ that will be used below to ensure that ratios of Dirichlet distributions remain bounded when showing that they provide differential privacy.



(a) Vectors in the unit simplex (b) Vectors after adding Gaussian noise (c) Projections of noisy vectors

Fig. 1: Left-hand figure: a random sampling of $n = 1,000$ vectors from the unit simplex in $\mathbb{R}^3$. Center figure: the privatized forms of all n data points after adding Gaussian noise from the distribution $\mathcal{N}(0, 0.25I)$. Right-hand figure: the results of projecting all privatized vectors back onto the unit simplex. Adding Gaussian noise and projecting back onto the simplex results in many points accumulating at the boundary of the simplex, which harms accuracy.

Assumption 1. *In* $\Delta_{n,W}^{(\eta,\bar{\eta})}$*,* $\eta > 0$*,* $\bar{\eta} > 0$*, and* $\eta + \bar{\eta} < \frac{1}{2}$*.*

Letting p be a vector in $\mathbb{R}^n$, we use the notation $p_{(i,j)}$ to denote the vector $(p_i, p_j)^T \in \mathbb{R}^2$, where $(\cdot)^T$ is the transpose of a vector, and $p_{-(i,j)} \in \mathbb{R}^{n-2}$ to denote the vector p with i^{th} and j^{th} entries removed. $\mathbb{P}[\cdot]$ denotes the probability of an event. For a random variable, $\mathbb{E}[\cdot]$ denotes its expectation and Var$[\cdot]$ denotes its variance. We use the notation $|\cdot|$ for the cardinality of a finite set. $||\cdot||_1$ denotes the 1-norm of a vector. We also use the special functions

$$\Gamma(x) = \int_0^\infty z^{x-1} \exp(-z) dz, \qquad x \in \mathbb{R}_+$$

$$\text{beta}(a, b) = \int_0^1 t^{a-1}(1-t)^{b-1} dt = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}, \qquad a, b \in \mathbb{R}_+$$

$$\psi^{(0)}(x) = \frac{d}{dx} \log\big(\Gamma(x)\big), \ \psi^{(1)}(x) = \frac{d^2}{dx^2} \log\big(\Gamma(x)\big), \qquad x \in \mathbb{R}_+$$

which are the gamma, beta, digamma, and trigamma functions, respectively.

B. Differential Privacy

Intuitively, differential privacy guarantees that two *nearby* pieces of sensitive data will have statistically similar privatized values. In differential privacy, the notion of "nearby" is formally defined by an adjacency relation, and we define adjacency over the unit simplex as follows.

Definition 1. *For a constant* $b \in (0, 1]$ *and fixed set* $W \subset [n-1]$*, two vectors* $p, q \in \Delta_{n,W}^{(\eta,\bar{\eta})}$ *are said to be* b*-adjacent if there exist indices* $i, j \in W$ *such that*

$$p_{-(i,j)} = q_{-(i,j)} \text{ and } ||p - q||_1 \leq b.$$

We express this condition with the binary symmetric adjacency relation

$$\mathrm{Adj}_b(p,q) = \begin{cases} 1 & p \text{ and } q \text{ are adjacent} \\ 0 & \text{otherwise} \end{cases}.$$

In words, two vectors are adjacent if they differ in two entries by an amount not more than b. Conventional differential privacy considers sensitive data differing in a single entry, *e.g.*, one entry in a database [14]. However, it is not possible to do so for an element of the unit simplex because changing only a single entry would violate the condition that vectors' entries sum to one. We therefore consider privacy with the above adjacency relation. Differential privacy itself is defined next.

Definition 2. *(Probabilistic differential privacy; [25]) Let* $b \in (0,1]$ *and* $W \subseteq [n-1]$ *be given. Fix a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$*. A mechanism* $\mathcal{M} : \Delta_{n,W}^{\eta,\bar{\eta}} \times \Omega \to \Delta_n$ *is said to be probabilistically* (ϵ,δ)*-differentially private if, for all* $p \in \Delta_{n,W}^{\eta,\bar{\eta}}$*, we can partition the output space* Δ_n *into two disjoint sets* Ω_1, Ω_2*, such that*

$$\mathbb{P}[\mathcal{M}(p) \in \Omega_2] \leq \delta,$$

and for all $q \in \Delta_{n,W}^{\eta,\bar{\eta}}$ *b-adjacent to* p *and for all* $x \in \Omega_1$*,*

$$\log\left(\frac{\mathbb{P}[\mathcal{M}(p) = x]}{\mathbb{P}[\mathcal{M}(q) = x]}\right) \leq \epsilon.$$

(ϵ,δ)-probabilistic differential privacy is known to imply conventional (ϵ,δ)-differential privacy [25], and in Sections III-V we prove that the Dirichlet mechanism satisfies Definition 2 as a means of showing that it provides conventional differential privacy.

C. On Additive Noise Approaches

It was noted in the Introduction that differential privacy (in both its probabilistic and ordinary forms) can be enforced with additive Gaussian or Laplacian noise. For simplex-valued data, adding noise with infinite support can perturb data outside the simplex, and some form of projection would be required to ensure membership in the unit simplex after privatization. However, we show in Figure 1 that adding Gaussian noise to elements of the simplex and projecting back onto the simplex results in biases in private data. Specifically, Figure 1 shows elements of the simplex to which Gaussian noise has been added, after which the projection algorithm in [26] is used to project them onto the simplex. The variance of noise added is $\sigma^2 = 0.25$. Using [27, Theorem 3], this level of noise provides $(1, 0.01)$-differential privacy (for an adjacency relation that classifies vectors as adjacent if their 2-norm distance is bounded above by 0.1).

It can be seen that the privatized-then-projected forms of many elements of the simplex are simply mapped to its boundary. The reason is that adding infinite-support noise to a point in the simplex will often move it out of the simplex, and the projection step maps such points to the boundary of the simplex. This both impairs the accuracy of the mechanism itself, because the privatized forms of data points can be far from their original, sensitive values, and harms any downstream uses of this data. Rather than taking this approach, accuracy could be improved by effectively utilizing the interior of the simplex to enforce the approximate indistinguishability required by differential privacy, and that is the subject of the next subsection.

D. Dirichlet Mechanism

One contribution of this paper is to present a differentially private mechanism that, without any need of projection, maps elements of Δ_n to Δ_n. In order to do so, we first introduce the Dirichlet mechanism. A Dirichlet mechanism with parameter $k \in \mathbb{R}_+$, denoted by $\mathcal{M}_D^{(k)}$, takes as input a vector $p \in \Delta_n^\circ$ and outputs $x \in \Delta_n$ according to the Dirichlet probability distribution function (PDF) centered on p, *i.e.*,

$$\mathbb{P}[\mathcal{M}_D^{(k)}(p) = x] = \frac{1}{\mathrm{B}(kp)} \prod_{i=1}^{n-1} x_i^{kp_i - 1} \left(1 - \sum_{i=1}^{n-1} x_i\right)^{kp_n - 1}, \tag{1}$$

where

$$\mathrm{B}(kp) := \frac{\prod_{i=1}^{n} \Gamma(kp_i)}{\Gamma\left(k \sum_{i=1}^{n} p_i\right)} \tag{2}$$

is the multi-variate beta function. For brevity, we will use the notation Dir_k to denote the PDF on the right-hand side of (1). We impose the following assumption on the parameter k.

Assumption 2. *For the Dirichlet mechanism $\mathcal{M}_D^{(k)}$, the parameter k satisfies*

$$k \geq \max\left\{\frac{1}{\eta}, \frac{1}{1 - \eta - \bar{\eta}}\right\}.$$

We later use the parameter k to adjust the trade-off that we establish between the accuracy and the privacy level of the Dirichlet mechanism. Next, we establish the privacy guarantees that the Dirichlet mechanism provides.

III. Dirichlet Mechanism for Differential Privacy of Identity Queries

We begin by analyzing identity queries under the Dirichlet mechanism. Here, a sensitive vector p is directly input to the Dirichlet mechanism to make it approximately indistinguishable from other adjacent sensitive vectors. The space of sensitive data of interest is $\Delta_{n,W}^{(\eta,\bar{\eta})}$, and it is over this space that we provide privacy. To show the level of privacy that holds, we first bound δ, then bound ϵ.

A. Computing δ

Fix $W \subseteq [n-1]$. In accordance with Definition 2, we partition the output space of the Dirichlet mechanism into two sets Ω_1, Ω_2 defined by

$$\Omega_1 := \{x \in \Delta_n \mid x_i \geq \gamma \text{ for all } i \in W\} \tag{3}$$

and $\Omega_2 := \Delta_n \backslash \Omega_1$, where $\gamma \in (0,1)$ is a parameter that defines these sets, upon which we impose the following.

Assumption 3. *Fix $W \subseteq [n-1]$. Then $\gamma \leq \frac{1}{|W|}$.*

Next, our goal is to show that the Dirichlet mechanism output belongs to Ω_1 with high probability. Let p be a vector in $\Delta_{n,W}^{(\eta,\bar{\eta})}$. In the next lemma we show how to calculate $\mathbb{P}[\mathcal{M}_D^{(k)}(p) \in \Omega_1]$. As in Definition 2, we will bound $\mathbb{P}[\mathbb{M}_D^{(k)}(p) \in \Omega_2]$ by δ and use ϵ to bound the ratios of distributions of outputs generated by inputs in Ω_1. Changing γ changes Ω_1 and Ω_2, and its role in determining ϵ and δ will be elaborated upon below.

Lemma 1. *Let Assumptions 1 and 3 hold. Let $W \subseteq [n-1]$, let $p \in \Delta_{n,W}^{(\eta,\bar{\eta})}$, and let*

$$\mathcal{A}_r := \left\{ x \in \mathbb{R}^{r-1} \mid \sum_{i\in[r-1]} x_i \leq 1, x_i \geq \gamma \text{ for all } i \in W \right\},$$

for all $r \geq |W|+1$. Then, for a Dirichlet mechanism with parameter $k \in \mathbb{R}_+$, we have that $\mathbb{P}[\mathcal{M}_D^{(k)}(p) \in \Omega_1]$ equals

$$\frac{\int_{\mathcal{A}_{|W|+1}} \prod_{i\in W} x_i^{kp_i-1}\left(1-\sum_{i\in W} x_i\right)^{k(1-\sum_{i\in W} p_i)-1} \prod_{i\in W} dx_i}{\mathrm{B}(k\tilde{p}_W)},$$

where $\tilde{p}_W \in \Delta_{|W|+1}$ is equal to p after removing entries with indices outside W and with an additional entry equal to $1-\sum_{i\in W} p_i$ appended as its new final entry.

Proof. For concreteness we set $W = [n-1]$, though the proof is identical for other cases. In order to find $\mathbb{P}[\mathcal{M}_D^{(k)}(p) \in \Omega_1]$, we need to integrate the Dirichlet PDF over the region $\mathcal{A}_n$. Therefore, we need to evaluate the $(n-1)$-fold integral

$$\frac{\int_{\mathcal{A}_n} \left(\prod_{i=1}^{n-1} x_i^{kp_i-1}\right)\left(1-\sum_{i=1}^{n-1} x_i\right)^{kp_n-1} dx_{n-1}\dots dx_1}{\mathrm{B}(kp)}. \tag{4}$$

Using a method similar to the one adopted in [28], let $y := \sum_{i=1}^{n-2} x_i$. Then we can rewrite (4) as

$$\frac{1}{\mathrm{B}(kp)}\int_{\mathcal{A}_{n-1}}\int_0^{1-y}\left(\prod_{i=1}^{n-1} x_i^{kp_i-1}\right)(1-y-x_{n-1})^{kp_n-1}dx_{n-1}\dots dx_1. \tag{5}$$

Now let $u := \frac{x_{n-1}}{1-y}$ and take the inner integral with respect to u. Then (5) becomes

$$\frac{1}{\mathrm{B}(kp)}\int_{\mathcal{A}_{n-1}}\prod_{i=1}^{n-2} x_i^{kp_i-1}(1-y)^{k(p_{n-1}+p_n)-1}\int_0^1 u^{kp_{n-1}-1}(1-u)^{kp_n-1}du\ dx_{n-2}\dots dx_1.$$

From the definition of the beta function, we have

$$\int_0^1 u^{kp_{n-1}-1}(1-u)^{kp_n-1}du = \text{beta}(kp_{n-1}, kp_n).$$

Using the gamma function representation of beta functions, *i.e.*,

$$\text{beta}(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)},\ a,b\in\mathbb{R}_+, \tag{6}$$

and (2), we find that $\mathbb{P}[\mathcal{M}_D^{(k)}(p) \in \Omega_1]$ is equal to

$$\frac{1}{\mathrm{B}(kp)}\frac{\Gamma(kp_{n-1})\Gamma(kp_n)}{\Gamma(k(p_{n-1}+p_n))}\int_{\mathcal{A}_{n-1}}\prod_{i=1}^{n-2} x_i^{kp_i-1}\left(1-\sum_{i=1}^{n-2} x_i\right)^{k(p_{n-1}+p_n)-1} dx_{n-2}\dots dx_1.$$

Using the same idea, for the next step, let $y := \sum_{i=1}^{n-3} x_i$ and $u := \frac{x_{n-2}}{1-y}$. Then $\mathbb{P}[\mathcal{M}_D^{(k)}(p) \in \Omega_1]$ is equal to

$$\frac{1}{\mathrm{B}(kp)}\frac{\Gamma(kp_{n-2})\Gamma(kp_{n-1})\Gamma(kp_n)}{\Gamma(k(p_{n-2}+p_{n-1}+p_n))}\int_{\mathcal{A}_{n-2}}\prod_{i=1}^{n-3} x_i^{kp_i-1}\left(1-\sum_{i=1}^{n-3} x_i\right)^{k(p_{n-2}+p_{n-1}+p_n)-1} dx_{n-3}\dots dx_1.$$

We continue to adopt the same change of variable strategy until we are left with an integral over the region $\mathcal{A}_{|W|+1}$, which concludes the proof. ■

Lemma 1 shows that instead of an $(n-1)$-fold integral of the Dirichlet PDF, the computations can be reduced to a $|W|$-fold integral. However, the expression still depends on the input vector p, which is undesirable and generally incompatible with differential privacy. The reason is that (ϵ,δ)-differential privacy must be a guarantee for all adjacent input data and not for a specific data point. In the next lemma, we show that $\mathbb{P}[\mathcal{M}_D^{(k)}(p) \in \Omega_1]$ is a log-concave function of p over $\Delta_{n,W}^{(\eta,\bar{\eta})}$, which we will use to derive a bound for δ that holds for all p of interest.

Lemma 2. *Let Assumption 1 hold, fix $W \subseteq [n-1]$, and let $\mathcal{M}_D^{(k)}$ be the Dirichlet mechanism with parameter k. Then $\mathbb{P}[\mathcal{M}_D^{(k)}(p) \in \Omega_1]$ is a log-concave function of p over the domain $\Delta_{n,W}^{(\eta,\bar{\eta})}$.*

Proof: See Appendix A. ■

Revisiting the definitions of Ω_1, Ω_2 above, we find that

$$\begin{aligned}\mathbb{P}[\mathcal{M}_D^{(k)}(p) \in \Omega_2] &= 1 - \mathbb{P}[\mathcal{M}_D^{(k)}(p) \in \Omega_1] \\ &\leq 1 - \min_p \mathbb{P}[\mathcal{M}_D^{(k)}(p) \in \Omega_1] = \delta.\end{aligned} \tag{7}$$

From this, we see that bounding δ can be done by minimizing $\mathbb{P}[\mathcal{M}_D^{(k)}(p) \in \Omega_1]$, an explicit form of which was given in Lemma 1. In Lemma 2, we established the log-concavity of the function that we seek to minimize. As a result, instead of minimizing $\mathbb{P}[\mathcal{M}_D^{(k)}(p) \in \Omega_1]$ over the entirety of $\Delta_{n,W}^{(\eta,\bar{\eta})}$, we can only consider the extreme points. Note that the points within $\Delta_{n,W}^{(\eta,\bar{\eta})}$ form a polyhedron with at most $|W|(|W|+1)/2$ vertices. As the minimum of an unsorted list of n entries can be found in linear time, the time complexity of finding $\min \mathbb{P}[\mathcal{M}_D^{(k)}(p) \in \Omega_1]$ is $\mathcal{O}(|W|^2)$. This analytical bound will be further explored through numerical results in Section VII. Next, we develop analogous bounds for ϵ.

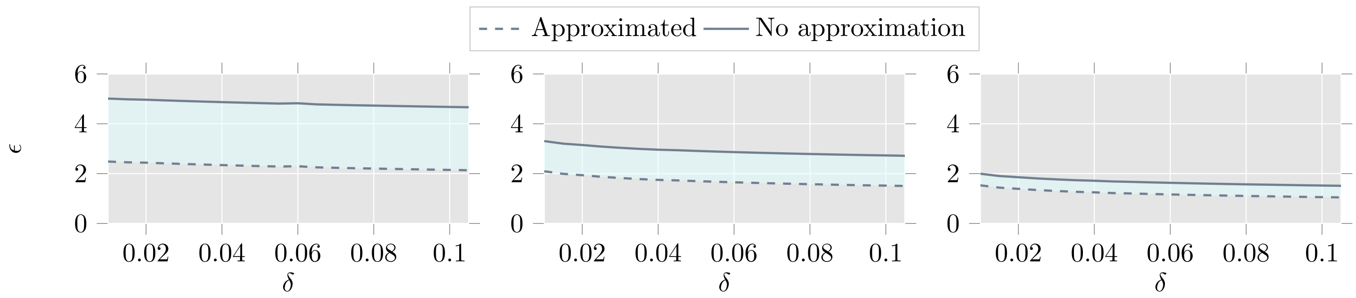
B. Computing ϵ

As above, fix $\eta, \bar{\eta} \in (0,1)$ satisfying Assumption 1, $b \in (0,1]$, and $W \subseteq [n-1]$. Then, for a given $k \in \mathbb{R}_+$, bounding ϵ requires evaluating the term

$$\log\left(\frac{\mathbb{P}[\mathcal{M}_D^{(k)}(p) = x]}{\mathbb{P}[\mathcal{M}_D^{(k)}(q) = x]}\right)$$

for all $x \in \Omega_1$, where p and q are any b-adjacent vectors in $\Delta_{n,W}^{(\eta,\bar{\eta})}$. Let $i,j \in W$ be the indices in which p and q differ. Using the definition of the Dirichlet mechanism, we find

$$\begin{aligned}\log\left(\frac{\mathbb{P}[\mathcal{M}_D^{(k)}(p) = x]}{\mathbb{P}[\mathcal{M}_D^{(k)}(q) = x]}\right) &= \log\left(\frac{\mathrm{B}(kq)\prod_{i=1}^{n} x_i^{kp_i-1}}{\mathrm{B}(kp)\prod_{i=1}^{n} x_i^{kq_i-1}}\right) \\ &= \log\left(\frac{\Gamma(kq_i)\Gamma(kq_j)x_i^{kp_i-1}x_j^{kp_j-1}}{\Gamma(kp_i)\Gamma(kp_j)x_i^{kq_i-1}x_j^{kq_j-1}}\right) \\ &= \log\left(\frac{\Gamma(kq_i)\Gamma(kq_j)}{\Gamma(kp_i)\Gamma(kp_j)}x_i^{k(p_i-q_i)}x_j^{k(p_j-q_j)}\right).\end{aligned}$$



(a): $\eta = 0.15$, $\bar{\eta} = 0.15$, $k = 6.7$ (b): $\eta = 0.20$, $\bar{\eta} = 0.20$, $k = 5.1$ (c): $\eta = 0.25$, $\bar{\eta} = 0.25$, $k = 4.1$

Fig. 2: An example where $|W| = 3$ to compare the values of ϵ in Theorem 1 with those computed numerically. At each level of δ, first γ is optimized according to the optimization problem in (14), then the optimal γ is substituted in the expressions for the original and approximated values.

Since p and q are b-adjacent, we have that $p_i + p_j = q_i + q_j$. Therefore, we can compute ϵ by evaluating the term

$$\log\left(\frac{\Gamma(kq_i)\Gamma(kq_j)}{\Gamma(kp_i)\Gamma(kp_j)}\left(\frac{x_i}{x_j}\right)^{k(p_i-q_i)}\right). \tag{8}$$

Note that if either x_i or x_j goes to 0, then the term in (8) would be unbounded. Recalling that the indices at which p and q can differ are restricted to the set W, we find that the values at these indices must be bounded below by η, and therefore the ratios of interest remain bounded as well.

Lemma 4 below will provide an explicit value of ϵ, aided in part by the following lemma.

Lemma 3. *Let Assumptions 1 and 2 hold. Let W be a given set of indices which is used to construct $\Delta_{n,W}^{(\eta,\bar{\eta})}$ and let p, q be any b-adjacent vectors in $\Delta_{n,W}^{(\eta,\bar{\eta})}$ with their i^{th} and j^{th} entries different. Then, for a constant $k \in \mathbb{R}_+$, we have that*

$$\frac{\text{beta}(kq_i, kq_j)}{\text{beta}(kp_i, kp_j)} \leq \frac{\text{beta}(kq_i, k(1-\bar{\eta}-q_i))}{\text{beta}(kp_i, k(1-\bar{\eta}-p_i))}.$$

Proof: See Appendix B. ■

Lemma 4. *Let Assumptions 1, 2, and 3 hold, let Ω_1 be as defined in (3), let $W \subseteq [n-1]$, and let $\mathcal{M}_D^{(k)}$ be a Dirichlet mechanism with parameter k. Then, for all $x \in \Omega_1$ we have that*

$$\log\left(\frac{\mathbb{P}[\mathcal{M}_D^{(k)}(p) = x]}{\mathbb{P}[\mathcal{M}_D^{(k)}(q) = x]}\right) \leq \log\left(\frac{\text{beta}(k\eta, k(1-\bar{\eta}-\eta))}{\text{beta}(k(\eta+\frac{b}{2}), k(1-\bar{\eta}-\eta-\frac{b}{2}))}\right) + \frac{kb}{2}\log\left(\frac{1-(|W|-1))\gamma}{\gamma}\right),$$

where the parameter $\gamma \in (0,1)$ takes the same value of γ used to compute δ in Section III-A.

Proof. From (8) we know that

$$\log\left(\frac{\mathbb{P}[\mathcal{M}_D^{(k)}(p)=x]}{\mathbb{P}[\mathcal{M}_D^{(k)}(q)=x]}\right) = \log\left(\frac{\Gamma(kq_i)\Gamma(kq_j)}{\Gamma(kp_i)\Gamma(kp_j)}\left(\frac{x_i}{x_j}\right)^{k(p_i-q_i)}\right).$$

Let

$$
\begin{aligned}
v := \quad \max_{p,q,x\in\mathbb{R}^n} \quad & \log\left(\frac{\Gamma(kq_i)\Gamma(kq_j)}{\Gamma(kp_i)\Gamma(kp_j)}\left(\frac{x_i}{x_j}\right)^{k(p_i-q_i)}\right) \\
\text{subject to} \quad & |p_i - q_i| \le \frac{b}{2}, \\
& p_i + p_j = q_i + q_j, \\
& p_i + p_j \le 1 - \bar{\eta}, \\
& p_{(i,j)} \in [\eta, 1-\bar{\eta}-\eta]^2, \\
& q_{(i,j)} \in [\eta, 1-\bar{\eta}-\eta]^2, \\
& x_{(i,j)} \in [\gamma, 1-(|W|-1)\gamma]^2,
\end{aligned} \tag{9}
$$

and let $\mathcal{C}$ denote the set of feasible points of the optimization problem in (9); we note that the first two constraints enforce adjacency, while the others encode $p, q \in \Delta_{n,W}^{(\eta,\bar{\eta})}$ and $x \in \Omega_1$. Assumptions 1-3 ensure that all intervals above are non-empty.

By sub-additivity of the maximum, we have

$$
v \le \max_{p,q,x\in\mathcal{C}} \log\left(\frac{\Gamma(kq_i)\Gamma(kq_j)}{\Gamma(kp_i)\Gamma(kp_j)}\right) + \max_{p,q,x\in\mathcal{C}} \log\left(\frac{x_i}{x_j}\right)^{k(p_i-q_i)}. \tag{10}
$$

Now, with

$$
v_1 := \max_{p,q,x\in\mathcal{C}} \log\left(\frac{x_i}{x_j}\right)^{k(p_i-q_i)},
$$

we find

$$
\begin{aligned}
v_1 &\le \max_{p,q,x\in\mathcal{C}} |k(p_i-q_i)| \left|\log\left(\frac{x_i}{x_j}\right)\right| \\
&= \frac{kb}{2}\log\left(\frac{1-(|W|-1)\gamma}{\gamma}\right),
\end{aligned}
$$

where Assumption 3 ensures that the argument of the logarithm is positive.

Next, let $c := p_i + p_j = q_i + q_j$ and substitute q_j, p_j with $c - q_i$ and $c - p_i$ respectively. Let

$$
\begin{aligned}
v_2 := \quad \max_{p_i,q_i,c\in\mathbb{R}} \quad & \log\left(\frac{\Gamma(kq_i)\Gamma(k(c-q_i))}{\Gamma(kp_i)\Gamma(k(c-p_i))}\right) \\
\text{subject to} \quad & |p_i - q_i| \le \frac{b}{2}, \\
& c \in [2\eta, 1-\bar{\eta}], \\
& p_i \in [\eta, 1-\bar{\eta}-\eta], \\
& q_i \in [\eta, 1-\bar{\eta}-\eta],
\end{aligned}
$$

where the constraints again encode adjacency of p and q and their containment in $\Delta_{n,W}^{(\eta,\bar{\eta})}$.

Next, either $q_i < p_i$ or $q_j < p_j$, and we assume without loss of generality that $q_i < p_i$. Then, from Lemma 3 and (6), we have that

$$
\begin{aligned}
v_2 \leq \quad & \max_{p_i, q_i \in \mathbb{R}} \quad \log\left(\frac{\text{beta}(kq_i, k(1-\bar{\eta}-q_i))}{\text{beta}(kp_i, k(1-\bar{\eta}-p_i))}\right) \\
& \text{subject to} \quad |p_i - q_i| \leq \frac{b}{2}, \\
& \qquad\qquad\quad p_i \in [\eta, 1-\bar{\eta}-\eta], \\
& \qquad\qquad\quad q_i \in [\eta, 1-\bar{\eta}-\eta].
\end{aligned}
\tag{11}
$$

Evaluating the gradient of the objective function in the optimization problem in (11), it can be shown that the Karush-Kuhn-Tucker (KKT) conditions of optimality are not satisfied in the interior of the set of feasible points except for points that lie on the line $p_i = q_i$, which are minima. Thus, since the KKT conditions are only necessary conditions (see Chapter 11 of [29]), satisfying them does not imply optimality, and we exclude points where $p_i = q_i$ from the set of possible maximizers.

Evaluating points on the boundary of the feasible region shows that KKT conditions are also not satisfied. Thus, we need only to consider the extreme (p_i, q_i)'s in the set

$$
\left\{\left(\eta+\frac{b}{2}, \eta\right), \left(1-\bar{\eta}-\eta-\frac{b}{2}, 1-\bar{\eta}-\eta\right), \left(\eta, \eta+\frac{b}{2}\right), \left(1-\bar{\eta}-\eta, 1-\bar{\eta}-\eta-\frac{b}{2}\right)\right\}, \tag{12}
$$

which are the vertices of the feasible region. Note that since $\text{beta}(a, b) = \text{beta}(b, a)$, the points in the first row give equal positive objectives and the points in the second row have equal negative objectives. Hence, we can choose the first point in (12) to find

$$
v_2 = \log\left(\frac{\text{beta}(k\eta, k(1-\bar{\eta}-\eta))}{\text{beta}(k(\eta+\frac{b}{2}), k(1-\bar{\eta}-\eta-\frac{b}{2}))}\right).
$$

Substituting v_1 and v_2 in (10) concludes the proof. ■

We now state the main theorem of this section, which formally establishes the (ϵ, δ)-differential privacy of the Dirichlet mechanism for identity queries.

Theorem 1. *Fix $\eta, \bar{\eta} \in (0, 1)$, $b \in (0, 1]$, and $W \subseteq [n-1]$, and let Assumptions 1-3 hold. Let the adjacency relation in Definition 1 hold. Then the Dirichlet mechanism with parameter $k \in \mathbb{R}_+$, defined as $\mathcal{M}_D^{(k)}(p) = \text{Dir}_k(p)$, is (ϵ, δ)-differentially private, where*

$$
\epsilon = \log\left(\frac{\text{beta}(k\eta, k(1-\bar{\eta}-\eta))}{\text{beta}(k(\eta+\frac{b}{2}), k(1-\bar{\eta}-\eta-\frac{b}{2}))}\right) + \frac{kb}{2}\log\left(\frac{1-(|W|-1)\gamma}{\gamma}\right),
$$

and

$$
\delta = 1 - \min_{p \in \Delta_{n,W}^{(\eta,\bar{\eta})}} \mathbb{P}[\mathcal{M}_D^{(k)}(p) \in \Omega_1].
$$

Proof. The expression for ϵ results immediately from Lemma 4 and the expression for δ is a direct result of (7). ■

In Figure 2, for three instances of $(\eta, \bar{\eta}, k)$ and $b = 0.1$, we show how Theorem 1 captures the behavior of ϵ. All three cases show that Theorem 1 causes an offset to the numerically computed value of ϵ, and the level of offset decreases with the value of the original ϵ.

Remark 1. *Note that if a mechanism is ϵ_1-differentially private, it is also ϵ_2-differentially private for all $\epsilon_2 \geq \epsilon_1$. Therefore, if the upper bound for ϵ after simplification of beta functions is still within an acceptable range,* e.g., *$\delta \leq 0.05$ and $\epsilon \leq 5$ [30], [31], [32], then using an over-approximation of ϵ does not substantially harm our interpretation of the Dirichlet mechanism's protections.*

The expression given for ϵ in Theorem 1 contains a logartihm of a ratio of beta functions, which can be difficult to reason about intuitively. In the following lemma we present upper and lower bounds for beta functions in terms of simpler functions, which we will use to provide a simplified upper bound for ϵ.

Lemma 5. *Let $a, b \in \mathbb{R}$. Then*

$$\exp(2 - a - b) \leq \text{beta}(a, b) \leq \frac{a + b - 1}{(2a - 1)(2b - 1)}.$$

Proof: See Appendix C. ■

Using Lemma 5, we can provide a simplified bound on ϵ in exchange for that bound being somewhat looser.

Corollary 1. *Let all conditions of Theorem 1 hold. Then, for identity queries over $\Delta_{n,W}^{(\eta,\bar{\eta})}$, the Dirichlet mechanism is (ϵ, δ)-differentially private, with*

$$\epsilon = 2k(1 - \bar{\eta}) - 3 + \frac{kb}{2} \log\left(\frac{1 - (|W| - 1)\gamma}{\gamma}\right)$$

and

$$\delta = 1 - \min_{p \in \Delta_{n,W}^{(\eta,\bar{\eta})}} \mathbb{P}[\mathcal{M}_D^{(k)}(p) \in \Omega_1].$$

Proof: The value of δ is the same as that in Theorem 1. For ϵ from Theorem 1, we need to upper bound the term $\text{beta}\big(k\eta, k(1 - \bar{\eta} - \eta)\big)$ and lower bound the term $\text{beta}\big(k(\eta + \frac{b}{2}), k(1 - \bar{\eta} - \eta - \frac{b}{2})\big)$. We thus apply Lemma 5 to find

$$\frac{\text{beta}(k\eta, k(1 - \bar{\eta} - \eta))}{\text{beta}(k(\eta + \frac{b}{2}), k(1 - \bar{\eta} - \eta - \frac{b}{2}))} \leq \frac{k - k\bar{\eta} - 1}{(2k\eta - 1)(2k - 2k\bar{\eta} - 2k\eta - 1)} \frac{1}{\exp(2 - k + k\bar{\eta})}.$$

Then, taking the logarithm of both sides, we find that

$$\begin{aligned}
\log\left(\frac{\text{beta}(k\eta, k(1 - \bar{\eta} - \eta))}{\text{beta}(k(\eta + \frac{b}{2}), k(1 - \bar{\eta} - \eta - \frac{b}{2}))}\right) &\leq \log\left(\frac{k - k\bar{\eta} - 1}{(2k\eta - 1)(2k - 2k\bar{\eta} - 2k\eta - 1)} \frac{1}{\exp(2 - k + k\bar{\eta})}\right) \\
&\leq \log(k - k\bar{\eta} - 1) - \log(2k\eta - 1) \\
&\quad - \log(2k - 2k\bar{\eta} - 2k\eta - 1) - 2 + k - k\bar{\eta}. \qquad (13)
\end{aligned}$$

Then, using the fact that $\log(x) < x$ in (13), the first term satisfies

$$\log(k - k\bar{\eta} - 1) \leq k - k\bar{\eta} - 1.$$

Next, using the fact that

$$k \geq \max\left\{\frac{1}{\eta}, \frac{1}{1 - \bar{\eta} - \eta}\right\}$$

from Assumption 1, the next two logarithms terms are both non-positive and an upper bound is furnished by eliminating them. Then we find

$$\log\left(\frac{\text{beta}(k\eta, k(1-\bar{\eta}-\eta))}{\text{beta}(k(\eta+\frac{b}{2}), k(1-\bar{\eta}-\eta-\frac{b}{2}))}\right) \leq 2k-3.$$

Combining this result with Theorem 1 completes the proof. ■

Because this bound on ϵ is linear in k, it offers a more intuitive understanding of how changing k affects privacy.

Next, we point out that the parameter γ, which is used in the definition of Ω_2, is not a parameter of the mechanism, in the sense that changing γ does not change the mechanism itself. Instead, γ balances the trade-off between privacy level and the probability of failing to guarantee that privacy level, *i.e.,* changing γ can decrease ϵ in exchange for increasing δ and vice versa.

In some cases, we are given the highest probability of privacy failure, δ, that is acceptable, and one must maximize the level of privacy, ϵ, subject to that upper bound. Let $\hat{\delta}$ denote the maximum admissible value of δ. Then we are interested in minimizing ϵ while obeying $\delta \leq \hat{\delta}$. Using Theorem 1, we note that ϵ is a strictly decreasing function of γ. Letting V be the set of vertices of $\Delta_{n,W}^{(\eta,\bar{\eta})}$, we then can minimize ϵ by solving the problem

$$\begin{aligned} &\min_{\gamma} \quad \gamma \\ &\text{subject to} \quad \mathbb{P}[\mathcal{M}_D^{(k)}(p) \in \Omega_1] \geq 1-\hat{\delta} \text{ for all } p \in V. \end{aligned} \tag{14}$$

Note that the feasible region of the optimization problem (14) is a convex set because the function $\mathbb{P}[\mathcal{M}_D^{(k)}(p) \in \Omega_1]$ is a strictly decreasing function of γ. Therefore, ϵ can be optimized for a given $\hat{\delta}$ using off-the-shelf convex optimization tool-boxes, and this will be done in Section VII. Next, we apply the Dirichlet mechanism to average queries.

IV. Dirichlet Mechanism for Differential Privacy of Average Queries

In this section we consider a collection of N vectors indexed over $i \in [N]$, with the i^{th} denoted $p^i \in \Delta_{n,W}^{(\eta,\bar{\eta})}$. The goal is to compute the average of the collection $\{p^i\}_{i\in[N]}$ while providing differential privacy. Accordingly, the space of sensitive data under consideration is now

$$\mathcal{S} := \left\{\{p^i\}_{i\in[N]} : N \in \mathbb{N} \text{ and } p^i \in \Delta_{n,W}^{(\eta,\bar{\eta})}\right\}.$$

Accordingly, we first re-define the adjacency relationship for the average query setting.

Definition 3. *Fix a scalar $b \in (0,1]$. Two collections in $\mathcal{S}$, denoted $\{p^i\}_{i\in[N]}$ and $\{q^i\}_{i\in[N]}$, are b-adjacent if there is some j such that*

1) *$p^i = q^i$ for all $j \neq i$,*
2) *there exist indices m and l such that $p^j_{-(m,l)} = q^j_{-(m,l)}$ and $||p^j_{(m,l)} - q^j_{-(m,l)}|| \leq b$.*

With adjacency defined over collections, we have a corresponding definition of probabilistic differential privacy.

Definition 4. *(Probabilistic differential privacy for collections) Let $b \in (0,1]$ and $W \subseteq [n-1]$ be given. Fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. A mechanism $\mathcal{M} : \left(\Delta_{n,W}^{\eta,\bar{\eta}}\right)^N \times \Omega \to \Delta_n$ is said to be probabilistically (ϵ,δ)-differentially private if, for all $\mathcal{P} := \{p^i\}_{i\in[N]} \in \mathcal{S}$, we can partition the output space Δ_n into two disjoint sets Ω_1, Ω_2, such that*

$$\mathbb{P}[\mathcal{M}(\mathcal{P}) \in \Omega_2] \le \delta,$$

and for all $\mathcal{Q} := \{q^i\}_{i\in[N]} \in \mathcal{S}$ b-adjacent to $\mathcal{P}$ and for all $x \in \Omega_1$,

$$\log\left(\frac{\mathbb{P}[\mathcal{M}(\mathcal{P}) = x]}{\mathbb{P}[\mathcal{M}(\mathcal{Q}) = x]}\right) \le \epsilon.$$

As with Definition 2, we use Definition 4 simply as a means to show that ordinary (ϵ,δ)-differential privacy holds.

The query we now consider is the average. Set $\mathcal{P} = \{p^i\}_{i\in[N]}$ and $\mathcal{Q} = \{q^i\}_{i\in[N]}$. Mathematically we define the average operator $\mathcal{A}$ via

$$\mathcal{A}(\mathcal{P}) := \frac{1}{N}\sum_{i=1}^{N} p^i,$$

with $\mathcal{A}(\mathcal{Q})$ defined analogously. The next theorem formalizes the privacy protections of the Dirichlet mechanism when applied to such averages.

Theorem 2. *Fix $\eta, \bar{\eta} \in (0,1)$, let $b \in (0,1]$, let $W \subseteq [n-1]$, and let Assumptions 1-3 hold. Let the adjacency relation in Definition 3 hold. Then the Dirichlet mechanism with parameter $k \in \mathbb{R}_+$ defined as $\mathcal{M}_D^{(k)}\big(\mathcal{P}\big) = \mathrm{Dir}_k\big(\mathcal{A}(\mathcal{P})\big)$ is (ϵ,δ)-differentially private, where*

$$\epsilon = \log\left(\frac{\mathrm{beta}(k\eta, k(1-\bar{\eta}-\eta))}{\mathrm{beta}(k(\eta+\frac{b}{2N}), k(1-\bar{\eta}-\eta-\frac{b}{2N}))}\right) + \frac{kb}{2N}\log\left(\frac{1-(|W|-1)\gamma}{\gamma}\right), \tag{15}$$

and

$$\delta = 1 - \min_{p\in\Delta_{n,W}^{(\eta,\bar{\eta})}} \mathbb{P}[\mathcal{M}_D^{(k)}(\mathcal{A}(\mathcal{P})) \in \Omega_1].$$

Proof. We proceed by showing that Definition 4 is satisfied. For all $i \in [n]$, let $A(p_i) := \mathcal{A}(\mathcal{P})_i$ and $x \in \Omega_1$. Then, we are interested in the quantity

$$\frac{\mathbb{P}[\mathcal{M}_D^{(k)}(\mathcal{A}(\mathcal{P})) = x]}{\mathbb{P}[\mathcal{M}_D^{(k)}(\mathcal{A}(\mathcal{Q})) = x]} = \frac{\mathrm{B}(k\mathcal{A}(\mathcal{Q}))\prod_{i=1}^{n} x_i^{kA(p_i)-1}}{\mathrm{B}(k\mathcal{A}(\mathcal{P}))\prod_{i=1}^{n} x_i^{kA(q_i)-1}}. \tag{16}$$

Based on the definition of the adjacency relationship for collections in Definition 3, $A(p)$ and $A(q)$ will differ only in their m^{th} and l^{th} entries. Taking the logarithm of both sides of (16) and using the same approach as in Theorem 1, we have that

$$\log\left(\frac{\mathbb{P}[\mathcal{M}_D^{(k)}(\mathcal{A}(\mathcal{P})) = x]}{\mathbb{P}[\mathcal{M}_D^{(k)}(\mathcal{A}(\mathcal{Q})) = x]}\right) \le \max_{\mathcal{A}(\mathcal{P}),\mathcal{A}(\mathcal{Q})} \log\left(\frac{\mathrm{B}(k\mathcal{A}(\mathcal{Q}))}{\mathrm{B}(k\mathcal{A}(\mathcal{P}))}\right) + \max_{\mathcal{A}(\mathcal{P}),\mathcal{A}(\mathcal{Q})} \log\left(\frac{1-(|W|-1)\gamma}{\gamma}\right)^{k|A(p_m)-A(q_m)|}. \tag{17}$$

Because $\mathcal{P}$ and $\mathcal{Q}$ are b-adjacent, and each entry of $\mathcal{A}(\cdot)$ represents the average of a component, we have that

$$|A(p_m) - A(q_m)| \leq \frac{b}{2N}. \tag{18}$$

Combining (17), (18) and Lemma 4 completes the proof for the value of ϵ. For δ, the same approach for calculating δ in identity queries applies to average queries. ■

Remark 2. *As seen in* (15)*, the level of privacy increases with the number of vectors present in the collection. In particular,* $\epsilon \to 0$ *as* $n \to \infty$*. This is consistent with the intuition that it should be harder to uncover the sensitive information of an individual in a population when their data is mixed together in an average.*

As with Corollary 1, we provide simplified bounds on the value of ϵ to ease the interpretation of each parameter's influence upon privacy.

Corollary 2. *Let all conditions of Theorem 2 hold. Then, for average queries, the Dirichlet mechanism* $\mathcal{M}_D^{(k)}\big(\mathcal{P}\big) = \mathrm{Dir}_k\big(\mathcal{A}(\mathcal{P})\big)$ *is* (ϵ, δ)*-differentially private with*

$$\epsilon = 2k(1-\bar{\eta}) - 3 + \frac{kb}{2n}\log\left(\frac{1-(|W|-1)\gamma}{\gamma}\right)$$

and

$$\delta = 1 - \min_{p \in \Delta_{n,W}^{(\eta,\bar{\eta})}} \mathbb{P}[\mathcal{M}_D^{(k)}(\mathcal{A}(\mathcal{P})) \in \Omega_1].$$

Proof: The proof is similar to that of Corollary 1 and is therefore omitted. ■

V. Differential Privacy for General Linear Queries

In this section, we derive privacy guarantees for arbitrary linear queries over collections of points in the unit simplex. Examples of such queries are weighted averages of vectors of transition probabilities, e.g., in the smart power grid. In particular, with a variety of smart devices and smart buildings modeled as MDPs, one may wish to compute average behaviors, with the weights encoding the importance of a device or size of a building. We begin by establishing the class of queries to be considered, then we derive privacy guarantees provided by the Dirichlet mechanism for this class.

A. General Linear Queries over the Simplex

As above, we consider privacy over the set $\mathcal{S}$, which contains N-element collections of vectors in the unit simplex. We again consider $\mathcal{P} = \{p^i\}_{i\in[N]}$, with $p^i \in \Delta_{n,W}^{(\eta,\bar{\eta})}$. The collection $\mathcal{P}$ can also be represented as an $n \times N$ matrix, where column i is equal to p^i. With an abuse of notation, we also use $\mathcal{P}$ to denote this matrix representation, and we note that $\mathcal{P}_{ij} = p_i^j$. With $\mathcal{P} \in \mathbb{R}^{n\times N}$, we can represent the linear queries of interest by vector multiplication on the right. Namely, a linear query $L : \mathcal{S} \to \Delta_n$ can be represented via

$$L\big(\mathcal{P}\big) = \mathcal{P}\ell = \begin{pmatrix} \sum_{j=1}^N p_1^j \ell_j \\ \vdots \\ \sum_{j=1}^N p_n^j \ell_j \end{pmatrix}, \tag{19}$$

where $\ell \in \mathbb{R}^N$. The following lemma establishes the stronger statement that ensuring $L(\mathcal{P}) \in \Delta_n$ for arbitrary $\mathcal{P} \in \mathcal{S}$ in fact requires $\ell \in \Delta_N$.

Lemma 6. *Let $L : \mathcal{S} \to \Delta_n$ be a linear query identified with the vector $\ell \in \mathbb{R}^N$. Then $\ell \in \Delta_N$.*

Proof: Using (19), for $L(\mathcal{P}) \in \Delta_n$, we require that

$$\sum_{i=1}^{n} \sum_{j=1}^{N} p_i^j \ell_j = \sum_{j=1}^{N} \sum_{i=1}^{n} p_i^j \ell_j = \sum_{j=1}^{N} \ell_j \sum_{i=1}^{n} p_i^j = \sum_{j=1}^{N} \ell_j = 1,$$

where the third equality follows from $p^j \in \Delta_n$. We also must have $\ell_j \geq 0$ for all $j \in [N]$. ■

We also incorporate additional boundedness of the entries of ℓ in the following definition.

Definition 5. *Fix $\alpha \in (0, 1]$. A vector $\ell \in \Delta_n$ is said to be α-bounded if $\|\ell\|_\infty \leq \alpha$.*

All vectors $\ell \in \Delta_n$ are trivially 1-bounded, though the inclusion of α-boundedness allows for additional bounds on the entries of ℓ to be used in our privacy analysis. This inclusion in turn enables us to make stronger statements about the Dirichlet mechanism's guarantees, which we explore further in the next section.

B. Privacy Guarantees for Linear Queries

With this characterization of ℓ in hand, we now quantify the privacy guarantees afforded to such queries by the Dirichlet mechanism. As above, determining these privacy guarantees will require bounding ratios of Dirichlet distributions, a component of which is a term involving gamma functions. The next lemma provides a bound on this term.

Lemma 7. *Fix $\eta, \bar{\eta} \in (0, 1)$ and $W \subseteq [n-1]$, and let Assumptions 1-3 hold. Fix $b \in (0, 1]$ and let $\mathcal{P}, \mathcal{Q} \in \mathcal{S}$ be adjacent according to Definition 3. Let $L : \mathcal{S} \to \Delta_n$ be a linear query over $\mathcal{S}$ identified with $\ell \in \Delta_n$. Then*

$$\frac{\mathrm{B}\big(k\mathcal{Q}\ell\big)}{\mathrm{B}\big(k\mathcal{P}\ell\big)} \leq \frac{\text{beta}(k\eta, k(1-\bar{\eta}-\eta)}{\text{beta}\big(k(\eta+\frac{b}{2}), k(1-\bar{\eta}-\eta-\frac{b}{2})\big)}.$$

Proof: The proof is similar to those of Lemmas 3 and 4 and is therefore omitted. ■

It is using this lemma that we next state the privacy guarantees afforded to arbitrary linear queries over $\mathcal{S}$.

Theorem 3. *Let $L : \mathcal{S} \to \Delta_n$ be α-bounded. Fix an adjacency parameter $b \in (0, 1]$ and let Assumptions 1-3 hold. Then, for adjacency as defined in Definition 3, the Dirichlet mechanism applied to L, denoted $\mathcal{M}_D\big(L(\cdot)\big) : \mathcal{S} \to \Delta_n$, is (ϵ, δ)-differentially private, where*

$$\epsilon = \log\left(\frac{\text{beta}\big(k\eta, k(1-\bar{\eta}-\eta)\big)}{\text{beta}\big(k(\eta+\frac{b}{2}), k(1-\bar{\eta}-\eta-\frac{b}{2})}\big)\right) + kb\alpha\big|\log(\gamma)\big|$$

and

$$\delta = 1 - \min_{p \in \Delta_{n,W}^{(\eta,\bar{\eta})}} \mathbb{P}\big[\mathcal{M}_D^{(k)}(p) \in \Omega_1\big].$$

Proof: We proceed by showing that Definition 4 is satisfied. Deriving δ follows the same steps as in Theorem 1 and 2 and these steps are omitted here. For ϵ, consider adjacent $\mathcal{P}, \mathcal{Q} \in \mathcal{S}$. Then we consider

$$\log\left(\frac{\mathrm{Dir}_k\big(L(\mathcal{P})\big)}{\mathrm{Dir}_k\big(L(\mathcal{Q})\big)}\right) = \log\left(\frac{\mathrm{B}\big(L(\mathcal{Q})\big)}{\mathrm{B}\big(L(\mathcal{P})\big)}\right) + \log\left(\prod_{i=1}^{n} \frac{x_i^{k\left(\sum_{j=1}^N p_i^j \ell_j\right)}}{x_i^{k\left(\sum_{j=1}^N q_i^j \ell_j\right)}}\right).$$

The bounds on the beta function term follow from Theorems 1 and 2.

For the product term, we note that

$$\log\left(\prod_{i=1}^{n} \frac{x_i^{k\left(\sum_{j=1}^N p_i^j \ell_j\right)}}{x_i^{k\left(\sum_{j=1}^N q_i^j \ell_j\right)}}\right) = \sum_{i=1}^{n} k\left(\sum_{j=1}^{N} p_i^j \ell_j - q_i^j \ell_j\right)\log(x_i).$$

Next, using adjacency, we know that there exists an index t such that, for all $j \neq t$, we have $p^j = q^j$. Moreover, we know that there are indices c and d such that $p^t_{-(c,d)} = q^t_{-(c,d)}$. Then we can simplify the above to find

$$\begin{aligned}
\log\left(\prod_{i=1}^{n} \frac{x_i^{k\left(\sum_{j=1}^N p_i^j \ell_j\right)}}{x_i^{k\left(\sum_{j=1}^N q_i^j \ell_j\right)}}\right) &= \sum_{i\in\{c,d\}} k\left(p_i^t \ell_t - q_i^t \ell_t\right)\log(x_i) \\
&= k\left(p_c^t \ell_t - q_c^t \ell_t\right)\log(x_c) + k\left(p_d^t \ell_t - q_d^t \ell_d\right)\log(x_d) \\
&\leq k\big|p_c^t - q_c^t\big| \cdot |\ell_t| \cdot \big|\log(x_c)\big| + k\big|p_d^t - q_d^t\big| \cdot |\ell_t| \cdot \big|\log(x_d)\big| \\
&\leq kb\alpha\big|\log(\gamma)\big|,
\end{aligned}$$

which follows from the α-boundedness of L. Combining this term with the aforementioned beta function bounds completes the proof. ■

As above, we provide a simplified bound on ϵ that offers a straightforward dependence of ϵ upon other parameters in the problem.

Corollary 3. *Let all conditions of Theorem 3 hold. Then, for an arbitrary α-bounded linear query L, the Dirichlet mechanism $\mathcal{M}_D\big(L(\cdot)\big) : \mathcal{S} \to \Delta_n$ is (ϵ, δ)-differentially private with*

$$\epsilon = 2k(1-\bar{\eta}) - 3 + kb\alpha\big|\log(\gamma)\big|$$

and

$$\delta = 1 - \min_{p\in\Delta_{n,W}^{(\eta,\bar{\eta})}} \mathbb{P}\big[\mathcal{M}_D^{(k)}(p) \in \Omega_1\big].$$

Proof: This proof is similar to Corollaries 1 and 2 and is therefore omitted. ■

VI. Accuracy Analysis

We analyze the accuracy of the Dirichlet mechanism by two metrics. First, in terms of its moments and second in terms of concentration of its outputs about the underlying sensitive probability vector.

Proposition 1. *Let $x \in \Delta_n$ be the output of a Dirichlet mechanism with input $p \in \Delta_{n,W}^{(\eta,\bar{\eta})}$ and parameter $k \in \mathbb{R}_+$. Then we have that $\mathbb{E}[x_i] = p_i$ and*

$$\mathrm{Var}[x_i] = \frac{p_i(1-p_i)}{k+1}. \tag{20}$$

Proof. Let $\bar{p} = \sum_{r=1}^{n} kp_r$. Equation (49.9) in [33] gives

$$\mathbb{E}[x_i] = \frac{kp_i}{\bar{p}} = p_i \quad \text{and} \quad \text{Var}[x_i] = \frac{kp_i(\bar{p} - kp_i)}{\bar{p}^2(\bar{p}+1)}.$$

Since the input p belongs to the unit simplex, we have that $\bar{p} = k$. Substituting $\bar{p}$ with k concludes the proof. ■

Remark 3. *As seen in* (20) *the variance of the output depends on the input data* p_i. *However, we can find the worst-case variance by maximizing the expression for the variance which occurs at* $p_i = 0.5$. *Hence, we have that*

$$\textit{Var}[x_i] \leq \frac{1}{4(k+1)}. \tag{21}$$

It is using this form of upper bound that we use to establish the concentration of the Dirichlet mechanism's output about its input. In particular, we bound the probability of a large deviation of the private output from the sensitive input.

Theorem 4. *Fix* $\eta, \bar{\eta} \in (0,1)$ *and* $W \subseteq [n-1]$, *and let Assumptions 1 and 2 hold. Let* $\mathcal{M}_D^{(k)}$ *denote a Dirichlet mechanism with parameter* $k \in \mathbb{R}_+$. *Let* $\mu \in (0,1)$ *and* $\theta \in (0, e^{-2\mu^2})$ *be given. Then a sufficient condition to ensure*

$$\mathbb{P}\Big[\big\|\mathcal{M}_D^{(k)}(p) - p\big\|_\infty \leq \mu\Big] \geq 1 - \theta$$

is to select $k = -\frac{\log(\theta)}{2\mu^2} - 1$.

Proof: Lemma 2 in [34] implies that, for any $\beta > 0$,

$$\mathbb{P}\left[\|\mathcal{M}_D^{(k)}(p) - p\|_\infty \geq \sqrt{\frac{\log(1/\beta)}{2(k+1)}}\right] \geq 1 - \beta.$$

The result follows by setting $\beta = \theta$, setting $\mu = \sqrt{\log(1/\beta)/2(k+1)}$ and solving for k. ■

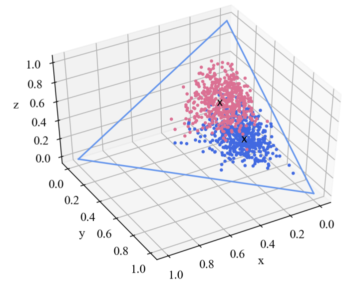
Theorem 4 provides both the means to assess accuracy of the Dirichlet mechanism for a given k, as well as a prescriptive tool for selecting k based on desired accuracy.

VII. Simulation Results

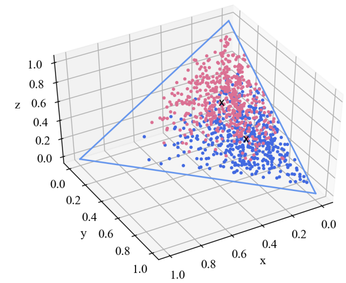
In this section, we simulate the output of the Dirichlet mechanism for an identity query and an average query. As an example, suppose we ask a number of experts for their opinion on the probability of certain events happening, which we can represent as a vector in the unit simplex. Taking an identity query of each one allows the private release of all opinions for further analysis. For other cases, to make a decision based on all opinions, we can integrate the opinions into one [35], for which the average can be suitable.

A. Simulation of Identity Queries

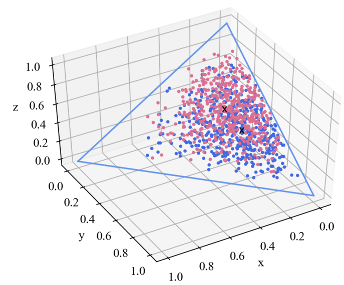
In Figure 3, we show an example of privatizing an identity query. We consider sensitive vectors $p, q \in \Delta_3$, where $p = (0.2, 0.3, 0.5)^T$ and $q = (0.2, 0.5, 0.3)^T$. We choose $\eta = \bar{\eta} = 0.05$, $W = \{2, 3\}$, and $b = 0.4$. We use the values $k = 3$, $k = 10$, and $k = 24$ to illustrate different levels of privacy. For $\hat{\delta} = 0.05$, Theorem 1 implies that these identity queries are provided with $(2.30, 0.05)$-differential privacy with $k = 3$, $(4.40, 0.05)$-differential



(a) Distribution of 500 outputs for identity queries with $k = 24$.



(b) Distribution of 500 outputs for identity queries with $k = 10$.



(c) Distribution of 500 outputs for identity queries with $k = 3$.

Fig. 3: Distributions of outputs for identity queries on 0.4-adjacent inputs in Δ_3. The top left figure uses $k = 24$, the top right figure uses $k = 10$, and the bottom figure uses $k = 3$.

privacy with $k = 10$, and $(8.53, 0.05)$-differential privacy with $k = 24$. Figure 3a shows the distribution of outputs generated for $k = 24$. As expected, a larger value of k shows a relatively concentrated distribution of outputs about inputs.

Figure 3b shows the distribution of outputs for $k = 10$, which provides stronger differential privacy than $k = 24$, and, as shown in Figure 3b, private outputs are not as tightly concentrated about private inputs. This illustrates that one must tradeoff stronger privacy for reduced accuracy and vice versa. This tradeoff is further shown for $k = 3$ in Figure 3c, where private outputs are less concentrated than for $k = 10$ and $k = 24$ in exchange for privacy protections being stronger.

B. Simulation of Average Queries

In Figure 4, we show an example of privatizing the average of 100 experts' opinions. In this example we have a collection of opinions $\mathcal{P}$, and we want to compare the output of the Dirichlet mechanism with the output of the mechanism when fed with a collection $\mathcal{Q}$ that is 1-adjacent to $\mathcal{P}$.

As above, we choose $k = 24$, which keeps the variance of the output below 0.01 according to (21). We have fixed W to be the set $\{1, 2\}$, set $\eta = 0.05$, and set $\bar{\eta} = 0.05$. Using Theorem 2, for $\hat{\delta} = 0.05$, we find that the

Statistics	Values
$\mathcal{A}(\mathcal{P})$	$(0.314923, 0.315923, 0.320923)$
$\mathcal{A}(\mathcal{Q})$	$(0.314923, 0.320923, 0.315923)$
$\mathcal{M}_D(\mathcal{A}(\mathcal{P}))$	$(0.327731, 0.336119, 0.336149)$
$\mathcal{M}_D(\mathcal{A}(\mathcal{Q}))$	$(0.326620, 0.338976, 0.334402)$
$Var[\mathcal{M}_D(\mathcal{A}(\mathcal{P})]$	$(0.00934632, 0.00983122, 0.0102117)$
$Var[\mathcal{M}_D(\mathcal{A}(\mathcal{Q})]$	$(0.00922426, 0.0100889, 0.0100757)$

TABLE I: Comparing the empirical average of the output of the mechanism with input collections $\mathcal{P}$ and $\mathcal{Q}$, alongside with their empirical variances. The values correspond to $k = 24$.

mechanism is $(1.18, 0.05)$-differentially private. Table I shows the empirical accuracy analysis of mechanism. In Figure 4, we can observe that, given the location of the mechanism output, it is not possible to determine with high probability whether the input is $\mathcal{P}$ or $\mathcal{Q}$. Table I also shows the desired variance is attained.

In order to illustrate the effect of changing k in the mechanism accuracy, Figure 4b compares the output of the mechanism when $k = 24$ and $k = 10$. As seen in the figure, the output when $k = 10$ is less concentrated around the average. It can also be seen that the probability that the output belongs to Ω_2 is higher when $k = 10$, which is consistent with the expressions derived in Theorems 1-8.

VIII. Conclusion

In this work we introduced a mechanism used for privatizing data inputs that belong to the unit simplex. We used the Dirichlet distribution to probabilistically map a vector within the unit simplex to itself. We proved that the Dirichlet mechanism is differentially private with high probability in identity, average, and linear queries. Our simulation results validated that the privacy bounds and the accuracy of the mechanism are within ranges typically considered in the differential privacy literature. As an extension to this work, we are interested in applying the Dirichlet mechanism to privatizing a policy in a Markov decision process. In particular, we are interested in showing how accurate the Dirichlet mechanism is in terms of the total accumulated rewards.

References

[1] G. W. Hart, "Nonintrusive appliance load monitoring," *Proceedings of the IEEE*, vol. 80, no. 12, pp. 1870–1891, 1992.

[2] Y. Wang, Z. Huang, S. Mitra, and G. E. Dullerud, "Differential privacy in linear distributed control systems: Entropy minimizing mechanisms and performance tradeoffs," *IEEE Transactions on Control of Network Systems*, vol. 4, no. 1, pp. 118–130, March 2017.

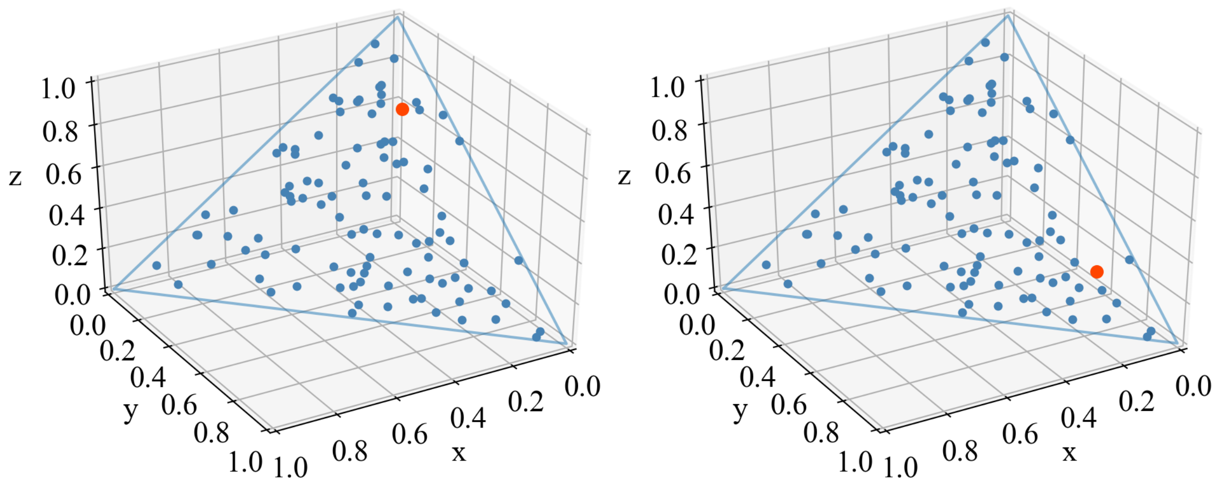
[3] E. Nozari, P. Tallapragada, and J. Cortés, "Differentially private distributed convex optimization via functional perturbation," *IEEE Transactions on Control of Network Systems*, vol. 5, no. 1, pp. 395–408, March 2018.

[4] M. Hale, A. Jones, and K. Leahy, "Privacy in feedback: The differentially private lqg," in *2018 Annual American Control Conference (ACC)*, June 2018, pp. 3386–3391.

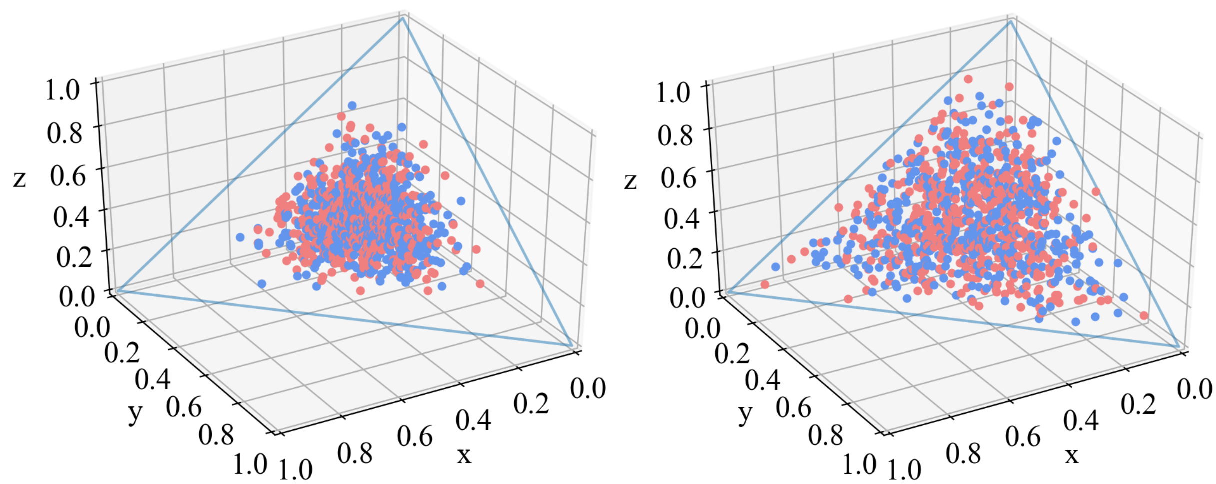
[5] A. Jones, K. Leahy, and M. Hale, "Towards differential privacy for symbolic systems," in *2019 American Control Conference (ACC)*, July 2019, pp. 372–377.

[6] M. L. Puterman, *Markov Decision Processes.: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, 2014.

[7] R. S. Sutton, A. G. Barto *et al.*, *Introduction to reinforcement learning*. MIT press Cambridge, 1998, vol. 2, no. 4, ch. 3.



(a) Visualization of two b-adjacent vector collections $\mathcal{P}$ and $\mathcal{Q}$. The left figure depicts $\mathcal{P}$ and the right figure corresponds to $\mathcal{Q}$. The data points with orange markers correspond to the vectors in which $\mathcal{P}$ and $\mathcal{Q}$ differ.



(b) The output of the Dirichlet mechanism when the input is $\mathcal{P}$ vs. $\mathcal{Q}$. The left plot shows the case where $k = 24$ and the right plot corresponds to $k = 10$. Each red data point corresponds to an independent run of $\mathcal{M}_D^{(k)}(\mathcal{A}(\mathcal{P}))$ and the blue data points correspond to $\mathcal{M}_D^{(k)}(\mathcal{A}(\mathcal{Q}))$.

Fig. 4: An average query on a collection of 100 vectors in Δ_3.

[8] Y. Savas, M. Ornik, M. Cubuktepe, M. O. Karabag, and U. Topcu, "Entropy maximization for markov decision processes under temporal logic constraints," *IEEE Transactions on Automatic Control*, 2019.

[9] B. Wu, M. Cubuktepe, and U. Topcu, "Switched linear systems meet markov decision processes: Stability guaranteed policy synthesis," in *2019 IEEE 58th Annual Conference on Decision and Control (CDC)*. IEEE, 2019, to appear, preprint arXiv:1904.11456.

[10] K. Chatterjee, R. Majumdar, and T. A. Henzinger, "Markov decision processes with multiple objectives," in *Annual Symposium on Theoretical Aspects of Computer Science*. Springer, 2006, pp. 325–336.

[11] S. Brechtel, T. Gindele, and R. Dillmann, "Probabilistic mdp-behavior planning for cars," in *2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, Oct 2011, pp. 1537–1542.

[12] S. Misra, A. Mondal, S. Banik, M. Khatua, S. Bera, and M. S. Obaidat, “Residential energy management in smart grid: A markov decision process-based approach,” in *2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of things and IEEE Cyber, Physical and Social Computing*. IEEE, 2013, pp. 1152–1157.

[13] C. Dwork, F. McSherry, K. Nissim, and A. Smith, “Calibrating noise to sensitivity in private data analysis,” in *Theory of cryptography conference*. Springer, 2006, pp. 265–284.

[14] 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.

[15] S. P. Kasiviswanathan and A. Smith, “On the’semantics’ of differential privacy: A bayesian formulation,” *Journal of Privacy and Confidentiality*, vol. 6, no. 1, 2014.

[16] K. Nissim and A. Wood, “Is privacy privacy?” *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 376, no. 2128, p. 20170358, 2018.

[17] J. Cortés, G. E. Dullerud, S. Han, J. Le Ny, S. Mitra, and G. J. Pappas, “Differential privacy in control and network systems,” in *2016 IEEE 55th Conference on Decision and Control (CDC)*. IEEE, 2016, pp. 4252–4272.

[18] S. Han and G. J. Pappas, “Privacy in control and dynamical systems,” *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 1, pp. 309–332, 2018.

[19] Z. Xu, K. Yazdani, M. T. Hale, and U. Topcu, “Differentially private controller synthesis with metric temporal logic specifications,” 2019.

[20] C. Hawkins and M. Hale, “Differentially private formation control,” 2020.

[21] M. Gotz, A. Machanavajjhala, G. Wang, X. Xiao, and J. Gehrke, “Publishing search logs—a comparative study of privacy guarantees,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 24, no. 3, pp. 520–532, 2011.

[22] D. B. Karp, “Normalized incomplete beta function: log-concavity in parameters and other properties,” *Journal of Mathematical Sciences*, vol. 217, no. 1, pp. 91–107, 2016.

[23] S. Vadhan, “The complexity of differential privacy,” in *Tutorials on the Foundations of Cryptography*. Springer, 2017, pp. 347–450.

[24] P. Gohari, B. Wu, M. T. Hale, and U. Topcu, “The dirichlet mechanism for differential privacy on the unit simplex,” in *Proceedings of the 2020 American Control Conference (ACC)*, 2020, accepted. To be presented. Preprint at: `https://arxiv.org/abs/1910.00043`.

[25] A. Machanavajjhala, D. Kifer, J. Abowd, J. Gehrke, and L. Vilhuber, “Privacy: Theory meets practice on the map,” in *Proceedings of the 2008 IEEE 24th International Conference on Data Engineering*. IEEE Computer Society, 2008, pp. 277–286.

[26] Y. Chen and X. Ye, “Projection Onto A Simplex,” *arXiv e-prints*, p. arXiv:1101.6081, Jan. 2011.

[27] J. Le Ny and G. J. Pappas, “Differentially private filtering,” *IEEE Transactions on Automatic Control*, vol. 59, no. 2, pp. 341–354, 2013.

[28] J. Rao and M. Sobel, “Incomplete dirichlet integrals with applications to ordered uniform spacings,” *Journal of Multivariate Analysis*, vol. 10, no. 4, pp. 603–610, 1980.

[29] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge University Press, 2004.

[30] F. McSherry and R. Mahajan, “Differentially-private network trace analysis,” *ACM SIGCOMM Computer Communication Review*, vol. 41, no. 4, pp. 123–134, 2011.

[31] L. Bonomi, L. Xiong, R. Chen, and B. Fung, “Privacy preserving record linkage via grams projections,” *arXiv preprint arXiv:1208.2773*, 2012.

[32] J. Hsu, M. Gaboardi, A. Haeberlen, S. Khanna, A. Narayan, B. C. Pierce, and A. Roth, “Differential privacy: An economic method for choosing epsilon,” in *2014 IEEE 27th Computer Security Foundations Symposium*. IEEE, 2014, pp. 398–410.

[33] S. Kotz, N. Balakrishnan, and N. L. Johnson, *Continuous multivariate distributions, Volume 1: Models and applications*. John Wiley & Sons, 2004, vol. 1.

[34] P. Gohari, M. Hale, and U. Topcu, “Privacy-preserving policy synthesis in markov decision processes,” 2020.

[35] R. T. Clemen and R. L. Winkler, “Combining probability distributions from experts in risk analysis,” *Risk analysis*, vol. 19, no. 2, pp. 187–203, 1999.

[36] A. Prékopa, “On logarithmic concave measures and functions,” *Acta Scientiarum Mathematicarum*, vol. 34, pp. 335–343, 1973.

[37] A. Prékopa, “Logarithmic concave measures with application to stochastic programming,” *Acta Scientiarum Mathematicarum*, vol. 32, pp. 301–316, 1971.

Appendix A: Proof of Lemma 2

We first state a result from [36].

Lemma 8 (Theorem 3 in [36]). *Let $f_1, \dots, f_k$ be non-negative and Borel measurable functions defined on $\mathbb{R}^n$ and let*

$$r(t) = \sup_{\substack{(x_1,\dots,x_k)\ s.t.\\ \lambda_1 x_1+\cdots+\lambda_k x_k = t}} f_1(x_1)\cdots f_k(x_k), \quad t \in \mathbb{R}^n,$$

where $\lambda_1, \dots, \lambda_k$ are positive constants satisfying the equality $\lambda_1 + \cdots + \lambda_k = 1$. Then, the function $r(t)$ also Borel measurable and

$$\int_{\mathbb{R}^n} r(t)dt \geq \left(\int_{\mathbb{R}^n} f_1(x_1)^{\frac{1}{\lambda_1}} dx_1\right)^{\lambda_1} \cdots \left(\int_{\mathbb{R}^n} f_k(x_k)^{\frac{1}{\lambda_k}} dx_k\right)^{\lambda_k}.$$

We next review the definition of log-concave functions. A function $g : \mathbb{R}^n \to \mathbb{R}$ is said to be log-concave if for all $x_1, x_2 \in \mathbb{R}^n$ and $\theta \in [0, 1]$, we have that

$$g(\theta x_1 + (1-\theta)x_2) \geq (g(x_1))^{\theta}(g(x_2))^{1-\theta}.$$

This condition is equivalent to

$$g(t) \geq \sup_{\theta u + (1-\theta)v = t} g(u)^{\theta} g(v)^{1-\theta}. \tag{22}$$

Note that g is log-concave if and only if $\log g$ is concave.

Next, for $x \in \mathbb{R}^n$ and $p \in \Delta_{n,W}^{(\eta,\bar{\eta})}$ let $f : \mathbb{R}^n \times \Delta_{n,W}^{(\eta,\bar{\eta})} \to [0, 1]$ be defined as

$$f(x, p) = \frac{\prod_{i\in W} x_i^{kp_i - 1}\left(1 - \sum_{i\in W} x_i\right)^{k(1-\sum_{i\in W} p_i)-1}}{B(k\tilde{p}_W)}.$$

For a fixed $p \in \Delta_{n,W}^{\eta,\bar{\eta}}$, let

$$f_1(x) := f(x, p).$$

The function $f_1(x)$ is the Dirichlet probability distribution function with parameter $\alpha \in \mathbb{R}^W$, where $\alpha := k\tilde{p}_W$. Since $p \in \Delta_{n,W}^{\eta,\bar{\eta}}$, we have that $\alpha_i \geq 1$, for all $i \in [|W|]$. Therefore f_1 is a log-concave function [37, Equation (4.4)]. Then, by (22),

$$f(t_x, p) \geq \sup_{\beta u_x + (1-\beta)v_x = t_x} f(u_x, p)^{\beta} f(v_x, p)^{1-\beta} \tag{23}$$

for all $p \in \Delta_{n,W}^{(\eta,\bar{\eta})}$, all $t_x, u_x, v_x \in \mathbb{R}^n$, and $\beta \in [0, 1]$.

Similarly, for a fixed $x \in \mathbb{R}^n$, let

$$f_2(p) := f(x, p).$$

Toward evaluating the Hessian of $\log f_2(p)$, recall the trigamma function $\psi^{(1)}(x) = \frac{d^2}{dz^2}\log\big(\Gamma(x)\big)$. Then

$$\left[\nabla^2 \log f_2(p)\right]_{i,j} = \begin{cases} -k^2\psi^{(1)}(kx) & i = j \text{ and } i, j \in W \\ 0 & \text{otherwise} \end{cases}.$$

The trigamma function is positive on the interval $(0, \infty)$. Therefore, the Hessian matrix of $\log\big(f_2(p)\big)$ is a diagonal matrix whose diagonal entries are either zero or negative. This provides log-concavity of $f_2(p)$. Therefore, using (22),

$$f(x, t_p) \geq \sup_{\beta u_p + (1-\beta)v_p = t_p} f(x, u_p)^{\beta} f(x, v_p)^{1-\beta} \tag{24}$$

for all $x \in \mathbb{R}^n$, all $t_p, u_p, v_p \in \Delta_{n,W}^{(\eta,\bar{\eta})}$, and all $\beta \in [0,1]$. Next, fix a choice of $\lambda \in [0,1]$, choose $\tilde{u}_x, \tilde{v}_x, \tilde{u}_p, \tilde{v}_p \in \mathbb{R}^n$ such that

$$\lambda \tilde{u}_x + (1-\lambda)\tilde{v}_x = t_x \quad \text{and} \quad \lambda \tilde{u}_p + (1-\lambda)\tilde{v}_p = p.$$

Assigning u_x to x in (24), we find

$$\begin{aligned} f(u_x, p) &\geq \sup_{\beta u_p + (1-\beta) v_p = p} f(u_x, u_p)^{\beta} f(u_x, v_p)^{1-\beta} \\ &\geq f(u_x, \tilde{u}_p)^{\lambda} f(u_x, \tilde{v}_p)^{1-\lambda}, \end{aligned} \tag{25}$$

where the second inequality follows by setting $\beta = \lambda$. Similarly, we can write

$$\begin{aligned} f(v_x, p) &\geq \sup_{\beta u_p + (1-\beta) v_p = p} f(v_x, u_p)^{\beta} f(v_x, v_p)^{1-\beta} \\ &\geq f(v_x, \tilde{u}_p)^{\lambda} f(v_x, \tilde{v}_p)^{1-\lambda}. \end{aligned} \tag{26}$$

Revisiting (23), using (25) and (26), we can write

$$\begin{aligned} f(t_x, p) &\geq \sup_{\beta u_x + (1-\beta) v_x = t_x} f(u_x, p)^{\beta} f(v_x, p)^{1-\beta} \\ &= \sup_{\lambda \tilde{u}_x + (1-\lambda) \tilde{v}_x = t_x} f(\tilde{u}_x, p)^{\lambda} f(\tilde{v}_x, p)^{1-\lambda} \\ &\geq \sup_{\lambda \tilde{u}_x + (1-\lambda) \tilde{v}_x = t_x} \left(f(\tilde{u}_x, \tilde{u}_p)^{\lambda^2} f(\tilde{v}_x, \tilde{v}_p)^{\lambda(1-\lambda)} f(\tilde{v}_x, \tilde{u}_p)^{(1-\lambda)\lambda} f(\tilde{v}_x, \tilde{v}_p)^{(1-\lambda)^2} \right). \end{aligned}$$

The first line holds for all $\beta \in [0,1]$, while the second follows from setting $\beta = \lambda$ for the choice of λ above. Note that

$$\lambda \tilde{u}_x + (1-\lambda)\tilde{v}_x = \lambda^2 \tilde{u}_x + \lambda(1-\lambda)\tilde{u}_x + \lambda(1-\lambda)\tilde{v}_x + (1-\lambda)^2 \tilde{v}_x.$$

Since $\lambda^2 + \lambda(1-\lambda) + \lambda(1-\lambda) + (1-\lambda)^2 = 1$, Theorem 8 applies. Therefore, we can write

$$\int_{\mathcal{A}_n} f(t_x, p) dt_x \geq \left(\int_{\mathcal{A}_n} f(u_x, \tilde{u}_p) du_x \right)^{\lambda^2} \left(\int_{\mathcal{A}_n} f(u_x, \tilde{v}_p) du_x \right)^{\lambda(1-\lambda)} \left(\int_{\mathcal{A}_n} f(v_x, \tilde{u}_p) dv_x \right)^{(1-\lambda)\lambda} \left(\int_{\mathcal{A}_n} f(v_x, \tilde{v}_p) dv_x \right)^{(1-\lambda)^2}.$$

By renaming the variables t_x, u_x and v_x to x inside the integrals and merging the similar terms into one, we find

$$\int_{\mathcal{A}_n} f(x, p) dx \geq \left(\int_{\mathcal{A}_n} f(x, \tilde{u}_p) dx \right)^{\lambda} \left(\int_{\mathcal{A}_n} f(x, \tilde{v}_p) dx \right)^{(1-\lambda)},$$

where $\lambda \tilde{u}_p + (1-\lambda)\tilde{v}_p = p$. Therefore, $\int_{\mathcal{A}_n} f(x,p) dx$ is log-concave in p. ■

APPENDIX B: PROOF OF LEMMA 3

Let $c = p_i + p_j = q_i + q_j$. Then using (6), we have that

$$\begin{aligned} \frac{\text{beta}(kq_i, kq_j)}{\text{beta}(kp_i, kp_j)} &= \frac{\Gamma(kq_i)\Gamma(k(c-q_i))}{\Gamma(kp_i)\Gamma(k(c-p_i))} && (27) \\ &= \frac{\Gamma(kq_j)\Gamma(k(c-q_j))}{\Gamma(kp_j)\Gamma(k(c-p_j))}. && (28) \end{aligned}$$

Using the definition of the digamma function, we have

$$\frac{\partial}{\partial x} \left[\frac{\Gamma(x-a)}{\Gamma(x-b)} \right] = \frac{\Gamma(x-a)[\psi^{(0)}(x-a) - \psi^{(0)}(x-b)]}{\Gamma(x-b)}. \tag{29}$$

Because the digamma function is strictly increasing on the interval $(0,\infty)$, the derivative in (29) is positive if and only if $x-b<x-a$, which is true if and only if $a<b$. Returning to (27) and (28), we see that (27) is increasing in c if $q_i<p_i$ and that (28) is increasing in c if $q_j<p_j$. Therefore, we will construct an upper bound using (27) if $q_i<p_i$ and we will construct an upper bound using (28) if $q_j<p_j$. For concreteness, suppose $q_i<p_i$. Then (27) is an increasing function of c. By definition, $c=q_i+q_j$ and $i,j\in W$. Then $c\leq 1-\bar{\eta}$ and we find

$$\begin{aligned}\frac{\text{beta}(kp_i,kp_j)}{\text{beta}(kq_i,kq_j)} &= \frac{\Gamma(kq_i)\Gamma(k(c-q_i))}{\Gamma(kp_i)\Gamma(k(c-p_i))}\\ &\leq \frac{\text{beta}(kq_i,k(1-\bar{\eta}-q_i))}{\text{beta}(kp_i,k(1-\bar{\eta}-p_i))}.\end{aligned}$$

The other case proceeds identically. ■

Appendix C: Proof of Lemma 5

Convexity of $\exp(\cdot)$ and Jensen's inequality imply

$$\text{beta}(a,b)\geq \exp\left(\int_0^1 \log\left(x^{a-1}(1-x)^{b-1}\right)dx\right)=2-(a+b).$$

The upper bound follows from

$$2\alpha\beta\leq\alpha^2+\beta^2,\ \text{for all } \alpha,\beta\in\mathbb{R}.$$

Substituting α,β with x^{a-1} and y^{b-1} in the integral representation of the beta function completes the proof. ■