
Asymptotic Optimality of the High-Dimensional Gaussian Mechanism and Improved Low-Dimensional Mechanisms for Differential Privacy

Yu Wei^{1,2} Alexander Bienstock³ Antigoni Polychroniadou³

Abstract

The additive noise mechanism is a foundational tool for differential privacy (DP) of T -dimensional real-valued vector queries. The Gaussian mechanism, utilizing Gaussian noise, is the mostly widely used such mechanism, due to its simplicity and strong privacy guarantees. In this work, we provide justification for this choice, showing that as the dimension $T \rightarrow \infty$, no additive-noise mechanism can asymptotically improve on the Gaussian mechanism’s privacy-utility tradeoff for the strong privacy settings typically used. We also develop a new family of *Spherical Generalized Gamma* DP mechanisms, which contains both the Gaussian mechanism and the recently studied ℓ_2 mechanism (Joseph *et al.*, ICML 2025). We identify members of this family that outperform both the Gaussian and ℓ_2 mechanisms in certain low-dimensional settings, and show tight composition of all mechanisms in this family, answering an open question of Joseph *et al.* regarding the ℓ_2 mechanism.

1. Introduction

Differential Privacy (DP) (Dwork *et al.*, 2006b) is a powerful tool that can be utilized to extract aggregate insights from data, while retaining individuals’ privacy. Differential privacy has myriad privacy-preserving applications, including training deep learning models (Abadi *et al.*, 2016), generating synthetic data (Kurakin *et al.*, 2024; Zhang *et al.*, 2021; Torkzadehmahani *et al.*, 2019; Yoon *et al.*, 2019; Mckenna *et al.*, 2019), publishing aggregate statistics (Dwork *et al.*, 2006b), and more.

¹Georgia Institute of Technology, Atlanta, Georgia, USA

²Work partially done while an intern at JPMorgan ³JPMorgan AI Research & AlgoCRYPT CoE, New York, New York, USA. Correspondence to: Alexander Bienstock <afb383@nyu.edu>, Yu Wei <ywei368@gatech.edu>.

Proceedings of the 43rd International Conference on Machine Learning, Seoul, South Korea. PMLR 306, 2026. Copyright 2026 by the author(s).

Many applications of DP, including the above, involve protecting real-valued T -dimensional vector queries $Q(D) \in \mathbb{R}^T$ over private datasets D . In this setting, we want to obfuscate any addition or removal of a single data record from D , which may perturb Q by any $\mu \in \mathbb{R}^T$. Typically, the size $\|\mu\|$ of μ is bounded by some scalar s , the *sensitivity* of the query. The standard way to achieve DP in this setting is by using the seminal additive noise mechanism (Dwork *et al.*, 2006b). This mechanism simply adds T -dimensional noise X sampled from some distribution \mathcal{D} to the query: $\hat{Q} = Q + X$. The typical distribution used is the multivariate Gaussian distribution (Dwork *et al.*, 2006a; Dwork & Roth, 2014), i.e., $X \sim \mathcal{N}(0, \sigma^2 I_T)$, where $\sigma \propto s$. This *Gaussian mechanism* has several appealing properties, including (i) privacy with respect to ℓ_2 -sensitivity, as opposed to e.g., ℓ_1 -sensitivity, which is important to retain utility of high-dimensional queries (ii) spherical symmetry, which means that privacy does not depend on the direction of the query sensitivity, (iii) tight, closed-form privacy analysis (Balle & Wang, 2018b), and (iv) tight composition guarantees (Dong *et al.*, 2022; Bun *et al.*, 2018).

Even though the Gaussian mechanism has these desirable properties and has found wide use in practice (Bittau *et al.*, 2017; Apple, 2016; Google, 2023), one may wonder whether it is the *best* mechanism for privately answering real-valued vector queries. Indeed, prior work has explored alternative mechanisms for such queries: It was recently shown that the so-called ℓ_2 -mechanism (Joseph *et al.*, 2025) (essentially, a high-dimensional version of the Laplace mechanism (Dwork *et al.*, 2006b) for ℓ_2 sensitivity) outperforms the Gaussian mechanism in certain low-dimensional settings, while appearing to obtain similar performance in higher dimensions.

However, these two questions remain to be answered:

- (Q1) Is the Gaussian mechanism *optimal* in high dimensions?
- (Q2) Are there mechanisms that outperform *both* the Gaussian and ℓ_2 mechanisms in low dimensions?

Our Contributions. In this paper, we answer both questions affirmatively, through the lens of (ϵ, δ) -DP (Dwork

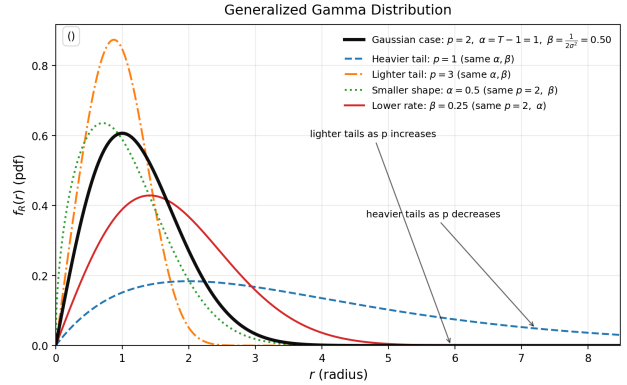
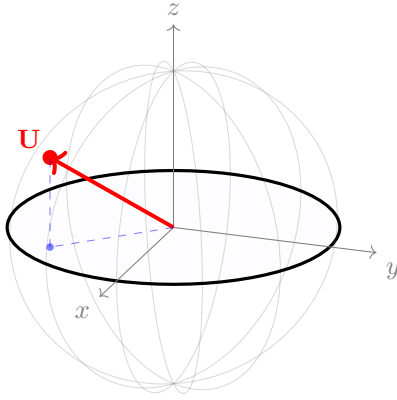


Figure 1. A Spherically Symmetric random variable $X = RU$ has independent (scalar) radial component R and directional component U . **Left:** The directional component U is uniformly distributed on the unit sphere. **Right:** The density functions of one family of radial components R that we study: the *Spherical Generalized Gamma* distribution with varying shape parameter α , rate β , and tail control p .

et al., 2006a). Along the way, we provide additional insights, including answering an open question from (Joseph et al., 2025).

To answer (Q1), we first show that one only need focus on *spherically symmetric* noise distributions (see Figure 1); they perform at least as well as all other additive noise distributions. Indeed, spherical symmetric noise distributions satisfy rotational invariance and are thus a natural choice for the worst-case nature of DP, in which a single data record could perturb the query in any direction; both the Gaussian and ℓ_2 mechanisms use spherical symmetric distributions. A bit more formally, spherical symmetric random variables are of the form $X = RU$, where $R \in \mathbb{R}_{\geq 0}$ is a radial random variable and $U \sim \text{Unif}(\mathcal{S}^{T-1})$ is an independent random T -dimensional unit vector.

Now, consider the Gaussian privacy function $g_{\varepsilon, s}(\sigma^2)$, which for fixed privacy level $\varepsilon > 0$ and ℓ_2 sensitivity $s > 0$, outputs the optimal δ_G of the Gaussian mechanism with per-coordinate variance σ^2 (Balle & Wang, 2018b; Lu et al., 2023). We show that for any fixed Mean Squared Error (MSE) constraint $e = \mathbf{E}x[R^2]$ and any spherical noise $X = RU$, as $T \rightarrow \infty$, the optimal δ that X can achieve for (ε, s) is lower bounded, up to an $o(1)$ term, by $\mathbf{E}x[g(R^2/T)]$, the average over R of the Gaussian privacy function evaluated at the normalized radial energy R^2/T . This value R^2/T for the Gaussian mechanism is exactly equal to the per-coordinate variance $\sigma^2 = e/T$ in the high-dimensional limit. Next, for large enough e , we show that g satisfies certain properties that allow for a Jensen-like inequality

$$\mathbf{E}x[g(R^2/T)] \geq g(\mathbf{E}x[R^2/T]) = g(e/T) = \delta_G.$$

Thus, the Gaussian mechanism cannot be asymptotically

improved upon in the sense of achieving a strictly smaller limiting worst-case optimal delta at the same MSE. We can thus deduce the following (since larger error yields stronger privacy):

Theorem 1.1 (Informal version of Theorem 3.1). *As the dimension of query $Q \in \mathbb{R}^T$, $T \rightarrow \infty$, using Gaussian noise X in the additive noise mechanism $\hat{Q} = Q + X$ yields lowest error among all noise distributions, for all strong enough privacy requirements.*

This result is of particular relevance to training deep learning models with DP (Abadi et al., 2016), in which T is the number of model parameters, which can be quite large.

To answer (Q2), we consider for the first time a broad and expressive subfamily of spherical symmetric distributions that contains those of both the Gaussian and ℓ_2 mechanisms, amongst many others: the *Spherical Generalized Gamma Distributions* (see right of Figure 1). We begin by demonstrating how to numerically evaluate the DP guarantees of these mechanisms. Indeed, by leveraging the radial decomposition perspective — where the spherical noise is separated into its magnitude and directional components — we reduce the privacy analysis to a one-dimensional integration over a smooth, absolutely continuous radial random variable. This structure allows the privacy loss to be evaluated efficiently and with arbitrary precision using standard numerical integration techniques. We furthermore bound the error that can arise from such numerical evaluation to obtain formal DP guarantees.

We leverage this ability in order to find settings and mechanisms for those settings that have less error than *both* the Gaussian and ℓ_2 mechanisms. The spherical generalized gamma distribution is characterized by three parameters:

the shape parameters α and p , and the scale parameter β . We provide an optimization algorithm that finds the parameters which minimize the error of the corresponding additive noise mechanism, for a particular privacy level. Our evaluation shows that in low dimensions, we are able to find α, p, β that result in a mechanism with error up to 15% less than *both* the Gaussian and ℓ_2 mechanisms.

As a bonus, we study tight composition of all Spherical Generalized Gamma Mechanisms using differential privacy accounting techniques. This demonstrates the compatibility of these mechanisms with existing DP frameworks and answers an open question regarding tight composition of the ℓ_2 mechanism (Joseph et al., 2025).¹

Related Work. Previous works have studied the optimality of the Gaussian mechanism: (Dong et al., 2021) show that a restricted class of additive noise mechanisms all converge to the Gaussian mechanism as $T \rightarrow \infty$ (in terms of Gaussian DP (Dong et al., 2022)), and that across all mechanisms in this class, the Gaussian mechanism performs best for *all* dimensions T . However, this class of additive noise mechanisms is quite restrictive, since we know, for example, that the ℓ_2 mechanism outperforms the Gaussian mechanism for low T . Moreover, Gaussian DP is tailored towards the Gaussian mechanism and is a “relaxation” (Dong et al., 2022) of the classical approximate DP that we study.

Works that study the *non*-optimality of the Gaussian mechanism include (Geng & Viswanath, 2016), who only study *integer* valued (vector) queries in terms of ℓ_1 sensitivity, and (Geng et al., 2020), who only study real valued *scalar* queries. The ℓ_2 mechanism extends (Hardt & Talwar, 2010), which only studies pure DP, to approximate DP. Building off the family of Generalized Gaussian Mechanisms from (Liu, 2019), (Rinberg et al., 2025) provide *empirical* evidence that the Gaussian is optimal within this family. However, we are able to show notable improvements of non-Gaussian mechanisms in our Spherical Generalized Gamma Mechanism over the Gaussian mechanism. Moreover, Generalized Gaussian Mechanisms, specified by only *two* parameters, are a proper subset of our Spherical Generalized Gamma Mechanisms, specified by *three* parameters.

The Rank-1 Singular Multivariate Gaussian (R1SMG) mechanism introduced in (Ji & Li, 2024), is a special case within our broader family of spherical generalized gamma mechanisms. The authors claim that R1SMG has error an order of magnitude lower than the Gaussian mechanism; however, we discover a bug in their proof (see Appendix E). Indeed, our numerical evaluation reveals that the privacy guarantees of R1SMG are substantially weaker than those of the standard multivariate Gaussian mechanism.

¹As per the review discussion: <https://openreview.net/forum?id=ypeeAHYK7W>.

The work of (Alghamdi et al., 2023) studies the (ϵ, δ) -DP composition of real valued vector queries. They show that as the number of sequential compositions, k , trends towards infinity, additive, spherical symmetric mechanisms are optimal for ℓ_2 sensitivity. They also provide an optimization algorithm to find the best such mechanism, but *only* in this asymptotic sense as $k \rightarrow \infty$. They provide some empirical evidence for small k of their mechanism’s superiority over the Gaussian mechanism for $T = 10$, but only for very small $\delta = 10^{-8}$ (note: it is already known that the Gaussian mechanism cannot perform well as $\delta \rightarrow 0$, and cannot achieve $\delta = 0$). We on the other hand study the standalone (no composition) setting and show superior mechanisms to the Gaussian for more realistic values of δ like $\delta = 10^{-3}$.

2. Preliminaries

2.1. Differential Privacy

We consider a dataset space \mathcal{X}^* equipped with a neighboring relation \simeq (add/remove of one record). A randomized mechanism $\mathcal{M} : \mathcal{X}^* \rightarrow \mathbb{R}^T$ is (ϵ, δ) -differentially private (Dwork & Roth, 2014) if for all neighboring $G \simeq G'$ and all set $\mathcal{S} \subseteq \mathbb{R}^T$,

$$\Pr[\mathcal{M}(G) \in \mathcal{S}] \leq e^\epsilon \Pr[\mathcal{M}(G') \in \mathcal{S}] + \delta \quad (1)$$

Hockey-stick divergence and optimal δ . Let $(t)_+ \stackrel{\text{def}}{=} \max\{t, 0\}$. For random variables X, Y taking values on the same support \mathbb{R}^T and $\epsilon \geq 0$, define the *hockey-stick divergence* at level ϵ (Balle et al., 2018)) (a.k.a., α -divergence with $\alpha = e^\epsilon$) by

$$H_\epsilon(X, Y) \stackrel{\text{def}}{=} \sup_{\mathcal{S} \subseteq \mathbb{R}^T} \left(\Pr[X \in \mathcal{S}] - e^{-\epsilon} \Pr[Y \in \mathcal{S}] \right)_+,$$

In the context of differential privacy, for a fixed neighboring pair $G \simeq G'$, we interpret $H_\epsilon(\mathcal{M}(G), \mathcal{M}(G'))$ as the *optimal (minimal) δ* for the (one-sided) DP inequality $\Pr[\mathcal{M}(G) \in \mathcal{S}] \leq e^\epsilon \Pr[\mathcal{M}(G') \in \mathcal{S}] + \delta$ holding for all measurable \mathcal{S} . Accordingly, we define the *(one-side) optimal delta* (Lu et al., 2024) at privacy level ϵ

$$\delta_{G, G'}(\epsilon) \stackrel{\text{def}}{=} H_\epsilon(\mathcal{M}(G), \mathcal{M}(G')) \quad (2)$$

and the *(worst-case) optimal delta* (a.k.a., privacy profile (Balle et al., 2018)) of the mechanism at privacy level ϵ as

$$\delta_{\mathcal{M}}(\epsilon) \stackrel{\text{def}}{=} \sup_{G \simeq G'} \delta_{G, G'}(\epsilon). \quad (3)$$

By construction, \mathcal{M} is (ϵ, δ) -DP if and only if $\delta_{\mathcal{M}}(\epsilon) \leq \delta$.

Radial r.v. R	Radial density $f_R(r)$	Noise density $f_X(x) = g(\ x\ _2^2)$ (with $y = \ x\ _2^2$)	Shape $f_X(x) \propto$
$\text{GGamma}(\alpha, \beta, p)$	$\frac{p\beta^{(\alpha+1)/p}}{\Gamma(\frac{\alpha+1}{p})} r^\alpha e^{-\beta r^p}$	$\frac{\Gamma(T/2)}{2\pi^{T/2}} \frac{p\beta^{(\alpha+1)/p}}{\Gamma(\frac{\alpha+1}{p})} y^{\alpha+1-T} e^{-\beta y^p}$	$y^{\alpha+1-T} e^{-\beta y^p}$
$\text{GGamma}(T-1, \frac{1}{2\sigma^2}, 2)$ (Gaussian in \mathbb{R}^T)	$\frac{r^{T-1}}{2^{\frac{T}{2}-1} \Gamma(\frac{T}{2}) \sigma^T} e^{-\frac{r^2}{2\sigma^2}}$	$\frac{1}{(2\pi\sigma^2)^{T/2}} e^{-\frac{y}{2\sigma^2}}$	$e^{-\frac{y}{2\sigma^2}}$
$\text{GGamma}(T-1, \frac{1}{\theta}, 1)$ (ℓ_2 -Laplace in \mathbb{R}^T)	$\frac{r^{T-1}}{\theta^T \Gamma(T)} e^{-\frac{r}{\theta}}$	$\frac{\Gamma(T/2)}{2\pi^{T/2} \theta^T \Gamma(T)} e^{-\frac{y}{\theta}}$	$e^{-\frac{y}{\theta}}$

Table 1. Radial random variables R and the induced spherically symmetric densities $f_X(x) = g(\|x\|_2^2)$ for noise random variable $X = RU$, obtained via Theorem 2.1. The last column reports $g(\|x\|_2^2)$ up to a normalization constant.

2.2. Gaussian Mechanism

We consider the T -dimensional vector-valued query $q : \mathcal{X}^* \rightarrow \mathbb{R}^T$ with ℓ_2 sensitivity

$$s \stackrel{\text{def}}{=} \sup_{G \simeq G'} \|q(G) - q(G')\|_2.$$

For $u > 0$, the Gaussian mechanism (Dwork et al., 2006b; Balle & Wang, 2018b) with per-coordinate variance u is

$$\mathcal{M}_{G,u}(G) \stackrel{\text{def}}{=} q(G) + G, \quad G \sim \mathcal{N}(0, uI_T). \quad (4)$$

Notice that the mechanism's mean-squared error is

$$\max_{G \in \mathcal{X}^*} \mathbf{E} \mathbf{x} [\|\mathcal{M}_{G,u}(G) - q(G)\|_2^2] = \mathbf{E} \mathbf{x} [\|G\|_2^2] = Tu. \quad (5)$$

Denote by $\Phi(\cdot)$ the standard normal CDF. According to the literature (Balle & Wang, 2018a; Lu et al., 2023), the optimal delta of $\mathcal{M}_{G,u}$ at privacy level ε is

$$\delta_{\mathcal{M}_{G,u}}(\varepsilon) = g(u) \stackrel{\text{def}}{=} \Phi\left(-\frac{\varepsilon\sqrt{u}}{s} + \frac{s}{2\sqrt{u}}\right) - e^\varepsilon \Phi\left(-\frac{\varepsilon\sqrt{u}}{s} - \frac{s}{2\sqrt{u}}\right). \quad (6)$$

In particular, for fixed (ε, δ, s) , the minimal Gaussian variance achieving (ε, δ) -DP is

$$u_0(\delta) \stackrel{\text{def}}{=} \inf\{u > 0 : g(u) \leq \delta\}. \quad (7)$$

We will repeatedly use Equation (6) as the baseline privacy-utility tradeoff to compare against other additive-noise mechanisms.

2.3. Spherical Symmetric Distribution

We begin by defining *spherically symmetric* random vectors $X \in \mathbb{R}^T$, which are the main noise distributions studied in this paper.

Definition 2.1 (Spherically Symmetric Distribution (Fang, 2018)²). A random vector $X \in \mathbb{R}^T$ is *spherically symmetric* if there exists a random variable $R \geq 0$ and a random

²The first characterization can be found in Thm 2.1 and the second appears in equation 2.7 of the book (Fang, 2018)

vector U uniformly distributed on the unit sphere S_{++}^T such that X can be decomposed as:

$$X = RU,$$

where R and U are independent. Equivalently, X is *spherically symmetric* if its density function $f_X(x)$ exists and depends only on the norm of x , that is,

$$f_X(x) = g(\|x\|_2^2),$$

for some deterministic function $g : [0, \infty) \rightarrow \mathbb{R}_{\geq 0}$.

The decomposition $X = RU$ separates *radial* (R) and *direction* (U) random variables. Theorem 2.1 shows how to convert a radial density into the induced spherically symmetric density. Table 1 lists some spherical random variables and the corresponding radial random variables studied in the paper.

Theorem 2.1 (Radial-to-Spherical Density Transformation (Fang, 2018)[Thm 2.9]). *Let $X = RU$ be a spherically symmetric random variable. If the random variable R has a density over $\mathbb{R}_{\geq 0}$ and denote its pdf as $f_R(\cdot)$, then,*

$$f_X(x) = g(\|x\|_2^2) = \frac{\Gamma(\frac{T}{2})}{2\pi^{\frac{T}{2}}} \|x\|_2^{1-T} f_R(\|x\|_2).$$

In Section A, we give Proposition A.1 that characterizes the distribution of $\cos \Theta$, where Θ is the angle between the random direction vector U and a fixed reference direction.

3. Gaussian Optimality in High Dimensions

We study the optimality of the Gaussian noise distribution for additive-noise differentially private mechanisms in high dimensions. A natural utility measure for this class is the mean-squared error (MSE), equivalently, the second moment of the additive noise (with zero-centered mean). Our notion of optimality asks which mechanism provides the strongest privacy guarantee under a fixed MSE budget. Concretely, we compare mechanisms at a fixed privacy level ε and sensitivity s by their *optimal delta* $\delta_{\mathcal{M}}(\varepsilon)$ (cf. Equation (3)); smaller $\delta_{\mathcal{M}}(\varepsilon)$ yields stronger privacy.

Theorem 3.1 states our main optimality result. Informally, for a fixed privacy level ε and sensitivity s and a target high-privacy regime with δ sufficiently small, Theorem 3.1 says that, as the query dimension $T \rightarrow \infty$, the Gaussian mechanism is asymptotically optimal among all additive-noise mechanisms with the same MSE: if an additive-noise mechanism matches the Gaussian error budget, then its optimal delta at privacy level ε cannot be smaller than δ .

Theorem 3.1 (Asymptotic Gaussian optimality among additive noises). *Fix $\varepsilon \geq 0$ and $s > 0$. There exists $\delta_\star \in (0, 1)$ such that for all $\delta \in (0, \delta_\star]$ the following holds.*

Let $u_0 = u_0(\delta)$ be the minimal Gaussian variance for which the Gaussian mechanism \mathcal{M}_{G, u_0} is (ε, δ) -DP. For each $T \geq 2$, let $\mathcal{M}_{T, u_0}(\mu_T) \stackrel{\text{def}}{=} \mu_T + X_T$ be any additive-noise mechanism with $\mathbf{E}\mathbf{x} [\|X_T\|_2^2] = Tu_0$. Then,

$$\liminf_{T \rightarrow \infty} \delta_{\mathcal{M}_{T, u_0}}(\varepsilon) \geq \delta_{\mathcal{M}_{G, u_0}}(\varepsilon) = \delta.$$

Equivalently, for every $\eta > 0$, for all sufficiently large T ,

$$\delta_{\mathcal{M}_{T, u_0}}(\varepsilon) \geq \delta - \eta.$$

Numerical lower bounds on the threshold δ_\star . Theorem 3.1 is stated in the privacy regime $\delta \leq \delta_\star$, where $\delta_\star \in (0, 1)$ is a constant. To demonstrate that this regime is non-vacuous and, in particular, includes δ -values typically used in practice, Table 2 reports numerically verified lower bounds on δ_\star for some ε (with sensitivity fixed to $s = 1$). The values in Table 2 provide evidence that the privacy regime $\delta \leq \delta_\star$ covers standard privacy choices such as $\delta \leq 10^{-3}$ (in fact, the range is much bigger), thereby highlighting the practical relevance of Theorem 3.1. Our computation is based on a numerical verification of the tangent-support conditions for the Gaussian privacy function g , explained below (cf. Proposition 3.1).

Finite-dimensional convergence. Theorem 3.1 is an asymptotic statement in the dimension T . In particular, it does not provide a concrete threshold T_0 or a convergence rate for how quickly the Gaussian’s asymptotic optimality kicks in. The only asymptotic loss in the proof enters through Lemma 3.1, where the distribution of a suitably scaled coordinate of a uniform random direction is approximated by a standard normal distribution. Making the result quantitative would require replacing this $o(1)$ approximation gap by an explicit normal-approximation bound and propagating that error through the distinguishing test used in the proof. We leave such finite- T rates as an interesting direction for future work.

3.1. Proof sketch of Theorem 3.1

In this section, we provide a proof sketch of Theorem 3.1. Full details are deferred to Section B.

Table 2. Numerical lower bounds on δ_\star in Theorem 3.1 (for some ε with fixed sensitivity $s = 1$).

ε	δ_\star	ε	δ_\star
0.25	0.736670	4.00	0.416972
0.50	0.706970	8.00	0.292170
1.00	0.649185	16.00	0.197615
2.00	0.549133		

We first focus our discussion on *spherically symmetric* noises, a natural yet broad class in the DP worst-case model: since neighboring datasets can induce sensitivity in arbitrary directions, it is natural to consider perturbations that treat all directions uniformly. We define the class of additive-noise mechanism adding *spherically symmetric* noises as *spherical additive-noise mechanisms*, and formally state it in Definition 3.1.

Definition 3.1 (Spherical additive-noise mechanism). Fix $T \geq 2$ and an ℓ_2 -sensitive query with sensitivity $s > 0$ under the add/remove neighboring relation. A *spherical additive-noise mechanism* is any mechanism of the form

$$\mathcal{M}_{R_T, T}(\mu_T) \stackrel{\text{def}}{=} \mu_T + R_T U_T,$$

where $\mu_T \in \mathbb{R}^T$ is the dataset query result, $U_T \sim \text{Unif}(S^{T-1})$ is uniform on the unit sphere, and $R_T \in \mathbb{R}_{>0}$ is independent of U_T .

Lemma 3.1 states our first observation that, when the query dimension T is sufficiently large, the optimal δ of any spherical additive-noise mechanism $\mathcal{M}_{R_T, T, u_0}$ can be lower bound, up to an asymptotically vanishing error term, in terms of the Gaussian privacy function $g(\cdot)$ (cf. Equation (6)). Specifically, letting R_T denote the radial random variable of the spherical noise, the optimal δ at privacy level ε is lower bounded by $\mathbf{E}\mathbf{x} \left[g \left(\frac{R_T^2}{T} \right) \right]$, which is the Gaussian privacy expression evaluated at the *normalized radial energy* R_T^2/T and averaged over its randomness. This reduction turns the comparison between the Gaussian mechanism and an arbitrary spherical mechanism into a one-dimensional question: under the mean-preserving constraint $\mathbf{E}\mathbf{x} [R_T^2/T] = u_0$ (the MSE budget), can any randomization R_T^2/T make $\mathbf{E}\mathbf{x} [g(R_T^2/T)]$ smaller than $g(\mathbf{E}\mathbf{x} [R_T^2/T]) = g(u_0)$? We note that for Gaussian benchmark, its corresponding R_T^2/T equals to $u_0 \cdot \chi_T^2/T$, which converges to u_0 as $T \rightarrow \infty$.

Lemma 3.1 (Proof is in Section B). *Fix $\varepsilon \geq 0$ and $s > 0$. For each $T \geq 2$, let $U_T \sim \text{Unif}(S^{T-1})$ and let $R_T \in \mathbb{R}_{>0}$ be independent of U_T . For any deterministic $\mu_T \in \mathbb{R}^T$ with $\|\mu_T\|_2 = s$, define $X_T = R_T U_T$ and $Y_T = X_T + \mu_T$. Let δ be the optimal delta with respect to X_T, Y_T at privacy*

level ε . Then, as $T \rightarrow \infty$, it holds that

$$\delta \geq \mathbf{E}x \left[g \left(\frac{R_T^2}{T} \right) \right] + o(1).$$

Intuitively, to prove Lemma 3.1, we lower bound the optimal δ between X_T and Y_T by constructing a set \mathcal{S}_T and bounding $\Pr[X \in \mathcal{S}_T] - e^\varepsilon \Pr[Y \in \mathcal{S}_T]$ from below. We choose \mathcal{S}_T to be a threshold region — an affine test with a mild radius-dependent correction — designed to mirror the affine test for distinguishing Gaussian shifts. Conditioning on $R_T = r$, we analyze the membership probabilities $\Pr[X_T \in \mathcal{S}_T \mid R_T = r]$ and $\Pr[Y_T \in \mathcal{S}_T \mid R_T = r]$, and show that as $T \rightarrow \infty$, they converge to the two Gaussian CDF terms that define the Gaussian privacy function evaluated at $u = r^2/T$. Averaging over R_T then yields the stated asymptotic lower bound in Lemma 3.1.

Proposition 3.1 answers the one-dimensional question induced by Lemma 3.1: if the Gaussian privacy function $g(\cdot)$ satisfies an appropriate supporting-line property at the MSE budget u_0 , then any mean-preserving randomization of the effective variance cannot improve privacy relative to the Gaussian benchmark. Concretely, the supporting-line property requires that g is convex on $[u_0, \infty)$ and that g lies above its tangent line at u_0 on $[0, u_0]$. This is a tailored Jensen-type condition without requiring global convexity: convexity on the upper side enables a Jensen argument, while the additional lower-side tangent bound ensures the same supporting line remains valid even where g need not be convex. A formal proof appears in Section B.

Proposition 3.1 (Proof is in Section B). *Let $\varepsilon \geq 0$, $\delta \in (0, 1)$ and $s > 0$, and recall the Gaussian privacy function $g(\cdot)$ defined in Equation (6). Let $u_0 > 0$ be the unique point such that $g(u_0) = \delta$ (i.e., the minimal Gaussian variance shown in Equation (7)). Define the tangent line function at u_0 by*

$$\ell_\delta(u) \stackrel{\text{def}}{=} g(u_0) + g'(u_0)(u - u_0). \quad (8)$$

If g is convex on $[u_0, \infty)$ and $g(u) \geq \ell(u)$ for all $u \in [0, u_0]$, then for every random variable $U \geq 0$ with $\mathbf{E}x[U] = u_0$, $\mathbf{E}x[g(U)] \geq g(u_0)$.

Lemma 3.2 asserts that, in the high-privacy regime (i.e., for sufficiently small δ), the Gaussian privacy function $g(\cdot)$ satisfies the supporting-line property required by Proposition 3.1. Concretely, for all δ in this regime, the corresponding minimal MSE level $u_0(\delta)$ lies in a well-behaved region of g : the function is convex on $[u_0(\delta), \infty)$ and, moreover, the tangent line at $u_0(\delta)$ lower bounds g throughout the interval $[0, u_0(\delta)]$.

Lemma 3.2 (Proof is in Section B). *Let $\varepsilon \geq 0$, $s > 0$, and recall the Gaussian privacy function $g(\cdot)$ defined in Equation (6). For each $\delta \in (0, 1)$, let $u_0(\delta) > 0$ be the unique*

point such that $g(u_0(\delta)) = \delta$ (i.e., the minimal Gaussian variance shown in Equation (7)). Denote the tangent line function at $u_0(\delta)$ by $\ell_\delta(u)$ (cf. Equation (8)).

Then there exists $\delta_\star \in (0, 1)$, such that for every $\delta \in (0, \delta_\star]$ the following hold:

1. *g is convex on $[u_0(\delta), \infty)$;*
2. *$g(u) \geq \ell_\delta(u)$ for all $u \in [0, u_0(\delta)]$.*

Intuitively, to prove Lemma 3.2, our first step is an asymptotic analysis of $g(u)$ as the MSE budget $u \rightarrow \infty$ (equivalently, $\delta = g(u) \rightarrow 0$): using Mills-type bounds for Φ , we obtain the expansion

$$g(u) = C_{\varepsilon, s} u^{-3/2} \exp\left(-\frac{\varepsilon^2}{2s^2}u\right) \left(1 + O\left(\frac{1}{u}\right)\right),$$

which implies that g is eventually convex, i.e., there exists u_{right} such that g is convex on $[u_{\text{right}}, \infty)$. To control the lower part, we study the tangent-line intercept $L(u) = g(u) - u g'(u)$. The same asymptotics show that $L(u) \rightarrow 0$ as $u \rightarrow \infty$. Since g is decreasing, we can choose a sufficiently large point $u_\star \geq u_{\text{right}}$ so that the tangent line at u_\star has intercept below $\min_{u \in [0, u_{\text{right}}]} g(u)$; this ensures the tangent line at u_\star lies below g on $[0, u_{\text{right}}]$, while convexity guarantees the supporting-line inequality on $[u_{\text{right}}, u_\star]$. Defining $\delta_\star \stackrel{\text{def}}{=} g(u_\star)$ yields the desired supporting-line property at (u_\star, δ_\star) . Finally, we use the downward-closure argument (Proposition B.1) to extend the property from δ_\star to all $\delta \leq \delta_\star$.

Combining Lemmas 3.1 and 3.2 and proposition 3.1 yields the conclusion that, in the high-privacy regime, the Gaussian mechanism is asymptotically optimal among all spherical additive-noise mechanisms with the same MSE. To extend this optimality beyond spherical noises, Lemma 3.3 formalizes the last piece. The idea is to reduce an arbitrary additive noise X to a spherically symmetric one X' via Haar symmetrization.

Lemma 3.3 says that the symmetrization preserves the MSE and cannot increase the worst-direction optimal delta over shifts of ℓ_2 norm at most s . The proof uses convexity of hockey-stick divergence under mixtures together with orthogonal invariance. Therefore, under a fixed MSE budget, the optimal privacy achievable by additive-noise mechanisms is no better than that achievable by spherical additive-noise mechanisms.

Lemma 3.3 (Haar symmetrization (Proof is in Section B)). *Fix $T \geq 2$, $\varepsilon \geq 0$, and $s > 0$. Let M be a Haar-uniform random matrix over the group of all $T \times T$ orthogonal matrices and independent of a random vector $X \in \mathbb{R}^T$. Define the symmetrized noise $X' \stackrel{\text{def}}{=} MX$. Let $\delta \stackrel{\text{def}}{=} \sup_{\|v\|_2 \leq s} H_\varepsilon(X, X + v)$ and $\delta' \stackrel{\text{def}}{=} \sup_{\|v\|_2 \leq s} H_\varepsilon(X', X' + v)$ denote the worst-direction op-*

timal deltas at privacy level ε over the ℓ_2 ball of radius s (cf. Equation (2)).

Then X' is spherically symmetric, and

$$\mathbf{E}x \left[\|X'\|_2^2 \right] = \mathbf{E}x \left[\|X\|_2^2 \right] \quad \delta' \leq \delta.$$

Lemmas 3.1 to 3.3 and proposition 3.1 yields Theorem 3.1 immediately. We defer a formal wrap-up to Section B.

4. Spherical Generalized Gamma Mechanism

In this section, we study the question of additive-noise design in finite (especially low) dimensions, where the *shape* of the noise can materially affect privacy–utility trade-off. Concretely, we formalize the Spherical Generalized Gamma (SGG) noise family and its use as an additive DP primitive for ℓ_2 -sensitive vector queries.

Definition 4.1 defines the three-parameter generalized-gamma family, denoted by $R \sim \text{GGamma}(\alpha, \beta, p)$, which we use to model the *radial* component of our noise. Intuitively, α controls the near-zero behavior (shape), β sets the overall scale (noise magnitude), and p tunes tail decay; see Figure 1 (right) for a visual.

Definition 4.1. A three-parameter **generalized-Gamma** $R \sim \text{GGamma}(\alpha, \beta, p)$ is defined with the density function:

$$f_{\alpha, \beta, p}(r) = \frac{p \beta^{(\alpha+1)/p}}{\Gamma\left(\frac{\alpha+1}{p}\right)} r^\alpha \exp(-\beta r^p), \quad \alpha > -1, p > 0, \beta > 0.$$

Definition 4.2 defines the *spherical generalized-gamma* (SGG) family, induced by a generalized-gamma *radial* distribution. The SGG family includes several standard spherical noises. For instance: (i) setting $(p, \alpha) = (2, T-1)$ and $\beta = 1/(2\sigma^2)$ yields $X \sim \mathcal{N}(0, \sigma^2 I_T)$ (Gaussian noise); (ii) setting $(p, \alpha) = (1, T-1)$ and $\beta = 1/\theta$ yields spherical (i.e., ℓ_2 -) Laplace noise with scale θ .

Definition 4.2 (Spherical generalized-gamma (SGG) noise). Fix a dimension $T \geq 2$ and parameters $\alpha > -1$, $\beta > 0$, $p > 0$. A random vector $X \in \mathbb{R}^T$ follows a *Spherical Generalized-Gamma* (SGG) distribution, denoted by $X \sim \text{SGG}(\alpha, \beta, p)$, if $X = RU$, where $R \sim \text{GGamma}(\alpha, \beta, p)$ and $U \sim \text{Unif}(S^{T-1})$ are independent.

4.1. Privacy Analysis

Lemma 4.1 shows that the optimal delta at privacy level ε between SGG noise X and its shift $X + \mu$ admits an explicit *one-dimensional* integral representation. In particular, the optimal $\delta_\mu^*(\varepsilon)$ can be evaluated by integrating a smooth integrand over $r \in (0, \infty)$, which yields an efficient and numerically stable procedure (in contrast to a direct exponentially expensive computation over \mathbb{R}^T).

Lemma 4.1 (Proof is in Section C.1). Fix $T \geq 2$ and parameters $\alpha \in (-1, T-1]$, $\beta > 0$, and $p > 0$. Let $X = RU \sim \text{SGG}(\alpha, \beta, p)$ and let $\mu \in \mathbb{R}^T$ be a shift vector with $s = \|\mu\|_2$. Let $W = \cos \Theta = \frac{\langle \mu, U \rangle}{\|\mu\|_2} \in [-1, 1]$, with CDF F_W . Then the optimal $\delta_\mu^*(\varepsilon)$ for the neighboring pair $(X, X + \mu)$ satisfies

$$\delta_\mu^*(\varepsilon) = \max \left\{ 0, \int_0^\infty f_R(r) (1 - F_W(w^*(r, -\varepsilon))) dr - e^\varepsilon \int_0^\infty f_R(r) F_W(w^*(r, \varepsilon)) dr \right\},$$

where f_R is the density of $R \sim \text{GGamma}(\alpha, \beta, p)$, and for each $y \in \mathbb{R}$ and $r > 0$, $w^*(r, y) \in [-1, 1]$ denotes the unique solution in w to

$$y = \frac{\alpha + 1 - T}{2} \ln \left(1 + \frac{2sw}{r} + \frac{s^2}{r^2} \right) + \beta \left[r^p - (r^2 + 2swr + s^2)^{p/2} \right].$$

Informally, the proof follows by observing that under the radial–directional decomposition $X = RU$, the privacy loss can be expressed as a function of two *independent* one-dimensional random variables: the radial component R and the directional statistic W (the cosine of the angle between U and μ). Moreover, monotonicity in W converts the resulting two-dimensional inequality involving (R, W) into a one-dimensional threshold condition on W given a fixed R , allowing us to integrate over R and obtain the claimed one-dimensional representation.

Proposition 4.1 shows that for $\alpha \leq T-1$ the SGG density is non-increasing with the radius: points farther from the origin receive no larger probability density.³

Proposition 4.1 (Proof is in Section C.1). Fix a dimension $T \geq 2$ and parameters $\alpha \in (-1, T-1]$, $\beta \geq 0$, $p > 0$. Let $X \sim \text{SGG}(\alpha, \beta, p)$, and let $g : [0, \infty) \rightarrow [0, \infty)$ denote the density generator such that $f_X(x) = g(\|x\|_2^2)$. Then $y \mapsto g(y^2)$ is non-increasing for $y > 0$.

Building on this, Lemma 4.2 shows that, for the neighboring pair $(X, X + \mu)$, the optimal delta $\delta_\mu^*(\varepsilon)$ is monotone in the shift magnitude $\|\mu\|_2$. In particular, larger shifts are easier to distinguish and hence induce larger divergence, so the worst-case divergence over all shifts is attained at the largest feasible $\|\mu\|_2$.

Lemma 4.2 (Proof is in Section C.1). Fix $T \geq 2$ and parameters $\alpha \in (-1, T-1]$, $\beta > 0$, and $p > 0$, and let $X \sim \text{SGG}(\alpha, \beta, p)$. For any shift vector $\mu \in \mathbb{R}^T$, let $\delta_\mu^*(\varepsilon)$ denote the optimal δ for the neighboring pair $(X, X + \mu)$ at point ε . Then for any $\mu_1, \mu_2 \in \mathbb{R}^T$ with $\|\mu_1\|_2 \leq \|\mu_2\|_2$ and any $\varepsilon > 0$, $\delta_{\mu_1}^*(\varepsilon) \leq \delta_{\mu_2}^*(\varepsilon)$.

³We note this concerns $f_X(x)$ over \mathbb{R}^T , not the radial density $f_R(r)$: due to the surface-area factor $r^{-(T-1)}$ in $f_X(x)$, $f_R(r)$ need not be decreasing even if $f_X(x)$ decreases with $r = \|x\|_2$.

Lemma 4.1, together with the monotonicity result in Lemma 4.2, indicates that calibrating an additive SGG mechanism for ℓ_2 -sensitive real-vector queries is, a numerical task: given parameters (α, p) and sensitivity s , one can evaluate $\delta_s^*(\varepsilon)$ as a function of the noise multiplier β , and then choose β to meet a target (ε, δ) guarantee.

However, a naïve implementation of Lemma 4.1 may incur numerical errors (e.g., from truncating an improper integral). Since our goal is to have a *provable* differentially private primitive, we abstract these numerical subtleties through the oracle interface in Definition 4.3. Concretely, given $(\alpha, p, \beta, \varepsilon)$ and a shift magnitude s , the oracle outputs an η -tight *upper bound* on $\delta_\mu^*(\varepsilon)$, for all shifts μ with $\|\mu\|_2 \leq s$. Lemma 4.3 says that such an oracle exists.

Definition 4.3. Fix a dimension $T \geq 2$ and parameters $\alpha \in (-1, T - 1]$, $\beta > 0$, $p > 0$, and $\varepsilon \geq 0$. Let $X \sim \text{SGG}(\alpha, \beta, p)$. Fix an ℓ_2 -sensitivity bound $s \geq 0$ and a target slack $\eta \geq 0$. An algorithm $\text{Oracles}_{\text{SGG}} : (\alpha, p, \beta, \varepsilon, s) \mapsto \hat{\delta}$ is called an η -tight *upper-bound oracle* (for $\delta_\mu^*(\varepsilon)$) if, for every $\varepsilon \geq 0$ and every $\mu \in \mathbb{R}^T$ with $\|\mu\|_2 \leq s$, it outputs $\hat{\delta}$ satisfying

$$\delta_\mu^*(\varepsilon) \leq \text{Oracles}_{\text{SGG}}(\alpha, p, \beta, \varepsilon, \|\mu\|_2) \leq \delta_\mu^*(\varepsilon) + \eta.$$

Lemma 4.3 (Proof is in Section C.2). *There exists a (deterministic) η -tight upper-bound oracle.*

Lemma 4.3 follows by an explicit construction. The oracle reduces the one-dimensional integrals defining $\delta_\mu^*(\varepsilon)$ to expectations under a standard $\Gamma(k, 1)$ measure via the change of variables $Z = \beta R^p$ (with $k = (\alpha + 1)/p$), so that Gamma tail probabilities can be computed in closed form. It then (i) truncates the integral at a data-independent threshold z_{\max} chosen to make the discarded tail contribute at most η_{tail} , and (ii) estimates the remaining truncated expectations using a standard binning scheme.

We then in Section C.2 provide a parameterized differentially private mechanism $\mathcal{M}_{\text{SGG}}^{\alpha, p}$ (cf. Algorithm 4) for vector-valued queries. For fixed (α, p) , the mechanism uses an intuitive bracketing-and-bisection search via $\text{Oracles}_{\text{SGG}}$ to calibrate the noise: it finds (up to tolerance τ) the largest feasible rate parameter β — equivalently, the smallest noise magnitude under the parameterization — subject to the target (ε, δ) constraint, and then outputs the noisy answer $q(G) + X$ with $X \sim \text{SGG}(\alpha, \beta, p)$.

4.2. Utility-Optimized SGG Mechanism

Algorithm 4 shows that for any fixed (α, p) , we can calibrate the rate parameter β so that $\mathcal{M}_{\text{SGG}}^{\alpha, p}$ satisfies the target (ε, δ) guarantee. However, the same privacy level can typically be achieved by many different pairs (α, p) , and these choices can lead to different accuracy. This motivates selecting (α, p) to minimize a prescribed error metric (MSE)

while preserving the same privacy constraint.

4.2.1. FINDING OPTIMAL PARAMETERS

We next provide Algorithm 5 (c.f. Section C.2), that numerically tunes the Spherical Generalized Gamma Mechanism by searching for parameters (α, β, p) that minimize the mean-squared error (MSE) c subject to a target (ε, δ) -privacy constraint. It proceeds by running a binary search over MSE values c . At each binary-search step, it solves the two-dimensional problem of selecting α^*, p^* to minimize the achieved optimal delta δ^* under the fixed noise budget c . For each candidate (α, p) , Algorithm 4 determines whether there exists a noise multiplier β^* that satisfies the MSE constraint c and returns the corresponding δ^* ; we implement the (α, p) search using standard Bayesian optimization. If $\delta^* \leq \delta$, the target parameter, then the noise is decreased in the next iteration; otherwise it is increased. The algorithm uses in the mechanism output the parameters α^*, β^*, p^* for the distribution of radial component R that minimize the MSE c subject to $\delta^* \leq \delta$.

4.3. Advantage over Gaussian and ℓ_2 Mechanisms

Based on Algorithm 5, we use the SGG family as a search space to identify selected low-dimensional regimes where an SGG mechanism outperforms both the Gaussian and ℓ_2 mechanisms. The purpose of this experiment is to demonstrate that there is a low dimensional regimens that we could have improvement over both gaussian and ℓ_2 ; it is not intended to show that SGG uniformly or systematically dominates Gaussian and ℓ_2 across privacy parameters or dimensions.

We know that the ℓ_2 mechanism (high-dimensional version of Laplace) is best suited for $\delta_{\ell_2} \rightarrow 0$, while the Gaussian mechanism is suited for $\delta_G > 0$. Thus, we hypothesize that the SGG Mechanism will have the best advantage at some δ^* that is not too small or too big, for fixed privacy parameter $\varepsilon > 0$, dimension T , and sensitivity s . Based on this intuition, we develop Algorithm 6 (in Section C.4) that takes in ε, T, s and performs a binary search (using Algorithm 5) to find the δ^* at which the MSE of the SGG mechanism has the biggest advantage compared to the smaller of the MSEs of the Gaussian and ℓ_2 mechanisms.

Using Algorithm 6, we identify realistic settings in which the optimal SGG distribution is much different from the Gaussian and ℓ_2 mechanisms (p far from 1 and 2), and has MSE up to 15% lower than them. In Figure 2 we demonstrate our findings. Unsurprisingly given Theorem 3.1, the improvements are larger for small dimension T and get smaller as T increases.

Scope of the low-dimensional improvement. The gains in Figure 2 should be interpreted as a principled existence

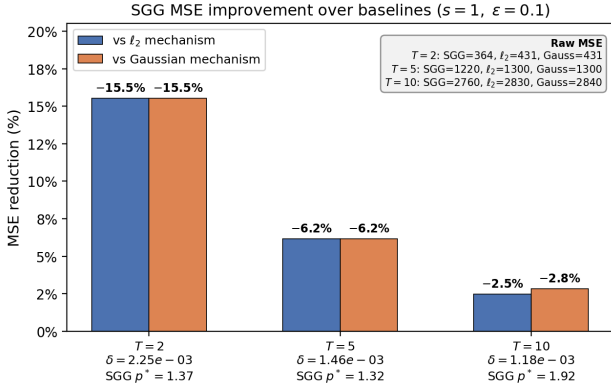


Figure 2. MSE reduction of the SGG mechanism over the ℓ_2 and Gaussian baselines with $\varepsilon = 0.1$ and sensitivity $s = 1$. Each group shows the dimension T , the δ at which this advantage occurs, and the optimal SGG shape parameter p^* .

result rather than as a uniform recommendation to replace the Gaussian mechanism in concrete dimensional mechanism design. Our study does not show strong evidence that non-Gaussian SGG mechanisms consistently yield significant improvements over Gaussian across all dimensions and privacy parameters. In our experiments, the advantage is concentrated in selected low-dimensional regimes and shrinks rapidly as the dimension increases. Thus, SGG is best viewed as a useful design space for finite-dimensional additive-noise mechanisms, while the Gaussian mechanism remains the asymptotically justified default in high dimensions.

This distinction is also important for interpreting the practical implications of the mechanism-level comparison. When the additive vector release itself is the private primitive, the MSE reduction in Figure 2 directly corresponds to improved accuracy in the stated parameter regimes. Examples include low-dimensional mean estimation and low-dimensional vector releases under composition. By contrast, when the additive noise mechanism is used as part in a larger algorithmic design, the same mechanism-level gain need not automatically imply an end-to-end improvement. For high-dimensional private learning methods such as DP-SGD, additional work is needed to determine whether any low-dimensional SGG gains translate into improved training utility

4.4. Privacy Accounting of SGG Mechanisms

Since additive-noise primitives are often invoked repeatedly in DP computations, we study tight composed privacy for multi-call SGG mechanisms. The key observation is that the privacy loss random variables (PRVs) add under independent composition. Thus, once the single-step PRV distribution of an SGG mechanism can be evaluated, the composed privacy profile can be computed by convolution.

Proposition 4.2 (Composition accounting for SGG mech-

anisms). Consider k independent invocations of SGG additive mechanisms, each calibrated for an ℓ_2 -sensitivity bound. Let Z_1, \dots, Z_k denote their worst-case one-step PRVs, as characterized in Lemma D.1. Then the composed privacy profile at level ε_{tot} is

$$\delta_k(\varepsilon_{\text{tot}}) = \mathbf{Ex} \left[\left(1 - \exp \left(\varepsilon_{\text{tot}} - \sum_{i=1}^k Z_i \right) \right)_+ \right].$$

Consequently, any upper bound $\widehat{\delta}_k(\varepsilon_{\text{tot}}) \geq \delta_k(\varepsilon_{\text{tot}})$ provides that the k -fold composition is $(\varepsilon_{\text{tot}}, \widehat{\delta}_k(\varepsilon_{\text{tot}}))$ -DP.

The proof and the SGG-specific reduction of the PRV CDF to a one-dimensional radial integral are given in Section D. Algorithm 7 implements the resulting accountant by discretizing the single-step PRV distribution and computing the k -fold convolution by FFT. Since the ℓ_2 mechanism is a special case of SGG, the same accounting framework also gives a tight accountant for composed ℓ_2 mechanisms.

5. Conclusion

We show that as $T \rightarrow \infty$, the standard Gaussian mechanism is asymptotically best among all additive noise mechanisms. Also, we provide a new family of SGG mechanisms and specific choices within this family that outperform both the Gaussian and recent ℓ_2 mechanism in certain low-dimensional settings. These low-dimensional improvements should be viewed as an existence result: they demonstrate nontrivial finite-dimensional design space, but do not show that non-Gaussian SGG consistently provide meaningful improvement over all parameters, dimensions, or downstream applications. Furthermore, we show tight composition of the SGG mechanisms under the (ε, δ) -differential privacy, and answering an open question of (Joseph et al., 2025) regarding the ℓ_2 mechanism. We view extending the studies beyond additive noise — for example, to data-dependent noise and to non-additive mechanisms — as an important direction for future work.

Acknowledgements

We thank the reviewers for their careful reading and insightful comments, which helped us clarify the scope and presentation of the paper. We are also grateful to Hubert Chan for several helpful discussions on the direction and formulation of our main asymptotic optimality result in Theorem 3.1.

Disclaimer

This paper was prepared for informational purposes in part by the Artificial Intelligence Research group of JPMorgan Chase & Co. and its affiliates ("JP Morgan") and is not

a product of the Research Department of JP Morgan. JP Morgan makes no representation and warranty whatsoever and disclaims all liability, for the completeness, accuracy or reliability of the information contained herein. This document is not intended as investment research or investment advice, or a recommendation, offer or solicitation for the purchase or sale of any security, financial instrument, financial product or service, or to be used in any way for evaluating the merits of participating in any transaction, and shall not constitute a solicitation under any jurisdiction or to any person, if such solicitation under such jurisdiction or to such person would be unlawful.

© 2026 JPMorgan Chase & Co. All rights reserved

Impact Statement

Our work studies theoretical and algorithmic tools for analyzing and calibrating additive-noise mechanisms under (ϵ, δ) -differential privacy. The primary intended impact is to enable practitioners to obtain *tighter* privacy accounting and, consequently, to add less noise for a given privacy target or to certify stronger privacy for a fixed utility level.

We do not anticipate negative societal impacts beyond the general risk that differential privacy may be misapplied in a manner that is inappropriate to the application, threat model, and sensitivity of the data. Additionally, we emphasize that our methods do not prescribe a universal “safe” choice of (ϵ, δ) ; users should select privacy parameters in consultation with domain requirements and established guidance, and should report these choices transparently.

References

Abadi, M., Chu, A., Goodfellow, I. J., McMahan, H. B., Mironov, I., Talwar, K., and Zhang, L. Deep learning with differential privacy. In Weippl, E. R., Katzenbeisser, S., Kruegel, C., Myers, A. C., and Halevi, S. (eds.), *ACM CCS 2016*, pp. 308–318. ACM Press, October 2016. doi: 10.1145/2976749.2978318.

Alghamdi, W., Asoodeh, S., Calmon, F. P., Felipe Gomez, J., Kosut, O., and Sankar, L. Optimal multidimensional differentially private mechanisms in the large-composition regime. In *2023 IEEE International Symposium on Information Theory (ISIT)*, pp. 2195–2200, 2023. doi: 10.1109/ISIT54713.2023.10206658.

Apple. Differential privacy, 2016.

Balle, B. and Wang, Y. Improving the gaussian mechanism for differential privacy: Analytical calibration and optimal denoising. In Dy, J. G. and Krause, A. (eds.), *Proceedings of the 35th International Conference on Machine Learning, ICML 2018, Stockholmsmässan, Stockholm, Sweden, July 10-15, 2018*, volume 80 of *Proceedings of Machine Learning Research*, pp. 403–412. PMLR, 2018a. URL <http://proceedings.mlr.press/v80/balle18a.html>.

Balle, B. and Wang, Y.-X. Improving the gaussian mechanism for differential privacy: Analytical calibration and optimal denoising. In *International Conference on Machine Learning (ICML)*, 2018b.

Balle, B., Barthe, G., and Gaboardi, M. Privacy amplification by subsampling: Tight analyses via couplings and divergences. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2018.

Bittau, A., Úlfar Erlingsson, Maniatis, P., Mironov, I., Raghunathan, A., Lie, D., Rudominer, M., Kode, U., Tinnes, J., and Seefeld, B. Prochlo: Strong privacy for analytics in the crowd. In *Proceedings of the Symposium on Operating Systems Principles (SOSP)*, pp. 441–459, 2017. URL <https://arxiv.org/abs/1710.00901>.

Bun, M., Dwork, C., Rothblum, G. N., and Steinke, T. Composable and versatile privacy via truncated CDP. In Diakonikolas, I., Kempe, D., and Henzinger, M. (eds.), *50th ACM STOC*, pp. 74–86. ACM Press, June 2018. doi: 10.1145/3188745.3188946.

Dong, J., Su, W. J., and Zhang, L. A central limit theorem for differentially private query answering. In *Proceedings of the 35th International Conference on Neural Information Processing Systems, NIPS '21*, Red Hook, NY, USA, 2021. Curran Associates Inc. ISBN 9781713845393.

Dong, J., Roth, A., and Su, W. J. Gaussian differential privacy. *Journal of the Royal Statistical Society Series B*, 84(1):3–37, February 2022. doi: 10.1111/rssb.12454. URL <https://ideas.repec.org/a/bla/jorssb/v84y2022i1p3-37.html>.

Dwork, C. and Roth, A. *The Algorithmic Foundations of Differential Privacy*. Foundations and Trends in Theoretical Computer Science, 2014.

Dwork, C., Kenthapadi, K., McSherry, F., Mironov, I., and Naor, M. Our data, ourselves: Privacy via distributed noise generation. In Vaudenay, S. (ed.), *EUROCRYPT 2006*, volume 4004 of *LNCS*, pp. 486–503. Springer, Berlin, Heidelberg, May / June 2006a. doi: 10.1007/11761679_29.

Dwork, C., McSherry, F., Nissim, K., and Smith, A. Calibrating noise to sensitivity in private data analysis. In Halevi, S. and Rabin, T. (eds.), *TCC 2006*, volume 3876 of *LNCS*, pp. 265–284. Springer, Berlin, Heidelberg, March 2006b. doi: 10.1007/11681878_14.

- Fang, K. W. *Symmetric multivariate and related distributions*. Chapman and Hall/CRC, 2018.
- Geng, Q. and Viswanath, P. Optimal noise adding mechanisms for approximate differential privacy. *IEEE Transactions on Information Theory*, 62(2):952–969, 2016. doi: 10.1109/TIT.2015.2504972.
- Geng, Q., Ding, W., Guo, R., and Kumar, S. Tight analysis of privacy and utility tradeoff in approximate differential privacy. In Chiappa, S. and Candalra, R. (eds.), *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*, volume 108 of *Proceedings of Machine Learning Research*, pp. 89–99. PMLR, 26–28 Aug 2020. URL <https://proceedings.mlr.press/v108/geng20a.html>.
- Google. Learn how gboard gets better, 2023.
- Gopi, S., Lee, Y. T., and Wutschitz, L. Numerical composition of differential privacy. *J. Priv. Confidentiality*, 14(1), 2024. doi: 10.29012/JPC.870. URL <https://doi.org/10.29012/jpc.870>.
- Hardt, M. and Talwar, K. On the geometry of differential privacy. In Schulman, L. J. (ed.), *42nd ACM STOC*, pp. 705–714. ACM Press, June 2010. doi: 10.1145/1806689.1806786.
- Ji, T. and Li, P. Less is more: Revisiting the gaussian mechanism for differential privacy. In Balzarotti, D. and Xu, W. (eds.), *USENIX Security 2024*. USENIX Association, August 2024. URL <https://www.usenix.org/conference/usenixsecurity24/presentation/ji>.
- Joseph, M., Kulesza, A., and Yu, A. Approximate differential privacy of the ℓ_2 mechanism. In *International Conference on Machine Learning (ICML)*, 2025.
- Kurakin, A., Ponomareva, N., Syed, U., MacDermed, L., and Terzis, A. Harnessing large-language models to generate private synthetic text, 2024. URL <https://arxiv.org/abs/2306.01684>.
- Liu, F. Generalized gaussian mechanism for differential privacy. *IEEE Transactions on Knowledge and Data Engineering*, 31(4):747–756, 2019. doi: 10.1109/TKDE.2018.2845388.
- Lu, Y., Magdon-Ismail, M., Wei, Y., and Zikas, V. The normal distributions indistinguishability spectrum and its application to privacy-preserving machine learning, 2023. URL <https://arxiv.org/abs/2309.01243>. arXiv:2309.01243.
- Lu, Y., Magdon-Ismail, M., Wei, Y., and Zikas, V. Eureka: A general framework for black-box differential privacy estimators. In *2024 IEEE Symposium on Security and Privacy (SP)*, pp. 913–931. IEEE, 2024.
- Mckenna, R., Sheldon, D., and Miklau, G. Graphical-model based estimation and inference for differential privacy. In Chaudhuri, K. and Salakhutdinov, R. (eds.), *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pp. 4435–4444. PMLR, 09–15 Jun 2019. URL <https://proceedings.mlr.press/v97/mckenna19a.html>.
- Rinberg, R., Shumailov, I., Singhal, V., Cummings, R., and Papernot, N. Beyond laplace and gaussian: Exploring the generalized gaussian mechanism for private machine learning, 2025. URL <https://arxiv.org/abs/2506.12553>.
- Torkzadehmahani, R., Kairouz, P., and Paten, B. DP-CGAN: Differentially Private Synthetic Data and Label Generation. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 98–104, Los Alamitos, CA, USA, June 2019. IEEE Computer Society. doi: 10.1109/CVPRW.2019.00018. URL <https://doi.ieeecomputersociety.org/10.1109/CVPRW.2019.00018>.
- Yoon, J., Jordon, J., and van der Schaar, M. PATE-GAN: Generating synthetic data with differential privacy guarantees. In *International Conference on Learning Representations*, 2019. URL <https://openreview.net/forum?id=S1zk9iRqF7>.
- Zhang, Z., Wang, T., Li, N., Honorio, J., Backes, M., He, S., Chen, J., and Zhang, Y. PrivSyn: Differentially Private Data Synthesis. In *30th USENIX Security Symposium (USENIX Security 21)*, pp. 929–946. USENIX Association, August 2021. ISBN 978-1-939133-24-3. URL <https://www.usenix.org/conference/usenixsecurity21/presentation/zhang-zhikun>.

A. Additional Preliminaries

Proposition A.1 characterizes the distribution of $\cos \Theta$, where Θ is the angle between the random direction vector U and a fixed reference direction. This random variable appears repeatedly in our analysis. We write $\text{Unif}(\mathcal{S}^{T-1})$ for the uniform distribution on the T -dimensional unit sphere.

Proposition A.1. *Let random variable Θ be the angle between a fixed T -dimensional vector and a T -dimensional random vector $U \sim \text{Unif}(\mathcal{S}^{T-1})$. The density function $f_{\cos \Theta}$ for the random variable $\cos \Theta$ is with the form*

$$f_{\cos \Theta}(u) = \frac{\Gamma(\frac{T}{2})}{\sqrt{\pi}\Gamma(\frac{T-1}{2})} (1-u^2)^{\frac{T-3}{2}}, \quad u \in [-1, 1],$$

and its CDF is with the form⁴

$$\Pr[\cos \Theta \leq u] = I_{\frac{u+1}{2}}\left(\frac{T-1}{2}, \frac{T-1}{2}\right).$$

Proof. Let random variable $X = \frac{\cos \Theta + 1}{2} \in [0, 1]$ so that $\cos \Theta = 2X - 1$. Applying the change of variable formula for pdf, we have

$$\Pr[X = x] = 2 \Pr[\cos \Theta = 2x - 1] = \frac{\Gamma(\frac{T}{2})2^{T-2}}{\sqrt{\pi}\Gamma(\frac{T-1}{2})} (x(1-x))^{\frac{T-3}{2}}.$$

Using Legendre's duplication formula, we observe that

$$\Gamma\left(\frac{T-1}{2}\right)\Gamma\left(\frac{T}{2}\right) = 2^{2-T}\sqrt{\pi}\Gamma(T-1), \quad \Gamma\left(\frac{T}{2}\right) = \frac{2^{2-T}\sqrt{\pi}\Gamma(T-1)}{\Gamma\left(\frac{T-1}{2}\right)}.$$

Plugging it in the expression of the PDF of X , we have

$$\begin{aligned} \Pr[X = x] &= \frac{2^{2-T}\sqrt{\pi}\Gamma(T-1)}{\Gamma\left(\frac{T-1}{2}\right)} 2^{T-2} x^{\frac{T-3}{2}} (1-x)^{\frac{T-3}{2}} \\ &= \frac{\Gamma(T-1)}{\Gamma\left(\frac{T-1}{2}\right)^2} x^{\frac{T-3}{2}} (1-x)^{\frac{T-3}{2}} \\ &= \frac{1}{B\left(\frac{T-1}{2}, \frac{T-1}{2}\right)} x^{\frac{T-1}{2}-1} (1-x)^{\frac{T-1}{2}-1}, \end{aligned}$$

which is the density function of beta distribution with parameter $\alpha = \frac{T-1}{2}, \beta = \frac{T-1}{2}$.

Finally, using the CDF of the beta distribution, we have

$$\Pr[\cos \Theta \leq u] = \Pr[2X - 1 \leq u] = \Pr\left[X \leq \frac{u+1}{2}\right] = I_{\frac{u+1}{2}}\left(\frac{T-1}{2}, \frac{T-1}{2}\right).$$

□

B. Proof of Theorem 3.1

Proof of Lemma 3.1. Recall the definition of optimal δ (Equation (2)) between X, Y at privacy level ε as

$$\delta = \sup_{\mathcal{S} \subseteq \mathbb{R}^T} \left(\Pr[X \in \mathcal{S}] - e^\varepsilon \Pr[Y \in \mathcal{S}] \right)_+.$$

We will explicitly find a set \mathcal{S}_T such that $\left(\Pr[X \in \mathcal{S}] - e^\varepsilon \Pr[Y \in \mathcal{S}] \right)_+$ approaches $\mathbf{Ex} \left[g\left(\frac{R_T^2}{T}\right) \right]$.

⁴ $I_x(a, b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^x t^{a-1} (1-t)^{b-1} dt$ is the regularised incomplete beta function.

By rotational invariance of $U_T \sim \text{Unif}(\mathcal{S}^{T-1})$, WLOG, set $\mu_T = se_1$ (where $e_1 = (1, 0, \dots, 0)$). Then, we set

$$X_T = R_T U_T, \quad Y_T = R_T U_T + se_1.$$

We then define the set \mathcal{S}_T as the following

$$\mathcal{S}_T = \left\{ x \in \mathbb{R}^T : \langle e_1, x \rangle \leq \frac{s}{2} - \frac{\varepsilon \|x\|_2^2}{sT} \right\}.$$

By definition of δ , we know that $\delta \geq \Pr[X \in \mathcal{S}_T] - e^\varepsilon \Pr[Y \in \mathcal{S}_T]$.

In the rest of the proof, we will study the asymptotics of the expression $\Pr[X \in \mathcal{S}_T] - e^\varepsilon \Pr[Y \in \mathcal{S}_T]$. First, look at the asymptotics of the first coordinate of a random direction $\langle e_1, U_T \rangle$. Define

$$W_T = \sqrt{T} \langle e_1, U_T \rangle.$$

Using the Gaussian representation $U_T \stackrel{d}{=} G/\|G\|_2$ with $G \sim \mathcal{N}(0, I_T)$, we have that $W_T = \frac{G_1}{\|G\|_2/\sqrt{T}}$ converges in distribution to a standard normal distribution $\mathcal{N}(0, 1)$. By Pólya's theorem, we have that

$$\lim_{T \rightarrow \infty} \sup_{t \in \mathbb{R}} |\Pr[W_T \leq t] - \Phi(t)| = 0, \quad (9)$$

where Φ is the standard normal cdf.

Then, we study the asymptotics of the term $\Pr[X_T \in \mathcal{S}_T \mid R_T = r]$. Condition on $R_T = r$, we observe that $X_T = rU_T$ and $\|X\|_2^2 = r^2$ is deterministic given r . Therefore, by the definition of \mathcal{S}_T ,

$$X_T \in \mathcal{S}_T \iff r \langle e_1, U \rangle \leq \frac{s}{2} - \frac{\varepsilon}{s} \cdot \frac{r^2}{T} \iff W_T \leq a\left(\frac{r^2}{T}\right),$$

where we denote function $a(u)$ for $u > 0$:

$$a(u) = -\frac{\varepsilon \sqrt{u}}{s} + \frac{s}{2\sqrt{u}}.$$

Using Equation (9), we have that

$$\Pr[X_T \in \mathcal{S}_T \mid R_T = r] = \Phi\left(a\left(\frac{r^2}{T}\right)\right) + o(1),$$

where the $o(1)$ is uniform in r .

Similarly, look at the asymptotics of the term $\Pr[Y_T \in \mathcal{S}_T \mid R_T = r]$. Recall $Y_T = rU_T + se_1$ and $\langle e_1, Y_T \rangle = rU_{T,1} + s$. Then,

$$\begin{aligned} \frac{\|Y\|_2^2}{T} &= \frac{r^2 + s^2 + 2rs \langle e_1, U_T \rangle}{T} \\ &= \frac{r^2}{T} + \frac{s^2}{T} + \frac{2rs}{T} \cdot \frac{W_T}{\sqrt{T}} \\ &= \frac{r^2}{T} + O_p\left(\frac{1}{T}\right) \\ &= \frac{r^2}{T} + o_p(1), \end{aligned}$$

where O_p, o_p is the standard notation for asymptotic stochastic boundedness. Therefore, by the definition of \mathcal{S}_T ,

$$\begin{aligned} Y_T \in \mathcal{S}_T &\iff r\langle e_1, U \rangle \leq -\frac{s}{2} - \frac{\varepsilon}{s} \cdot \frac{r^2}{T} + O_p\left(\frac{1}{T}\right) \\ &\iff W_T \leq b\left(\frac{r^2}{T}\right) + O_p\left(\frac{1}{\sqrt{T}}\right) \\ &\iff W_T \leq b\left(\frac{r^2}{T}\right) + o_p(1), \end{aligned}$$

where we denote function $b(u)$ for $u > 0$:

$$b(u) = -\frac{\varepsilon\sqrt{u}}{s} - \frac{s}{2\sqrt{u}}.$$

Using Equation (9), we have that

$$\Pr[Y_T \in \mathcal{S}_T \mid R_T = r] = \Phi\left(b\left(\frac{r^2}{T}\right)\right) + o(1),$$

where the $o(1)$ is uniform in r .

Taking expectation of $\Pr[Y_T \in \mathcal{S}_T \mid R_T = r]$ and $\Pr[X_T \in \mathcal{S}_T \mid R_T = r]$ over R_T and plugging into $\Pr[X \in \mathcal{S}_T] - e^\varepsilon \Pr[Y \in \mathcal{S}_T]$ gives

$$\begin{aligned} &\Pr[X \in \mathcal{S}_T] - e^\varepsilon \Pr[Y \in \mathcal{S}_T] \\ &= \mathbf{E}\mathbf{x} \left[\Phi\left(a\left(\frac{R_T^2}{T}\right)\right) - e^\varepsilon \Phi\left(b\left(\frac{R_T^2}{T}\right)\right) \right] + o(1), \end{aligned}$$

where the expression $\Phi\left(a\left(\frac{R_T^2}{T}\right)\right) - e^\varepsilon \Phi\left(b\left(\frac{R_T^2}{T}\right)\right)$ is exactly the optimal delta of the Gaussian mechanism $\mathcal{M}_{G,u}$ at privacy level ε (cf. Equation (6)).

Finally, recall that $\delta \geq \left(\Pr[X \in \mathcal{S}_T] - e^\varepsilon \Pr[Y \in \mathcal{S}_T]\right)_+$, we finish the proof. \square

Proof of Proposition 3.1. By convexity of g on $[u_0, \infty)$, the tangent line at u_0 is a global under-estimator on the right, namely, for all $u \geq u_0$,

$$g(u) \geq g(u_0) + g'(u_0)(u - u_0) = \ell_\delta(u),$$

Since we assume that $g(u) \geq \ell_\delta(u)$ for all $u \in [0, u_0]$ in the statement, we have $g(u) \geq \ell_\delta(u)$ for all $u \geq 0$.

Taking expectation with respect to any $U \geq 0$ yields

$$\begin{aligned} \mathbf{E}\mathbf{x} [g(U)] &\geq \mathbf{E}\mathbf{x} [\ell_\delta(U)] \\ &= \mathbf{E}\mathbf{x} [g(u_0) + g'(u_0)(U - u_0)] \\ &= g(u_0) + g'(u_0)(\mathbf{E}\mathbf{x} [U] - u_0) \\ &= g(u_0). \end{aligned}$$

\square

Proposition B.1 (Downward closure of the tangent-support conditions). *Fix $\varepsilon \geq 0$, $\delta \in (0, 1)$, $s > 0$, and let $g : (0, \infty) \rightarrow \mathbb{R}$ be differentiable and strictly decreasing. For any $\delta \in (0, 1)$, Let $u(\delta) > 0$ be the unique point such that $g(u(\delta)) = \delta$ and define the tangent line function at $t > 0$ by*

$$\ell_t(u) = g(t) + g'(t)(u - t).$$

If there exists a $\delta_0 \in (0, 1)$ and $u_0 = u(\delta_0)$ such that: g is convex on $[u_0, \infty)$; and $g(u) \geq \ell_{u_0}(u)$ for all $u \in [0, u_0]$. Then the two conditions hold is downward closed for all $\delta \in (0, \delta_0)$.

Proof. Fix any $\delta \in (0, \delta_0)$ and let $u_1 = u(\delta)$. Since g is strictly decreasing and $g(u_0) = \delta_0 > \delta = g(u_1)$, it follows that $u_1 > u_0$.

It is easy to see that the first condition holds on u_1 . Because $[u_1, \infty) \subseteq [u_0, \infty)$ and g is convex on $[u_0, \infty)$, it is convex on the subset $[u_1, \infty)$ as well.

To show the second condition holds, we split the interval $[0, u_1]$ into two parts and consider each case separately. For any fixed $u \in [0, u_0]$, define function $F(t) = l_t(u)$ for $t \geq u$, whose output is the tangent line function at t evaluating at a fixed point u . Differentiating $F(t)$ with respect to t , we have

$$F'(t) = g'(t) + g''(t)(u - t) - g'(t) = g''(t)(u - t).$$

On $[u_0, \infty)$ we have $g''(t) \geq 0$ by convexity, and since $t \geq u$ we have $u - t \leq 0$. Hence $F'(t) \leq 0$ for all $t \geq u_0$, so F is nonincreasing on $[u_0, \infty)$. Because $u_1 \geq u_0$, $l_{u_1}(u) = F(u_1) \leq F(u_0) = l_{u_0}(u)$. By assumption, $g(u) \geq l_{u_0}(u)$ for all $u \in [0, u_0]$, thus

$$g(u) \geq l_{u_0}(u) \geq l_{u_1}(u), \quad \forall u \in [0, u_0].$$

For the second case, where $u \in [u_0, u_1]$, $h(u) = g(u) - l_{u_1}(u)$. We know that $h'(u) = g'(u) - g'(u_1)$. Since g is decreasing and convex on $[u_0, \infty)$, its derivative g' is nondecreasing there. For any $u_0 \leq u \leq u_1$, this implies $g'(u) \leq g'(u_1)$, hence $h'(u) \leq 0$ on $[u_0, u_1]$. Therefore h is nonincreasing on $[u_0, u_1]$, and for all $u \in [u_0, u_1]$ we have

$$h(u) \geq h(u_1) = 0,$$

i.e., $g(u) \geq l_{u_1}(u)$.

Combining both cases yields $g(u) \geq l_{u_1}(u)$ for all $u \in [0, u_1]$, completing the proof for the second condition. \square

Proof is in Lemma 3.2. Denote function $\phi(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ as the standard normal density. We start from the case where $\varepsilon > 0$. Our first step is to give an asymptotic form for function $g(u)$, in which the asymptotics is in terms of u . Concretely,

$$g(u) = C_{\varepsilon, s} u^{-3/2} \exp\left(-\frac{\varepsilon^2}{2s^2}u\right) \left(1 + O\left(\frac{1}{u}\right)\right), \quad (10)$$

with $C_{\varepsilon, s} = \frac{\varepsilon^{\varepsilon/2} s^3}{\sqrt{2\pi} \varepsilon^2}$.

Let $x = x(u) = \frac{\varepsilon\sqrt{u}}{s}$, and $d = d(u) = \frac{s}{2\sqrt{u}} = \frac{\varepsilon}{2x}$. We rewrite $g(u)$ as $g(u) = \Phi(-(x-d)) - e^\varepsilon \Phi(-(x+d))$. The standard Mills bounds says that for all $z > 0$,

$$\frac{1}{z + z^{-1}} \leq \frac{\Phi(-z)}{\phi(z)} \leq \frac{1}{z}.$$

Then we have

$$0 \leq \frac{1}{z} - \frac{\Phi(-z)}{\phi(z)} \leq \frac{1}{z} - \frac{1}{z + z^{-1}} = \frac{1}{z^3 + z} \leq \frac{1}{z^3},$$

which gives us $\frac{\Phi(-z)}{\phi(z)} = \frac{1}{z} + O(z^{-3})$ as $z \rightarrow \infty$.

Define $F(z) = \frac{\Phi(-z)}{\phi(z)}$. Using $\Phi(-z) = \phi(z)F(z)$ and the identity $\phi(x-d) = e^\varepsilon \phi(x+d)$, we express $g(u)$ as

$$g(u) = \phi(x-d)(F(x-d) - F(x+d)).$$

Applying the asymptotic form of $F(z)$ at $z = x \pm d$, and use that $d = \varepsilon/(2x) = O(1/x)$ as $x \rightarrow \infty$:

$$\begin{aligned} F(x-d) - F(x+d) &= \left(\frac{1}{x-d} - \frac{1}{x+d}\right) + O\left(\frac{1}{x^3} \cdot \frac{d}{x}\right) \\ &= \frac{2d}{x^2 - d^2} + O\left(\frac{1}{x^5}\right) \\ &= \frac{\varepsilon}{x^3} + O\left(\frac{1}{x^5}\right). \end{aligned}$$

Moreover,

$$\begin{aligned}\phi(x-d) &= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x-d)^2}{2}\right) \\ &= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2} + \frac{\varepsilon}{2} - \frac{d^2}{2}\right) \\ &= \frac{e^{\varepsilon/2}}{\sqrt{2\pi}} e^{-x^2/2} \left(1 + O\left(\frac{1}{x^2}\right)\right).\end{aligned}$$

Plugging these into $g(u) = \phi(x-d)(F(x-d) - F(x+d))$ yields

$$g(u) = \frac{e^{\varepsilon/2}}{\sqrt{2\pi}} e^{-x^2/2} \left(\frac{\varepsilon}{x^3} + O\left(\frac{1}{x^5}\right)\right).$$

Finally, substitute x and d in terms of ε, u, s back, we obtain the asymptotic form of $g(u)$ as expressed in Equation (10).

We then show that for the case $\varepsilon > 0$, there exists $u_{\text{right}} > 0$ such that g is convex on $[u_{\text{right}}, \infty)$. Let $k = \varepsilon^2/(2s^2) > 0$, we first rewrite $g(u) = f(u)(1 + \eta(u))$, where $f(u) = C_{\varepsilon,s} u^{-3/2} e^{-ku}$ and $\eta(u) = O(1/u)$ by Equation (10).

Differentiating $(\log g)(u) = (\log f)(u) + \log(1 + \eta(u))$ with respect to u , we have

$$\begin{aligned}(\log g)'(u) &= -k - \frac{3}{2u} + O\left(\frac{1}{u^2}\right) \\ (\log g)''(u) &= \frac{3}{2u^2} + O\left(\frac{1}{u^3}\right).\end{aligned}\tag{11}$$

Using the identity $g(u) = \exp(\log g(u))$, and differentiate both side with respect to u twice, we have

$$g''(u) = g(u) \left((\log g)'(u)^2 + (\log g)''(u) \right).$$

Plugging Equation (11) into the expression of $g''(u)$ above, we have

$$g''(u) = g(u) \left(k^2 + O\left(\frac{1}{u}\right) \right).$$

Since $g(u) > 0$ for all $u > 0$, it is clear to see that there exists $u_{\text{right}} > 0$ such that $g''(u) > 0$ for all $u \geq u_{\text{right}}$, i.e. g is convex on $[u_{\text{right}}, \infty)$.

Next, we define the function $L(u)$ as the y-intercept of the tangent line at u (recall the tangent line is defined in the statement $\ell_\delta(u)$, and u is the $u_0(\delta)$ as defined). Concretely,

$$L(u) = g(u) - u g'(u) = g(u) (1 - u (\log g)'(u)).$$

Plugging Equation (11) into the above, we have $L(u) = g(u)(1 + ku + O(1))$. Since $g(u)$ decays like $u^{-3/2} e^{-ku}$, we have

$$\lim_{u \rightarrow \infty} L(u) = 0.$$

We now are ready to show that for the case $\varepsilon > 0$, there exists a u_\star that satisfies both conditions. Let $m = g(u_{\text{right}})$, and since g is a decreasing function, we know that $m \leq g(u)$ for all $u \in [0, u_{\text{right}}]$. Choose $u_\star \geq u_{\text{right}}$ large enough that

$$L(u_\star) \leq m.$$

Let $\delta_\star = g(u_\star) \in (0, 1)$ and define the intercept function at u_\star as $\ell_\star(u) = g(u_\star) + g'(u_\star)(u - u_\star)$.

We claim that $g(u) \geq \ell_\star(u)$ for all $u \in [0, u_\star]$. First, for $u \in [0, u_{\text{right}}]$, since function $\ell_\star(u)$ is decreasing in u , we have that

$$\ell_\star(u) \leq \ell_\star(0) = g(u_\star) - u_\star g'(u_\star) \leq m \leq g(u).$$

Second, for $u \in [u_{\text{right}}, u_\star]$, convexity of g on $[u_{\text{right}}, \infty)$ implies that $g(u) \geq \ell_\star(u)$. Therefore we prove that for the case $\varepsilon > 0$, there exists a point u_\star such that $g(u) \geq \ell_\star(u)$ on $[0, u_\star]$, and g is convex on $[u_\star, \infty)$.

We notice that the existence also holds for the case $\varepsilon = 0$. Using Taylor expansion of Φ at 0, we can derive an asymptotic form for function $g(u)$ when $\varepsilon = 0$:

$$g(u) = 2 \left(\frac{1}{2} + \phi(0)t + O(t^3) \right) - 1 = \frac{s}{\sqrt{2\pi}} u^{-1/2} + O(u^{-3/2}).$$

Also, we have $L(u) = g(u) - u g'(u) = \frac{3s}{2\sqrt{2\pi}} u^{-1/2} + o(u^{-1/2}) \rightarrow 0$ as $u \rightarrow \infty$. The rest will follow directly from the proof for the case $\varepsilon > 0$.

To finish the proof, we then apply the downward-closure argument (Proposition B.1), and observe that the two conditions that hold at u_\star also hold at $u_0 \geq u_\star$: g is convex on $[u_0, \infty)$ and $g(u) \geq \ell_\delta(u)$ for all $u \in [0, u_0]$.

This proves the lemma. □

Combining Lemmas 3.1 and 3.2 and proposition 3.1 yields the conclusion that, in the high-privacy regime, the Gaussian mechanism is asymptotically optimal among all *spherical* additive-noise mechanisms with the same MSE. Theorem B.1 formalizes this.

Theorem B.1 (Asymptotic Gaussian optimality among spherical noises). *Fix $\varepsilon \geq 0$ and $s > 0$. There exists $\delta_\star \in (0, 1)$ such that for every $\delta \in (0, \delta_\star]$ the following holds. Let $u_0 = u_0(\delta)$ be the minimal Gaussian variance for which the Gaussian mechanism $\mathcal{M}_{\mathcal{G}, u_0}$ is (ε, δ) -DP. For each $T \geq 2$, let $\mathcal{M}_{R_T, T, u_0}$ be any spherically symmetric additive-noise mechanism with*

$$\mathbf{E} \mathbf{x} [R_T^2] = T u_0.$$

Then

$$\liminf_{T \rightarrow \infty} \delta_{\mathcal{M}_{R_T, T, u_0}}(\varepsilon) \geq \delta_{\mathcal{M}_{\mathcal{G}, u_0}}(\varepsilon) = \delta.$$

Equivalently, for every $\eta > 0$, for all sufficiently large T ,

$$\delta_{\mathcal{M}_{R_T, T, u_0}}(\varepsilon) \geq \delta - \eta.$$

Proof. Fix $\varepsilon \geq 0$, $s > 0$, and let δ_\star be as in Lemma 3.2. Fix any $\delta \in (0, \delta_\star]$, and let $u_0 = u_0(\delta) > 0$ be the (unique) solution to $g(u_0) = \delta$. Recall that, we define Gaussian privacy function g for $u > 0$ (cf. Equation (6)),

$$g(u) = \Phi \left(-\frac{\varepsilon\sqrt{u}}{s} + \frac{s}{2\sqrt{u}} \right) - e^\varepsilon \Phi \left(-\frac{\varepsilon\sqrt{u}}{s} - \frac{s}{2\sqrt{u}} \right).$$

Now consider the spherical mechanism $\mathcal{M}_{R_T, T, u_0}(\mu_T) = \mu_T + R_T U_T$ with $\mathbf{E} \mathbf{x} [R_T^2] = T u_0$. Let $\mu_T \in \mathbb{R}^T$ be any shift with $\|\mu_T\|_2 = s$, and define neighboring output distribution

$$X_T = R_T U_T, \quad Y_T = X_T + \mu_T.$$

Let $\delta_T^{\text{hid}}(\varepsilon)$ denote the optimal delta at level ε for the pair (X_T, Y_T) . By definition of $\delta_\varepsilon(\mathcal{M}_{R_T, T, u_0})$ as the worst-case delta over neighboring databases, we know that for every choice of μ_T with $\|\mu_T\|_2 = s$, we have that

$$\delta_\varepsilon(\mathcal{M}_{R_T, T, u_0}) \geq \delta_T^{\text{hid}}(\varepsilon).$$

We compare the privacy parameter $\delta_T^{\text{hid}}(\varepsilon)$ to that of the Gaussian benchmark $\mathcal{M}_{\mathcal{G}}$.

Let $U = R_T^2/T$. By Lemma 3.1, we have

$$\begin{aligned} \delta_T^{\text{hid}}(\varepsilon) &\geq \mathbf{E}\mathbf{x} [g(U)] + o(1) \\ &= \mathbf{E}\mathbf{x} \left[g \left(\frac{R_T^2}{T} \right) \right] + o(1), \quad T \rightarrow \infty, \end{aligned}$$

where $o(1) \rightarrow 0$ as $T \rightarrow \infty$ for fixed (ε, s) .

It remains to lower bound $\mathbf{E}\mathbf{x} \left[g \left(\frac{R_T^2}{T} \right) \right]$ under the second-moment constraint $\mathbf{E}\mathbf{x} \left[\frac{R_T^2}{T} \right] = u_0$. Apply Proposition 3.1 with the random variable U . By Lemma 3.2 (with $\delta \leq \delta_*$), the function g satisfies the convexity and tangent-line conditions at u_0 , hence Proposition 3.1 yields

$$\mathbf{E}\mathbf{x} [g(U)] \geq g(u_0) = \delta. \quad (12)$$

Combining the above two inequalities gives

$$\delta_T^{\text{hid}}(\varepsilon) \geq \mathbf{E}\mathbf{x} [g(U)] + o(1) \geq g(u_0) + o(1) = \delta + o(1).$$

Therefore,

$$\liminf_{T \rightarrow \infty} \delta_\varepsilon(\mathcal{M}_{R_T, T, u_0}) \geq \liminf_{T \rightarrow \infty} \delta_T^{\text{hid}}(\varepsilon) \geq \delta,$$

which is exactly the claimed statement. □

Proof of Lemma 3.3. Recall that, by Lemma 3.3's definition, M is a Haar-uniform random matrix over the group of all $T \times T$ orthogonal matrices and independent of a random vector $X \in \mathbb{R}^T$. Let $X' = MX$.

We first show that X' is spherically symmetric. Fix any deterministic $T \times T$ orthogonal matrix Q . We know that $QM \stackrel{d}{=} M$ by Haar-uniform random matrix's property. Therefore, we have

$$QX' = QMX \stackrel{d}{=} MX = X',$$

which, by definition, says that X' is spherically symmetric.

Since orthogonal transformations preserve the Euclidean norm, so $\|X'\|_2 = \|MX\|_2 = \|X\|_2$, and therefore

$$\mathbf{E}\mathbf{x} \left[\|X'\|_2^2 \right] = \mathbf{E}\mathbf{x} \left[\|X\|_2^2 \right].$$

It remains to prove $\delta' \leq \delta$. Fix any $\mu \in \mathbb{R}^T$ with $\|\mu\|_2 \leq s$. Let ν denote Haar measure over the orthogonal group. For each orthogonal matrix O , define P_O, Q_O as the distribution of OX and $OX + \mu$, respectively. By convexity of hockey-stick divergence under mixtures,

$$\mathbf{H}_\varepsilon(X', X' + \mu) \leq \int \mathbf{H}_\varepsilon(P_O, Q_O) d\nu(O).$$

For every fixed orthogonal matrix O , applying the bijection $x \mapsto O^\top x$ to both distributions gives

$$\mathbf{H}_\varepsilon(OX, OX + \mu) = \mathbf{H}_\varepsilon(X, X + O^\top \mu).$$

Since $\|O^\top \mu\|_2 = \|\mu\|_2 \leq s$, the definition of δ gives $\mathbf{H}_\varepsilon(X, X + O^\top \mu) \leq \delta$. Therefore,

$$\mathbf{H}_\varepsilon(X', X' + \mu) \leq \delta.$$

Taking the supremum over all μ with $\|\mu\|_2 \leq s$ gives $\delta' \leq \delta$. □

To finish the proof of Theorem 3.1, we further use Lemma 3.3, which formalize that, under a fixed MSE budget, the optimal privacy achievable by additive-noise mechanisms is no better than that achievable by spherical additive-noise mechanisms.

Proof of Theorem 3.1. Fix $\varepsilon \geq 0$ and $s > 0$, and let δ_* be as in Lemma 3.2. Fix any $\delta \in (0, \delta_*]$, and let $u_0 = u_0(\delta) > 0$ be the unique solution to $g(u_0) = \delta$, where (cf. Equation (6))

$$g(u) = \Phi\left(-\frac{\varepsilon\sqrt{u}}{s} + \frac{s}{2\sqrt{u}}\right) - e^\varepsilon \Phi\left(-\frac{\varepsilon\sqrt{u}}{s} - \frac{s}{2\sqrt{u}}\right).$$

Now fix any $T \geq 2$ and consider an arbitrary additive-noise mechanism

$$\mathcal{M}_{T,u_0}(\mu_T) = \mu_T + X_T,$$

where $X_T \in \mathbb{R}^T$ is a random vector satisfying the MSE constraint $\mathbf{E}x[\|X_T\|_2^2] = Tu_0$. Let

$$\delta_T(\varepsilon) \stackrel{\text{def}}{=} \sup_{\|v\|_2 \leq s} \mathbf{H}_\varepsilon(X_T, X_T + v).$$

By definition of the worst-case optimal delta for the additive-noise mechanism, $\delta_{\mathcal{M}_{T,u_0}}(\varepsilon) = \delta_T(\varepsilon)$. Let $X'_T = M_T X_T$ be the symmetrized noise of X_T , where M_T is the Haar-uniform random matrix over the group of all $T \times T$ orthogonal matrices. Define

$$\delta'_T(\varepsilon) \stackrel{\text{def}}{=} \sup_{\|v\|_2 \leq s} \mathbf{H}_\varepsilon(X'_T, X'_T + v).$$

By Lemma 3.3, we know X'_T is spherically symmetric, has the same MSE as X_T , and moreover

$$\delta'_T(\varepsilon) \leq \delta_T(\varepsilon). \tag{13}$$

Since X'_T is spherically symmetric, there exists a nonnegative random variable $R_T > 0$ and an independent $U_T \sim \text{Unif}(\mathcal{S}^{T-1})$ such that $X'_T \stackrel{d}{=} R_T U_T$, and $\mathbf{E}x[R_T^2] = Tu_0$. Therefore, the spherical optimality result Theorem B.1 applies to the spherical mechanism $\mu_T + X'_T$ and yields

$$\liminf_{T \rightarrow \infty} \delta'_T(\varepsilon) \geq \delta_{\mathcal{M}_{G,u_0}}(\varepsilon) = \delta. \tag{14}$$

Combining Equation (13) and Equation (14) gives

$$\liminf_{T \rightarrow \infty} \delta_{\mathcal{M}_{T,u_0}}(\varepsilon) = \liminf_{T \rightarrow \infty} \delta_T(\varepsilon) \geq \liminf_{T \rightarrow \infty} \delta'_T(\varepsilon) \geq \delta.$$

This is exactly the claimed statement. □

C. Supplementary Material for Spherical Generalized Gamma Mechanism

C.1. Numerical Bound

Proof of Lemma 4.1. Let Y_0, Y_1, X be i.i.d. with $X = RU \sim \text{SGG}(\alpha, \beta, p)$, where $R \sim \text{GGamma}(\alpha, \beta, p)$ and $U \sim \text{Unif}(\mathcal{S}^{T-1})$ are independent. Define Θ the angle between μ and U , and recall $W = \cos \Theta \in [-1, 1]$, whose CDF F_W given in Proposition A.1.

Let f_X denote the density of X . Since $Y_0 \stackrel{d}{=} X$ and $Y_1 + \mu \stackrel{d}{=} X + \mu$, we have $f_{Y_0}(x) = f_X(x)$ and $f_{Y_1 + \mu}(x) = f_X(x - \mu)$. By the privacy-loss random variable characterization of the optimal $\delta(\varepsilon)$ by a fixed ε ((Balle & Wang, 2018a)),

$$\delta^*(\varepsilon) = \left[\Pr(L^{\text{neg}}(X) \leq -\varepsilon) - e^\varepsilon \Pr(L^{\text{pos}}(X) \geq \varepsilon) \right]_+,$$

where

$$\begin{aligned} L^{\text{pos}}(X) &= \ln \frac{f_{Y_1 + \mu}(X)}{f_{Y_0}(X)} = \ln \frac{f_X(X - \mu)}{f_X(X)}, \\ L^{\text{neg}}(X) &= \ln \frac{f_{Y_0}(X + \mu)}{f_{Y_1 + \mu}(X + \mu)} = \ln \frac{f_X(X + \mu)}{f_X(X)}. \end{aligned}$$

Spherical symmetry reduces the PLRV to a function of (R, W) . By Theorem 2.1, the SGG density is spherically symmetric, i.e., $f_X(x) = g(\|x\|_2^2)$ for a suitable generator g . Therefore,

$$\begin{aligned} L^{\text{pos}}(X) &= \ln \frac{g(\|X - \mu\|_2^2)}{g(\|X\|_2^2)} = \ln \frac{g(R^2 - 2RsW + s^2)}{g(R^2)}, \\ L^{\text{neg}}(X) &= \ln \frac{g(\|X + \mu\|_2^2)}{g(\|X\|_2^2)} = \ln \frac{g(R^2 + 2RsW + s^2)}{g(R^2)}. \end{aligned}$$

Moreover, since $U \sim \text{Unif}(\mathcal{S}^{T-1})$ implies $W \stackrel{d}{=} -W$, we have $L^{\text{pos}}(R, W) = L^{\text{neg}}(R, -W)$ and hence $L^{\text{pos}}(X) \stackrel{d}{=} L^{\text{neg}}(X)$. Consequently, we may work with the single privacy loss

$$L(X) = L^{\text{neg}}(X),$$

and write

$$\delta^*(\varepsilon) = \left[\Pr[L(X) \leq -\varepsilon] - e^\varepsilon \Pr[L(X) \geq \varepsilon] \right]_+.$$

For $X \sim \text{SGG}(\alpha, \beta, p)$, Table 1 yields

$$f_X(x) \propto \|x\|_2^{\alpha+1-T} e^{-\beta\|x\|_2^p}, \quad \text{so} \quad g(t) \propto t^{\frac{\alpha+1-T}{2}} e^{-\beta t^{p/2}}.$$

Substituting this g into $L(X) = \ln(g(\|X + \mu\|_2^2)/g(\|X\|_2^2))$ gives

$$L(X) = \ell(R, W),$$

where for $r > 0$ and $w \in [-1, 1]$,

$$\ell(r, w) = \frac{\alpha + 1 - T}{2} \ln\left(1 + \frac{2sw}{r} + \frac{s^2}{r^2}\right) + \beta \left[r^p - (r^2 + 2swr + s^2)^{p/2} \right].$$

Monotonicity in w (for $\alpha \leq T - 1$). Since the lemma assumes $\alpha \in (-1, T - 1]$, we have $\alpha + 1 - T \leq 0$. Let $q(r, w) = r^2 + 2swr + s^2 > 0$. A direct calculation yields

$$\frac{\partial \ell}{\partial w}(r, w) = sr \left[\frac{\alpha + 1 - T}{q(r, w)} - \beta p q(r, w)^{\frac{p}{2}-1} \right] = \frac{sr}{q(r, w)} \left[(\alpha + 1 - T) - \beta p q(r, w)^{p/2} \right].$$

Because $sr/q(r, w) > 0$, $\beta p q(r, w)^{p/2} > 0$, and $(\alpha + 1 - T) \leq 0$, the bracketed term is strictly negative for all $r > 0$ and $w \in [-1, 1]$. Hence $\partial \ell / \partial w(r, w) < 0$ on $(0, \infty) \times [-1, 1]$, so for each fixed $r > 0$ the map $w \mapsto \ell(r, w)$ is strictly decreasing on $[-1, 1]$. Therefore, for any $y \in \mathbb{R}$ and $r > 0$ there is a unique threshold $w^*(r, y) \in [-1, 1]$ satisfying $\ell(r, w^*(r, y)) = y$.

One-dimensional integral expressions. Fix $y \in \mathbb{R}$. By strict monotonicity, conditional on $R = r$,

$$\{\ell(r, W) \leq y\} \iff \{W \geq w^*(r, y)\}, \quad \{\ell(r, W) \geq y\} \iff \{W \leq w^*(r, y)\}.$$

Using independence of R and W ,

$$\begin{aligned} \Pr(L(X) \leq -\varepsilon) &= \int_0^\infty f_R(r) (1 - F_W(w^*(r, -\varepsilon))) \, dr, \\ \Pr(L(X) \geq \varepsilon) &= \int_0^\infty f_R(r) F_W(w^*(r, \varepsilon)) \, dr. \end{aligned}$$

Substituting into $\delta^*(\varepsilon) = [\Pr(L \leq -\varepsilon) - e^\varepsilon \Pr(L \geq \varepsilon)]_+$ yields the desired one-dimensional representation stated in Lemma 4.1. \square

Proof of Proposition 4.1. By Theorem 2.1 and Definition 4.1, for $y = \|x\|_2 > 0$ we have

$$f_X(x) = g(y^2) = C_{\alpha,\beta,p,T} y^{\alpha+1-T} \exp(-\beta y^p),$$

for a constant $C_{\alpha,\beta,p,T} > 0$ that does not depend on y . Differentiate with respect to y :

$$\frac{d}{dy}g(y^2) = g(y^2) \left(\frac{\alpha + 1 - T}{y} - \beta p y^{p-1} \right).$$

Since $\alpha \leq T - 1$ implies $\alpha + 1 - T \leq 0$, and since $\beta p y^{p-1} \geq 0$ for $y > 0$, the bracketed term is non-positive for all $y > 0$. Therefore $\frac{d}{dy}g(y^2) \leq 0$ for all $y > 0$, i.e., $y \mapsto g(y^2)$ is non-increasing. This is equivalent to $t \mapsto g(t)$ being non-increasing on $(0, \infty)$ under the change of variables $t = y^2$. \square

Proof of Lemma 4.2. Write P_μ for the distribution of $X + \mu$ and P_0 for the distribution of X . Since X is spherically symmetric, the quantity $\delta_\mu^*(\varepsilon)$ depends on μ only through $s = \|\mu\|_2$; hence it suffices to prove monotonicity in s . Let f denote the density of X and note that P_μ has density $f_\mu(x) = f(x - \mu)$. Recall that the optimal δ at level ε can be expressed as

$$\delta_\mu^*(\varepsilon) = \int_{\mathbb{R}^T} (f(x) - e^\varepsilon f(x - \mu))_+ dx.$$

Equivalently, let $c = e^\varepsilon$ and using $(a - b)_+ = a - \min\{a, b\}$ for $a, b \geq 0$, we may rewrite

$$\delta_\mu^*(\varepsilon) = 1 - \int_{\mathbb{R}^T} \min\{f(x), cf(x - \mu)\} dx.$$

Using the identity $\min\{a, b\} = \int_0^\infty \mathbf{1}\{a \geq t\} \mathbf{1}\{b \geq t\} dt$ for $a, b \geq 0$, we have

$$\begin{aligned} \int_{\mathbb{R}^T} \min\{f(x), cf(x - \mu)\} dx &= \int_0^\infty \lambda(\{x : f(x) \geq t\} \cap \{x : cf(x - \mu) \geq t\}) dt \\ &= \int_0^\infty \lambda(A_t \cap (\mu + A_{t/c})) dt, \end{aligned}$$

where $\lambda(\cdot)$ denotes Lebesgue measure and $A_t = \{x : f(x) \geq t\}$.

By Proposition 4.1, the radial profile $y \mapsto g(y^2)$ is non-increasing for $\alpha \in (-1, T - 1]$. Since $f(x) = g(\|x\|_2^2)$ depends only on $\|x\|_2$ and is non-increasing in the radius, each superlevel set A_t is a (possibly empty) Euclidean ball centered at the origin: there exists a radius $\rho(t) \in [0, \infty]$ such that $A_t = B(0, \rho(t))$. Likewise, $A_{t/c} = B(0, \rho(t/c))$. Therefore, for each $t > 0$,

$$\lambda(A_t \cap (\mu + A_{t/c})) = \lambda(B(0, \rho(t)) \cap B(\mu, \rho(t/c))).$$

A standard geometric fact is that, overlap of two balls decreases as their center separate. Formally, for fixed radii $r_1, r_2 \geq 0$, the function

$$\lambda(B(0, r_1) \cap B(se_1, r_2))$$

is non-increasing in $s \geq 0$, where $e_1 = \frac{\mu}{\|\mu\|_2}$ is a fixed reference direction. Since $\lambda(B(0, r_1) \cap B(\mu, r_2))$ depends on μ only through $s = \|\mu\|_2$, it follows that $\lambda(A_t \cap (\mu + A_{t/c}))$ is non-increasing in $s = \|\mu\|_2$ for every t . Plugging this shows that the overlap integral $\int \min\{f(x), cf(x - \mu)\} dx$ is non-increasing in s .

Finally, $\delta_\mu^*(\varepsilon)$ is therefore non-decreasing in $s = \|\mu\|_2$. Hence, if $\|\mu_1\|_2 \leq \|\mu_2\|_2$, then $\delta_{\mu_1}^*(\varepsilon) \leq \delta_{\mu_2}^*(\varepsilon)$, as claimed. \square

Algorithm 1 η -tight upper-bound oracle $\text{Oracles}_{\text{SGG}}$

```

1: Input: Parameters  $(T, \alpha, p, \beta, \varepsilon)$  with  $\alpha \in (-1, T - 1]$ ,  $p > 0$ ,  $\beta > 0$ , sensitivity bound  $s \geq 0$ , target slack  $\eta \geq 0$ .
2: Output:  $\hat{\delta} = \text{Oracles}_{\text{SGG}}(\alpha, p, \beta, \varepsilon, s)$ .
3:  $k \leftarrow (\alpha + 1)/p$ ,  $\eta_{\text{tail}} \leftarrow \eta/3$ ,  $\eta_{\text{int}} \leftarrow \eta - \eta_{\text{tail}}$ 
4: Compute (by bisection) any  $z_{\text{max}} > 0$  such that  $(1 + e^\varepsilon)Q(k, z_{\text{max}}) \leq \eta_{\text{tail}}$ .

    Adaptive probability binning on  $[0, z_{\text{max}}]$  to bracket the truncated integrals.
5:  $\mathcal{P} \leftarrow \{[0, z_{\text{max}}]\}$  // current partition into bins
6: while true do
7: // For each bin  $I_i = [z_i, z_{i+1}]$ , compute its Gamma mass and a binwise bracket on  $\tilde{g}_\pm$ .
8: Initialize accumulators:  $\underline{A} \leftarrow 0$ ,  $\overline{A} \leftarrow 0$ ,  $\underline{B} \leftarrow 0$ ,  $\overline{B} \leftarrow 0$ .
9: Initialize bin-score list (for refinement):  $\text{score}(I) \leftarrow 0$  for all  $I \in \mathcal{P}$ .
10: for each  $I_i = [z_i, z_{i+1}] \in \mathcal{P}$  do
11:  $\pi_i \leftarrow P(k, z_{i+1}) - P(k, z_i)$  //  $P$  denotes regularized lower incomplete gamma function
12:  $[\underline{g}_{-,i}, \overline{g}_{-,i}] \leftarrow \text{EncloseG}(I_i, -\varepsilon; T, \alpha, p, \beta, \varepsilon, s)$ 
13:  $[\underline{g}_{+,i}, \overline{g}_{+,i}] \leftarrow \text{EncloseG}(I_i, \varepsilon; T, \alpha, p, \beta, \varepsilon, s)$ 
14:  $\underline{A} \leftarrow \underline{A} + \pi_i \underline{g}_{-,i}$ ;  $\overline{A} \leftarrow \overline{A} + \pi_i \overline{g}_{-,i}$ 
15:  $\underline{B} \leftarrow \underline{B} + \pi_i \underline{g}_{+,i}$ ;  $\overline{B} \leftarrow \overline{B} + \pi_i \overline{g}_{+,i}$ 
16:  $\text{score}(I_i) \leftarrow \pi_i(\overline{g}_{-,i} - \underline{g}_{-,i}) + e^\varepsilon \pi_i(\overline{g}_{+,i} - \underline{g}_{+,i})$ 
17: end for
18: if  $(\overline{A} - \underline{A}) + e^\varepsilon(\overline{B} - \underline{B}) \leq \eta_{\text{int}}$  then break
19: // Refine the partition by splitting the bin with largest contribution to the bracket width.
20: Let  $I^* = [z_L, z_U] \in \mathcal{P}$  be a bin maximizing  $\text{score}(I)$ .
21:  $z_M \leftarrow (z_L + z_U)/2$ .
22:  $\mathcal{P} \leftarrow (\mathcal{P} \setminus \{I^*\}) \cup \{[z_L, z_M], [z_M, z_U]\}$ .
23: end while

    // Add the tail bound and output an upper bound on  $\delta_\mu^*(\varepsilon)$ .
24:  $\tau \leftarrow Q(k, z_{\text{max}})$  //  $Q$  denotes regularized upper incomplete gamma function
25:  $\hat{\delta} \leftarrow \max\{0, (\overline{A} + \tau) - e^\varepsilon \underline{B}\}$ 
26: Output:  $\hat{\delta}$ .
    
```

C.2. Provable DP Guarantee and Mechanism

We first provide intuition for the proof of Lemma 4.3 (existence of an η -tight upper-bound oracle $\text{Oracles}_{\text{SGG}}$). It follows by an explicit construction. At a high level, the oracle reduces the one-dimensional integrals defining $\delta_\mu^*(\varepsilon)$ to expectations under a standard $\Gamma(k, 1)$ measure via the change of variables $Z = \beta R^p$ (with $k = (\alpha + 1)/p$), so that Gamma tail probabilities can be computed in closed form. It then (i) truncates the integral at a data-independent threshold z_{max} chosen to make the discarded tail contribute at most η_{tail} , and (ii) brackets the remaining truncated expectations by an adaptive probability-binning scheme: on each bin, the integrand is provably upper and lower bounded by an interval $[\underline{g}, \overline{g}]$, and these bounds are aggregated with the bin masses π_i to obtain global lower/upper bounds whose total width is provably controlled. We now give the formal proof.

Proof of Lemma 4.3. We prove the lemma constructively by showing that Algorithm 1 implements an η -tight upper-bound oracle defined in Definition 4.3.

Fix any dimension $T \geq 2$, parameters $\alpha \in (-1, T - 1]$, $p > 0$, $\beta > 0$, privacy level $\varepsilon \geq 0$, and any shift vector $\mu \in \mathbb{R}^T$ with $\|\mu\|_2 = s$. Let $X \sim \text{SGG}(\alpha, \beta, p)$. By Lemma 4.1, the optimal δ for the neighboring pair $(X, X + \mu)$ can be written as

$$\delta_\mu^*(\varepsilon) = \max\{0, A - e^\varepsilon B\}, \quad (15)$$

where $A = \mathbf{Ex}[g_-(R)]$ and $B = \mathbf{Ex}[g_+(R)]$ are expectations under $R \sim \text{Gamma}(\alpha, \beta, p)$, and $g_\pm(\cdot) \in [0, 1]$ are the

Algorithm 2 ENCLOSEG: Binwise bounds on $\tilde{g}_{\pm}(z)$ over $I = [z_L, z_U]$

```

1: Input: Bin  $I = [z_L, z_U]$ , target  $y \in \{\varepsilon, -\varepsilon\}$ , parameters  $(T, \alpha, p, \beta, \varepsilon)$ , sensitivity bound  $s$ 
2: Parameter: bisection tolerance  $\tau_w > 0$ .
3: Output: Bounds  $[\underline{g}, \bar{g}]$  such that  $\tilde{g}(z) \in [\underline{g}, \bar{g}]$  for all  $z \in I$ .
4:  $r_L \leftarrow (z_L/\beta)^{1/p}$ ,  $r_U \leftarrow (z_U/\beta)^{1/p}$ ,  $w^{\text{lo}} \leftarrow -1$ ,  $w^{\text{hi}} \leftarrow 1$ ,  $\text{ok} \leftarrow \text{true}$ 
   // Interval-bisection in  $w$  for  $\Phi(r, w) = \frac{\alpha+1-T}{2} \ln\left(1 + \frac{2sw}{r} + \frac{s^2}{r^2}\right) + \beta \left[r^p - (r^2 + 2swr + s^2)^{p/2}\right]$  on  $r \in [r_L, r_U]$ .
5: while  $w^{\text{hi}} - w^{\text{lo}} > \tau_w$  do
6:    $w^{\text{mid}} \leftarrow (w^{\text{lo}} + w^{\text{hi}})/2$ .
7:    $[\underline{\Phi}, \bar{\Phi}] \leftarrow \text{EnclosePhi}([r_L, r_U], w^{\text{mid}}; T, \alpha, p, \beta, s)$ .
8:   if  $\underline{\Phi} \leq y \leq \bar{\Phi}$  then  $\text{ok} \leftarrow \text{false}$ ; break // Cannot decide the sign of  $\Phi(r, w^{\text{mid}}) - y$  uniformly on this bin within  $\tau_w$ .
9:   if  $\underline{\Phi} > y$  then  $w^{\text{lo}} \leftarrow w^{\text{mid}}$  else  $w^{\text{hi}} \leftarrow w^{\text{mid}}$ 
10: end while
11: if  $\text{ok} = \text{false}$  then Return:  $[0, 1]$ 
12: if  $y = \varepsilon$  then  $\underline{g} \leftarrow F_W(w^{\text{lo}})$ ,  $\bar{g} \leftarrow F_W(w^{\text{hi}})$  else  $\underline{g} \leftarrow 1 - F_W(w^{\text{hi}})$ ,  $\bar{g} \leftarrow 1 - F_W(w^{\text{lo}})$ 
13: Output:  $[\underline{g}, \bar{g}]$ .

```

bounded angle-probability functions defined in Lemma 4.1, i.e.

$$g_-(R) = 1 - F_W(w^*(R, -\varepsilon)), \quad g_+(R) = F_W(w^*(R, \varepsilon)).$$

We now show that Algorithm 1 outputs a value $\hat{\delta}$ satisfying

$$\delta_{\mu}^*(\varepsilon) \leq \hat{\delta} \leq \delta_{\mu}^*(\varepsilon) + \eta. \quad (16)$$

Rewriting Under a Standard Gamma Measure. Let $k = (\alpha + 1)/p$ and define the change of variables $Z = \beta R^p$. Then $Z \sim \text{Gamma}(k, 1)$ (shape k , unit scale), with density

$$f_Z(z) = \frac{1}{\Gamma(k)} z^{k-1} e^{-z}, \quad z > 0.$$

Define $\tilde{g}_{\pm}(z) \in [0, 1]$ by $\tilde{g}_{\pm}(z) = g_{\pm}\left(\left(\frac{z}{\beta}\right)^{1/p}\right)$. A direct substitution yields

$$A = \mathbf{Ex}[\tilde{g}_-(Z)], \quad B = \mathbf{Ex}[\tilde{g}_+(Z)].$$

Truncation and a Closed-form Tail Bound. Let $z_{\max} > 0$ and set $\tau = \Pr[Z > z_{\max}] = Q(k, z_{\max})$, where $Q(\cdot, \cdot)$ is the regularized upper incomplete gamma function. Define the truncated integrals

$$A_{\leq z_{\max}} = \int_0^{z_{\max}} f_Z(z) \tilde{g}_-(z) dz, \quad B_{\leq z_{\max}} = \int_0^{z_{\max}} f_Z(z) \tilde{g}_+(z) dz.$$

Since $0 \leq \tilde{g}_{\pm} \leq 1$, the discarded tails satisfy $0 \leq A - A_{\leq z_{\max}} \leq \tau$, $0 \leq B - B_{\leq z_{\max}} \leq \tau$. Consequently,

$$|(A - e^{\varepsilon} B) - (A_{\leq z_{\max}} - e^{\varepsilon} B_{\leq z_{\max}})| \leq (1 + e^{\varepsilon})\tau.$$

In Algorithm 1, we choose z_{\max} (by bisection on $Q(k, \cdot)$) so that $(1 + e^{\varepsilon})\tau \leq \eta_{\text{tail}}$. This guarantees that truncation contributes at most η_{tail} additive error to the linear form $A - e^{\varepsilon} B$.

Probability binning and Brackets on the Truncated Integrals. Fix any partition $0 = z_0 < z_1 < \dots < z_m = z_{\max}$, $I_i = [z_i, z_{i+1}]$. Let $\pi_i = \Pr[Z \in I_i] = P(k, z_{i+1}) - P(k, z_i)$, where $P(\cdot, \cdot)$ is the regularized lower incomplete gamma function. Suppose that for each bin I_i we have a provable lower and upper bound

$$\tilde{g}_{\pm}(z) \in [\underline{g}_{\pm, i}, \bar{g}_{\pm, i}], \quad \forall z \in I_i.$$

Algorithm 3 ENCLOSEPHI: Interval bound for $\Phi(r, w)$ over $r \in [r_L, r_U]$

- 1: **Input:** interval $[r_L, r_U]$ with $0 < r_L \leq r_U$, scalar $w \in [-1, 1]$, parameters (T, α, p, β, s)
 - 2: **Output:** $[\underline{\Phi}, \overline{\Phi}]$ s.t. $\Phi(r, w) \in [\underline{\Phi}, \overline{\Phi}]$ for all $r \in [r_L, r_U]$.
 - 3: $c_1 \leftarrow (\alpha + 1 - T)/2$
// Bound $q(r) = r^2 + 2swr + s^2$ on $[r_L, r_U]$ (convex quadratic).
 - 4: $r^* \leftarrow \min\{r_U, \max\{r_L, -sw\}\}$ {projection of the vertex $-sw$ }
 - 5: $q_L \leftarrow r_L^2 + 2swr_L + s^2$, $q_U \leftarrow r_U^2 + 2swr_U + s^2$, $q_* \leftarrow (r^*)^2 + 2swr^* + s^2$
 - 6: $\underline{q} \leftarrow \min\{q_L, q_U, q_*\}$, $\overline{q} \leftarrow \max\{q_L, q_U\}$
// Bound $c_1 \ln(q(r)/r^2)$ using $\underline{q}, \overline{q}$ and $r^2 \in [r_L^2, r_U^2]$.
 - 7: $\underline{u} \leftarrow \underline{q}/r_U^2$, $\overline{u} \leftarrow \overline{q}/r_L^2$
 - 8: **if** $\underline{u} \leq 0$ **then** $\underline{\ell} \leftarrow -\infty$ **else** $\underline{\ell} \leftarrow \ln(\underline{u})$
 - 9: $\overline{\ell} \leftarrow \ln(\overline{u})$
 - 10: **if** $c_1 \geq 0$ **then** $\underline{L} \leftarrow c_1 \underline{\ell}$, $\overline{L} \leftarrow c_1 \overline{\ell}$ **else** $\underline{L} \leftarrow c_1 \overline{\ell}$, $\overline{L} \leftarrow c_1 \underline{\ell}$
// Bound $r^p - q(r)^{p/2}$ (monotone for $r \geq 0, q \geq 0$).
 - 11: $\underline{a} \leftarrow r_L^p$, $\overline{a} \leftarrow r_U^p$
 - 12: $\underline{b} \leftarrow (\underline{q})^{p/2}$, $\overline{b} \leftarrow (\overline{q})^{p/2}$
 - 13: $\underline{D} \leftarrow \underline{a} - \underline{b}$, $\overline{D} \leftarrow \overline{a} - \overline{b}$
// Combine: $\Phi = L + \beta \cdot D$ (with $\beta > 0$).
 - 14: $\underline{\Phi} \leftarrow \underline{L} + \beta \underline{D}$
 - 15: $\overline{\Phi} \leftarrow \overline{L} + \beta \overline{D}$
 - 16: **Output:** $[\underline{\Phi}, \overline{\Phi}]$.
-

Multiplying by the nonnegative density and integrating over I_i yields

$$\begin{aligned} \underline{A} &= \sum_{i=0}^{m-1} \pi_i \underline{g}_{-,i} \leq A_{\leq z_{\max}} \leq \sum_{i=0}^{m-1} \pi_i \overline{g}_{-,i} = \overline{A}, \\ \underline{B} &= \sum_{i=0}^{m-1} \pi_i \underline{g}_{+,i} \leq B_{\leq z_{\max}} \leq \sum_{i=0}^{m-1} \pi_i \overline{g}_{+,i} = \overline{B}. \end{aligned}$$

Therefore the truncated linear form is bracketed as

$$\underline{A} - e^\varepsilon \overline{B} \leq A_{\leq z_{\max}} - e^\varepsilon B_{\leq z_{\max}} \leq \overline{A} - e^\varepsilon \underline{B},$$

and the bracket width satisfies

$$\left(\overline{A} - \underline{A}\right) + e^\varepsilon \left(\overline{B} - \underline{B}\right) = \sum_{i=0}^{m-1} \pi_i \left[(\overline{g}_{-,i} - \underline{g}_{-,i}) + e^\varepsilon (\overline{g}_{+,i} - \underline{g}_{+,i}) \right],$$

which is exactly the quantity tracked by line 18 in Algorithm 1.

It remains to justify that $\tilde{g}_\pm(z) \in [g_{\pm,i}, \overline{g}_{\pm,i}]$ is true for all $i \in [m]$, $z \in I_i$, and that the adaptive refinement loop terminates once $\left(\overline{A} - \underline{A}\right) + e^\varepsilon \left(\overline{B} - \underline{B}\right)$ falls below η_{int} .

Correctness of ENCLOSEPHI and ENCLOSEG. We first recall the (binwise) privacy-loss threshold function used in Lemma 4.1. For $r > 0$ and $w \in [-1, 1]$, define

$$\Phi(r, w) = \frac{\alpha + 1 - T}{2} \ln\left(1 + \frac{2sw}{r} + \frac{s^2}{r^2}\right) + \beta \left[r^p - (r^2 + 2swr + s^2)^{p/2} \right].$$

Because $\alpha \leq T - 1$, the coefficient $(\alpha + 1 - T)/2 \leq 0$, and since $q(r, w) = r^2 + 2swr + s^2$ increases with w , both terms in $\Phi(r, w)$ are non-increasing in w . Hence, for each fixed $r > 0$, the map $w \mapsto \Phi(r, w)$ is non-increasing on $[-1, 1]$.

For a target level $y \in \{\varepsilon, -\varepsilon\}$, define the threshold

$$w^*(r, y) = \sup\{w \in [-1, 1] : \Phi(r, w) \geq y\},$$

with the convention that $\sup \emptyset = -1$. By monotonicity in w , we have that for every $r > 0$,

$$\Phi(r, w) \geq y \iff w \leq w^*(r, y).$$

Recall that $g_-(r) = 1 - F_W(w^*(r, -\varepsilon))$, $g_+(R) = F_W(w^*(r, \varepsilon))$, and $\tilde{g}_\pm(z) = g_\pm\left(\left(\frac{z}{\beta}\right)^{1/p}\right)$.

Correctness of ENCLOSEPHI. Fix $w \in [-1, 1]$ and an interval $r \in [r_L, r_U]$ with $0 < r_L \leq r_U$. Algorithm 3 computes bounds for $\Phi(r, w)$ by interval arithmetic: (1) it bounds the quadratic $q(r, w) = r^2 + 2swr + s^2$ on $[r_L, r_U]$ using convexity, (2) it bounds $\ln(q(r, w)/r^2)$ using monotonicity of \ln together with $q(r, w) \in [q, \bar{q}]$ and $r^2 \in [r_L^2, r_U^2]$, and (3) it bounds $r^p - q(r, w)^{p/2}$ using monotonicity of $x \mapsto x^{p/2}$ on $\mathbb{R}_{\geq 0}$ and $r^p \in [r_L^p, r_U^p]$. Combining these bounds linearly yields an interval $[\underline{\Phi}, \bar{\Phi}]$ satisfying

$$\Phi(r, w) \in [\underline{\Phi}, \bar{\Phi}] \quad \forall r \in [r_L, r_U].$$

Correctness of ENCLOSEG. Fix a z -bin $I = [z_L, z_U]$ and let $r \in [r_L, r_U]$ be its corresponding r -range with $r = (z/\beta)^{1/p}$. WLOG, assume now $0 < r_L \leq r_U$. Algorithm 2 performs bisection on $w \in [-1, 1]$ to enclose $w^*(r, y)$ for all $r \in [r_L, r_U]$. At each bisection step with midpoint w^{mid} , it calls ENCLOSEPHI to obtain an interval $[\underline{\Phi}, \bar{\Phi}]$ satisfying $\Phi(r, w^{\text{mid}}) \in [\underline{\Phi}, \bar{\Phi}]$. If $\underline{\Phi} > y$, then $\Phi(r, w^{\text{mid}}) > y$ for all r in the bin; since $\Phi(r, \cdot)$ is non-increasing, this implies $w^*(r, y) \geq w^{\text{mid}}$ for all r , so updating $w^{\text{lo}} \leftarrow w^{\text{mid}}$ is sound. Similarly, if $\bar{\Phi} < y$. If neither condition holds (i.e., the interval straddles y), the routine returns $[0, 1]$, which again is sound.

When the routine terminates without the fallback, it has produced $w^{\text{lo}} \leq w^{\text{hi}}$ such that

$$w^*(r, y) \in [w^{\text{lo}}, w^{\text{hi}}], \quad \forall r \in [r_L, r_U].$$

Applying the monotonicity of the CDF F_W yields

$$F_W(w^*(r, \varepsilon)) \in [F_W(w^{\text{lo}}), F_W(w^{\text{hi}})], \quad 1 - F_W(w^*(r, -\varepsilon)) \in [1 - F_W(w^{\text{hi}}), 1 - F_W(w^{\text{lo}})],$$

which is exactly what Algorithm 2 outputs. Composing with $r = (z/\beta)^{1/p}$, we justify that $\tilde{g}_\pm(z) \in [\underline{g}_{\pm, i}, \bar{g}_{\pm, i}]$ is true for all $i \in [m], z \in I_i$.

Termination of the Adaptive Refinement Loop. We show that the refinement loop in Algorithm 1 terminates, i.e., it eventually constructs a partition such that

$$(\bar{A} - \underline{A}) + e^\varepsilon(\bar{B} - \underline{B}) \leq \eta_{\text{int}}.$$

Over the compact interval $[0, z_{\text{max}}]$, the functions $\tilde{g}_\pm(z)$ take values in $[0, 1]$ and are Borel-measurable (indeed, they are continuous except possibly at $z = 0$, and bounded everywhere). Hence, for every error bound $\eta > 0$ there exists a partition of $[0, z_{\text{max}}]$ such that the oscillation of \tilde{g}_\pm on each bin is at most η , except possibly on bins that touch 0. For bins touching 0, even the trivial enclosure $[0, 1]$ contributes at most their Gamma mass to the error accumulator; since $k > 0$ the Gamma measure of $[0, \zeta]$ tends to 0 as $\zeta \downarrow 0$, so by refining near 0 we can make the total contribution of such bins arbitrarily small. We also note that the accumulated bracket width is a *probability-weighted* sum of binwise oscillations, and hence does *not* scale with the number of bins.

Formally, pick $\eta > 0$ such that $(1 + e^\varepsilon)\eta \leq \eta_{\text{int}}/2$. Choose $\zeta \in (0, z_{\text{max}})$ so that $(1 + e^\varepsilon) \Pr[Z \in [0, \zeta]] \leq \eta_{\text{int}}/2$. On $[\zeta, z_{\text{max}}]$, bounded (and continuous) functions are uniformly continuous, so there exists a finite partition of $[\zeta, z_{\text{max}}]$ in which each bin has oscillation at most η for both \tilde{g}_- and \tilde{g}_+ . Using the correctness property of ENCLOSEG on those bins then yields

$$\sum_{I \subseteq [\zeta, z_{\text{max}}]} \pi_I [(\bar{g}_{-, I} - \underline{g}_{-, I}) + e^\varepsilon(\bar{g}_{+, I} - \underline{g}_{+, I})] \leq (1 + e^\varepsilon)\delta \leq \eta_{\text{int}}/2.$$

For bins contained in $[0, \zeta]$, even the trivial enclosure $[0, 1]$ gives

$$\sum_{I \subseteq [0, \zeta]} \pi_I [(\bar{g}_{-, I} - \underline{g}_{-, I}) + e^\varepsilon(\bar{g}_{+, I} - \underline{g}_{+, I})] \leq (1 + e^\varepsilon) \Pr[Z \in [0, \zeta]] \leq \eta_{\text{int}}/2.$$

Algorithm 4 Parameterized SGG Mechanism $\mathcal{M}_{\text{SGG}}^{\alpha, p}$

1: **Input:** A vector-valued query $q(G) \in \mathbb{R}^T$ with ℓ_2 sensitivity $s > 0$, privacy parameters (ε, δ) , fixed (α, p) with $\alpha \in (-1, T - 1]$, $p > 0$, η -tight upper-bound oracle $\text{Oracles}_{\text{SGG}}$ (Definition 4.3), tolerance $\tau > 0$.

2: $\beta^{\text{lo}} \leftarrow 1$

3: **while** $\text{Oracles}_{\text{SGG}}(\alpha, p, \beta^{\text{lo}}, \varepsilon, s) > \delta$ **do** $\beta^{\text{lo}} \leftarrow \beta^{\text{lo}}/2$

4: $\beta^{\text{hi}} \leftarrow \beta^{\text{lo}}$

5: **while** $\text{Oracles}_{\text{SGG}}(\alpha, p, \beta^{\text{hi}}, \varepsilon, s) \leq \delta$ **do** $\beta^{\text{hi}} \leftarrow 2\beta^{\text{hi}}$

6: **while** $\beta^{\text{hi}}/\beta^{\text{lo}} > 1 + \tau$ **do**

7: $\beta^{\text{mid}} \leftarrow (\beta^{\text{lo}} + \beta^{\text{hi}})/2$.

8: **if** $\text{Oracles}_{\text{SGG}}(\alpha, p, \beta^{\text{mid}}, \varepsilon, s) \leq \delta$ **then** $\beta^{\text{lo}} \leftarrow \beta^{\text{mid}}$ **else** $\beta^{\text{hi}} \leftarrow \beta^{\text{mid}}$

9: **end while**

10: $\beta \leftarrow \beta^{\text{lo}}$.

11: **Output:** $q(G) + X$, where $X \sim \text{SGG}(\alpha, \beta, p)$.

Adding the two contributions gives $(\bar{A} - \underline{A}) + e^\varepsilon(\bar{B} - \underline{B}) \leq \eta_{\text{int}}$. Since Algorithm 1 refines by repeatedly splitting the bin with largest contribution to $(\bar{A} - \underline{A}) + e^\varepsilon(\bar{B} - \underline{B})$, its mesh size tends to 0 and it must eventually reach a refinement at least as fine as the existence argument above; therefore it terminates.

Conclusion. Let $D = A - e^\varepsilon B$ and let

$$\hat{D} = (\bar{A} + \tau) - e^\varepsilon \underline{B}, \quad \hat{\delta} = \max\{0, \hat{D}\}, \quad \delta_\mu^*(\varepsilon) = \max\{0, D\}.$$

We first show $\hat{D} \geq D$ (hence $\hat{\delta} \geq \delta_\mu^*(\varepsilon)$). This is because

$$A = A_{\leq z_{\max}} + (A - A_{\leq z_{\max}}) \leq \bar{A} + \tau, \quad B \geq B_{\leq z_{\max}} \geq \underline{B},$$

and therefore $D = A - e^\varepsilon B \leq (\bar{A} + \tau) - e^\varepsilon \underline{B} = \hat{D}$.

Next, we upper bound $\hat{D} - D$. Using the decompositions $A = A_{\leq z_{\max}} + A_{> z_{\max}}$ and $B = B_{\leq z_{\max}} + B_{> z_{\max}}$ with $0 \leq A_{> z_{\max}}, B_{> z_{\max}} \leq \tau$, we have

$$\begin{aligned} \hat{D} - D &= (\bar{A} - A_{\leq z_{\max}}) + (\tau - A_{> z_{\max}}) + e^\varepsilon(B_{\leq z_{\max}} - \underline{B}) + e^\varepsilon B_{> z_{\max}} \\ &\leq (\bar{A} - \underline{A}) + e^\varepsilon(\bar{B} - \underline{B}) + (1 + e^\varepsilon)\tau \\ &\leq \eta_{\text{int}} + \eta_{\text{tail}} = \eta. \end{aligned}$$

Finally, since $x \mapsto \max\{0, x\}$ is 1-Lipschitz, $0 \leq \hat{\delta} - \delta_\mu^*(\varepsilon) \leq \hat{D} - D \leq \eta$, which is exactly Equation (16). So we conclude that Algorithm 1 is an η -tight upper-bound oracle, and such an oracle exists. \square

With Lemma 4.3 in hand, we show some supplementary statements needed for our mechanism in Algorithm 4.

Lemma C.1 states that, for fixed (α, p) , changing the rate parameter β simply rescales the SGG noise by a factor $\lambda = \beta^{-1/p}$. Equivalently, $\text{SGG}(\alpha, \beta, p)$ is the distribution of λX_1 for $X_1 \sim \text{SGG}(\alpha, 1, p)$.

Lemma C.1. Fix $T \geq 2$, $\alpha > -1$, $p > 0$, and $\beta > 0$, and define $\lambda = \beta^{-1/p}$. Let $R_1 \sim \text{GGamma}(\alpha, 1, p)$ and $U \sim \text{Unif}(\mathcal{S}^{T-1})$ be independent, and let $X_1 \triangleq R_1 U \sim \text{SGG}(\alpha, 1, p)$. Then

$$R \sim \text{GGamma}(\alpha, \beta, p) \iff R \stackrel{d}{=} \lambda R_1,$$

and consequently,

$$X \sim \text{SGG}(\alpha, \beta, p) \iff X \stackrel{d}{=} \lambda X_1.$$

Proposition C.1 (Monotonicity in the rate parameter β). Fix $T \geq 2$, $\alpha \in (-1, T - 1]$, $p > 0$, and $\varepsilon \geq 0$. For $\beta > 0$, let $X_\beta \sim \text{SGG}(\alpha, \beta, p)$ and let $\delta_\mu^*(\varepsilon; \beta)$ denote the optimal delta for the neighboring pair $(X_\beta, X_\beta + \mu)$ at level ε . Then, for any fixed shift $\mu \in \mathbb{R}^T$, the mapping $\beta \mapsto \delta_\mu^*(\varepsilon; \beta)$ is non-decreasing. Equivalently, for $0 < \beta_1 \leq \beta_2$,

$$\delta_\mu^*(\varepsilon; \beta_1) \leq \delta_\mu^*(\varepsilon; \beta_2).$$

Proof. Let $\lambda(\beta) = \beta^{-1/p}$. By Lemma C.1, we may write $X_\beta \stackrel{d}{=} \lambda(\beta) X_1$ where $X_1 \sim \text{SGG}(\alpha, 1, p)$. Consider the invertible map $\varphi_\lambda(x) = x/\lambda$. Since hockey-stick divergence (and hence $\delta_\mu^*(\varepsilon)$) is invariant under bijective transformations, we have

$$\begin{aligned} \delta_\mu^*(\varepsilon; \beta) &= \delta^*(\varepsilon; X_\beta, X_\beta + \mu) \\ &= \delta^*(\varepsilon; \varphi_{\lambda(\beta)}(X_\beta), \varphi_{\lambda(\beta)}(X_\beta + \mu)) \\ &= \delta^*(\varepsilon; X_1, X_1 + \mu/\lambda(\beta)) = \delta_{\mu/\lambda(\beta)}^*(\varepsilon; 1). \end{aligned}$$

Now fix $0 < \beta_1 \leq \beta_2$. Then $\lambda(\beta_1) \geq \lambda(\beta_2)$, hence

$$\left\| \frac{\mu}{\lambda(\beta_1)} \right\|_2 \leq \left\| \frac{\mu}{\lambda(\beta_2)} \right\|_2.$$

Applying the shift monotonicity result Lemma 4.2 to the base distribution $X_1 \sim \text{SGG}(\alpha, 1, p)$ yields

$$\delta_{\mu/\lambda(\beta_1)}^*(\varepsilon; 1) \leq \delta_{\mu/\lambda(\beta_2)}^*(\varepsilon; 1).$$

Substituting back proves $\delta_\mu^*(\varepsilon; \beta_1) \leq \delta_\mu^*(\varepsilon; \beta_2)$. □

Algorithm 5 Spherical Generalized Gamma Mechanism

1: **Input:**

- $T \in \mathcal{N}_{>0}$, the query dimension, $s > 0$, the query's ℓ_2 sensitivity
- $\varepsilon > 0, \delta \in [0, 1]$, the desired privacy parameters
- $q(G) \in \mathbb{R}^T$, query result
- $\text{Oracles}_{\text{SGG}}$ which on input $T, \alpha, p, \beta, \varepsilon$ outputs an upper bound for δ .
- $\text{err} : [0, T-1] \times \mathbb{R}_{>0} \times \mathbb{R}_{>0} \mapsto \mathbb{R}_{>0}$, the error function for generalized gamma distribution

2: **Parameters:**

- $\alpha \in [0, T-1], \beta > 0, p > 0$ the parameters of the generalized gamma distribution

3: Set the noise budget upper bound c^{upp} such that the standard Gaussian mechanism that achieves (ε, δ) -DP introduces error c^{upp} . Also set $c^{\text{mid}} \leftarrow 1$.

4: $\delta^* \leftarrow \infty$

5: // Perform a binary search over $c \in [1, c^{\text{upp}}]$ to find the smallest feasible noise level

6: **while** $\delta^* > \delta \vee (|\delta^* - \delta| > \text{atol} \wedge |c^{\text{low}} - c^{\text{upp}}| > \text{atol})$ **do**

7: **if** $\delta^* \leq \delta$ **then** set $c^{\text{upp}} \leftarrow c^{\text{mid}}$ **else** set $c^{\text{low}} \leftarrow c^{\text{mid}}$

8: $c^{\text{mid}} \leftarrow (c^{\text{low}} + c^{\text{upp}})/2$

9: Solve the 2-dimensional optimization problem

$$\alpha^*, p^* \leftarrow \arg \min_{\alpha, p} \text{Oracles}_{\text{SGG}}(T, \alpha, \beta, p, \varepsilon)$$

// β and β^* below computed via Algorithm 4 on input $s, \varepsilon, \alpha, p$ and target $\hat{\delta}$, constrained to noise budget c^{mid}

10: $\delta^* \leftarrow \text{Oracles}_{\text{SGG}}(T, \alpha^*, \beta^*, p^*, \varepsilon)$

11: **end while**

12: **Output:** $q(G) + RU$, where $R \sim \text{GGamma}(\alpha^*, h(\alpha^*, p^*, c^{\text{mid}}), p^*)$ and $U \sim \text{Unif}(S^{T-1})$.

Theorem C.1. The mechanism $\mathcal{M}_{\text{SGG}}^{\alpha, p}$ (Algorithm 4) is (ε, δ) -differentially private.

Proof. Fix neighboring databases $G \sim G'$. Let $\mu = q(G) - q(G') \in \mathbb{R}^T$, and we know $\|\mu\|_2 \leq s$ by the assumption on $q(\cdot)$. The mechanism in Algorithm 4 outputs

$$\mathcal{M}_{\text{SGG}}^{\alpha, p}(G) = q(G) + X_\beta, \quad \mathcal{M}_{\text{SGG}}^{\alpha, p}(G') = q(G') + X_\beta,$$

where $X_\beta \sim \text{SGG}(\alpha, \beta, p)$ and β is the value returned by the bracketing-and-bisection procedure.

By translation invariance, distinguishing $\mathcal{M}_{\text{SGG}}^{\alpha,p}(G)$ from $\mathcal{M}_{\text{SGG}}^{\alpha,p}(G')$ is equivalent to distinguishing X_β from $X_\beta + \mu$. Thus, the optimal δ at level ε for the pair of the mechanism outputs equals $\delta_\mu^*(\varepsilon; \beta)$, the optimal $\delta(\varepsilon)$ between X_β and $X_\beta + \mu$.

By the shift monotonicity result Lemma 4.2, the map $\mu \mapsto \delta_\mu^*(\varepsilon; \beta)$ is non-decreasing in $\|\mu\|_2$, hence for any μ with $\|\mu\|_2 \leq s$ and any fixed β ,

$$\delta_\mu^*(\varepsilon; \beta) \leq \delta_{\mu_s}^*(\varepsilon; \beta) \quad \text{for any } \mu_s \in \mathbb{R}^T \text{ with } \|\mu_s\|_2 = s.$$

Applying the oracle guarantee in Definition 4.3 to μ_s gives

$$\delta_{\mu_s}^*(\varepsilon; \beta) \leq \text{Oracles}_{\text{SGG}}(\alpha, p, \beta, \varepsilon, \|\mu_s\|_2) = \text{Oracles}_{\text{SGG}}(\alpha, p, \beta, \varepsilon, s).$$

The chosen β is feasible for the target (ε, δ) . By construction, the algorithm maintains a feasible lower bracket β^{lo} with $\text{Oracles}_{\text{SGG}}(\alpha, p, \beta^{\text{lo}}, \varepsilon, s) \leq \delta$ and an infeasible upper bracket β^{hi} with $\text{Oracles}_{\text{SGG}}(\alpha, p, \beta^{\text{hi}}, \varepsilon, s) > \delta$. The update rule in the bisection step relies on the monotonicity in β (Proposition C.1) to preserve feasibility/infeasibility of the brackets. At termination the algorithm sets $\beta \leftarrow \beta^{\text{lo}}$, and therefore

$$\text{Oracles}_{\text{SGG}}(\alpha, p, \beta, \varepsilon, s) \leq \delta.$$

Combining the last two inequalities, we conclude that for every neighboring $G \sim G'$, the output distributions satisfy the (ε, δ) -DP inequality. Formally,

$$\delta_\mu^*(\varepsilon; \beta) \leq \text{Oracles}_{\text{SGG}}(\alpha, p, \beta, \varepsilon, s) \leq \delta,$$

i.e., $\mathcal{M}_{\text{SGG}}^{\alpha,p}$ is (ε, δ) -differentially private. □

C.3. Finding Optimal Parameters of the SGG Mechanism

Finally, based on Algorithm 4 that finds best β for fixed $(\varepsilon, \delta, \alpha, p)$, we provide Algorithm 5, which finds the parameters (α^*, β^*, p^*) with minimal MSE given target privacy (ε, δ) .

C.4. Finding Settings where SGG Dominates

Here we present Algorithm 6, which for a given T, s, ε , finds the δ in which the SGG has the biggest advantage over both the Gaussian and ℓ_2 mechanisms.

Algorithm 6 Finding Optimal Advantage

- 1: **Input:** $T \in \mathcal{N}_{>0}$, the query dimension, $s > 0$, the query's ℓ_2 sensitivity, and $\varepsilon > 0$, the desired privacy parameter.
 - 2: Set $\delta^{\text{low}} \leftarrow 0, \delta^{\text{upp}} \leftarrow 0.1$.
 - 3: // Perform a binary search over $\delta \in [\delta^{\text{low}}, \delta^{\text{upp}}]$ to find the largest advantage
 - 4: **while** $\delta^{\text{upp}} - \delta^{\text{low}} > \text{atol}$ **do**
 - 5: $\delta^{\text{mid}} \leftarrow (\delta^{\text{low}} + \delta^{\text{upp}})/2$
 - 6: Use Algorithm 5 to find (α^*, β^*, p^*) achieving minimal MSE c^* of the SGG for T -dimensional query with sensitivity s and privacy parameters (ε, δ) .
 - 7: Compute MSE c_{ℓ_2} and c_G of ℓ_2 and Gaussian mechanisms, respectively for the above setting.
 - 8: **if** $(c_G - c^*)/c_G < (c_{\ell_2} - c^*)/c_{\ell_2}$ **then** set $\delta^{\text{upp}} \leftarrow \delta^{\text{mid}}$ **else** set $\delta^{\text{low}} \leftarrow \delta^{\text{mid}}$
 - 9: **end while**
 - 10: **Output:** $(\alpha^*, \beta^*, p^*), c^*$.
-

D. Privacy Accounting of SGG Mechanisms

In this section, we study tight composed privacy for SGG mechanisms under multiple invocations.

Our approach follows the FFT-based accountant developed by (Gopi et al., 2024). Their work casts privacy accounting in terms of the *privacy-loss random variable* (PRV). The central benefit is that privacy losses add: the PRV of a k -fold

Algorithm 7 Privacy Accounting for $\mathcal{M}_{\text{SGG}}^{\alpha, p}$ under k invocations

- 1: **Input:** Dimension T , sensitivity s , SGG parameters (α, β, p) , composition count k , target ε
- 2: **Discretization parameters:** truncation $L > 0$, step size $h > 0$, odd grid size $M = 1 + \lceil 2L/h \rceil$
- 3: **Numerical parameters:** radial quadrature size K , angular grid size N_w , arithmetic precision prec
 // Discretize the single-step PRV
- 4: Define grid centers $z_i \leftarrow -L + ih$ for $i = 0, 1, \dots, M - 1$ and bin edges $b_i \leftarrow z_i - \frac{h}{2}$.
- 5: For each edge b_i , approximate $\Pr[Z \leq b_i]$ using a K -node Gauss–Laguerre quadrature (cf. Equation (17)).
- 6: Set the pmf $\pi_i \leftarrow \Pr[Z \leq b_{i+1}] - \Pr[Z \leq b_i]$
 // Compose by FFT convolution
- 7: Compute the k -fold convolution $\pi^{(*k)}$ using FFTs, cropping back to the support $[-L, L]$ after each multiplication (per (Gopi et al., 2024))
 // Output the composed privacy guarantee
- 8: Compute $\delta_k(\varepsilon) \leftarrow \sum_{i=0}^{M-1} \pi_i^{(*k)} \cdot (1 - e^{\varepsilon - z_i})_+$
- 9: Return $(\varepsilon, \delta_k(\varepsilon))$

composition is the sum of k single-step PRV. This reduces tight (ε, δ) accounting to (i) characterizing the distribution of the single-step PRV and (ii) computing the distribution of its k -fold sum via repeated convolution. These two observations allow computing tight composed privacy efficiently by using FFTs after discretizations of single-step PRV. We follow (Gopi et al., 2024) along its theoretical analysis, and tailor the missing SGG-specific component: an efficient discretization of the single-step PRV using the one-dimensional integral reduction from Section 4.1.

To discretize the single-step PRV Z for an SGG mechanism, we use the standard approach of placing a grid on the privacy-loss axis and computing bin masses from CDF differences at grid edges. Thus, efficient discretization amounts to evaluating the PRV CDF $\Pr[Z \leq z]$ at many grid points. Lemma D.1 makes this feasible: it expresses $\Pr[Z \leq z]$ as a one-dimensional radial integral via the same technique used in Section 4.1.

Lemma D.1. (Proof is below) *Adopt the setup and notation of Lemma 4.1. Let $X \sim \text{SGG}(\alpha, \beta, p)$ and define the PRV for the pair $(X, X + \mu)$ by*

$$Z \stackrel{\text{def}}{=} \log \frac{f_X(X)}{f_X(X + \mu)}.$$

Then, for every $z \in \mathbb{R}$,

$$\Pr[Z \leq z] = \int_0^\infty f_R(r) F_W(w^*(r, -z)) dr,$$

where f_R and F_W are as in Lemma 4.1, and $w^*(r, \cdot)$ denotes the corresponding threshold map.

The next step is to increase the efficiency of evaluations of $\Pr[Z \leq z]$ over the discretization grid. The efficiency tricks exploit the independence between random variable R and the monotonicity of $w \mapsto \ell(r, w)$. First, we approximate the one-dimensional radial (r) integral by a K -node Gauss–Laguerre quadrature after the change of variables $t = \beta r^p$, leading fixed radii $\{r_k\}_{k=1}^K$ and weights $\{\omega_k\}_{k=1}^K$ such that

$$\Pr[Z \leq z] \approx \sum_{k=1}^K \omega_k F_W(w^*(r_k, -z)). \quad (17)$$

Second, for each quadrature node r_k we precompute a dense lookup table of the strictly decreasing map $w \mapsto \ell(r_k, w)$ on a grid over $w \in [-1, 1]$; thereafter, each threshold $w^*(r_k, -z)$ is computed as a table lookup followed by linear interpolation. Third, we evaluate $F_W(\cdot)$ in vectorized form and complete each CDF query with a single dot product against the cached weights. With these caches, computing $\Pr[Z \leq z]$ costs essentially $O(K)$ floating-point operations per z . We also (re)use parallel workers to amortize overhead in practice.

Algorithm 7 describes the full SGG-tailored accounting algorithm. We also note that as the ℓ_2 mechanism is a special case of SGG (e.g., $(\alpha, p) = (T - 1, 1)$ with an appropriate scale), this yields a practical tight accountant for its composed approximate-DP guarantees as well; more broadly, the same pipeline applies to any (α, β, p) and to mixed compositions by convolving the corresponding discretized PRVs.

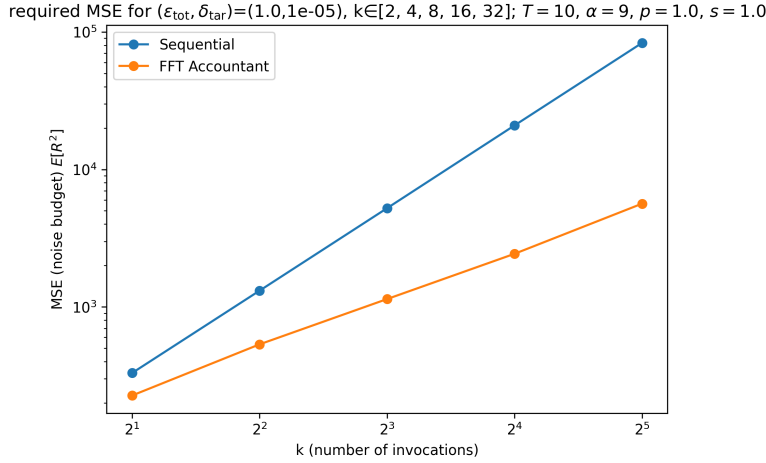


Figure 3. MSE Required for privacy target $(\varepsilon_{\text{tot}}, \delta_{\text{tar}}) = (1, 10^{-5})$ under k invocations: sequential composition versus Algorithm 7.

Experimental Evaluation. Figure 3 compares the noise level (measured by the mean-squared error (MSE) of the SGG mechanisms) required to meet a fixed composed privacy target under two accounting methods: (i) the baseline *sequential composition* bound and (ii) our SGG-tailored *FFT accountant* in Algorithm 7. Concretely, we fix a total privacy target $(\varepsilon_{\text{tot}}, \delta_{\text{tar}}) = (1, 10^{-5})$ and consider $k \in \{2, 4, 8, 16, 32\}$ invocations of the same SGG mechanism $\mathcal{M}_{\text{SGG}}^{\alpha, p}$ in dimension $T = 10$ (with fixed shape parameter $\alpha = 9$ and $p = 1$). For each k and for each accountant, we compute the *minimal* per-invocation MSE such that the k -fold composition satisfies

$$\delta_k(\varepsilon_{\text{tot}}) \leq \delta_{\text{tar}}.$$

For the sequential baseline, we allocate a per-step budget $(\varepsilon_{\text{tot}}/k, \delta_{\text{tar}}/k)$ and solve for the smallest noise meeting this single-step guarantee (cf. Algorithm 4); for the FFT accountant, we instead solve for the smallest noise such that the computed composed optimal delta from Algorithm 7 meeting δ_{tar} at the privacy level ε_{tot} .

Figure 3 shows a clear and growing gap as k increases between the the baseline accountant and our SGG-tailored FFT accountant. It illustrates the practical impact of tight accounting

Proof of Lemma D.1. Write $X = RU$ with $R \sim \text{GGamma}(\alpha, \beta, p)$ independent of $U \sim \text{Unif}(\mathcal{S}^{T-1})$, and let

$$W = \frac{\langle \mu, U \rangle}{\|\mu\|_2} \in [-1, 1].$$

By the spherical form of the SGG density (see the derivation in the proof of Lemma 4.1), the log-density ratio $(-Z)$

$$L^{\text{neg}}(X) = \log \frac{f_X(X + \mu)}{f_X(X)}$$

can be written as a deterministic function of (R, W) :

$$L^{\text{neg}}(X) = \ell(R, W),$$

where for each $r > 0$ the map $w \mapsto \ell(r, w)$ is strictly decreasing on $[-1, 1]$ (again, proved in Section C.1). Hence, for every $y \in \mathbb{R}$ and $r > 0$, there is a unique threshold $w^*(r, y) \in [-1, 1]$ such that $\ell(r, w^*(r, y)) = y$, and

$$\{\ell(r, W) \geq y\} \iff \{W \leq w^*(r, y)\}.$$

Now note that the PRV in this lemma is the negative of L^{neg} ($Z = -L^{\text{neg}}(X) = -\ell(R, W)$). Therefore, conditioning on $R = r$ and using the monotonicity property,

$$\Pr[Z \leq z \mid R = r] = \Pr[\ell(r, W) \geq -z] = \Pr[W \leq w^*(r, -z)] = F_W(w^*(r, -z)).$$

ε	δ_{claimed}	ψ	σ_*	$\delta_{\text{true}}(\varepsilon)$
0.1	10^{-5}	0.798721	25.040031	0.813284
1.0	10^{-5}	0.798721	2.504003	0.983594
2.0	10^{-5}	0.798721	1.252002	0.995020
4.0	10^{-5}	0.798721	0.626001	0.998804
8.0	10^{-5}	0.798721	0.313000	0.999755

Table 3. Testing (Ji & Li, 2024, Theorem 5) calibration for R1SMG at $T = 128$ and $s = 1$ with $\delta_{\text{claimed}} = 10^{-5}$. We compute ψ and the prescribed σ_* from the theorem, then evaluate the resulting privacy $\delta_{\text{true}}(\varepsilon)$ via Lemma 4.1. The computed $\delta_{\text{true}}(\varepsilon)$ is vastly larger than 10^{-5} , contradicting the claimed guarantee.

Finally, since R is independent of W , integrating over R yields

$$\Pr[Z \leq z] = \int_0^\infty f_R(r) \Pr[Z \leq z \mid R = r] dr = \int_0^\infty f_R(r) F_W(w^*(r, -z)) dr,$$

which is exactly the claim. □

E. Error in Proof of (Ji & Li, 2024, Theorem 5)

Here we identify an error in the proof of Theorem 5 in (Ji & Li, 2024). The proof attempts to analyze the Privacy Loss Random Variable (PRV) of their (spherical) R1SMG mechanism. One important fact about the PRV is that it is defined over the randomness of *only* the random variable $\mathcal{M}(q(G))$, and not the randomness of the random variable $\mathcal{M}(q(G'))$ (where G' is an allowed neighbor of G). Indeed, consider random variable $Y \stackrel{\text{def}}{=} \mathcal{M}(q(G))$; the PRV is the function

$$L(Y) \stackrel{\text{def}}{=} \ln \left(\frac{f_{\mathcal{M}(q(G))}(Y)}{f_{\mathcal{M}(q(G'))}(Y)} \right),$$

where $f_{\mathcal{M}(q(G))}$ is the density of $\mathcal{M}(q(G))$ and likewise for $f_{\mathcal{M}(q(G'))}$. However, the proof of (Ji & Li, 2024) defines the PRV L as a function of some unspecified vector s and analyzes the PRV over the randomness of both $\mathcal{M}(q(G))$ and $\mathcal{M}(q(G'))$. Indeed, they bound the distribution of the angle between the uniformly random directions of the noises used in $\mathcal{M}(q(G))$ and $\mathcal{M}(q(G'))$, and conclude that this angle must be close to $\pi/2$. However, the more appropriate distribution to analyze would be the angle between the noise used in $\mathcal{M}(q(G))$ and the difference vector $\mu \stackrel{\text{def}}{=} q(G) - q(G')$. By similar analysis, it is actually *this* angle that is close to $\pi/2$. This is much different than the angle between the directions in which the noises in $\mathcal{M}(q(G))$ and $\mathcal{M}(q(G'))$ must point to hit Y , which the authors seem to attempt to study when considering the ratios of the densities $f_{\mathcal{M}(q(G))}$ and $f_{\mathcal{M}(q(G'))}$ on Y ; this angle actually can be far from $\pi/2$ since the direction of the noise from $q(G')$ to hit Y will be quite different than the direction of the difference vector μ .

E.1. Numerical contradiction of the claimed R1SMG calibration

According to (Ji & Li, 2024, Theorem 5), the R1SMG mechanism is (ε, δ) -DP if its corresponding rank-1 noise parameter σ_* satisfies

$$\psi(T, \delta) = \left(\frac{\delta_{\text{claimed}} \Gamma(\frac{T-1}{2})}{\sqrt{\pi} \Gamma(\frac{T}{2})} \right)^{\frac{2}{T-2}},$$

$$\sigma_*(\varepsilon) \geq \frac{2s^2}{\varepsilon \psi(T, \delta_{\text{claimed}})}.$$

We test this claim in dimension $T = 128$ with sensitivity $s = 1$ at target $\delta = 10^{-5}$, for $\varepsilon \in \{0.1, 1, 2, 4, 8\}$. For each ε , we instantiate the R1SMG mechanism using the *minimal* value $\sigma_*(\varepsilon) = \frac{2s^2}{\varepsilon \psi(T, \delta)}$ defined above, and then compute the instantiated mechanism's *actual* optimal delta $\delta_{\text{true}}(\varepsilon)$ via Lemma 4.1. This is because the R1SMG noise is a member of the SGG noise family, so we can use Lemma 4.1 to analyze its privacy. Concretely, the R1SMG noise is rank-1 Gaussian along a random direction; equivalently, it is spherical with half-normal radius $R = \sqrt{\sigma_*} |Z|$ for $Z \sim \mathcal{N}(0, 1)$. In SGG's

parameterization, this corresponds to an SGG mechanism with parameters

$$\alpha = 0, \quad p = 2, \quad \beta = \frac{1}{2\sigma_*}.$$

Table 3 show a dramatic violation: across $\varepsilon \in \{0.1, 1, 2, 4, 8\}$, the computed $\delta_{\text{true}}(\varepsilon)$ ranges from 0.813 to 0.9998, vastly larger than the claimed $\delta = 10^{-5}$.