THE CONVEX GEOMETRY OF PRIVATE SELECTION

Anonymous authors
Paper under double-blind review

ABSTRACT

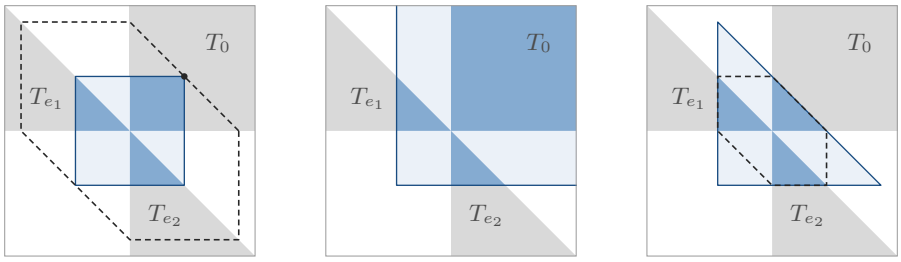
Many private learning systems spend their privacy budget on a winner, adding noise to a vector of scores and publishing the candidate with the largest noisy score. For every log-concave noise and every polytope of candidates we prove that the exact pure privacy of this mechanism is the largest change that an allowed score shift makes on one convex set of the noise, its privacy body, inside the tangent cones of the polytope. The exponential mechanism and permute-and-flip have the same body and hence exactly the same privacy on every readout, and the cone identity behind this result shows that the standard calibrations of the sparse vector technique are exact for long streams. Calibrating a noise thus becomes choosing a body. After the argmax, the K-norm noise that is optimal for releasing the scores has the largest expected spread among the K-norm noise distributions whose bodies are feasible and contain its body. On selection with abstention and on assignment, mechanisms already in use attain the optimal worst-case regret at high privacy, up to a constant factor.

1 INTRODUCTION

Differential privacy (Dwork et al., 2006) has two primitive readouts: publish a noisy value, or publish a noisy winner. Machine learning spends much of its privacy budget on the second. Top-k selection publishes the k candidates with the largest noisy scores (Durfee & Rogers, 2019; Qiao et al., 2021); PATE labels data with the argmax of the noisy votes of an ensemble of teachers (Papernot et al., 2017; 2018); private text generation samples tokens with the exponential mechanism, whose softmax is the argmax of the logits plus Gumbel noise (Amin et al., 2024; Vinod et al., 2025); private best-of-N alignment adds Gumbel noise to reward scores before it chooses a response (Jain et al., 2026); and the post-processing approach to private minimum spanning trees adds noise to every edge weight before one call to a spanning-tree algorithm (Sealfon, 2016; Pagh & Retschmeier, 2024). Each of these adds noise to a vector of scores and publishes the maximizer over a set of candidates, with Gumbel, exponential, Laplace or Gaussian noise, and the candidates may be single elements, subsets, permutations, trees, or an option to abstain. Hyperparameter tuning and model selection publish a winner as well (Liu & Talwar, 2019; Papernot & Steinke, 2022).

Releasing a value has had a geometry since Hardt & Talwar (2010) calibrated the noise to the sensitivity polytope of a query workload, and sharpening that geometry has been a sustained program: K-norm mechanisms, their optimal norm ball, and workload-adaptive noise (Bhaskara et al., 2012; Nikolov et al., 2013; Li et al., 2010; McKenna et al., 2018; Awan & Slavković, 2021). Selection has had privacy analyses of individual mechanisms: the exponential mechanism (McSherry & Talwar, 2007), report noisy max (Dwork & Roth, 2014), permute-and-flip (McKenna & Sheldon, 2020) and its identity with report noisy max under exponential noise (Ding et al., 2021), Gaussian noise for bounded scores (Lebensold et al., 2024), and one sufficient condition on independent noise that covers the first three and their one-shot top-k versions (Shekelyan & Loukides, 2022). These analyses concern particular mechanisms and readouts, and the per-mechanism mode of analysis has been fragile: published variants of the sparse vector technique repeatedly came with incorrect privacy proofs (Lyu et al., 2017). Post-processing a private release of the scores bounds the privacy of the winner by that of the release, and gives nothing more. What is missing is the exact privacy of the maximization, a comparison between noise distributions that upper bounds cannot make, and a rule for which noise suits which readout. McKenna and Sheldon close their paper with an open direction, "to understand more fully the nature of optimal mechanisms on more general domains or with other ways of averaging or aggregating over score vectors" (McKenna & Sheldon, 2020). *We ask: for*

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.



Laplace: $2\lambda\Delta$ exponential: ∞ Gumbel $G_i - G_0$: $\lambda\Delta$

Figure 1: Exact privacy on $Q_2 = \mathrm{conv}\{0, e_1, e_2\}$, two candidates and the option to abstain, under score-range adjacency, for three noise distributions of the same rate λ. Gray: the tangent cones T_0, T_{e_1}, T_{e_2}, the directions from each vertex into Q_2. Blue: the privacy body of the noise, dark inside the cones. Dashed: the smallest multiple of $D = Q_2 - Q_2$ that contains the dark part; its scale times Δ is $\varepsilon_{\text{exact}}$, the value under each panel. The three bodies agree in the cones of the two candidates and differ in the abstain cone T_0: there iid Laplace noise attains the value at one corner (dot), iid exponential noise fills the cone, and Gumbel noise on all three options, $Z_i = G_i - G_0$, reaches exactly to λD, while its body extends beyond λD outside the cones.

every log-concave noise and every polytope of candidates, what is the exact pure privacy of adding the noise to the scores and publishing the maximizer, and how does the readout constrain the choice of noise?

1.1 Our Results

The answer is one support-function computation. Call two score vectors adjacent when their difference lies in a set A of allowed shifts, for instance when every score moves by at most Δ_∞ between neighboring datasets. Writing $\varepsilon_{\text{exact}}$ for the least ε at which the mechanism is ε-differentially private, we prove in Theorem 1 that for every log-concave noise Z and every polytope Q

$$\varepsilon_{\text{exact}} = \max_{v \in \mathrm{vert}\, Q} \sup_{z \in T_v \cap B(Z)} h_A(z). \tag{1}$$

Here T_v is the tangent cone of Q at the vertex v, the directions from v into Q; $h_A(z) = \sup_{d \in A} \langle d, z \rangle$ is the largest change that an allowed shift makes to $\langle z, \cdot \rangle$; and $B(Z)$, the *privacy body* of the noise, is a convex set built from the domain of its cumulant generating function, which records the exponential rates at which its tails decay. For independent coordinates the body is a box of tail decay rates: $[-\lambda, \lambda]$ per coordinate for Laplace noise of rate λ, $[-\lambda, +\infty)$ for exponential and Gumbel noise of rate λ, and all of $\mathbb{R}$ for Gaussian noise (Lemma A.2). The formula is an equality, so it separates mechanisms as well as it bounds them.

Figure 1 shows the formula on two candidates with an option to abstain, $Q_2 = \mathrm{conv}\{0, e_1, e_2\}$, under score-range adjacency, where no comparison between two outputs moves by more than Δ and h_A is Δ times the gauge of the difference body $D = Q_2 - Q_2$. The exact privacy is then Δ times the smallest multiple of D that contains the part of the body inside the tangent cones. It is $2\lambda\Delta$ for Laplace noise, infinite for exponential noise, and $\lambda\Delta$ for Gumbel noise added to all three options, $Z_i = G_i - G_0$, whose body extends far beyond λD outside the cones.

Behind (1) is an identity for a single cone: for every closed convex cone C with nonempty interior, the *extreme loss* of C under a shift d, the supremum over score vectors w of $\log \Pr[w + Z \in C] / \Pr[w + d + Z \in C]$, is the support function of $B(Z) \cap C^\circ$ at d (Lemma 1). The argmax over a polytope is one readout of this kind. AboveThreshold, the core of the sparse vector technique (Dwork et al., 2009; Dwork & Roth, 2014), is another. It reports the first of $n \geq 2$ queries whose noisy answer reaches a noisy threshold, its cones are the normal cones of no polytope, and private text generation uses it to spend no privacy on tokens that a public model predicts (Amin et al., 2024). With noise of rate λ on the answers and μ on the threshold, its exact privacy is $\Delta_\infty \min\{n\lambda, 2\lambda + \mu\}$ for Laplace noise and $\Delta_\infty(2\lambda + \mu)$ for exponential and Gumbel noise (Corollary 2). The textbook calibration is therefore exactly ε once $n \geq 4$, and the bounds of Lyu et al. (2017) and Durfee (2023) are exact for long streams.

The noise enters (1) only through its body, and nothing finer about the noise enters the extreme loss of any vertex under any shift: two noise distributions have the same such losses on every polytope exactly when their bodies coincide, and two-candidate instances already tell distinct bodies apart (Theorem B.3). Exponential and Gumbel noise of the same rate have the same body, so the exponential mechanism and permute-and-flip have exactly the same privacy on every readout under every adjacency (Corollary 1), an equality between two different densities that upper bounds cannot establish. Gaussian noise has infinite exact privacy on unrestricted scores.

For one-sided noise, with independent coordinates whose right tails decay at an exponential rate and whose left tails are lighter than every exponential, the formula determines when the exact privacy is finite: exactly when every vertex of Q is Pareto-maximal (Theorem 2). An option to abstain, or a threshold without noise, rules one-sided noise out at every rate. Adding noise of the same distribution to the threshold makes the exact privacy finite again, and the resulting mechanisms are the exponential mechanism and permute-and-flip with the abstain option as one more candidate.

Calibrating a noise to a budget therefore means choosing a body. Under score-range adjacency a mechanism is ε-differentially private exactly when its body B is feasible, $T_v \cap B \subseteq \tau D$ at every vertex, with $\tau = \varepsilon/\Delta$ and $D = Q - Q$. The body τD belongs to the K-norm noise of Hardt and Talwar, the optimal K-norm noise for releasing the score vector (Awan & Slavković, 2021). The feasible bodies that contain it have a greatest element H_τ, obtained from τD by deleting the constraints of the facets that no tangent cone reaches (Theorem 3). Across the argmax the ranking that holds for release reverses: the K-norm noise with body τD has expected spread, the expected width of the noise across Q, exactly m/τ on every polytope of dimension m, and the K-norm noise of every other feasible body that contains τD has strictly less (Theorem 4). On the release side the same body is the optimal one, so the reversal cannot be read off from there.

On two readouts we compare mechanisms by their worst-case regret. With n candidates and an option to abstain, at high privacy ($\varepsilon \leq 1$), noise with independent coordinates has minimax regret $\Theta(n\Delta/\varepsilon)$, while the noise $G_i - G_0$, whose body is H_τ, attains $\Theta(\Delta \log n/\varepsilon)$, the best order of any ε-differentially private selection mechanism (Theorem 5). Independent noise is thus worse by a factor $\Theta(n/\log n)$, a statement about a class of mechanisms that no analysis of single mechanisms can make. On assignment, independent exponential noise on the score matrix followed by one call to the Hungarian algorithm has exact privacy ε and minimax-optimal regret $\Theta(n \log n\, \Delta/\varepsilon)$ at high privacy, while releasing the scores of all assignments has minimax error $\Theta(n^2\Delta/\varepsilon)$ (Theorem 6).

For the direction that McKenna and Sheldon name, the paper establishes four facts. On every polytope the exact privacy of every additive-noise selection mechanism is one support-function computation. The K-norm noise that is optimal for releasing the scores has, after the argmax, the largest expected spread among the K-norm noise distributions whose bodies are feasible and contain its body. On the simplex the greatest such body is the body of permute-and-flip. And on selection with abstention and on assignment, the optimal worst-case regret at high privacy is determined up to a constant factor and attained by mechanisms already in use. Every proof is in the appendix.

1.2 Related Work

Permute-and-flip matches or beats the exponential mechanism in expected error, by up to a factor of two (McKenna & Sheldon, 2020), and Gaussian-based selection reaches nearly optimal error (Leeman & Manurangsi, 2026). Closest to this paper, Shekelyan & Loukides (2022) prove one sufficient condition for the exponential mechanism, report noisy max and permute-and-flip and for their one-shot top-k versions: iid noise whose distribution function F makes $\log(1-F)$ 1-Lipschitz, a hazard rate of at most one everywhere after scaling. For log-concave noise this condition bounds the right-tail decay rate, one coordinate of the privacy body per candidate. Their result is an upper bound for independent noise on k-subsets, and Theorem 1 is an equality for every log-concave noise, independent or not, on every polytope, where the tangent cones of the readout determine which decay rates count.

On the release side, Hardt & Talwar (2010) introduced the K-norm mechanism together with packing lower bounds, and among K-norm mechanisms the optimal norm ball is the smallest valid one, the sensitivity hull (Awan & Slavković, 2021), whose noise has privacy body τD under score-range adjacency. Theorem 4 takes this geometry across the argmax, where the same noise has the largest expected spread among the K-norm noise distributions whose bodies are feasible and contain

τD. The packing technique, which also gives the pure lower bound for top-k selection (Steinke & Ullman, 2017), reappears in Theorems 5 and 6.

For the sparse vector technique, Lyu et al. (2017) proved the bound $\varepsilon_1 + \varepsilon_2$ for a corrected version that spends ε_1 on the threshold and ε_2 on the answers, and Durfee (2023) extended the analysis to exponential and Gumbel noise and to privacy losses that depend on the data. Their worst-case bounds are exact for long streams (Corollary 2 and Proposition B.6).

2 Preliminaries

Definition 1 (Additive-noise selection mechanism)**.** Let $Q \subset \mathbb{R}^m$ be a polytope with at least two vertices, which are the outputs. An *additive-noise selection mechanism* M_Z is specified by a noise vector Z with a log-concave density on $\mathbb{R}^m$. On input a score vector $w \in \mathbb{R}^m$ it draws Z and outputs the vertex that maximizes the noisy scores,

$$M_Z(w) \coloneqq \operatorname*{arg\,max}_{x \in \operatorname{vert} Q} \langle w + Z, x \rangle.$$

The polytope is the readout. On the simplex $\operatorname{conv}\{e_1, \dots, e_n\}$ the output is the candidate with the largest noisy score; on the k-subset polytope, whose vertices are the $0/1$ vectors with k ones, it is the set of the k candidates with the largest noisy scores; $Q_n = \operatorname{conv}\{0, e_1, \dots, e_n\}$ adds an option to abstain, the vertex 0; and on the Birkhoff polytope of doubly stochastic matrices the output is a maximum-weight assignment.

For a set $K \subseteq \mathbb{R}^m$ we write $h_K(y) = \sup_{x \in K} \langle x, y \rangle$ for its support function and $K^\circ = \{y : \langle x, y \rangle \le 1 \text{ for all } x \in K\}$ for its polar, and for a convex K containing the origin $\gamma_K(y) = \inf\{t > 0 : y \in tK\}$ for its gauge.

Definition 2 (Adjacency and exact privacy)**.** Let $A \subset \mathbb{R}^m$ be nonempty with $A = -A$ and h_A finite on $Q - Q$. Score vectors w and w' are *adjacent* if $w' - w \in A$, and the elements of A are the *allowed shifts*. The *exact privacy* of M_Z is

$$\varepsilon_{\text{exact}}(M_Z) \coloneqq \sup\Big\{ \log \frac{\Pr[M_Z(w) \in S]}{\Pr[M_Z(w') \in S]} \;:\; w, w' \text{ adjacent},\ S \subseteq \operatorname{vert} Q,\ \Pr[M_Z(w) \in S] > 0 \Big\},$$

with $\log(p/0) = +\infty$ for $p > 0$. The mechanism is ε-differentially private (Dwork et al., 2006) exactly when $\varepsilon_{\text{exact}}(M_Z) \le \varepsilon$.

When the scores are computed from a dataset and neighboring datasets give adjacent score vectors, this is ε-differential privacy of the published winner. Only differences of scores matter, so every instance may be read within $V = \operatorname{span}(Q - Q)$ (Appendix A), and the design theory of Section 4 takes Q full-dimensional.

Three sets of allowed shifts cover the practice. *Per-score adjacency*, $A = \Delta_\infty[-1, 1]^m$, lets every score move by at most Δ_∞, the sensitivity condition of the exponential mechanism and of report noisy max; *monotone per-score adjacency*, $A = \Delta_\infty\big([0, 1]^m \cup [-1, 0]^m\big)$, asks moreover that all scores move in the same direction. *Score-range adjacency*, $A = \Delta D^\circ$ with $D \coloneqq Q - Q$ the difference body of Q, asks that no comparison between two outputs move by more than Δ, since $h_D(d) = \max_{x, x' \in \operatorname{vert} Q} \langle d, x - x' \rangle$ and $D^\circ = \{d : h_D(d) \le 1\}$. Their support functions are $h_A(z) = \Delta_\infty \|z\|_1$, $h_A(z) = \Delta_\infty \max\{\sum_i z_i^+, \sum_i z_i^-\}$ with $z^\pm = \max\{\pm z, 0\}$ taken coordinate-wise, and $h_A(z) = \Delta \gamma_D(z)$.

For a vertex v of Q, $T_v = \operatorname{cone}(Q - v)$ is the *tangent cone* of Q at v, the set of directions from v into Q. Noise distributions are parametrized by their *rate* $\lambda > 0$: Laplace noise has the density $\frac{\lambda}{2} e^{-\lambda |x|}$, exponential noise the density $\lambda e^{-\lambda x}$ on $x \ge 0$, and Gumbel noise the distribution function $\exp(-e^{-\lambda x})$, of scale $1/\lambda$. The running example is $Q_2 = \operatorname{conv}\{0, e_1, e_2\}$: two candidates e_1 and e_2 and the abstain vertex 0 (Figure 1).

3 The Privacy Body

One convex set, determined by the tails of the noise, determines the exact privacy of every mechanism of Definition 1.

Definition 3 (Privacy body)**.** The cumulant generating function of Z is $\Lambda_Z(\theta) = \log \mathbb{E}\, e^{\langle \theta, Z\rangle}$, with the effective domain $\operatorname{dom}\Lambda_Z = \{\theta : \Lambda_Z(\theta) < +\infty\}$. The *privacy body* of Z is

$$B(Z) := -\operatorname{cl}\operatorname{dom}\Lambda_Z.$$

A vector θ lies in $\operatorname{dom}\Lambda_Z$ when the density of Z tilted by $e^{\langle \theta, z\rangle}$ stays integrable, so the body records the exponential rates at which the tails of the noise decay in every direction. It is a closed convex set with the origin in its interior (Lemma A.1). For independent coordinates it is a box, $B(Z) = \prod_i [-\beta_i, \alpha_i]$, where β_i and α_i are the exponential decay rates of the right and left tails of Z_i, an infinite rate meaning a tail lighter than every exponential (Lemma A.2).

The event that the argmax returns a given vertex is the event that the noisy scores lie in a convex cone, so the log-ratios in Definition 2 reduce to comparisons of the probability of one cone at two translates. For a closed convex cone C, $C^\circ = \{y : \langle x, y\rangle \le 0 \text{ for all } x \in C\}$ is its polar cone.

Lemma 1 (Cone identity)**.** *Let $C \subseteq \mathbb{R}^m$ be a closed convex cone with nonempty interior, let $d \in \mathbb{R}^m$, and put $p_C(w) := \Pr[w + Z \in C]$. Then*

$$\sup_{w\,:\,p_C(w)>0} \log \frac{p_C(w)}{p_C(w+d)} \;=\; h_{B(Z)\cap C^\circ}(d),$$

a value in $[0, +\infty]$.

The supremum in Lemma 1, the *extreme loss* of C under d, arises arbitrarily far out, along the directions that stay inside the cone. For the upper bound, translating the integration variable by a point of the cone keeps it in the cone and changes the log-density by at most its slope at infinity, a support function of the body by convex duality. For the lower bound, Markov's inequality gives $p_C(w) \le e^{\langle \theta, w\rangle + \Lambda_Z(\theta)}$ for $-\theta \in C^\circ$, so along the ray $w + td$ the log-ratio grows at a slope of at least $\langle -\theta, d\rangle$, and these slopes approach the same support function (Appendix B.1). The identity applies to every readout that reports which of finitely many closed convex cones contains the noisy scores, and the argmax over a polytope is the first.

Theorem 1 (Exact privacy)**.** *Every mechanism of Definition 1 satisfies*

$$\varepsilon_{\text{exact}}(M_Z) \;=\; \max_{v \in \operatorname{vert} Q} \; \sup_{z \in T_v \cap B(Z)} h_A(z).$$

Up to ties, the vertex v wins exactly when the noisy scores lie in the normal cone of Q at v, a closed convex cone whose polar cone is T_v; Lemma 1 therefore gives the extreme loss of v under a shift d as $h_{B(Z)\cap T_v}(d)$, the losses of larger events reduce to those of single vertices, and exchanging the suprema over $d \in A$ and over $z \in T_v \cap B(Z)$ gives the formula (Appendix B.1). Under score-range adjacency $h_A = \Delta\gamma_D$, and

$$\varepsilon_{\text{exact}}(M_Z) \;=\; \Delta \cdot \max_{v \in \operatorname{vert} Q} \; \sup_{z \in T_v \cap B(Z)} \gamma_D(z), \tag{2}$$

Δ times the smallest multiple of D that contains the part of the body inside the tangent cones, the quantity drawn in Figure 1 (computed in Appendix B.2).

Corollary 1 (The exponential mechanism and permute-and-flip)**.** *Noise vectors with the same privacy body have the same exact privacy on every polytope under every adjacency. Iid exponential noise and iid Gumbel noise of rate λ both have the body $[-\lambda, +\infty)^m$, so the exponential mechanism and permute-and-flip have the same exact privacy, and so do their versions on every other polytope.*

On the simplex, Gumbel noise makes the argmax sample from the softmax distribution, which is the exponential mechanism (McFadden, 1974; McSherry & Talwar, 2007), and exponential noise gives permute-and-flip (Ding et al., 2021). The two mechanisms were designed to share their guarantee (McKenna & Sheldon, 2020); the corollary makes the guarantee an equality and gives its reason. One noise is supported on a half-line and the other on all of $\mathbb{R}$, and both have right tails of rate λ and left tails lighter than every exponential. In Figure 1, independent Gumbel noise has the middle body and $Z_i = E_i - E_0$ with iid exponential E_i has the right one, so they have the same exact privacy as the noise drawn there. Theorem 1 gives the exact privacy of the familiar readouts in one line each (Proposition B.2). Under per-score adjacency a single winner has exact privacy $2\lambda\Delta_\infty$ for Laplace, exponential and Gumbel noise alike. One-shot top-k selection, which reports the k candidates with

the largest noisy scores, has exact privacy $2k\lambda\Delta_\infty$ for exponential and Gumbel noise, which equals the composition bound for k sequential draws of the exponential mechanism, the distribution of one-shot Gumbel selection (Durfee & Rogers, 2019), and $2\min\{k, n-k\}\lambda\Delta_\infty$ for Laplace noise; monotone per-score adjacency halves every value. At equal rates the formula thus separates Laplace noise from the other two once $k > n/2$.

Gaussian noise has the body $\mathbb{R}^m$, so its exact privacy is infinite as soon as some allowed shift moves the comparison between two outputs (Proposition B.1). No finer feature of the noise than its body matters: two noise distributions have the same extreme loss at every vertex of every polytope under every shift exactly when their bodies coincide, and two-candidate instances already separate distinct bodies (Theorem B.3).

The formula also determines when the exact privacy is finite. Exponential noise, Gamma noise of shape at least one, and Gumbel noise have a right tail of exponential rate and a left tail lighter than every exponential. We call noise with independent coordinates of this kind *one-sided*; its body is $\prod_i[-\beta_i, +\infty)$. For a closed convex set B, $\operatorname{rec} B = \{y : B + y \subseteq B\}$ is its recession cone.

Theorem 2 (One-sided noise)**.** *Let the set A of allowed shifts detect every direction: every $z \neq 0$ has $\langle d, z\rangle \neq 0$ for some $d \in A$, as under per-score, monotone per-score and score-range adjacency.*

(a) *$\varepsilon_{\text{exact}}(M_Z)$ is finite if and only if $T_v \cap \operatorname{rec} B(Z) = \{0\}$ for every vertex v of Q.*
(b) *Let Z be one-sided, with right-tail rates $0 < \beta_i < +\infty$. Then $\varepsilon_{\text{exact}}(M_Z)$ is finite if and only if every vertex of Q is Pareto-maximal: no $x \in Q$ has $x \geq v$ coordinatewise with $x \neq v$.*

Under such an A the support function h_A is a norm on the span of $Q - Q$, which contains every tangent cone, so the exact privacy is finite exactly when the body meets every tangent cone in a bounded set; the body of one-sided noise recedes along every direction of $\mathbb{R}^m_+$, and T_v contains such a direction exactly when a point of Q dominates v (Appendix B.4). A polytope in a level set of a linear functional with strictly positive coefficients has only Pareto-maximal vertices, and the simplex, the k-subset polytope, the Birkhoff polytope and the base polytope of a matroid all lie on level sets of the coordinate sum. An abstain vertex is the obstruction on the other side: if 0 is a vertex of $Q \subseteq \mathbb{R}^m_+$, every other vertex dominates it, and one-sided noise has infinite exact privacy at every rate, as in the middle panel of Figure 1. On Q_n, with n candidates and the option to abstain, independent Laplace noise has exact privacy $n\lambda\Delta_\infty$ under per-score adjacency (Proposition B.2), because the abstain cone $T_0 = \mathbb{R}^n_+$ adds the left-tail rates of all coordinates.

Noise on the threshold removes the obstruction. The abstain option is a threshold that every candidate must beat, and adding X_0 to it and X_i to the score of candidate i, with $X_0, \ldots, X_n$ iid, is the noise $Z_i = X_i - X_0$. For Gumbel and exponential noise of rate λ its body is $\{z : z_i \geq -\lambda \text{ for all } i,\ \sum_i z_i \leq \lambda\}$, and the resulting mechanisms are the exponential mechanism and permute-and-flip with the abstain option as one more candidate (Proposition B.5). Under per-score adjacency the noise $X_i - X_0$ has exact privacy $2\lambda\Delta_\infty$, the value of a single winner, for Laplace, exponential and Gumbel X_i alike (Proposition B.2).

AboveThreshold (Dwork & Roth, 2014, Algorithm 1) adds noise to a threshold once and to the answer of every query, and reports the first of n queries whose noisy answer is at least the noisy threshold, or $\perp$ if there is none. Each answer has sensitivity Δ_∞, and adjacency is per-query: every answer moves by at most Δ_∞ while the threshold stays fixed. With the threshold appended as a coordinate, each output occurs exactly when the noisy vector lies in a closed convex cone, but for $n \geq 2$ these $n+1$ cones are not the normal cones of any polytope (Appendix B.5), and Lemma 1 applies to them.

Corollary 2 (AboveThreshold)**.** *Let AboveThreshold run on $n \geq 2$ queries with independent noise of rate λ on each answer and of rate μ on the threshold, under per-query adjacency. If the noise on the answers and on the threshold is Laplace, $\varepsilon_{\text{exact}} = \Delta_\infty \min\{n\lambda,\ 2\lambda + \mu\}$; if it is exponential throughout or Gumbel throughout, $\varepsilon_{\text{exact}} = \Delta_\infty(2\lambda + \mu)$. Under monotone per-query adjacency, where all answers move in the same direction, the exact privacy is $\Delta_\infty(\lambda + \mu)$, for Laplace noise when $n \geq 2 + \mu/\lambda$ and for exponential and Gumbel noise for every $n \geq 2$.*

The body is a box, the polar cone of each output is generated by the differences of unit vectors that define its cone, and Lemma 1 turns the extreme loss of each output into a linear maximization over the weights of these generators; halting at the last query gives the largest value (Appendix B.5). Since the algorithm halts at its first positive answer, the values hold for adaptively chosen queries as

well. The calibration of Dwork & Roth (2014, Theorem 3.23), $\lambda = \varepsilon/(4\Delta_\infty)$ and $\mu = \varepsilon/(2\Delta_\infty)$, therefore has exact privacy $\min\{n/4, 1\} \cdot \varepsilon$, and the bounds of Lyu et al. (2017) and Durfee (2023) are exact for long streams (Appendix B.5).

4 Feasible Bodies

By Theorem 1 the exact privacy depends on the noise only through its body, so calibrating a mechanism to a budget means choosing a convex body. The design theory is stated for score-range adjacency, under which privacy and regret are measured by the same body D; per-score shifts are covered with Δ equal to their largest value of h_D, which is $2\Delta_\infty$ on the simplex and on Q_n and $2n\Delta_\infty$ on the Birkhoff polytope (Appendix C). For a budget $\varepsilon > 0$ we write $\tau = \varepsilon/\Delta$.

Definition 4 (Feasible body)**.** A closed convex set $B \subseteq \mathbb{R}^m$ with $0 \in \operatorname{int} B$ is *feasible* if

$$T_v \cap B \;\subseteq\; \tau D \qquad \text{for every vertex } v \text{ of } Q.$$

Since $\tau D = \{\gamma_D \le \tau\}$, (2) says that M_Z is ε-differentially private exactly when $B(Z)$ is feasible. The feasible bodies are ordered by inclusion, and every bounded one is the body of a noise distribution.

Definition 5 (K-norm noise with a given body)**.** For a bounded closed convex B with $0 \in \operatorname{int} B$, the *K-norm noise with body B* has the density proportional to $e^{-h_{-B}}$.

The exponent h_{-B} is the gauge γ_K of $K = (-B)^\circ$. When $B = -B$ it is a norm, the K-norm noise with body B is the noise of the K-norm mechanism of Hardt & Talwar (2010) for that K, and these are the noise distributions among which Awan & Slavković (2021) optimize the release of a vector. The privacy body of the K-norm noise with body B is B itself (Appendix C). Taking $B = \tau D$ gives the density proportional to $e^{-\tau h_D}$, the K-norm noise calibrated to the set ΔD° of allowed shifts. This *baseline* is feasible, since $T_v \cap \tau D \subseteq \tau D$, and it is the optimal K-norm noise for releasing the score vector (Awan & Slavković, 2021).

Privacy of the argmax requires less. Releasing $w + Z$ is ε-differentially private exactly when the whole body lies in τD, since the extreme log-density ratio at a shift d is $h_{B(Z)}(d)$ (Proposition C.1). The argmax depends on the scores only through the winning events, and its privacy requires only $T_v \cap B \subseteq \tau D$ at every vertex. A feasible body may therefore reach beyond τD outside the tangent cones, and a larger body gives the K-norm noise an exponent h_{-B} at least as large at every point. How far a feasible body that contains τD may reach is determined by which facets of D the tangent cones reach. Each facet F of D is $F = \{y \in D : \langle \nu_F, y\rangle = 1\}$ for a vector ν_F with $\langle \nu_F, \cdot\rangle \le 1$ on D, so that $\tau D = \{z : \langle \nu_F, z\rangle \le \tau \text{ for every facet } F\}$. We call a facet F *active* if its relative interior meets the interior of some tangent cone T_v, and we put

$$H_\tau \coloneqq \{z : \langle \nu_F, z\rangle \le \tau \text{ for every active facet } F \text{ of } D\},$$

the polyhedron obtained from τD by deleting the constraints of the facets that no tangent cone reaches.

Theorem 3 (The greatest feasible body containing the baseline)**.** *H_τ is feasible, contains τD, contains every feasible body that contains τD, and is bounded.*

On a tangent cone the gauge of D is attained at an active facet, since tilting a point into the interior of the cone breaks the ties between facets, and this makes H_τ feasible. A body that contains τD and violates the constraint of an active facet contains, near a point of that facet interior to a tangent cone, a point of the cone outside τD. Every facet normal of Q gives an active facet of D, and these normals positively span $\mathbb{R}^m$, which bounds H_τ (Appendix C.2). The baseline is already the greatest exactly when every facet of D is active (Proposition C.3), and $|\operatorname{vert} Q|$ times the number of facets of D linear programs compute H_τ. In the right panel of Figure 1, at $\tau = \lambda$, the three dashed edges inside the gray cones are the active facets and continue as the sides of the triangle H_λ, and the three dashed edges in the white wedges are the facets that no tangent cone reaches.

We measure the noise that a mechanism adds by its *expected spread* $\mathbb{E}\, h_D(Z)$, the expected width of $\langle Z, \cdot\rangle$ across Q. It bounds the regret, the expected gap between the optimal value and the value of the output, at every score vector: writing x^* for a maximizer of $\langle w, \cdot\rangle$ over Q and $\hat{x}$ for the output, the choice of $\hat{x}$ gives $\langle w, x^* - \hat{x}\rangle \le \langle Z, \hat{x} - x^*\rangle \le h_D(Z)$.

Theorem 4 (Expected spread). *The K-norm noise with body τD has expected spread m/τ for every polytope Q. For every feasible body B that contains τD, the K-norm noise with body B has expected spread at most m/τ, with equality only for $B = \tau D$.*

For K-norm noise with body B the sublevel sets of the exponent are the multiples of $K = (-B)^\circ$, and homogeneity gives $\mathbb{E}\,\psi(Z) = (m+1)\int_K \psi/|K|$ for every continuous nonnegative ψ that is positively homogeneous of degree one (Lemma C.4). With $\psi = h_D$ the expected spread becomes m/τ times an average of a ratio of the radial functions of K and $(-\tau D)^\circ$, which is at most one and equals one exactly at $B = \tau D$ (Appendix C.3). The release-optimal noise thus has the largest spread after the argmax, because Definition 4 constrains a body only inside the tangent cones, and the containment of τD cannot be dropped (Example C.5).

The greatest body is explicit on the readouts of this paper (Proposition C.6, Example C.7). On the simplex, $H_\tau = \{z \in V : z_i \geq -\tau \text{ for all } i\}$, the body that Corollary 1 computes for the exponential mechanism and permute-and-flip at rate τ. On Q_n, $H_\tau = \{z : z_i \geq -\tau \text{ for all } i,\ \sum_i z_i \leq \tau\}$, the body of the noise $G_i - G_0$ and of $E_i - E_0$ at rate τ. On the cube $[0,1]^m$ every facet of D is active and $H_\tau = \tau D$. On the first two readouts, the greatest feasible body that contains the baseline belongs to a mechanism already in use.

5 Abstention and Assignment

Privacy constrains the body of the noise, and regret measures the accuracy of the resulting mechanism. This section computes both on selection with abstention and on assignment, under score-range adjacency, and compares the regret with the least regret of any private selection mechanism. A *selection mechanism* on Q is a randomized map M from score vectors to $\operatorname{vert} Q$; it is ε-differentially private when $\Pr[M(w) \in S] \leq e^\varepsilon \Pr[M(w') \in S]$ for every ordered adjacent pair (w, w') and every $S \subseteq \operatorname{vert} Q$, which for the mechanisms of Definition 1 is $\varepsilon_{\text{exact}} \leq \varepsilon$. The regret of M at a score vector w is $\operatorname{Regret}(M, w) = \max_{x \in Q}\langle w, x\rangle - \mathbb{E}\,\langle w, M(w)\rangle$, and mechanisms are compared by their worst-case regret $\sup_w \operatorname{Regret}(M, w)$.

Theorem 5 (Selection with abstention). *On Q_n under score-range adjacency, let $0 < \varepsilon \leq 1$ and $n \geq 30$. The least worst-case regret of an ε-differentially private mechanism M_Z whose noise has independent coordinates is $\Theta(n\Delta/\varepsilon)$, and the least worst-case regret of an ε-differentially private selection mechanism is $\Theta(\Delta \log n/\varepsilon)$. The second is attained by the noise $Z_i = G_i - G_0$ with $G_0, \ldots, G_n$ iid Gumbel of rate τ, whose body is H_τ and whose exact privacy is ε.*

For independent coordinates Theorem 1 gives $\varepsilon_{\text{exact}} = \Delta \max\{\sum_i \alpha_i, \max_j \beta_j\}$, since the abstain cone adds the left-tail rates of all coordinates. Privacy therefore forces some $\alpha_j \leq \tau/n$, so the density of Z_j is at most τ/n, and at one of two score vectors the mechanism misses the optimum by order $n\Delta/\varepsilon$ with constant probability (Theorem D.1). The noise $G_i - G_0$ makes M_Z the exponential mechanism over the n candidates and the abstain option, with regret at most $(1 + \log(n+1))\Delta/\varepsilon$, and packing the multiples of the unit vectors e_j shows that no ε-differentially private selection mechanism improves the order (Theorem D.2).

The gap between the two classes is structural: for noise with independent coordinates, $\varepsilon_{\text{exact}} \leq \varepsilon$ holds exactly when $B(Z) \subseteq H_\tau$ (Appendix D.1), so every body of the first class lies inside H_τ, the body of the noise $G_i - G_0$, and at the abstain vertex the n left-tail rates of independent noise must share one budget.

On assignment, an instance is a score matrix $w \in \mathbb{R}^{n \times n}$, where w_{ij} is the value of assigning i to j, and the readout is the maximum-weight assignment, so Q is the Birkhoff polytope B_n, whose vertices are the permutation matrices X_σ. Adjacency asks that no individual move the difference between the values of two assignments by more than Δ; adding a constant to a row or a column of the score matrix changes no such difference. The structure is public and the scores are private, as in the private-weights model of Sealfon (2016), whose ℓ_1 edge-weight adjacency is covered, since $h_D(d) \leq \|d\|_1$ (Appendix D.2). Releasing the scores is a different task: a mechanism may output anything from which a fixed rule decodes a value for every assignment, and its range error is the expected oscillation of the decoded values around the true ones (Appendix D.2).

Theorem 6 (Private assignment). *On B_n under score-range adjacency, iid exponential noise of rate λ on the entries followed by one maximum-weight assignment has exact privacy $\lambda\Delta$. For $0 < \varepsilon \leq 1$*

and $n \geq 50$, at $\lambda = \tau$ it attains the least worst-case regret of an ε-differentially private selection mechanism, $\Theta(n \log n\, \Delta/\varepsilon)$, while the least worst-case range error of an ε-differentially private mechanism that releases all assignment scores is $\Theta(n^2 \Delta/\varepsilon)$.

Read within the space V of matrices with zero row and column sums, the body of the noise is $\{z \in V : z_{ij} \geq -\lambda\}$, and on the tangent cone at X_π the gauge of D is the largest of the entries $-z_{i\pi(i)}$, so Theorem 1 gives $\lambda\Delta$. The regret is at most the expected sum of the row maxima of the noise minus n/λ. Packings of permutations in Hamming distance, and of the ball of h_D for the release, give the lower bounds, and the K-norm noise with body τD attains the release rate (Theorem D.3, Theorem D.4). The mechanism makes one call to the Hungarian algorithm (Kuhn, 1955) on the noisy matrix and runs in $O(n^3)$ time with the standard implementation. It is the template already in use for private weights, now calibrated exactly. A release of the scores also selects, with regret at most its range error (Appendix D.2), so releasing the optimized object in place of the scores (Gupta et al., 2010) lowers the minimax error on assignment from $\Theta(n^2\Delta/\varepsilon)$ to $\Theta(n \log n\, \Delta/\varepsilon)$. Under per-score adjacency with sensitivity Δ_∞, iid Laplace, exponential or Gumbel noise of rate λ followed by one maximum-weight assignment has exact privacy $2n\lambda\Delta_\infty$, while releasing the score matrix with iid Laplace noise of the same rate has exact privacy $n^2\lambda\Delta_\infty$ (Proposition D.5).

6 Discussion

The Birkhoff polytope entered Theorem 6 only through its vertices, its tangent cones and its difference body, and Theorem 1 and Section 4 take no more than that from any readout. Over the permutohedron the vertex that maximizes $\langle w + Z, \cdot \rangle$ is the permutation that sorts the noisy scores, so a full noisy ranking is an instance of Definition 1, and over the base polytope of a matroid the mechanism returns a maximum-weight basis. Both lie on level sets of the coordinate sum, so one-sided noise has finite exact privacy on them (Theorem 2), and the post-processing approach to private minimum spanning trees (Sealfon, 2016; Pagh & Retschmeier, 2024) is the mechanism on the base polytope of a graphic matroid. Private matching has also been studied under joint differential privacy (Hsu et al., 2014) and with the graph itself as the sensitive input (Dinitz et al., 2025).

For a released value, the geometric mechanism is universally optimal for a single counting query (Ghosh et al., 2009), staircase noise is optimal for a single real query (Geng & Viswanath, 2016), and no universally optimal mechanism exists beyond counting queries (Brenner & Nissim, 2010). Within f-differential privacy (Dong et al., 2022), a canonical noise distribution matches the tradeoff function of a guarantee exactly, one exists for every guarantee (Awan & Vadhan, 2023), and for pure differential privacy none is log-concave (Awan & Dong, 2022). In this paper exactness concerns the privacy parameter $\varepsilon_{\text{exact}}$, where the privacy profile of a mechanism (Koskela et al., 2024) reaches zero, and exact characterizations are also known for zero-concentrated differential privacy (Harrison & Manurangsi, 2025). When the scores are bounded by design, as the vote counts of PATE or validation accuracies in $[0, 1]$ are, Theorem 1 remains an upper bound, and Gaussian noise followed by the argmax is private (Lebensold et al., 2024).

Adding noise to scores and publishing the maximizer is a primitive of several fields besides privacy. Perturb-and-MAP samples structured models with noisy maximization (Papandreou & Yuille, 2011) and bounds partition functions with random maxima (Hazan & Jaakkola, 2012); the Gumbel-max trick and its relaxations are standard components of deep learning (Jang et al., 2017; Maddison et al., 2017; Huijben et al., 2023); follow-the-perturbed-leader is the same maximization in online learning (Kalai & Vempala, 2005), where the stability that differential privacy provides is used to analyze perturbed leaders (Abernethy et al., 2019); random-utility maximization is the same maximization in econometrics (McFadden, 1974); and permute-and-flip has become a decoding and watermarking scheme for language models, valued there for its stability (Zhao et al., 2025). Whenever the noise has a log-concave density, the exact pure privacy of the maximization is given by Theorem 1.

Releasing a score vector requires the whole body of its noise to lie in τD (Proposition C.1), and publishing its argmax only the part of the body inside the tangent cones, a difference that Section 5 measures on two readouts. The exact privacy of a noisy winner depends on the noise only through its body, and the readout determines which part of the body it depends on.

AI USE STATEMENT

The authors used OpenAI's ChatGPT and Anthropic's Claude to help organize and polish the exposition of this paper, based on the authors' own drafts. The same tools assisted in surveying the related literature. The authors formulated and proved all mathematical claims in this paper, and checked, revised, and finalized every statement, proof, and reference. They take full responsibility for the content of the submission.

REPRODUCIBILITY STATEMENT

All results are theorems with complete proofs. The proofs for Section 3 are in Appendix B, those for Section 4 in Appendix C, and those for Section 5 in Appendix D; Appendix A collects the facts of convex analysis that they use and the properties of the privacy body. Every value quoted in the main text is proved in the appendix, and the three values of Figure 1 are computed in Appendix B.2.

REFERENCES

Jacob Abernethy, Young Hun Jung, Chansoo Lee, Audra McMillan, and Ambuj Tewari. Online learning via the differential privacy lens. In *Advances in Neural Information Processing Systems (NeurIPS)*, pp. 8892–8902, 2019.

Kareem Amin, Alex Bie, Weiwei Kong, Alexey Kurakin, Natalia Ponomareva, Umar Syed, Andreas Terzis, and Sergei Vassilvitskii. Private prediction for large-scale synthetic text generation. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pp. 7244–7262, 2024.

Jordan Awan and Jinshuo Dong. Log-concave and multivariate canonical noise distributions for differential privacy. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2022.

Jordan Awan and Aleksandra Slavković. Structure and sensitivity in differential privacy: Comparing K-norm mechanisms. *Journal of the American Statistical Association*, 116(534):935–954, 2021.

Jordan Awan and Salil Vadhan. Canonical noise distributions and private hypothesis tests. *The Annals of Statistics*, 51(2):547–572, 2023.

Aditya Bhaskara, Daniel Dadush, Ravishankar Krishnaswamy, and Kunal Talwar. Unconditional differentially private mechanisms for linear queries. In *ACM Symposium on Theory of Computing (STOC)*, pp. 1269–1284, 2012.

Hai Brenner and Kobbi Nissim. Impossibility of differentially private universally optimal mechanisms. In *IEEE Symposium on Foundations of Computer Science (FOCS)*, pp. 71–80, 2010.

Zeyu Ding, Daniel Kifer, Sayed M. Saghaian N. E., Thomas Steinke, Yuxin Wang, Yingtai Xiao, and Danfeng Zhang. The Permute-and-Flip mechanism is identical to Report-Noisy-Max with exponential noise, 2021. arXiv:2105.07260.

Michael Dinitz, George Z. Li, Quanquan C. Liu, and Felix Zhou. Differentially private matchings, 2025. arXiv:2501.00926.

Jinshuo Dong, Aaron Roth, and Weijie J. Su. Gaussian differential privacy. *Journal of the Royal Statistical Society, Series B*, 84(1):3–37, 2022.

David Durfee. Unbounded differentially private quantile and maximum estimation. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2023.

David Durfee and Ryan Rogers. Practical differentially private top-k selection with pay-what-you-get composition. In *Advances in Neural Information Processing Systems (NeurIPS)*, pp. 3527–3537, 2019.

Cynthia Dwork and Aaron Roth. The algorithmic foundations of differential privacy. *Foundations and Trends in Theoretical Computer Science*, 9(3–4):211–407, 2014.

Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. In *Theory of Cryptography Conference (TCC)*, pp. 265–284, 2006.

Cynthia Dwork, Moni Naor, Omer Reingold, Guy N. Rothblum, and Salil Vadhan. On the complexity of differentially private data release: efficient algorithms and hardness results. In *Proceedings of the 41st Annual ACM Symposium on Theory of Computing (STOC)*, pp. 381–390, 2009.

Quan Geng and Pramod Viswanath. The optimal noise-adding mechanism in differential privacy. *IEEE Transactions on Information Theory*, 62(2):925–951, 2016.

Arpita Ghosh, Tim Roughgarden, and Mukund Sundararajan. Universally utility-maximizing privacy mechanisms. In *ACM Symposium on Theory of Computing (STOC)*, pp. 351–360, 2009.

Anupam Gupta, Katrina Ligett, Frank McSherry, Aaron Roth, and Kunal Talwar. Differentially private combinatorial optimization. In *ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pp. 1106–1125, 2010.

Moritz Hardt and Kunal Talwar. On the geometry of differential privacy. In *ACM Symposium on Theory of Computing (STOC)*, pp. 705–714, 2010.

Charlie Harrison and Pasin Manurangsi. Exact zCDP characterizations for fundamental differentially private mechanisms, 2025. arXiv:2510.25746.

Tamir Hazan and Tommi Jaakkola. On the partition function and random maximum a-posteriori perturbations. In *International Conference on Machine Learning (ICML)*, 2012.

Justin Hsu, Zhiyi Huang, Aaron Roth, Tim Roughgarden, and Zhiwei Steven Wu. Private matchings and allocations. In *ACM Symposium on Theory of Computing (STOC)*, pp. 21–30, 2014.

Iris A. M. Huijben, Wouter Kool, Max B. Paulus, and Ruud J. G. van Sloun. A review of the Gumbel-max trick and its extensions for discrete stochasticity in machine learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(2):1353–1371, 2023.

Ishi Jain, Nandini Bhattad, and Sayak Ray Chowdhury. Privacy without regret: differentially private inference-time alignment, 2026. arXiv:2608.26324.

Eric Jang, Shixiang Gu, and Ben Poole. Categorical reparameterization with Gumbel-softmax. In *International Conference on Learning Representations (ICLR)*, 2017.

Adam Tauman Kalai and Santosh S. Vempala. Efficient algorithms for online decision problems. *Journal of Computer and System Sciences*, 71(3):291–307, 2005.

Antti Koskela, Rachel Redberg, and Yu-Xiang Wang. Privacy profiles for private selection. In *Proceedings of the 41st International Conference on Machine Learning (ICML)*, volume 235 of *PMLR*, pp. 25313–25332, 2024.

Harold W. Kuhn. The Hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, 2(1–2):83–97, 1955.

Jonathan Lebensold, Doina Precup, and Borja Balle. On the privacy of selection mechanisms with Gaussian noise. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, pp. 1495–1503, 2024.

Ethan Leeman and Pasin Manurangsi. Nearly-optimal private selection via Gaussian mechanism. In *Symposium on Foundations of Responsible Computing (FORC)*, pp. 4:1–4:13, 2026.

Chao Li, Michael Hay, Vibhor Rastogi, Gerome Miklau, and Andrew McGregor. Optimizing linear counting queries under differential privacy. In *ACM Symposium on Principles of Database Systems (PODS)*, pp. 123–134, 2010.

Jingcheng Liu and Kunal Talwar. Private selection from private candidates. In *ACM Symposium on Theory of Computing (STOC)*, pp. 298–309, 2019.

Min Lyu, Dong Su, and Ninghui Li. Understanding the sparse vector technique for differential privacy. *Proceedings of the VLDB Endowment*, 10(6):637–648, 2017.

Chris J. Maddison, Andriy Mnih, and Yee Whye Teh. The Concrete distribution: a continuous relaxation of discrete random variables. In *International Conference on Learning Representations (ICLR)*, 2017.

Daniel McFadden. Conditional logit analysis of qualitative choice behavior. In Paul Zarembka (ed.), *Frontiers in Econometrics*, pp. 105–142. Academic Press, 1974.

Ryan McKenna and Daniel R. Sheldon. Permute-and-Flip: A new mechanism for differentially private selection. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2020.

Ryan McKenna, Gerome Miklau, Michael Hay, and Ashwin Machanavajjhala. Optimizing error of high-dimensional statistical queries under differential privacy. *Proceedings of the VLDB Endowment*, 11(10):1206–1219, 2018.

Frank McSherry and Kunal Talwar. Mechanism design via differential privacy. In *IEEE Symposium on Foundations of Computer Science (FOCS)*, pp. 94–103, 2007.

Aleksandar Nikolov, Kunal Talwar, and Li Zhang. The geometry of differential privacy: the sparse and approximate cases. In *ACM Symposium on Theory of Computing (STOC)*, pp. 351–360, 2013.

Rasmus Pagh and Lukas Retschmeier. Faster private minimum spanning trees, 2024. arXiv:2408.06997.

George Papandreou and Alan L. Yuille. Perturb-and-MAP random fields: Using discrete optimization to learn and sample from energy models. In *IEEE International Conference on Computer Vision (ICCV)*, pp. 193–200, 2011.

Nicolas Papernot and Thomas Steinke. Hyperparameter tuning with Rényi differential privacy. In *International Conference on Learning Representations (ICLR)*, 2022.

Nicolas Papernot, Martín Abadi, Úlfar Erlingsson, Ian Goodfellow, and Kunal Talwar. Semi-supervised knowledge transfer for deep learning from private training data. In *International Conference on Learning Representations (ICLR)*, 2017.

Nicolas Papernot, Shuang Song, Ilya Mironov, Ananth Raghunathan, Kunal Talwar, and Úlfar Erlingsson. Scalable private learning with PATE. In *International Conference on Learning Representations (ICLR)*, 2018.

András Prékopa. On logarithmic concave measures and functions. *Acta Scientiarum Mathematicarum (Szeged)*, 34:335–343, 1973.

Gang Qiao, Weijie Su, and Li Zhang. Oneshot differentially private top-k selection. In *Proceedings of the 38th International Conference on Machine Learning (ICML)*, volume 139 of *PMLR*, pp. 8672–8681, 2021.

R. Tyrrell Rockafellar. *Convex Analysis*. Princeton University Press, 1970.

Adam Sealfon. Shortest paths and distances with differential privacy. In *ACM Symposium on Principles of Database Systems (PODS)*, pp. 29–41, 2016.

Michael Shekelyan and Grigorios Loukides. Differentially private top-k selection via canonical Lipschitz mechanism, 2022. arXiv:2201.13376.

Thomas Steinke and Jonathan Ullman. Tight lower bounds for differentially private selection. In *IEEE Symposium on Foundations of Computer Science (FOCS)*, pp. 552–563, 2017.

Vishnu Vinod, Krishna Pillutla, and Abhradeep Guha Thakurta. InvisibleInk: high-utility and low-cost text generation with differential privacy. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2025.

Xuandong Zhao, Lei Li, and Yu-Xiang Wang. Permute-and-Flip: an optimally stable and watermarkable decoder for LLMs. In *International Conference on Learning Representations (ICLR)*, 2025.

A Convex Analysis and the Privacy Body

Notation. Notation is standard (Rockafellar, 1970). For a nonempty set $K \subseteq \mathbb{R}^m$, $h_K(y) = \sup_{x\in K}\langle x, y\rangle$ is its support function and $K^\circ = \{y : \langle x, y\rangle \le 1 \text{ for all } x \in K\}$ its polar; for a convex K containing 0, $\gamma_K(y) = \inf\{t > 0 : y \in tK\}$ is its gauge, and $\rho_K(y) = \sup\{t \ge 0 : ty \in K\}$ its radial function. For a closed convex K with $0 \in \operatorname{int} K$ the gauge is the support function of the polar, $\gamma_K = h_{K^\circ}$, and $h_K = \gamma_{K^\circ}$. For a cone C the polar becomes $C^\circ = \{y : \langle x, y\rangle \le 0 \text{ for all } x \in C\}$, and $C^{\circ\circ} = C$ when C is closed and convex. For a nonempty closed convex K, $\operatorname{rec} K = \{y : K + y \subseteq K\}$ is its recession cone. For a vertex v of Q, $T_v = \operatorname{cone}(Q - v)$ is the tangent cone and $N_v = \{y : \langle y, x - v\rangle \le 0 \text{ for all } x \in Q\}$ the normal cone of Q at v. For a closed, proper, convex $\phi\colon \mathbb{R}^m \to (-\infty, +\infty]$ we write $\phi^*(y) = \sup_x\big(\langle x, y\rangle - \phi(x)\big)$ for the conjugate, $\operatorname{dom}\phi = \{\phi < +\infty\}$ for the effective domain, and ϕ^∞ for the recession function, $\phi^\infty(y) = \lim_{t\to\infty}\big(\phi(z_0 + ty) - \phi(z_0)\big)/t$ with $z_0 \in \operatorname{dom}\phi$ arbitrary. Throughout the appendix the noise Z of Definition 1 has the density $e^{-\phi}/\int e^{-\phi}$ for a closed, proper, convex ϕ with $0 < \int e^{-\phi} < \infty$, and $\Lambda_Z(\theta) = \log \mathbb{E}\, e^{\langle\theta, Z\rangle}$ is its cumulant generating function.

Beyond these definitions the paper uses two classical facts. For closed, proper, convex ϕ the recession function is the support function of the domain of the conjugate (Rockafellar, 1970, Theorem 13.3),

$$\phi^\infty \;=\; h_{\operatorname{cl}\operatorname{dom}\phi^*}, \tag{A.1}$$

and integrating out variables preserves log-concavity (Prékopa, 1973), so the orthogonal projection of a random vector with a log-concave density has a log-concave density on its range.

The reduction. The mechanism depends only on score differences. Let $V = \operatorname{span}(Q - Q)$ and let P_V be the orthogonal projection onto V. For $x, x' \in Q$ and any y we have $\langle y, x - x'\rangle = \langle P_V y, x - x'\rangle$, so the output distribution of $M_Z(w)$ depends on the scores and the noise only through $P_V w$ and $P_V Z$, and a shift d acts only through $P_V d$. The allowed shifts become $P_V A$, which is bounded because h_A is finite on $Q - Q$ and $Q - Q$ spans V, and for $z \in V$, $h_{P_V A}(z) = \sup_{d\in A}\langle P_V d, z\rangle = \sup_{d\in A}\langle d, z\rangle = h_A(z)$. Since P_V is self-adjoint, $\Lambda_{P_V Z} = \Lambda_Z$ on V, and $P_V Z$ has a log-concave density with respect to Lebesgue measure on V (Prékopa, 1973). Its privacy body, taken within V, is $-\operatorname{cl}_V(\operatorname{dom}\Lambda_Z \cap V) = B(Z) \cap V$: a point y of $\operatorname{cl}\operatorname{dom}\Lambda_Z \cap V$ is the limit, as $s \downarrow 0$, of the points $(1 - s)y$, which lie in $\operatorname{dom}\Lambda_Z \cap V$ because the origin is interior to $\operatorname{dom}\Lambda_Z$ (Lemma A.1). Translating a vertex of Q to the origin places Q inside V and merely relabels the outputs. Since every tangent cone lies in V, reading an instance within V leaves every set $T_v \cap B(Z)$, and with it the value of Theorem 1, unchanged.

The privacy body. Its basic properties are those of the domain of the conjugate ϕ^*.

Lemma A.1. *The domains satisfy*

$$\operatorname{int}\operatorname{dom}\phi^* \;\subseteq\; \operatorname{dom}\Lambda_Z \;\subseteq\; \operatorname{dom}\phi^*, \qquad 0 \in \operatorname{int}\operatorname{dom}\phi^*.$$

Consequently $B(Z) = -\operatorname{cl}\operatorname{dom}\phi^*$ is a closed convex set with $0 \in \operatorname{int} B(Z)$, the potential grows as $\phi^\infty = h_{-B(Z)}$, and for all $z, u \in \mathbb{R}^m$

$$\phi(z + u) \;\le\; \phi(z) + \phi^\infty(u). \tag{A.2}$$

Proof. With the sign reversed, the body records the exponential tilts the noise sustains: $\theta \in \operatorname{dom}\Lambda_Z$ says that the tilted density $e^{\langle\theta,z\rangle - \phi(z)}$ is integrable, and $\theta \in \operatorname{dom}\phi^*$ that it is bounded. For $\theta \in \operatorname{int}\operatorname{dom}\phi^*$, choose $\delta > 0$ so that $\|\eta - \theta\| \le \delta$ implies $\eta \in \operatorname{int}\operatorname{dom}\phi^*$, and let b be the maximum on that ball of ϕ^*, which is convex and hence continuous on $\operatorname{int}\operatorname{dom}\phi^*$; for $z \ne 0$ and $\eta = \theta + \delta z/\|z\|$,

$$\langle\theta, z\rangle - \phi(z) = \langle\eta, z\rangle - \phi(z) - \delta\|z\| \le b - \delta\|z\|,$$

so $\mathbb{E}\, e^{\langle\theta,Z\rangle} < \infty$, which is the first inclusion. For the second, an integrable $q(z) = e^{\langle\theta,z\rangle - \phi(z)}$ is bounded: fix a closed ball U in $\operatorname{int}\operatorname{dom}\phi$, nonempty because $\operatorname{dom}\phi$ is convex with positive measure, on which $q \ge a > 0$; log-concavity of q gives $q\big(\frac{z+y}{2}\big) \ge \sqrt{a\, q(z)}$ for $y \in U$, so integrating over the ball $\frac{z+U}{2}$, whose volume does not depend on z, bounds q uniformly, and $\phi^*(\theta) = \log\sup q < \infty$. Finally, were the origin not interior to $\operatorname{dom}\phi^*$, a hyperplane through the origin would support that set: some $u \ne 0$ would have $\langle\theta, u\rangle \le 0$ for all $\theta \in \operatorname{dom}\phi^*$, so $\phi^\infty(u) \le 0$

by (A.1) and ϕ would be nonincreasing along u. Then $e^{-\phi}$ would be at least as large on every translate $U + tu$, $t \geq 0$, as on U, and infinitely many of those translates are disjoint, contradicting $\int e^{-\phi} < \infty$.

A convex set with nonempty interior has the same closure as its interior, so the two inclusions give $\operatorname{cl}\operatorname{dom}\Lambda_Z = \operatorname{cl}\operatorname{dom}\phi^*$, and $B(Z) = -\operatorname{cl}\operatorname{dom}\phi^*$ contains the origin in its interior. By (A.1), $\phi^\infty = h_{\operatorname{cl}\operatorname{dom}\phi^*} = h_{-B(Z)}$. The inequality (A.2) holds because the difference quotient $t \mapsto (\phi(z+tu) - \phi(z))/t$ is nondecreasing with limit $\phi^\infty(u)$. □

For independent coordinates the body is a box of tail decay rates, and the four noise distributions of Section 3 give the following bodies.

Lemma A.2. *If Z has independent coordinates and $\operatorname{dom}\Lambda_{Z_i}$ has the endpoints $-\alpha_i$ and β_i, with $\alpha_i, \beta_i \in (0, +\infty]$, then $B(Z) = \prod_i [-\beta_i, \alpha_i]$, where an infinite endpoint is omitted. For a coordinate of rate λ, Laplace noise has $[-\lambda, \lambda]$, exponential noise and Gumbel noise have $[-\lambda, +\infty)$, and Gaussian noise has $\mathbb{R}$.*

Proof. Independence gives $\Lambda_Z(\theta) = \sum_i \Lambda_{Z_i}(\theta_i)$, so $\operatorname{dom}\Lambda_Z = \prod_i \operatorname{dom}\Lambda_{Z_i}$, and closing and negating the product of intervals gives the box; the endpoints are positive because $0 \in \operatorname{int} B(Z)$ by Lemma A.1. For the examples, the Laplace density $\frac{\lambda}{2}e^{-\lambda|z|}$ has $\mathbb{E}\, e^{\theta Z_i} = \lambda^2/(\lambda^2 - \theta^2)$ for $|\theta| < \lambda$ and $+\infty$ otherwise; the exponential density $\lambda e^{-\lambda z}\mathbb{1}\{z \geq 0\}$ has $\mathbb{E}\, e^{\theta Z_i} = \lambda/(\lambda - \theta)$ for $\theta < \lambda$ and $+\infty$ otherwise; the Gumbel distribution of scale $1/\lambda$ has $\mathbb{E}\, e^{\theta Z_i} = \Gamma(1 - \theta/\lambda)$ for $\theta < \lambda$ and $+\infty$ otherwise; and a Gaussian has a finite moment generating function everywhere. □

The endpoint β_i is the exponential rate at which the right tail of Z_i decays and α_i the rate of the left tail, an infinite rate meaning a tail lighter than every exponential.

B Proofs for Section 3

B.1 The Cone Identity and Exact Privacy

Lemma 1 (Cone identity)**.** *Let $C \subseteq \mathbb{R}^m$ be a closed convex cone with nonempty interior, let $d \in \mathbb{R}^m$, and put $p_C(w) := \Pr[w + Z \in C]$. Then*

$$\sup_{w\,:\,p_C(w)>0} \log \frac{p_C(w)}{p_C(w+d)} \;=\; h_{B(Z)\cap C^\circ}(d),$$

a value in $[0, +\infty]$.

Proof. The upper bound compares the two cone probabilities after a translation inside the cone, the lower bound tilts the density and reads off a slope along a ray, and a conjugacy computation identifies both answers with $h_{B(Z)\cap C^\circ}(d)$; the proof shows moreover that the common value equals $\inf_{a\in C} \phi^\infty(a - d)$. Write $B = B(Z)$.

We first locate where the ratio is defined. We have $p_C(w) = \int_C e^{-\phi(y-w)}\,dy \big/ \int e^{-\phi}$, and the integrand is positive precisely on the convex set $C \cap (w + \operatorname{dom}\phi)$, which has positive measure exactly when its interior $\operatorname{int} C \cap (w + \operatorname{int}\operatorname{dom}\phi)$ is nonempty. Hence $p_C(w) > 0$ if and only if w lies in the nonempty open convex set $W := \operatorname{int} C - \operatorname{int}\operatorname{dom}\phi$. The function $(w, y) \mapsto e^{-\phi(y-w)}\mathbb{1}\{y \in C\}$ is log-concave, so integrating out y leaves p_C log-concave (Prékopa, 1973), and $g := -\log p_C$ is finite and convex on W.

For the upper bound, let $a \in C$, so that $C + a \subseteq C$ because C is a convex cone. Restricting the integral to $C + a$, substituting $y \mapsto y + a$, and applying (A.2) with $u = a - d$,

$$p_C(w + d) \;\geq\; \frac{\int_C e^{-\phi((y-w)+(a-d))}\,dy}{\int e^{-\phi}} \;\geq\; e^{-\phi^\infty(a-d)}\,p_C(w) \qquad \text{for every } w.$$

Taking the infimum over a bounds the supremum of the lemma by $\inf_{a\in C}\phi^\infty(a - d)$. The convention $\log(p/0) = +\infty$ is covered: if $p_C(w) > 0 = p_C(w + d)$ for some w, the display forces $\phi^\infty(a - d) = +\infty$ for every a, so both sides are $+\infty$.

The infimum is a support function. By Lemma A.1, $\phi^\infty = h_{-B}$, and $h_{-B}(a-d) = h_B(d-a)$, so the infimum is $\Psi(d) \coloneqq \inf_{a\in C} h_B(d-a)$, the infimal convolution of h_B with the indicator function of C, a convex nonnegative function with $\Psi(0) = 0$. Its sublevel sets are closed: $0 \in \operatorname{int} B$ makes h_B coercive, so every finite infimum is attained and $\{\Psi \le \kappa\} = C + \{h_B \le \kappa\}$ is the sum of a closed and a compact set. The substitution $d = a + y$ computes the conjugate,

$$\Psi^*(\theta) = \sup_{a\in C} \langle \theta, a\rangle \;+\; \sup_{y} \big(\langle \theta, y\rangle - h_B(y)\big).$$

The first supremum is 0 for $\theta \in C^\circ$ and, C being a cone, $+\infty$ otherwise; the second is 0 for $\theta \in B$, where h_B dominates $\langle \theta, \cdot\rangle$, and $+\infty$ otherwise, by scaling a separating functional. So Ψ^* is the indicator function of $B \cap C^\circ$, and biconjugation (Rockafellar, 1970, Theorem 12.2) gives $\Psi = \Psi^{**} = h_{B\cap C^\circ}$.

For the lower bound, if some $w \in W$ has $w+d \notin W$, the supremum is $+\infty$. Otherwise $W+d \subseteq W$, so $\bar w + rd \in W$ for every $\bar w \in W$ and every integer $r \ge 0$, and convexity puts the whole ray $\{\bar w + td : t \ge 0\}$ in W. Fix such a $\bar w$. The function $t \mapsto g(\bar w + td)$ is finite and convex on $[0,\infty)$, so its unit increments $g(\bar w + (r+1)d) - g(\bar w + rd)$, each a value of $\log \frac{p_C(w)}{p_C(w+d)}$ at some $w \in W$ and hence at most the supremum, are nondecreasing with limit $\lim_{t\to\infty} g(\bar w + td)/t$; it suffices to bound this limit below by $h_{B\cap C^\circ}(d)$. Let $\theta \in \operatorname{dom}\Lambda_Z$ with $-\theta \in C^\circ$. On the event $w + Z \in C$ we have $\langle \theta, w + Z\rangle \ge 0$, so Markov's inequality gives

$$p_C(w) \;\le\; \mathbb{E}\, e^{\langle \theta, w+Z\rangle} \;=\; e^{\langle \theta, w\rangle + \Lambda_Z(\theta)},$$

whence $g(\bar w+td) \ge -\langle \theta, \bar w+td\rangle - \Lambda_Z(\theta)$ and $\lim_{t\to\infty} g(\bar w+td)/t \ge \langle -\theta, d\rangle$. For $\eta \in B\cap C^\circ$ and $0 < s < 1$ the choice $\theta = -s\eta$ qualifies: the segment from $0 \in \operatorname{int}\operatorname{dom}\Lambda_Z$ to $-\eta \in \operatorname{cl}\operatorname{dom}\Lambda_Z$ lies in $\operatorname{dom}\Lambda_Z$ short of its endpoint, and C° is a cone. Letting $s \uparrow 1$ gives $\lim_{t\to\infty} g(\bar w+td)/t \ge \langle \eta, d\rangle$ for every $\eta \in B \cap C^\circ$, that is, at least $h_{B\cap C^\circ}(d)$. Finally $0 \in B \cap C^\circ$ makes the common value nonnegative. □

Theorem 1 (Exact privacy). *Every mechanism of Definition 1 satisfies*

$$\varepsilon_{\text{exact}}(M_Z) \;=\; \max_{v\in \operatorname{vert} Q} \;\sup_{z\in T_v\cap B(Z)} h_A(z).$$

Proof. Event losses reduce to vertex losses, each winning event is a cone event, Lemma 1 evaluates each vertex under each shift, and two suprema are exchanged.

The event $\{M_Z(w) \in S\}$ is the disjoint union of the events $\{M_Z(w) = v\}$, $v \in S$, and $\sum_v p_v \big/ \sum_v p'_v \le \max_{v:\, p_v>0} p_v/p'_v$ for nonnegative numbers with $\sum_v p_v > 0$. Singletons are themselves events, so the supremum in Definition 2 is unchanged when S ranges over single vertices. Up to the null tie event, $\{M_Z(w) = v\} = \{w + Z \in N_v\}$, since a linear functional is maximized over $\operatorname{vert} Q$ at v exactly when it is maximized over Q at v. The normal cone N_v is closed, convex and full-dimensional: strictly separating v from the convex hull of the remaining vertices produces y with $\langle y, x - v\rangle < 0$ for every other vertex x, and the finitely many strict inequalities survive small perturbations of y. A functional nonpositive on $Q - v$ is nonpositive on the cone it generates, so $N_v = T_v^\circ$, and T_v is finitely generated, hence closed, so $N_v^\circ = T_v$. Lemma 1 applied to $C = N_v$ gives the extreme loss of the vertex v under the shift d,

$$\sup_{w\,:\,\Pr[M_Z(w)=v]>0} \log \frac{\Pr[M_Z(w)=v]}{\Pr[M_Z(w+d)=v]} \;=\; h_{B(Z)\cap T_v}(d). \tag{B.1}$$

By Definition 2, $\varepsilon_{\text{exact}}(M_Z)$ is the supremum of (B.1) over the vertices v and the shifts $d \in A$, and for each vertex

$$\sup_{d\in A} h_{B(Z)\cap T_v}(d) = \sup_{d\in A}\;\sup_{z\in T_v\cap B(Z)} \langle d, z\rangle = \sup_{z\in T_v\cap B(Z)} h_A(z). \qquad \square$$

Under score-range adjacency $A = \Delta D^\circ$, and $h_{\Delta D^\circ} = \Delta h_{D^\circ} = \Delta\gamma_D$ because D is closed and convex and contains the origin, which gives the form (2).

B.2 The Running Example, Gaussian Noise and the Values of Section 3

The three panels of Figure 1. On $Q_2 = \mathrm{conv}\{0, e_1, e_2\}$ the tangent cones are $T_0 = \mathbb{R}^2_+$ and $T_{e_1} = \{z : z_2 \geq 0,\ z_1 + z_2 \leq 0\}$, with T_{e_2} symmetric. On T_0 the gauge is $\gamma_D(z) = z_1 + z_2$: the point $z/(z_1 + z_2)$ lies on the segment from e_1 to e_2 inside D, and no smaller multiple of D contains z because $h_D(e_1 + e_2) = 1$. On T_{e_1} the gauge is $\gamma_D(z) = -z_1$, by the same two steps with $z = (-z_1)(x - e_1)$ for $x = (0,\ z_2/(-z_1)) \in Q_2$, and $h_D(-e_1) = 1$. For iid Laplace noise of rate λ the body is $[-\lambda, \lambda]^2$ (Lemma A.2), and the suprema of the gauge over $T_0 \cap B$, $T_{e_1} \cap B$ and $T_{e_2} \cap B$ are 2λ, λ and λ, attained at the corner (λ, λ) and at $-\lambda e_1$ and $-\lambda e_2$, so $\varepsilon_{\text{exact}} = 2\lambda\Delta$. For iid exponential noise the body is $[-\lambda, +\infty)^2$, which contains all of T_0, so $\varepsilon_{\text{exact}} = +\infty$. For $Z_i = G_i - G_0$ with iid Gumbel G_0, G_1, G_2 of rate λ the body is $\{z : z_1, z_2 \geq -\lambda,\ z_1 + z_2 \leq \lambda\}$ (Proposition B.5), on which the gauge is at most λ in each of the three cones, with equality along the edge $z_1 + z_2 = \lambda$ of T_0, so $\varepsilon_{\text{exact}} = \lambda\Delta$. In the candidate cones the three bodies have the same intersection, the triangle with vertices 0, $-\lambda e_1$ and $-\lambda e_1 + \lambda e_2$ and its mirror image.

Gaussian noise. A noise distribution whose cumulant generating function is finite everywhere has no exact privacy at all, once a single allowed shift moves the comparison between two outputs.

Proposition B.1. *If Λ_Z is finite on all of $\mathbb{R}^m$, as it is for Gaussian Z, then $B(Z) = \mathbb{R}^m$, and $\varepsilon_{\text{exact}}(M_Z) = +\infty$ as soon as some $d \in A$ and some vertices $x \neq v$ of Q have $\langle d, x - v\rangle \neq 0$.*

Proof. $B(Z) = -\operatorname{cl}\mathbb{R}^m = \mathbb{R}^m$. The set $T_v \cap B(Z) = T_v$ contains the ray $\mathbb{R}_+(x - v)$, and since $A = -A$, $h_A(t(x - v)) = t\, h_A(x - v) \geq t|\langle d, x - v\rangle| \to \infty$; apply Theorem 1. □

Corollary 1 (The exponential mechanism and permute-and-flip)**.** *Noise vectors with the same privacy body have the same exact privacy on every polytope under every adjacency. Iid exponential noise and iid Gumbel noise of rate λ both have the body $[-\lambda, +\infty)^m$, so the exponential mechanism and permute-and-flip have the same exact privacy, and so do their versions on every other polytope.*

Proof. By Theorem 1 the exact privacy depends on the noise only through its body, so equal bodies give equal exact privacy on every polytope under every adjacency; by Lemma 1 the same holds for every readout that reports which of finitely many closed convex cones contains the noisy scores. By Lemma A.2, iid exponential noise and iid Gumbel noise of rate λ both have the body $[-\lambda, +\infty)^m$. On the simplex, iid Gumbel noise makes M_Z sample from the softmax distribution, by the Gumbel-max identity of random-utility theory (McFadden, 1974), so that M_Z is the exponential mechanism (McSherry & Talwar, 2007), and iid exponential noise makes M_Z report noisy max with exponential noise, which is identical in distribution to permute-and-flip (Ding et al., 2021). □

On the simplex $Q = \mathrm{conv}\{e_1, \ldots, e_n\}$, read within $V = \{z : \sum_i z_i = 0\}$, the common body is $\{z \in V : z_i \geq -\lambda \text{ for all } i\}$, the tangent cone at e_j is $T_{e_j} = \{z \in V : z_i \geq 0 \text{ for } i \neq j\}$, and the gauge on it is $\gamma_D(z) = -z_j$: for $z \neq 0$ the point $x = z/(-z_j) + e_j$ lies in Q, so $z \in (-z_j)D$, and $h_D(-e_j) = 1$ shows that no smaller multiple works. The supremum of $-z_j$ over $T_{e_j} \cap B$ is λ, attained at $-\lambda e_j + \frac{\lambda}{n-1}\sum_{i \neq j} e_i$, so under score-range adjacency both mechanisms have $\varepsilon_{\text{exact}} = \lambda\Delta$ exactly. The values under per-score adjacency follow.

Proposition B.2. *Let the coordinates of the noise be iid, Laplace, exponential or Gumbel of rate λ, and let adjacency be per-score with sensitivity Δ_∞.*

(a) *On the k-subset polytope, $1 \leq k \leq n - 1$, $\varepsilon_{\text{exact}} = 2k\lambda\Delta_\infty$ for exponential and Gumbel noise and $\varepsilon_{\text{exact}} = 2\min\{k, n - k\}\lambda\Delta_\infty$ for Laplace noise; under monotone per-score adjacency both values are halved. In particular a single winner ($k = 1$) has $\varepsilon_{\text{exact}} = 2\lambda\Delta_\infty$ for all three.*

(b) *On Q_n, $n \geq 2$, $\varepsilon_{\text{exact}} = n\lambda\Delta_\infty$ for Laplace noise and $\varepsilon_{\text{exact}} = +\infty$ for exponential and Gumbel noise. For the noise $Z_i = X_i - X_0$ instead, with $X_0, \ldots, X_n$ iid of one of the three distributions, $\varepsilon_{\text{exact}} = 2\lambda\Delta_\infty$.*

Proof. Per-score adjacency has $A = \Delta_\infty[-1, 1]^m$ and $h_A(z) = \Delta_\infty\|z\|_1$; monotone per-score adjacency has $h_A(z) = \Delta_\infty \max\{\sum_i z_i^+, \sum_i z_i^-\}$.

(a) The vertices are $e_J = \sum_{i \in J} e_i$ for the k-subsets J of $[n]$, the polytope lies in the hyperplane $\sum_i x_i = k$, and by the reduction of Appendix A the body is read within $V = \{z : \sum_i z_i = 0\}$:

$\{z \in V : z_i \geq -\lambda\}$ for exponential and Gumbel noise and $\{z \in V : |z_i| \leq \lambda\}$ for Laplace noise, since in both cases $\mathrm{cl}_V(\mathrm{dom}\,\Lambda_Z \cap V) = \mathrm{cl}\,\mathrm{dom}\,\Lambda_Z \cap V$. The tangent cone at e_J is $T_J = \{z \in V : z_i \leq 0\ (i \in J),\ z_i \geq 0\ (i \notin J)\}$: it is generated by the differences $e_j - e_i$ with $i \in J$ and $j \notin J$, and every z in the displayed set is a nonnegative combination of them, with the weight $(-z_i)z_j / \sum_{l \notin J} z_l$ on $e_j - e_i$ when $z \neq 0$. On T_J the sum of the coordinates vanishes, so

$$\|z\|_1 = 2\sum_{i \in J}(-z_i) = 2\sum_{j \notin J} z_j, \qquad \max\Big\{\sum_i z_i^+, \sum_i z_i^-\Big\} = \sum_{i \in J}(-z_i).$$

For exponential and Gumbel noise $\sum_{i \in J}(-z_i) \leq k\lambda$ on the body, with equality at the point with $z_i = -\lambda$ on J and $z_j = k\lambda/(n-k)$ off J. For Laplace noise the two sums are also bounded by $(n-k)\lambda$, so $\sum_{i \in J}(-z_i) \leq \min\{k, n-k\}\lambda$, with equality at a point with $\min\{k, n-k\}$ coordinates $-\lambda$ in J and as many coordinates λ off J. Theorem 1 gives the values, the same at every vertex.

(b) The tangent cones of Q_n are $T_0 = \mathbb{R}^n_+$ and $T_{e_j} = \{z : z_i \geq 0\ (i \neq j),\ \sum_i z_i \leq 0\}$ (Proposition C.6), and on them $\|z\|_1 = \sum_i z_i$ and $\|z\|_1 = \sum_i z_i - 2z_j$. For iid Laplace noise the body is $[-\lambda, \lambda]^n$, the supremum over T_0 is $n\lambda$, and over T_{e_j} it is 2λ, since there $\sum_{i \neq j} z_i \leq -z_j \leq \lambda$; so $\varepsilon_{\mathrm{exact}} = n\lambda\Delta_\infty$ for $n \geq 2$. For exponential and Gumbel noise the body $[-\lambda, +\infty)^n$ contains T_0, so $\varepsilon_{\mathrm{exact}} = +\infty$. For $Z_i = X_i - X_0$ the body is $\{z : z_i \geq -\lambda,\ \sum_i z_i \leq \lambda\}$ for exponential and Gumbel noise and $\{z : |z_i| \leq \lambda,\ |\sum_i z_i| \leq \lambda\}$ for Laplace noise (Proposition B.5). In both cases the supremum over T_0 is λ, and over T_{e_j} it is 2λ: there $\sum_i z_i - 2z_j \leq 0 + 2\lambda$, with equality when $z_j = -\lambda$ and the other coordinates are nonnegative with sum λ. So $\varepsilon_{\mathrm{exact}} = 2\lambda\Delta_\infty$. □

B.3 Completeness

Theorem 1 shows that the exact privacy depends on the noise only through its body. No finer invariant of the noise enters the extreme losses of the individual vertices either. For a polytope Q with at least two vertices, a vertex v and a shift $d \in \mathbb{R}^m$, write $\ell_Z(Q, v, d)$ for the left side of (B.1).

Theorem B.3 (Completeness)**.** *Noise vectors Z and Z' have $\ell_Z(Q, v, d) = \ell_{Z'}(Q, v, d)$ for every polytope Q, vertex v and shift d if and only if $B(Z) = B(Z')$. Two candidates suffice: if the two values agree whenever $Q = \mathrm{conv}\{0, u\}$ with $u \neq 0$ and $v = u$, then $B(Z) = B(Z')$.*

Proof. Equal bodies give equal values, and segments recover the body. For any instance and any vertex, (B.1) gives $\ell_Z(Q, v, d) = h_{B(Z) \cap T_v}(d)$, a function of $B(Z)$ alone.

For the converse, fix $u \neq 0$ and consider $Q = \mathrm{conv}\{0, u\}$ with the event that u wins. Up to ties, this event is $\{\langle u, w + Z\rangle \geq 0\} = \{w + Z \in C\}$ for the halfspace $C = \{y : \langle u, y\rangle \geq 0\}$, whose polar cone is $C^\circ = -\mathbb{R}_+ u$: a functional nonpositive on C vanishes on $u^\perp$ and pairs nonpositively with u. Write $B = B(Z)$; its radial function is positive in every direction because $0 \in \mathrm{int}\, B$. Lemma 1 gives, for every $s > 0$,

$$\ell_Z\big(\mathrm{conv}\{0, u\},\, u,\, -su\big) \;=\; h_{B \cap (-\mathbb{R}_+ u)}(-su) \;=\; \sup_{0 \leq t \leq \rho_B(-u)} t\, s\|u\|^2 \;=\; s\|u\|^2 \rho_B(-u).$$

These values therefore determine the radial function of B in every direction. Since each radial section of the closed set B is closed, $z \in B$ exactly when $z = 0$ or $\rho_B(z) \geq 1$, so the radial function determines B. Hence $B(Z) = B(Z')$. □

The aggregate $\varepsilon_{\mathrm{exact}}$ does not separate bodies in this way.

Remark B.4. At $m = 1$, the asymmetric Laplace noise distributions with densities proportional to $\min\{e^{2z}, e^{-z}\}$ and $\min\{e^{2z}, e^{-z/2}\}$ have the privacy bodies $[-1, 2]$ and $[-\frac{1}{2}, 2]$. By Theorem 1 they have the same exact privacy on every one-dimensional instance under every adjacency; on $Q = [0, 1]$ under score-range adjacency the common value is 2Δ, the shared left-tail rate dominating both right tails. Their extreme losses differ already there: $\ell_Z(Q, 1, -s) = s\,\rho_{B(Z)}(-1)$ equals s for the first and $s/2$ for the second.

B.4 One-Sided Noise and Noisy Thresholds

Theorem 2 (One-sided noise)**.** *Let the set A of allowed shifts detect every direction: every $z \neq 0$ has $\langle d, z\rangle \neq 0$ for some $d \in A$, as under per-score, monotone per-score and score-range adjacency.*

(a) $\varepsilon_{\text{exact}}(M_Z)$ *is finite if and only if* $T_v \cap \operatorname{rec} B(Z) = \{0\}$ *for every vertex* v *of* Q.
(b) *Let* Z *be one-sided, with right-tail rates* $0 < \beta_i < +\infty$. *Then* $\varepsilon_{\text{exact}}(M_Z)$ *is finite if and only if every vertex of* Q *is Pareto-maximal: no* $x \in Q$ *has* $x \geq v$ *coordinatewise with* $x \neq v$.

Proof. (a) Since $A = -A$, $h_A(z) = \sup_{d \in A} |\langle d, z \rangle|$, which is positive for $z \neq 0$ because A detects every direction, symmetric, subadditive and positively homogeneous; it is finite on $Q - Q$, hence on the space V that $Q - Q$ spans. So h_A is a norm on V, bounded above and below there by positive multiples of the Euclidean norm, and since every T_v lies in V, Theorem 1 shows that the exact privacy is finite exactly when every set $T_v \cap B$ is bounded, where $B = B(Z)$.

If some $z \neq 0$ lies in $T_v \cap \operatorname{rec} B$, then $rz \in B$ for every integer $r \geq 0$, iterating $B + z \subseteq B$ from $0 \in B$, and $rz \in T_v$ because T_v is a cone: the intersection is unbounded. Conversely, let $T_v \cap B$ be unbounded and choose z_r in it with $\|z_r\| \to \infty$ and, after passing to a subsequence, $z_r/\|z_r\| \to u$, a unit vector of the closed cone T_v. For fixed $t \geq 0$ the points $(t/\|z_r\|)z_r$ lie on the segments from 0 to z_r, hence in B, and converge to tu, so $\mathbb{R}_+ u \subseteq B$; and then $y + u = \lim_{r \to \infty} \big((1 - \frac{1}{r})y + \frac{1}{r}(ru)\big) \in B$ for every $y \in B$. Thus $u \in T_v \cap \operatorname{rec} B$.

(b) By Lemma A.2, $B(Z) = \prod_i [-\beta_i, +\infty)$, so $\operatorname{rec} B(Z) = \mathbb{R}^m_+$. Every element of $T_v = \mathbb{R}_+(Q - v)$ has the form $t(x - v)$ with $t \geq 0$ and $x \in Q$, so a nonzero element of $T_v \cap \mathbb{R}^m_+$ exists exactly when some $x \in Q$ dominates v coordinatewise without equaling it. By (a), that is the failure of finiteness at v. □

Noise on the threshold, of the same distribution as the noise on the scores, gives one-sided noise a finite body on Q_n and recovers two mechanisms of the literature.

Proposition B.5. *Let* $X_0, X_1, \ldots, X_n$ *be iid with a log-concave density, and let* $Z_i = X_i - X_0$ *for* $i \in [n]$. *Then* Z *has a log-concave density. If the* X_i *are Gumbel or exponential of rate* λ, *then*

$$B(Z) = \Big\{ z \in \mathbb{R}^n : z_i \geq -\lambda \text{ for all } i,\ \textstyle\sum_i z_i \leq \lambda \Big\},$$

and on Q_n *the mechanism* M_Z *is the exponential mechanism, respectively permute-and-flip, over the* $n + 1$ *options with the abstain option scored* 0. *If the* X_i *are Laplace of rate* λ, *then* $B(Z) = \{z : |z_i| \leq \lambda \text{ for all } i,\ |\sum_i z_i| \leq \lambda\}$.

Proof. Writing f for the density of X_0, the pair (X_0, Z) has the density $f(x_0) \prod_i f(x_0 + z_i)$, log-concave in (x_0, z), so integrating out x_0 leaves Z with a log-concave density (Prékopa, 1973). Since $\langle \theta, Z \rangle = \sum_i \theta_i X_i - (\sum_i \theta_i) X_0$ with independent terms,

$$\Lambda_Z(\theta) = \sum_i \Lambda_X(\theta_i) + \Lambda_X\Big(-\textstyle\sum_i \theta_i\Big),$$

where Λ_X is the cumulant generating function of one X_i. For Gumbel and exponential noise $\Lambda_X(t)$ is finite exactly for $t < \lambda$ (Lemma A.2), so $\operatorname{dom} \Lambda_Z = \{\theta : \theta_i < \lambda,\ \sum_i \theta_i > -\lambda\}$; this set is open, convex and nonempty, so its closure is obtained by relaxing the strict inequalities, and negating it gives the displayed body. For Laplace noise $\Lambda_X(t)$ is finite exactly for $|t| < \lambda$, and the same steps give the second body. On Q_n the noisy values are 0 at the abstain vertex and $w_i + X_i - X_0$ at e_i; adding X_0 to all of them does not change the maximizer, which becomes the maximizer of $w_i + X_i$ over $i \in \{0, 1, \ldots, n\}$ with $w_0 = 0$. With Gumbel noise this maximizer is i with probability proportional to $e^{\lambda w_i}$ (McFadden, 1974), the exponential mechanism (McSherry & Talwar, 2007); with exponential noise it is report noisy max with exponential noise, which is permute-and-flip (Ding et al., 2021). □

B.5 AboveThreshold

We write AboveThreshold on n queries with answers $q \in \mathbb{R}^n$ and threshold $b \in \mathbb{R}$ as follows: independent noise Z_0 of rate μ and $Z_1, \ldots, Z_n$ of rate λ is drawn, and the output is the first $j \in [n]$ with $q_j + Z_j \geq b + Z_0$, or $\perp$ if there is none. Index the coordinates of $\mathbb{R}^{n+1}$ by $0, 1, \ldots, n$ and put $y = w + Z$ with $w = (b, q)$. Up to null events of ties, the output is j exactly when y lies in

$$C_j = \{y : y_i \leq y_0 \text{ for } 1 \leq i < j,\ y_j \geq y_0\},$$

and $\perp$ exactly when y lies in $C_\perp = \{y : y_i \le y_0 \text{ for all } i \in [n]\}$. These are closed convex cones with nonempty interior. For $n \ge 2$ they are not the normal cones of the vertices of any polytope: C_1 is a halfspace, and a normal cone that is a halfspace makes the tangent cone at that vertex a ray, which forces the polytope to be a segment with two vertices, while there are $n+1 \ge 3$ outputs.

Corollary 2 (AboveThreshold)**.** *Let AboveThreshold run on $n \ge 2$ queries with independent noise of rate λ on each answer and of rate μ on the threshold, under per-query adjacency. If the noise on the answers and on the threshold is Laplace, $\varepsilon_{\text{exact}} = \Delta_\infty \min\{n\lambda,\, 2\lambda + \mu\}$; if it is exponential throughout or Gumbel throughout, $\varepsilon_{\text{exact}} = \Delta_\infty(2\lambda + \mu)$. Under monotone per-query adjacency, where all answers move in the same direction, the exact privacy is $\Delta_\infty(\lambda + \mu)$, for Laplace noise when $n \ge 2 + \mu/\lambda$ and for exponential and Gumbel noise for every $n \ge 2$.*

Proof. Adjacent inputs have the same threshold and answers that differ by some $\delta \in A$, so the shift is $d = (0, \delta)$. Every cell is invariant under adding a multiple of $\mathbf{1}$ to y, so $p_C(w + t\mathbf{1}) = p_C(w)$ and the supremum in Lemma 1 over all $w \in \mathbb{R}^{n+1}$ equals the supremum over the inputs with threshold b. As in the proof of Theorem 1, event losses reduce to single outputs, and Lemma 1 gives the extreme loss of the output o under the shift $(0, \delta)$ as $h_{B \cap C_o^\circ}((0, \delta))$, where $B = B(Z)$ is the box of Lemma A.2.

The polar of C_j is generated by $e_i - e_0$ for $1 \le i < j$ and by $e_0 - e_j$, so its elements are $\theta = \sum_{i<j} \omega_i(e_i - e_0) + \omega_j(e_0 - e_j)$ with $\omega \ge 0$, whose coordinates are $\theta_i = \omega_i$ for $1 \le i < j$, $\theta_j = -\omega_j$, $\theta_0 = \omega_j - \Omega$ with $\Omega = \sum_{i<j} \omega_i$, and 0 elsewhere; and $\langle \theta, (0, \delta) \rangle = \sum_{i<j} \omega_i \delta_i - \omega_j \delta_j$. Likewise $C_\perp^\circ$ consists of $\sum_{i \in [n]} \omega_i(e_i - e_0)$ with $\omega \ge 0$, whose threshold coordinate is $-\Omega$ with $\Omega = \sum_i \omega_i$. Under per-query adjacency the supremum over δ is $\Delta_\infty(\Omega + \omega_j)$ for C_j and $\Delta_\infty \Omega$ for $C_\perp$; under monotone per-query adjacency it is $\Delta_\infty \max\{\Omega, \omega_j\}$ for C_j and $\Delta_\infty \Omega$ for $C_\perp$.

Laplace noise has $B = [-\mu, \mu] \times [-\lambda, \lambda]^n$, so $\theta \in B$ asks $\omega_i \le \lambda$ for $i \le j$ and $|\omega_j - \Omega| \le \mu$. Then $\Omega \le (j-1)\lambda$ and $\Omega + \omega_j \le \min\{\Omega + \lambda,\, 2\omega_j + \mu\} \le \min\{j\lambda,\, 2\lambda + \mu\}$, with equality for $j \ge 2$ at $\omega_j = \lambda$ and $\Omega = \min\{(j-1)\lambda, \lambda + \mu\}$; for $j = 1$ the value is $\min\{\lambda, \mu\}$, and for $\perp$ it is $\min\{n\lambda, \mu\}$. The largest of these is $\min\{n\lambda, 2\lambda + \mu\}$. Under monotone per-query adjacency the largest Ω for C_j is $\min\{(j-1)\lambda, \lambda + \mu\}$ and the largest ω_j is λ for $j \ge 2$, so the value of C_n is $\lambda + \mu$ once $(n-1)\lambda \ge \lambda + \mu$, and no output exceeds it.

Exponential and Gumbel noise have $B = [-\mu, +\infty) \times [-\lambda, +\infty)^n$, so $\theta \in B$ asks only $\omega_j \le \lambda$ and $\Omega \le \omega_j + \mu$. Then $\Omega + \omega_j \le 2\omega_j + \mu \le 2\lambda + \mu$, with equality for $j \ge 2$ at $\omega_j = \lambda$ and $\omega_1 = \lambda + \mu$; the value is λ for $j = 1$ and μ for $\perp$. Under monotone per-query adjacency the largest Ω is $\lambda + \mu$ and the largest ω_j is λ, so the value is $\lambda + \mu$ for every $n \ge 2$.

Finally, the queries may be chosen adaptively. Until the algorithm halts every answer is $\perp$, so the queries it asks are determined by the number of answers already given, and for each output the probability is the one computed above for that fixed sequence of queries. □

By Corollary 2, the calibration of Dwork & Roth (2014, Theorem 3.23), $\lambda = \varepsilon/(4\Delta_\infty)$ and $\mu = \varepsilon/(2\Delta_\infty)$, has exact privacy $\Delta_\infty \min\{n\lambda, 2\lambda + \mu\} = \min\{n/4, 1\} \cdot \varepsilon$, exactly ε once $n \ge 4$. Durfee (2023) proves the bounds $\varepsilon_1 + 2\varepsilon_2$, and $\varepsilon_1 + \varepsilon_2$ for monotone queries, for threshold noise of rate $\varepsilon_1/\Delta_\infty$ and answer noise of rate $\varepsilon_2/\Delta_\infty$, Laplace, exponential or Gumbel, the Gumbel case at $\varepsilon_1 = \varepsilon_2$. By Corollary 2 these bounds are exact at all rates, for Laplace noise once $n \ge 2 + \varepsilon_1/\varepsilon_2$. The calibration of Lyu et al. (2017) halves the answer rate and has exact privacy $\varepsilon_1 + \varepsilon_2$ once $n \ge 2 + 2\varepsilon_1/\varepsilon_2$, and so does their version with c positive answers and one noisy threshold once $n \ge 2c + \mu/\lambda$ (Proposition B.6).

The same computation gives the exact privacy of the version of the sparse vector technique that continues after a positive answer with the same noisy threshold and halts after c positive answers or after n queries, the algorithm of Lyu et al. (2017, Algorithm 1). An output is now a pattern of answers; if J_+ and J_- are the positions of its positive and negative answers and the last answer is at position j, the output occurs exactly when y lies in the closed convex cone $\{y : y_i \ge y_0 \ (i \in J_+),\ y_i \le y_0 \ (i \in J_-)\}$.

Proposition B.6. *With Laplace noise of rate λ on the answers and μ on the threshold, the version with c positive answers has $\varepsilon_{\text{exact}} \le \Delta_\infty(2c\lambda + \mu)$ under per-query adjacency and $\varepsilon_{\text{exact}} \le \Delta_\infty(c\lambda + \mu)$ under monotone per-query adjacency, with equality in both when $n \ge 2c + \mu/\lambda$.*

Proof. The polar of the cone of an output is generated by $e_i - e_0$ for $i \in J_-$ and $e_0 - e_i$ for $i \in J_+$, so its elements have $\theta_i = \omega_i$ on J_-, $\theta_i = -\omega_i$ on J_+ and $\theta_0 = \Omega_+ - \Omega_-$, with $\omega \geq 0$ and Ω_+, Ω_- the sums of the weights on J_+ and on J_-. The Laplace box asks $\omega_i \leq \lambda$ and $|\Omega_+ - \Omega_-| \leq \mu$, and, as in the proof of Corollary 2, the extreme loss of the output is Δ_∞ times the largest $\Omega_+ + \Omega_-$ under per-query adjacency and the largest $\max\{\Omega_+, \Omega_-\}$ under monotone per-query adjacency. With $c' = |J_+| \leq c$,

$$\Omega_+ + \Omega_- \;\leq\; \min\bigl\{j\lambda,\; 2c'\lambda + \mu,\; 2(j - c')\lambda + \mu\bigr\} \;\leq\; 2c\lambda + \mu,$$

because $\Omega_+ + \Omega_- = 2\min\{\Omega_+, \Omega_-\} + |\Omega_+ - \Omega_-|$ with $\Omega_+ \leq c'\lambda$ and $\Omega_- \leq (j - c')\lambda$; and $\max\{\Omega_+, \Omega_-\} \leq c'\lambda + \mu$, because $\Omega_+ \leq c'\lambda$ and $\Omega_- \leq \Omega_+ + \mu$. For an output with c positive answers ending at $j \geq 2c + \mu/\lambda$, the choice $\omega_i = \lambda$ on J_+ and total weight $c\lambda + \mu$ spread over J_-, at most λ per position, attains both bounds; such outputs exist when $n \geq 2c + \mu/\lambda$. □

C Proofs for Section 4

Per-score adjacency with sensitivity Δ_∞ is contained in score-range adjacency with $\Delta = \sup_{d \in \Delta_\infty [-1,1]^m} h_D(d) = \Delta_\infty \max_{x,x' \in \operatorname{vert} Q} \|x - x'\|_1$, which is $2\Delta_\infty$ on the simplex and on Q_n with $n \geq 2$, and $2n\Delta_\infty$ on B_n, where two permutation matrices can disagree in every row.

The K-norm noise with body B of Definition 5 is well defined and has privacy body B. Its exponent h_{-B} is finite because B is bounded and at least a positive multiple of the Euclidean norm because $0 \in \operatorname{int} B$, so the density is positive and integrable, and the conjugate of h_{-B} is the indicator function of $-B$, so Lemma A.1 gives the body B.

C.1 Release

Releasing the noisy score vector itself constrains the whole body.

Proposition C.1. *For every noise Z as in Definition 1, releasing $w + Z$ has exact privacy $\sup_{z \in B(Z)} h_A(z)$. Under score-range adjacency the release is ε-differentially private exactly when $B(Z) \subseteq \tau D$.*

Proof. At a shift d the log-ratio of the densities of $w + Z$ and $w + d + Z$ at the point $w + z$ is $\phi(z - d) - \phi(z)$, which is at most $\phi^\infty(-d) = h_{-B(Z)}(-d) = h_{B(Z)}(d)$ by (A.2) and Lemma A.1; so no event has a larger log-ratio of probabilities. Conversely, fix $\bar{z} \in \operatorname{int} \operatorname{dom} \phi$. If the ray $\{\bar{z} - sd : s \geq 0\}$ stays in $\operatorname{dom} \phi$, it stays in $\operatorname{int} \operatorname{dom} \phi$, where ϕ is continuous, and the increments $\phi(\bar{z} - (t+1)d) - \phi(\bar{z} - td)$ are nondecreasing in $t \geq 0$ with limit $\phi^\infty(-d)$; each is the value at $\bar{z} - td$ of the function $z \mapsto \phi(z - d) - \phi(z)$, continuous there, so the events $w + U$ for small balls U around $\bar{z} - td$ have log-ratios of probabilities close to it. If the ray leaves $\operatorname{dom} \phi$ at some $s^* < \infty$, then $\phi^\infty(-d) = +\infty$, and for $t \geq 0$ with $t < s^* < t + 1$ the point $\bar{z} - td$ is interior to $\operatorname{dom} \phi$ while $\bar{z} - (t+1)d$ lies outside its closure, so a small ball U around $\bar{z} - td$ gives an event $w + U$ of positive probability at w and probability zero at $w + d$. Taking the supremum over $d \in A$ gives $\sup_{d \in A} h_{B(Z)}(d) = \sup_{z \in B(Z)} h_A(z)$. Under score-range adjacency this is $\Delta \sup_{z \in B(Z)} \gamma_D(z)$, which is at most ε exactly when $B(Z) \subseteq \tau D$. □

C.2 Active Facets and the Greatest Feasible Body

Polarity matches the facets of D with the vertices of D°: every facet F of D is $\{y \in D : \langle \nu_F, y \rangle = 1\}$ for exactly one $\nu_F \in D^\circ$, the vector ν_F is a vertex of D°, and every vertex of D° arises in this way. For $z \neq 0$ the maximum of $\langle \cdot, z \rangle$ over D° is $h_{D^\circ}(z) = \gamma_D(z)$, and the vertices of D° that attain it are the vectors ν_F of the facets F that contain the boundary point $z/\gamma_D(z)$ of D. A boundary point of D lies on exactly one facet precisely when it lies in the relative interior of that facet, so ν_F is the unique maximizer of $\langle \cdot, z \rangle$ over D° exactly when $z/\gamma_D(z) \in \operatorname{relint} F$. Since tangent cones are cones, a facet F is active, that is, $\operatorname{relint} F$ meets $\operatorname{int} T_v$ for some vertex v of Q, exactly when ν_F is the unique maximizer of $\langle \cdot, z \rangle$ over D° at some $z \in \operatorname{int} T_v$. The gauge is attained at an active facet on every tangent cone, including its boundary, where several facets may tie.

Lemma C.2. *For every vertex v of Q and every $z \in T_v$ there is an active facet F with $\langle \nu_F, z \rangle = \gamma_D(z)$.*

Proof. Tilting z toward the interior of the cone breaks the tie. Let L be the set of vertices of D° that attain $\gamma_D(z) = \max_{\nu \in D^\circ} \langle \nu, z \rangle$. The nonempty open set $\operatorname{int} T_v$ is not contained in the union of the finitely many hyperplanes $\{y : \langle \nu - \nu', y \rangle = 0\}$ with $\nu \neq \nu'$ in L, so choose y in $\operatorname{int} T_v$ off all of them, and let ν be the maximizer of $\langle \cdot, y \rangle$ over L, unique by that choice.

Put $z_s = (1-s)z + sy$, a point of $\operatorname{int} T_v$ for $s \in (0, 1]$. Each $\nu' \in L \setminus \{\nu\}$ has $\langle \nu', z_s \rangle = (1-s)\gamma_D(z) + s\langle \nu', y \rangle < \langle \nu, z_s \rangle$, and each vertex of D° outside L satisfies the same strict inequality at $s = 0$, hence for all small s. So ν is the unique maximizer at z_s, which makes its facet active, and it attains $\gamma_D(z)$ by membership in L. □

Theorem 3 (The greatest feasible body containing the baseline)**.** *H_τ is feasible, contains τD, contains every feasible body that contains τD, and is bounded.*

Proof. Containment is one inequality and feasibility is Lemma C.2; a step from the boundary of τD toward a violating point puts every feasible body containing τD inside H_τ; boundedness comes from the facets of Q.

Every $\nu \in D^\circ$ and $z \in \tau D$ satisfy $\langle \nu, z \rangle \leq \gamma_D(z) \leq \tau$, so $\tau D \subseteq H_\tau$. If $z \in T_v \cap H_\tau$, Lemma C.2 supplies an active facet F with $\gamma_D(z) = \langle \nu_F, z \rangle \leq \tau$, so $z \in \tau D$: H_τ is feasible.

Let B be feasible with $\tau D \subseteq B$, let F be active, witnessed by a vertex v and a point of $\operatorname{int} T_v$ at which ν_F is the unique maximizer, and suppose $\langle \nu_F, z \rangle > \tau$ for some $z \in B$. The witness may be scaled to a point $z_0 \in \operatorname{int} T_v$ with $\langle \nu_F, z_0 \rangle = \gamma_D(z_0) = \tau$, which puts z_0 in $\tau D \subseteq B$. For small $s > 0$ the point $(1-s)z_0 + sz$ still lies in $\operatorname{int} T_v \cap B$, and $\langle \nu_F, \cdot \rangle$ takes the value $\tau + s(\langle \nu_F, z \rangle - \tau) > \tau$ there, so its gauge exceeds τ, contradicting feasibility. Hence $B \subseteq H_\tau$.

For boundedness we show that every facet normal of Q gives an active facet of D. Let ν be an outward normal of a facet of Q, scaled so that $h_D(\nu) = 1$, let v be a vertex of Q minimizing $\langle \nu, \cdot \rangle$, and fix x in the relative interior of that facet. The point $z := x - v$ lies in $\operatorname{int} T_v$. Indeed, for $t \in (0, 1)$ the combination $(1-t)v + tx$ satisfies every facet inequality of Q strictly, since x lies on no other facet and $\langle \nu, v \rangle < h_Q(\nu)$; so it lies in $\operatorname{int} Q$, and $\operatorname{int} Q - v$ is an open subset of T_v, which places $t(x - v)$ and therefore z in $\operatorname{int} T_v$. Its pairing with ν is $\langle \nu, z \rangle = h_Q(\nu) + h_Q(-\nu) = h_D(\nu) = 1$ while $z \in D$, so $\gamma_D(z) = 1$ and ν maximizes $\langle \cdot, z \rangle$ over D°. It does so uniquely. If $\nu' \in D^\circ$ also has $\langle \nu', z \rangle = 1$, then $1 = \langle \nu', x \rangle - \langle \nu', v \rangle \leq h_Q(\nu') + h_Q(-\nu') = h_D(\nu') \leq 1$ forces $\langle \nu', x \rangle = h_Q(\nu')$, so the face of Q maximizing ν' meets the relative interior of the facet and therefore contains it. Being a proper face it equals the facet, which makes ν' a positive multiple of ν, and $\langle \nu', z \rangle = 1$ then forces $\nu' = \nu$. So ν is an exposed point of D°, hence a vertex, and its facet of D is active. The facet normals of Q positively span $\mathbb{R}^m$ because Q is bounded, so the constraints they contribute already bound H_τ. □

The facets of D also determine whether H_τ exceeds τD at all.

Proposition C.3. *$H_\tau = \tau D$ if and only if every facet of D is active.*

Proof. The body τD is $\{z : \langle \nu_F, z \rangle \leq \tau \text{ for every facet } F \text{ of } D\}$. If every facet is active, H_τ is cut out by all of them and equals τD. If a facet F_0 is inactive, take $y \in \operatorname{relint} F_0$, so that $\langle \nu_{F_0}, y \rangle = 1$ while $\langle \nu_F, y \rangle < 1$ for every other facet F. For small $\delta > 0$ the point $z_\delta = \tau(y + \delta \nu_{F_0})$ still satisfies $\langle \nu_F, z_\delta \rangle \leq \tau$ for every facet $F \neq F_0$, hence lies in H_τ, and $\langle \nu_{F_0}, z_\delta \rangle > \tau$ puts it outside τD. □

Every facet normal of Q gives an active facet by the proof of Theorem 3, so H_τ can exceed τD only through facets F of D for which $\langle \nu_F, \cdot \rangle$ attains its maximum over Q on a face of dimension at most $m - 2$. Both H_τ and the criterion are computable. A facet F is active exactly when, for some vertex v, the strict system $z \in \operatorname{int} T_v$ and $\langle \nu_F - \nu_{F'}, z \rangle > 0$ for every other facet F' has a solution. All these inequalities are homogeneous in z, so the strict signs may be replaced by unit margins, and one linear program determines whether the system has a solution. Hence $|\operatorname{vert} Q|$ times the number of facets of D linear programs return H_τ and determine whether it equals τD.

C.3 Expected Spread

One identity computes the expected spread of every K-norm noise.

Lemma C.4 (Moment identity). *Let Z be the K-norm noise with a body B, put $K = (-B)^\circ$, and let ψ be continuous, nonnegative, and positively homogeneous of degree one. Then*

$$\mathbb{E}\,\psi(Z) = (m+1)\,\frac{\int_K \psi}{|K|}.$$

Proof. The sublevel sets of the potential are the multiples of one body, and homogeneity turns both integrals into moments of the scale. The set K is a bounded convex body with the origin in its interior, by polarity. The potential is $h_{-B} = \gamma_K$, whose sublevel sets are $\{\gamma_K \le t\} = tK$, and homogeneity gives $\int_{tK} \psi = t^{m+1} \int_K \psi$ and $|tK| = t^m |K|$. Writing $e^{-\gamma_K(z)} = \int_{\gamma_K(z)}^{\infty} e^{-t}\,dt$ and exchanging the integrals,

$$\int \psi\, e^{-\gamma_K} = \int_0^\infty e^{-t} \Big(\int_{tK} \psi \Big) dt = (m+1)! \int_K \psi,$$

and the same computation with ψ replaced by the constant 1, where $\int_{tK} 1 = t^m |K|$, gives $\int e^{-\gamma_K} = m!\,|K|$. Divide. □

Theorem 4 (Expected spread). *The K-norm noise with body τD has expected spread m/τ for every polytope Q. For every feasible body B that contains τD, the K-norm noise with body B has expected spread at most m/τ, with equality only for $B = \tau D$.*

Proof. Apply Lemma C.4 and compare the two polar bodies. Let B be feasible with $\tau D \subseteq B$, bounded by Theorem 3. Write $K = (-B)^\circ$ and $K_0 = (-\tau D)^\circ = \tau^{-1} D^\circ$, so that $K \subseteq K_0$ and $\gamma_{K_0} = \tau \gamma_{D^\circ} = \tau h_D$. By Lemma C.4,

$$\mathbb{E}\, h_D(Z) = \frac{m+1}{\tau} \cdot \frac{\int_K \gamma_{K_0}}{|K|}.$$

Integrating in polar coordinates, with radial functions $\rho_K \le \rho_{K_0}$ and $\gamma_{K_0} = 1/\rho_{K_0}$,

$$\frac{\int_K \gamma_{K_0}}{|K|} = \frac{m}{m+1} \cdot \frac{\int \rho_K^{m+1}/\rho_{K_0}}{\int \rho_K^m},$$

both integrals over the unit sphere. When $B = \tau D$ we have $K = K_0$, the last fraction is 1, and the expected spread is m/τ for every Q. In general $\rho_K \le \rho_{K_0}$ makes $\rho_K^{m+1}/\rho_{K_0} \le \rho_K^m$, so the fraction is at most 1; equality forces $\rho_K = \rho_{K_0}$ almost everywhere, since $\rho_K > 0$, hence everywhere by continuity, so $K = K_0$ and $B = \tau D$. □

The containment $\tau D \subseteq B$ cannot be dropped.

Example C.5. On the running example Q_2 at $\tau = 1$, the body $B = \{z : z_1 \ge -1,\ z_2 \ge -1,\ 2z_1 + z_2 \le 1\}$ is feasible: on $T_0 = \mathbb{R}^2_+$ its points satisfy $\gamma_D(z) = z_1 + z_2 \le 2z_1 + z_2 \le 1$, and on T_{e_1} and T_{e_2} the gauge is $-z_1$ and $-z_2$, at most 1 on B. It omits the point e_1 of τD. Its polar body $(-B)^\circ$ is the triangle $\mathrm{conv}\{(1,0),(0,1),(-2,-1)\}$, of area 2. Here $h_D(y) = \max\{|y_1|, |y_2|, |y_1 - y_2|\}$, which is linear on each of the six cones over the edges of D°, so cutting the triangle along those cones and integrating a linear function on each of the resulting pieces gives $\int_{(-B)^\circ} h_D = \frac{25}{18}$. By Lemma C.4 the K-norm noise with body B has expected spread $3 \cdot \frac{25}{18}/2 = \frac{25}{12}$, above the value 2 of the baseline.

C.4 The Instances

Proposition C.6. *For the simplex $Q = \mathrm{conv}\{e_1, \dots, e_n\}$ with $n \ge 2$, read within $V = \{z : \sum_i z_i = 0\}$,*

$$H_\tau = \{z \in V : z_i \ge -\tau \text{ for all } i\},$$

and for $Q_n = \mathrm{conv}\{0, e_1, \dots, e_n\}$,

$$H_\tau = \Big\{ z : z_i \ge -\tau \text{ for all } i,\ \textstyle\sum_i z_i \le \tau \Big\}.$$

Proof. In each case the displayed set is feasible and contains τD, so it is contained in H_τ by Theorem 3; the reverse inclusion repeats the step from that proof.

On the simplex, $T_{e_j} = \{z \in V : z_i \geq 0 \text{ for } i \neq j\}$ and $\gamma_D(z) = -z_j$ on it, as computed in Appendix B.2. The displayed set is therefore feasible, and it contains τD because every $z \in D$ has $z_i \geq -1$. Now let B be feasible with $\tau D \subseteq B$ and suppose $z \in B$ has $z_j < -\tau$. The point $y = -\tau e_j + \frac{\tau}{n-1}\sum_{i\neq j} e_i$ lies in $\operatorname{int} T_{e_j}$ and has $\gamma_D(y) = \tau$, so $y \in \tau D \subseteq B$; for small $s > 0$ the point $(1-s)y + sz$ lies in $T_{e_j} \cap B$ with j-th coordinate below $-\tau$, hence with gauge above τ, contradicting feasibility.

On Q_n the tangent cones and the gauge on them are

$$T_0 = \mathbb{R}^n_+, \quad \gamma_D(z) = \textstyle\sum_i z_i; \qquad T_{e_j} = \big\{z : z_i \geq 0\ (i \neq j),\ \textstyle\sum_i z_i \leq 0\big\}, \quad \gamma_D(z) = -z_j.$$

Each gauge value follows from one representation and one dual point: on T_0 from $z = (\sum_i z_i)(x - 0)$ with $x = z/\sum_i z_i \in Q_n$ and $h_D(\mathbf{1}) = 1$, and on T_{e_j} from $z = (-z_j)(x - e_j)$ with $x = z/(-z_j) + e_j \in Q_n$ and $h_D(-e_j) = 1$. The displayed set is therefore feasible, and it contains τD because every $z \in D$ has $z_i \geq -1$ and $\sum_i z_i \leq 1$. For the reverse inclusion let B be feasible with $\tau D \subseteq B$ and let $z \in B$. If $\sum_i z_i > \tau$, step from $y = \frac{\tau}{n}\mathbf{1} \in \tau D \cap \operatorname{int} T_0$ toward z; if $z_j < -\tau$, step from $y = -\tau e_j + \delta \sum_{i\neq j} e_i$ with $\delta > 0$ small, a point of τD interior to T_{e_j}. Either step produces a point of $T_v \cap B$ with gauge above τ. □

Example C.7. For the cube $Q = [0,1]^m$ the baseline is already the greatest feasible body containing it. Its difference body is $D = [-1,1]^m$, whose facets have the normals $\pm e_i$; these are also facet normals of Q, so the facets are active by the proof of Theorem 3, and Proposition C.3 gives $H_\tau = \tau D$ for every m.

D Proofs for Section 5

D.1 Selection with Abstention

Independence makes Theorem 1 explicit on Q_n: the body is the box of the tail decay rates, and the abstain cone adds the left-tail rates of all coordinates.

Theorem D.1 (Independent coordinates)**.** *On Q_n under score-range adjacency, let the noise Z have independent coordinates with the tail decay rates $\alpha_i, \beta_i \in (0, +\infty]$ of Lemma A.2.*

(a) $\varepsilon_{\text{exact}}(M_Z) = \Delta \max\{\sum_i \alpha_i,\ \max_j \beta_j\}$.
(b) *If M_Z is ε-differentially private, then* $\sup_w \operatorname{Regret}(M_Z, w) \geq n\Delta/(16\varepsilon)$.
(c) *The order is attained: iid coordinates with density proportional to $e^{-\max\{\tau z, -\tau z/n\}}$, so that $\alpha_i = \tau/n$ and $\beta_i = \tau$, give $\varepsilon_{\text{exact}}(M_Z) = \varepsilon$ and* $\sup_w \operatorname{Regret}(M_Z, w) \leq n\Delta/\varepsilon$.

Proof. The formula is Theorem 1 on the cones of Q_n; the lower bound turns the budget into a heavy left tail, the tail into a bound on the density, and the density bound into a constant-probability error; the upper bound is a direct computation.

(a) By Lemma A.2, $B(Z) = \prod_i [-\beta_i, \alpha_i]$. The tangent cones of Q_n and the gauge on each of them are computed in Proposition C.6. On $T_0 = \mathbb{R}^n_+$ the gauge is $\gamma_D(z) = \sum_i z_i$ and $T_0 \cap B(Z) = \prod_i [0, \alpha_i]$, so the supremum there is $\sum_i \alpha_i$. On T_{e_j} the gauge is $\gamma_D(z) = -z_j$, at most β_j on $B(Z)$, with the supremum β_j approached along the ray $\{-se_j : s \geq 0\}$ of T_{e_j}. Theorem 1 under score-range adjacency gives the formula.

(b) By (a), privacy forces $\sum_i \alpha_i \leq \tau$, so some coordinate j has $\alpha_j \leq \tau/n$. Write the density of Z_j as $e^{-\phi_j}$, log-concave by Prékopa (1973). The privacy body of Z_j is $[-\beta_j, \alpha_j]$, so Lemma A.1 gives $\phi_j^\infty = h_{[-\alpha_j, \beta_j]}$. This support function is positive away from the origin, so ϕ_j is coercive and attains its minimum at some z_0; (A.2) then gives $\phi_j(z_0 - s) \leq \phi_j(z_0) + s\alpha_j$ for $s \geq 0$, whence

$$1 \geq \int_{-\infty}^{z_0} e^{-\phi_j} \geq e^{-\phi_j(z_0)} \int_0^\infty e^{-\alpha_j s}\, ds = \frac{e^{-\phi_j(z_0)}}{\alpha_j},$$

so the density of Z_j is bounded by α_j. Set $t = 1/(4\alpha_j)$. The interval $[-t, t]$ then has probability at most $2t\alpha_j = 1/2$, so one of the tail events $\{Z_j < -t\}$ and $\{Z_j > t\}$ has probability at least $1/4$.

One score vector turns either tail into regret. If $\Pr[Z_j < -t] \geq 1/4$, take $w = t\,e_j$: the optimum is e_j with value t, and every other vertex has value 0. On $\{Z_j < -t\}$ the noisy score $t + Z_j$ of e_j is negative while the abstain vertex scores 0, so the output is not e_j. If instead $\Pr[Z_j > t] \geq 1/4$, take $w = -t\mathbf{1}$: the optimum is the abstain vertex with value 0, and every other vertex has value $-t$. On $\{Z_j > t\}$ the noisy score $-t + Z_j$ of e_j is positive, so the output is not the abstain vertex. In both cases the output falls t below the optimum on an event of probability at least $1/4$, so

$$\sup_w \mathrm{Regret}(M_Z, w) \;\geq\; \frac{t}{4} \;=\; \frac{1}{16\alpha_j} \;\geq\; \frac{n}{16\tau} \;=\; \frac{n\Delta}{16\varepsilon}.$$

(c) Write $\alpha = \tau/n$ and $\beta = \tau$. The density of Z_i is then $\frac{\alpha\beta}{\alpha+\beta} e^{-\max\{\beta z, -\alpha z\}}$, with $\mathbb{E}\, e^{\theta Z_i}$ finite exactly for $-\alpha < \theta < \beta$, so part (a) gives $\varepsilon_{\mathrm{exact}}(M_Z) = \Delta \max\{n\alpha, \beta\} = \varepsilon$. Fix w, let x^* maximize $\langle w, \cdot\rangle$ over Q_n, and let $\hat{x} = M_Z(w)$. Optimality of $\hat{x}$ for $w + Z$ gives

$$\langle w, x^* - \hat{x}\rangle \;\leq\; \langle Z, \hat{x} - x^*\rangle \;\leq\; h_{Q_n}(Z) - \langle Z, x^*\rangle.$$

The right tail is $\Pr[Z_i > s] = \frac{\alpha}{\alpha+\beta} e^{-\beta s}$ for $s \geq 0$, and the mean is $\mathbb{E}Z_i = 1/\beta - 1/\alpha$. Since $h_{Q_n}(Z) = \max\{0, \max_i Z_i\}$, a union bound gives $\Pr[\max_i Z_i > s] \leq \frac{n\alpha}{\alpha+\beta} e^{-\beta s} \leq e^{-\beta s}$ and hence $\mathbb{E}\, h_{Q_n}(Z) \leq 1/\beta$. Since x^* does not depend on Z, $-\mathbb{E}\langle Z, x^*\rangle \leq 1/\alpha - 1/\beta$. The two bounds add to $1/\alpha = n\Delta/\varepsilon$. □

Read as a constraint, Theorem D.1(a) is a containment. By Proposition C.6 the greatest feasible body containing τD on Q_n is $H_\tau = \{z : z_i \geq -\tau \text{ for all } i,\ \sum_i z_i \leq \tau\}$, and a box $\prod_i [-\beta_i, \alpha_i]$ lies inside it exactly when $\beta_j \leq \tau$ for every j and $\sum_i \alpha_i \leq \tau$. Those two conditions are exactly $\varepsilon_{\mathrm{exact}}(M_Z) \leq \varepsilon$ by Theorem D.1(a), so for noise with independent coordinates

$$\varepsilon_{\mathrm{exact}}(M_Z) \leq \varepsilon \qquad \Longleftrightarrow \qquad B(Z) \subseteq H_\tau.$$

The next theorem gives a noise with the body H_τ itself, and a lower bound that every selection mechanism obeys.

Theorem D.2 (The noise $G_i - G_0$ and the selection rate)**.** *On Q_n under score-range adjacency, let $G_0, G_1, \ldots, G_n$ be iid Gumbel of rate τ and let $Z_i = G_i - G_0$.*

(a) *The noise Z has privacy body H_τ, exact privacy $\varepsilon_{\mathrm{exact}}(M_Z) = \varepsilon$, and worst-case regret $\sup_w \mathrm{Regret}(M_Z, w) \leq \big(1 + \log(n+1)\big)\Delta/\varepsilon$.*

(b) *Every ε-differentially private selection mechanism M on Q_n satisfies*

$$\sup_w \mathrm{Regret}(M, w) \;\geq\; \frac{\Delta}{2}\Big(\frac{\log(n/2)}{\varepsilon} - 1\Big).$$

Proof. The body is a cumulant computation, the upper bound identifies M_Z as the exponential mechanism, and the lower bound is a standard packing argument, as in Hardt & Talwar (2010).

(a) By Proposition B.5 at $\lambda = \tau$, Z has a log-concave density and the body $\{z : z_i \geq -\tau,\ \sum_i z_i \leq \tau\}$, which is H_τ by Proposition C.6. This body is feasible by Theorem 3, so $\varepsilon_{\mathrm{exact}}(M_Z) \leq \varepsilon$. Since τe_1 lies in $T_0 \cap H_\tau$ with $\gamma_D(\tau e_1) = \tau$, Theorem 1 gives equality.

By Proposition B.5, $M_Z(w)$ is the maximizer of $w_i + G_i$ over $i \in \{0, 1, \ldots, n\}$, where $w_0 = 0$, and by the Gumbel-max identity (McFadden, 1974) the maximizer is i with probability proportional to $e^{\tau w_i}$: M_Z is the exponential mechanism (McSherry & Talwar, 2007) over these $n+1$ options. Writing $w^* = \max_i w_i$, the optimal value on Q_n, and I for the index selected,

$$\Pr[w_I \leq w^* - s] \;\leq\; \frac{(n+1)e^{\tau(w^* - s)}}{e^{\tau w^*}} \;=\; (n+1)e^{-\tau s},$$

so, integrating the tail of the nonnegative random variable $w^* - w_I$,

$$\mathrm{Regret}(M_Z, w) \;=\; \int_0^\infty \Pr[w^* - w_I > s]\, ds \;\leq\; \int_0^\infty \min\{1, (n+1)e^{-\tau s}\}\, ds \;=\; \frac{1 + \log(n+1)}{\tau}.$$

(b) Put $g = \Delta\big(\log(n/2)/\varepsilon - 1\big)$. For $g \leq 0$ the bound is nonpositive, so assume $g > 0$ and suppose that $\sup_w \mathrm{Regret}(M, w) < g/2$. At $w^{(j)} = g\,e_j$ the vertex e_j has value g and every other vertex

has value 0, so $\mathrm{Regret}(M, w^{(j)}) = g \Pr[M(w^{(j)}) \neq e_j]$ and therefore $\Pr[M(w^{(j)}) = e_j] > 1/2$. Since $h_D(y) = h_{Q_n}(y) + h_{Q_n}(-y)$ gives $h_D(g\, e_j) = g$ and h_D is positively homogeneous, the points $\frac{\ell}{r} g\, e_j$ for $\ell = 0, \dots, r$ with $r = \lceil g/\Delta \rceil$ join 0 to $w^{(j)}$ in r adjacent steps. Privacy along the chain gives $\Pr[M(0) = e_j] > \frac{1}{2} e^{-r\varepsilon}$. The n events $\{M(0) = e_j\}$ are disjoint, so $\frac{n}{2} e^{-r\varepsilon} < 1$ and $r > \log(n/2)/\varepsilon$. But $r = \lceil g/\Delta \rceil < g/\Delta + 1 = \log(n/2)/\varepsilon$, a contradiction. □

Theorem 5 (Selection with abstention). *On Q_n under score-range adjacency, let $0 < \varepsilon \leq 1$ and $n \geq 30$. The least worst-case regret of an ε-differentially private mechanism M_Z whose noise has independent coordinates is $\Theta(n\Delta/\varepsilon)$, and the least worst-case regret of an ε-differentially private selection mechanism is $\Theta(\Delta \log n/\varepsilon)$. The second is attained by the noise $Z_i = G_i - G_0$ with $G_0, \dots, G_n$ iid Gumbel of rate τ, whose body is H_τ and whose exact privacy is ε.*

Proof. For the first class, Theorem D.1(b) gives the lower bound $n\Delta/(16\varepsilon)$ and Theorem D.1(c) the upper bound $n\Delta/\varepsilon$. For the second, Theorem D.2(a) gives the upper bound $(1 + \log(n+1))\Delta/\varepsilon$, attained by the noise $G_i - G_0$ with the stated body and exact privacy, and Theorem D.2(b) is at least $\Delta \log n/(4\varepsilon)$ for $0 < \varepsilon \leq 1$ and $n \geq 30$, since $\varepsilon \leq 1$ and $\log(n/2) - 1 \geq \frac{1}{2} \log n$ there. □

D.2 Assignment

An instance is a score matrix $w \in \mathbb{R}^{n \times n}$ with $n \geq 2$, where w_{ij} is the value of assigning i to j, and the readout is the maximum-weight assignment, so Q is the Birkhoff polytope B_n with the permutation matrices X_σ as vertices, and

$$h_D(d) = \max_\sigma \sum_i d_{i\sigma(i)} - \min_\sigma \sum_i d_{i\sigma(i)}.$$

Row and column potentials do not move any comparison: if $d_{ij} = a_i + b_j$, every assignment receives the same shift $\sum_i a_i + \sum_j b_j$, so $h_D(d) = 0$. For a maximizing σ and a minimizing π, $h_D(d) = \sum_{i:\sigma(i)\neq\pi(i)} (d_{i\sigma(i)} - d_{i\pi(i)})$ is a signed sum of distinct entries of d, so $h_D(d) \leq \|d\|_1$. Differences of assignments span $V = \{z \in \mathbb{R}^{n\times n} : z\mathbf{1} = 0,\ z^\top \mathbf{1} = 0\}$, a space of dimension $m = (n-1)^2$ whose orthogonal complement is exactly the space of potentials, and by Appendix A only the V-components of the scores and the noise reach the readout or the adjacency. Every vertex of B_n is Pareto-maximal, since the polytope lies on a level set of the sum of the entries.

Theorem D.3 (Exponential noise on the Birkhoff polytope). *Let $n \geq 2$, let $\lambda > 0$, let Z have iid exponential entries of rate λ, and let adjacency be score-range.*

(a) *The projected noise $P_V Z$ has privacy body $\{z \in V : z_{ij} \geq -\lambda \text{ for all } i, j\}$, and $\varepsilon_{\mathrm{exact}}(M_Z) = \lambda\Delta$.*
(b) *At $\lambda = \tau$, so that M_Z is ε-differentially private, $\sup_w \mathrm{Regret}(M_Z, w) \leq n \log n\, \Delta/\varepsilon$.*
(c) *Every ε-differentially private selection mechanism M on B_n satisfies*

$$\sup_w \mathrm{Regret}(M, w) \ \geq\ \frac{\Delta}{8}\Big(\frac{1}{\varepsilon} \log \frac{\lceil n/2 \rceil !}{2n} - 1\Big).$$

Proof. The body comes from the projected cumulant domain; the regret bound needs only the row maxima of the noise; the lower bound packs permutations in Hamming distance.

(a) Each entry has $\mathrm{dom}\, \Lambda_{Z_{ij}} = (-\infty, \lambda)$, so $\Lambda_{P_V Z}$, which agrees with Λ_Z on V, has the domain $\{\theta \in V : \theta_{ij} < \lambda\}$, an open convex set containing the origin, whose closure in V relaxes the inequalities; negating it gives the stated body B.

Fix a vertex X_π. The tangent cone there and the gauge on it are

$$T_{X_\pi} = \{z \in V : z_{ij} \geq 0 \text{ for } j \neq \pi(i)\}, \qquad \gamma_D(z) = \max_i\big(-z_{i\pi(i)}\big) \quad (z \in T_{X_\pi}),$$

and one construction gives both. Let $z \in V$ be nonzero with $z_{ij} \geq 0$ off the support of X_π, and put $s = \max_i(-z_{i\pi(i)})$; since the rows of z sum to zero, some entry on the support is negative and $s > 0$. The matrix $x = X_\pi + z/s$ has row and column sums 1, its entries off the support of X_π are $z_{ij}/s \geq 0$, and its entries on the support are $1 + z_{i\pi(i)}/s \geq 0$. Hence $x \in \mathsf{B}_n$, and $z = s(x - X_\pi)$ lies in T_{X_π} and in sD, giving $\gamma_D(z) \leq s$. Conversely every element of T_{X_π} has the form $t(x - X_\pi)$ with

$t \geq 0$ and $x \in \mathsf{B}_n$, hence is nonnegative off the support. For each row i the functional $z \mapsto -z_{i\pi(i)}$ is at most 1 on D and so lies in D°, which gives $\gamma_D(z) \geq -z_{i\pi(i)}$ and hence $\gamma_D(z) \geq s$.

On $T_{X_\pi} \cap B$ the entries $z_{i\pi(i)}$ are at least $-\lambda$, so γ_D is at most λ there. Choosing σ with $\sigma(i) \neq \pi(i)$ for every i makes $z = \lambda(X_\sigma - X_\pi)$, a point of $T_{X_\pi} \cap B$, attain that value. The supremum is therefore λ at every vertex, and Theorem 1 gives $\varepsilon_{\text{exact}}(M_Z) = \lambda\Delta$.

(b) Fix w, let x^* maximize $\langle w, \cdot \rangle$ over B_n, and let $\hat{x} = M_Z(w)$. Optimality of $\hat{x}$ for $w + Z$ gives

$$\langle w, x^* - \hat{x} \rangle \;\leq\; \langle Z, \hat{x} - x^* \rangle \;\leq\; h_{\mathsf{B}_n}(Z) - \langle Z, x^* \rangle,$$

so $\operatorname{Regret}(M_Z, w) \leq \mathbb{E}\, h_{\mathsf{B}_n}(Z) - \mathbb{E}\langle Z, x^* \rangle$. Since x^* does not depend on Z and takes one entry from each row, $\mathbb{E}\langle Z, x^* \rangle = n/\lambda$. An assignment likewise takes one entry from each row, so $h_{\mathsf{B}_n}(Z) \leq \sum_i \max_j Z_{ij}$, and the maximum of n independent exponential random variables of rate λ has mean $\lambda^{-1} \sum_{j \leq n} j^{-1}$, at most $(1 + \log n)/\lambda$. Hence $\mathbb{E}\, h_{\mathsf{B}_n}(Z) \leq n(1 + \log n)/\lambda$ and the regret is at most $n \log n/\lambda$, which is $n \log n\, \Delta/\varepsilon$ at $\lambda = \tau$.

(c) Permutations separated in Hamming distance replace the n candidates of Theorem D.2. Write $\delta(\sigma, \pi) = \#\{i : \sigma(i) \neq \pi(i)\}$ and let Π be a maximal set of permutations of $[n]$ with pairwise distance exceeding $n/2$, so that the balls of radius $n/2$ about its members cover all permutations of $[n]$. There are $\binom{n}{j}$ choices for a set of j disagreements and at most $j!$ ways to fill it, so $\#\{\sigma : \delta(\sigma, \pi) = j\} \leq n!/(n-j)!$; summing at most n such terms, each at most $n!/\lceil n/2 \rceil!$, bounds every such ball by $n \cdot n!/\lceil n/2 \rceil!$, so $|\Pi| \geq \lceil n/2 \rceil!/n$. The claimed bound increases with this size, so it suffices to prove it with $\lceil n/2 \rceil!/n$ replaced by $|\Pi|$.

Put $g = \frac{\Delta}{n}\big(\frac{1}{\varepsilon} \log \frac{|\Pi|}{2} - 1\big)$. For $g \leq 0$ the bound is nonpositive, so assume $g > 0$ and suppose $\sup_w \operatorname{Regret}(M, w) < gn/8$. At the instance $w^{(\pi)} = g\, X_\pi$,

$$\langle w^{(\pi)}, X_\sigma \rangle = g\big(n - \delta(\sigma, \pi)\big), \qquad h_D\big(w^{(\pi)}\big) = gn,$$

the second because the value runs from gn at π down to 0 at any σ that disagrees with π everywhere. So X_π is optimal and $\operatorname{Regret}(M, w^{(\pi)}) = g\, \mathbb{E}\, \delta(\hat{\sigma}, \pi)$ for the output $X_{\hat{\sigma}}$. The regret assumption therefore gives $\mathbb{E}\, \delta(\hat{\sigma}, \pi) < n/8$, and Markov's inequality gives $\Pr[M(w^{(\pi)}) \in Y_\pi] > 1/2$ for $Y_\pi = \{X_\sigma : \delta(\sigma, \pi) \leq n/4\}$. These sets are disjoint over $\pi \in \Pi$, two members being more than $n/2$ apart.

The points $\frac{\ell}{r} g\, X_\pi$ for $\ell = 0, \ldots, r$ with $r = \lceil gn/\Delta \rceil$ join 0 to $w^{(\pi)}$ in r adjacent steps, and privacy along the chain gives $\Pr[M(0) \in Y_\pi] > \frac{1}{2} e^{-r\varepsilon}$. Summing over Π gives $\frac{|\Pi|}{2} e^{-r\varepsilon} < 1$, so $r > \frac{1}{\varepsilon} \log \frac{|\Pi|}{2}$, while $r < gn/\Delta + 1 = \frac{1}{\varepsilon} \log \frac{|\Pi|}{2}$, a contradiction. Hence $\sup_w \operatorname{Regret}(M, w) \geq gn/8$, which is the stated bound. □

Releasing the scores is a different task, and the adjacency fixes its error measure. A mechanism M may output anything from which a fixed rule decodes a value y_x for every vertex x; since w and $w + d$ with $d_{ij} = a_i + b_j$ are adjacent however large d is, no such mechanism can determine a common shift of the decoded values. Write osc for the difference between the largest and the smallest entry of a vector indexed by the vertices, and $\langle w, \cdot \rangle$ for the vector of true scores. The expected range error of M at w is

$$\operatorname{Range}(M, w) = \mathbb{E}\, \operatorname{osc}\big(y - \langle w, \cdot \rangle\big).$$

A release also selects: if $\hat{x}$ maximizes y_x and x^* maximizes $\langle w, x \rangle$, then $\langle w, x^* - \hat{x} \rangle \leq (y_{\hat{x}} - \langle w, \hat{x} \rangle) - (y_{x^*} - \langle w, x^* \rangle) \leq \operatorname{osc}(y - \langle w, \cdot \rangle)$, so the regret of $\hat{x}$ is at most the range error. The baseline body supplies the release, and a packing argument shows that it is optimal.

Theorem D.4 (Releasing the scores)**.** *Let $n \geq 2$, and let adjacency be score-range.*

(a) *The mechanism that releases $P_V w + Z$ for the K-norm noise Z with body τD and decodes $y_x = \langle P_V w + Z, x \rangle$ is ε-differentially private, with range error $m\Delta/\varepsilon$ at every w.*

(b) *For $0 < \varepsilon \leq 1$, every ε-differentially private mechanism M whose output decodes all assignment scores satisfies $\sup_w \operatorname{Range}(M, w) \geq m\Delta/(16e^2\varepsilon)$.*

Proof. The upper bound releases the noise of the baseline body instead of passing it through an argmax; the lower bound turns a decoding into an estimate of w and packs balls in the norm h_D.

(a) The noise has density proportional to $e^{-\tau h_D}$ on V, so the density of $P_V w + Z$ at $u \in V$ is proportional to $e^{-\tau h_D(u - P_V w)}$. Since h_D is subadditive and vanishes on $V^\perp$, the ratio at adjacent w, w' is at most $e^{\tau h_D(w-w')} \le e^{\varepsilon}$. The discrepancy is $y_x - \langle w, x\rangle = \langle Z, x\rangle + \langle P_V w - w, x\rangle$, and $P_V w - w$ is a potential, so the second term is the same at every vertex. The oscillation is therefore $h_D(Z)$, and Theorem 4 gives $\mathbb{E}\, h_D(Z) = m/\tau$.

(b) On V the support function h_D is a norm, and $h_D(u) = \operatorname{osc}(\langle u, \cdot\rangle)$. For a decoded vector y let $\widehat{w}(y)$ be the minimizer of least Euclidean norm of $\operatorname{osc}(y - \langle u, \cdot\rangle)$ over $u \in V$. The objective is continuous and, being at least $h_D(u) - \operatorname{osc}(y)$, coercive, so its minimizers form a nonempty compact convex set and the one of least norm is unique and depends measurably on y. Minimality at w gives

$$h_D(\widehat{w} - w) = \operatorname{osc}\big(\langle \widehat{w} - w, \cdot\rangle\big) \le \operatorname{osc}\big(y - \langle \widehat{w}, \cdot\rangle\big) + \operatorname{osc}\big(y - \langle w, \cdot\rangle\big) \le 2 \operatorname{osc}\big(y - \langle w, \cdot\rangle\big),$$

so for $w \in V$ the vector $\widehat{w}$ is an ε-differentially private estimate of w with $\mathbb{E}\, h_D(\widehat{w} - w) \le 2 \operatorname{Range}(M, w)$.

Write $\kappa = \sup_{w \in V} \operatorname{Range}(M, w)$, which may be assumed finite, and let $t > 2\kappa$. Put $R = m\Delta/\varepsilon$, and let $w^{(1)}, \dots, w^{(L)}$ be a maximal $4t$-separated subset of the h_D-ball of radius R in V. Maximality makes the balls of radius $4t$ about the $w^{(i)}$ cover that ball, so $L \ge (R/4t)^m$. Markov's inequality gives $\Pr_{w^{(i)}}[\widehat{w} \in U_i] > 1/2$ for the open balls U_i of radius $2t$ about $w^{(i)}$, which are disjoint. Joining 0 to $w^{(i)}$ by $r = \lceil R/\Delta \rceil$ equal steps, each of h_D-length at most Δ, privacy gives $\Pr_0[\widehat{w} \in U_i] > \frac{1}{2} e^{-r\varepsilon}$, and disjointness gives $L < 2e^{r\varepsilon}$. Since $r < R/\Delta + 1 = m/\varepsilon + 1$, these two bounds on L give $m \log \frac{R}{4t} < \log 2 + m + \varepsilon$, so

$$t \;>\; \frac{R}{4} e^{-1-(\log 2 + \varepsilon)/m} \;\ge\; \frac{m\Delta}{8e^2\varepsilon}$$

for $\varepsilon \le 1$. Letting t decrease to 2κ gives $\kappa \ge m\Delta/(16e^2\varepsilon)$, and $\sup_w \operatorname{Range}(M, w) \ge \kappa$. □

Theorem 6 (Private assignment)**.** *On B_n under score-range adjacency, iid exponential noise of rate λ on the entries followed by one maximum-weight assignment has exact privacy $\lambda\Delta$. For $0 < \varepsilon \le 1$ and $n \ge 50$, at $\lambda = \tau$ it attains the least worst-case regret of an ε-differentially private selection mechanism, $\Theta(n \log n\, \Delta/\varepsilon)$, while the least worst-case range error of an ε-differentially private mechanism that releases all assignment scores is $\Theta(n^2\Delta/\varepsilon)$.*

Proof. The exact privacy is Theorem D.3(a). The upper bound on the regret is Theorem D.3(b), and Theorem D.3(c) is at least $n \log n\, \Delta/(40\varepsilon)$ for $0 < \varepsilon \le 1$ and $n \ge 50$, since $\log\lceil n/2\rceil! \ge \frac{n}{2} \log \frac{n}{2e}$. Theorem D.4 gives the two bounds on the range error, with $m = (n-1)^2$. □

Under per-score adjacency the same bodies give the values quoted in Section 5.

Proposition D.5. *Let adjacency be per-score on the $n \times n$ entries, with sensitivity Δ_∞. With iid Laplace, exponential or Gumbel entries of rate λ, the mechanism on B_n has $\varepsilon_{\text{exact}} = 2n\lambda\Delta_\infty$, while releasing the score matrix with iid Laplace noise of rate λ has exact privacy $n^2\lambda\Delta_\infty$.*

Proof. Here $h_A(z) = \Delta_\infty \|z\|_1$. Read within V, the body is $\{z \in V : z_{ij} \ge -\lambda\}$ for exponential and Gumbel noise and $\{z \in V : |z_{ij}| \le \lambda\}$ for Laplace noise, as in the proof of Theorem D.3(a). On T_{X_π} every row has $z_{i\pi(i)} = -\sum_{j \ne \pi(i)} z_{ij} \le 0$, so $\|z\|_1 = 2\sum_i (-z_{i\pi(i)}) \le 2n\lambda$ on the body, with equality at $\lambda(X_\sigma - X_\pi)$ for a σ that disagrees with π everywhere; Theorem 1 gives $2n\lambda\Delta_\infty$. Releasing $w + Z$ with iid Laplace noise, the log-density ratio between the inputs w and $w + d$ at the released value $w + u$ is $\lambda(\|u - d\|_1 - \|u\|_1) \le \lambda\|d\|_1$, with equality when $u_{ij} d_{ij} \le 0$ for every entry, an event of positive probability; the largest allowed shift has $\|d\|_1 = n^2\Delta_\infty$. □