On Inherent Privacy of Posterior Sampling: A Unified Rényi-Divergence Framework

Ruyu Zhou
Department of Computer Science and Statistics
University of Rhode Island
ruyu.zhou@uri.edu

Fang Liu
Applied & Computational Math and Statistics
University of Notre Dame
fang.liu.131@nd.edu

Abstract

Differential Privacy (DP) provides a rigorous framework for quantifying privacy guarantees. While most methods achieve DP by injecting calibrated random noise, the intrinsic randomness of certain procedures such as sampling can yield DP guarantees "for free". In this work, we develop a unified Rényi divergence framework to characterize the inherent DP guarantees of sampling from a Bayesian posterior distribution. We show that the privacy loss incurred by One-Posterior-Sampling (OPS) is governed by posterior exponential moments of the single-record log-likelihood ratio, and establish explicit DP guarantees under uniformly bounded, sub-Gaussian, or sub-exponential tail regimes. We apply the framework to representative Bayesian models, including categorical likelihoods with arbitrary priors, Gaussian-Gaussian conjugate pair, generalized linear models from the exponential family and linear regression with Gaussian likelihood, both with Gaussian priors on the regression coefficients. Our results elucidate sample size, model structure, neighboring relations in DP (bounded or unbounded), and how prior if applicable, jointly determine the inherent DP guarantees of OPS. Furthermore, our results recover the existing DP guarantees for OPS as special cases – often with tighter privacy loss bounds – and substantially broaden the class of Bayesian models for which the inherent DP guarantees of OPS can be rigorously characterized.

1 Introduction

Background and Motivation. Differential Privacy (DP) has emerged as the gold standard for providing rigorous privacy guarantees when releasing information from sensitive databases since its introduction in 2006 [Dwork et al., 2006a,b]. By bounding the change in a randomized algorithm's output distribution induced by any single individual's data, classical DP notion ensures that no individual can exert a disproportionate influence on the outcome. However, the worst-case nature of pure DP can be overly conservative for practical applications. This has motivated several influential relaxations that enable more refined privacy accounting, such as Rényi DP [Mironov, 2017], zero-concentrated DP [Bun and Steinke, 2016], Gaussian DP [Dong et al., 2022], among others.

While a target DP level is typically achieved by explicitly injecting calibrated noise – a paradigm underpinning most DP methods – many learning procedures are intrinsically stochastic. Accounting for privacy protection provided by such inherent stochasticity would help meet DP requirements while reducing, or even eliminating, the need for external perturbation, thereby better preserving model utility. Bayesian posterior sampling naturally falls into this category and offers a principled setting for exploiting this form of "free" privacy. In particular, a single posterior draw $\boldsymbol{\theta}^*$ defines a complete randomized predictor or synthetic-data generator; any number of predictions or synthetic records computed from $\boldsymbol{\theta}^*$ inherit its privacy guarantee by post-processing, without composition.

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

Despite growing interest, existing work on DP guarantees for One-Posterior-Sampling (OPS) remain limited (see Related Work): many require uniformly bounded likelihood ratios, bounded sample or parameter spaces, or modified (pseudo-)posterior mechanisms; or yield loose privacy loss bounds. Recent work on posterior sampling under contamination and pseudo-posterior mechanisms (see Related Work) shows that this problem remains active and practically relevant. We revisit OPS and develop sharper DP guarantees under weaker and more broadly applicable conditions.

Our Contributions.

- We establish an *exact characterization* of the Rényi divergence between the posteriors formulated over two neighboring datasets, showing that the privacy loss of OPS is fully determined by two specific *exponential moments* of the single-record log-likelihood ratio.
- We derive explicit DP guarantees for OPS under three log-likelihood ratio regimes: *uniformly bounded*, *sub-Gaussian*, and *sub-exponential*. These results recover existing OPS DP guarantees as special cases and are compatible with various DP notions.
- We apply the framework to a broader class of Bayesian models than previously studied. The resulting bounds show how prior, sample size, model structure, and neighboring relation jointly determine the inherent DP guarantees of OPS; in particular, under unbounded neighboring, the two directions of the Rényi divergence can differ even in their range of finite orders.

Related Work. Chaudhuri and Mishra [2006] initiated the idea that random sampling itself can provide DP guarantees. Building on this insight, Li et al. [2012] showed that random sampling before k-anonymization [Sweeney, 2002] yields approximate DP. Subsequent work has mainly followed two directions: 1) studying DP guarantee induced by algorithmic randomness, including uncertainty and noise in control systems [Giraldo et al., 2017] and stochastic generative models such as GANs [Lin et al., 2021]; 2) studying the privacy amplification effect, where subsampling randomness amplifies the privacy guarantees of a base mechanism. Foundational results include Balle et al. [2018] and survey by Steinke [2022], with recent extensions to multistage sampling schemes [Zhao et al., 2024].

Privacy guarantees from posterior sampling in the Bayesian framework have also garnered significant research attention, especially for OPS. Early work Mir [2013], Dimitrakakis et al. [2014, 2017], Wang et al. [2015], Zhao and Liu [2022] established pure DP guarantees for general OPS by assuming uniformly bounded posterior distribution, log-likelihood ratio, or log-likelihood between neighboring datasets – an assumption that is often too restrictive. The resulting DP guarantees scale with the assumed uniform bound and can be conservative due to triangle-inequality-based arguments. Subsequent work examined OPS in specific Bayesian models, including discrete Bayesian networks [Zhang et al., 2016] and exponential family models with conjugate priors [Foulds et al., 2016]. The work most closely related to ours is Geumlek et al. [2017], which studies OPS under Rényi DP [Mironov, 2017], but is limited to exponential family and Bayesian logistic regression. In contrast, our work develops a unified Rényi DP characterization for OPS that does not require an exponential-family likelihood or a conjugate prior, and recovers existing results as special cases.

Beyond OPS, intrinsic DP guarantees have also been studied for related posterior-sampling-based inference procedures, including collapsed Gibbs sampling for latent Dirichlet allocation [Zhao et al., 2020, Blei et al., 2003] and stochastic gradient MCMC [Li et al., 2019]. Related directions include combining OPS with β divergences to obtain bounded sensitivity [Jewson et al., 2023], per-instance DP analyses for linear regression [Minami et al., 2016, Wang, 2019], and posterior sampling under contamination models without boundedness assumptions [Hu et al., 2024]. Our work differs by focusing on unmodified OPS and deriving sharper privacy-loss guarantees under weaker assumptions.

2 Preliminaries

Definition 1 (Differential Privacy (DP))**.** *[Dwork et al., 2006a] A randomized mechanism $\mathcal{M}$ is said to satisfy (ε, δ)-DP if for all pairs of neighboring datasets $(\mathbf{Y}, \mathbf{Y}')$ differing by one record, denoted by $d(\mathbf{Y}, \mathbf{Y}') = 1$, and for all subsets $\mathcal{S} \subseteq$ image$(\mathcal{M})$, it holds that*

$$\Pr(\mathcal{M}(\mathbf{Y}) \in \mathcal{S}) \leq e^{\varepsilon} \Pr(\mathcal{M}(\mathbf{Y}') \in \mathcal{S}) + \delta. \tag{1}$$

$d(\mathbf{Y}, \mathbf{Y}') = 1$ refers to either *bounded* neighboring [Kifer and Machanavajjhala, 2011], where $\mathbf{Y}$ becomes $\mathbf{Y}'$ by substituting exactly one record, or *unbounded* neighboring, where $\mathbf{Y}$ becomes $\mathbf{Y}'$ by

adding or removing a single record. $\varepsilon > 0$ and $\delta \geq 0$ in Eq. (1) are privacy loss parameters, with smaller values indicating stronger privacy protection. When $\delta = 0$, (ε, δ)-DP becomes pure ε-DP.

DP can also be characterized using a privacy loss random variable that quantifies the privacy loss incurred by releasing a particular output (Definition 2).

Definition 2 (Privacy Loss Random Variable)**.** *[Bun and Steinke, 2016, Sommer et al., 2018] For a randomized mechanism $\mathcal{M}$, let $o \in image(\mathcal{M})$ be any potential output of $\mathcal{M}$ and let $\mathbf{X}_0, \mathbf{X}_1 \in \mathcal{X}$ be two inputs. The privacy loss random variable $\mathcal{L}$ associated with releasing o is defined as*

$$\mathcal{L}_{\mathcal{M}(\mathbf{X}_0)\|\mathcal{M}(\mathbf{X}_1)}(o) = \log(\Pr(\mathcal{M}(\mathbf{X}_0) = o) / \Pr(\mathcal{M}(\mathbf{X}_1) = o)) . \tag{2}$$

For any two neighboring datasets $\mathbf{Y}$ and $\mathbf{Y}'$, a mechanism $\mathcal{M}$ satisfies pure ε-DP if and only if $\Pr(\mathcal{L}_{\mathcal{M}(\mathbf{Y})\|\mathcal{M}(\mathbf{Y}')} > \varepsilon) = 0$, whereas (ε, δ)-DP is equivalent, up to a small loss in parameters, to the requirement that $\Pr(\mathcal{L}_{\mathcal{M}(\mathbf{Y})\|\mathcal{M}(\mathbf{Y}')} > \varepsilon) \leq \delta$. A relaxation of DP based on the Rényi divergence is:

Definition 3 (Rényi Differential Privacy (RDP))**.** *[Mironov, 2017] A randomized mechanism $\mathcal{M}$ is said to have (α, ε)-RDP if for all pairs of neighboring datasets $\mathbf{Y}, \mathbf{Y}'$, it holds that*

$$D_\alpha(\mathcal{M}(\mathbf{Y}) \| \mathcal{M}(\mathbf{Y}')) \leq \varepsilon, \tag{3}$$

where $D_\alpha(P\|Q)$ is the Rényi divergence [Rényi, 1961] of order $\alpha > 1$ between distributions P, Q:

$$D_\alpha(P\|Q) = (\alpha - 1)^{-1} \log \mathbb{E}_{x\sim Q} \left(P/Q\right)^\alpha = (\alpha - 1)^{-1} \log \mathbb{E}_{x\sim P} \left(P/Q\right)^{\alpha-1} . \tag{4}$$

Beyond defining RDP, Rényi divergence is also linked to several popular DP notions, such as (ε, δ)-DP, ρ zero-concentrated DP (zCDP), and Gaussian DP (GDP), in the sense that they can be expressed via Rényi divergence. The following proposition summarizes these relationships.

Proposition 1 (Relations between Rényi divergence and DP notions)**.** *Let $\mathcal{M}$ be a randomized mechanism, and $\mathbf{Y}, \mathbf{Y}'$ be any pair of neighboring datasets satisfying $d(\mathbf{Y}, \mathbf{Y}') = 1$. Then,*

- *$\mathcal{M}$ is of ε-DP if and only if $D_\infty(\mathcal{M}(\mathbf{Y})\|\mathcal{M}(\mathbf{Y}')) \leq \varepsilon$ [Mironov, 2017]*[1].
- *$\mathcal{M}$ is of ρ-zCDP if and only if $\forall \alpha \in (1, \infty)$, $D_\alpha(\mathcal{M}(\mathbf{Y})\|\mathcal{M}(\mathbf{Y}')) \leq \rho\alpha$ [Bun and Steinke, 2016].*
- *If $\mathcal{M}$ is of (α, ε)-RDP, it is also of $(\varepsilon + (\alpha-1)^{-1} \log 1/\delta, \delta)$-DP for $\forall\ 0 < \delta < 1$ [Mironov, 2017].*
- *If $\mathcal{M}$ is of μ-GDP, it is also of $\mu^2/2$-zCDP [Dong et al., 2022].*

3 A Unified Rényi Framework for Privacy Loss from Posterior Sampling

We develop a unified Rényi-divergence framework for characterizing the inherent privacy loss of posterior sampling, based on a decomposition of the privacy loss random variable between the posterior distributions formulated over two neighboring datasets.

3.1 Privacy Loss Decomposition for Neighboring Posteriors

Let $\mathbf{y} \in \mathcal{Y}$ denote a single observation, its distribution is $p(\mathbf{y}|\boldsymbol{\theta})$ with parameter $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, and $\pi(\boldsymbol{\theta})$ is the prior distribution. Given a dataset $\mathbf{Y} = \{\mathbf{y}_i\}_{i=1}^n$ and assuming independence across the n observations, the posterior distribution of $\boldsymbol{\theta}$ is

$$\pi(\boldsymbol{\theta}|\mathbf{Y}) = \frac{\pi(\boldsymbol{\theta}) \prod_{i=1}^n p(\mathbf{y}_i|\boldsymbol{\theta})}{\int \prod_{i=1}^n p(\mathbf{y}_i|\boldsymbol{\theta}) \pi(\boldsymbol{\theta}) d\boldsymbol{\theta}} = \frac{\pi(\boldsymbol{\theta}) \ell(\boldsymbol{\theta}; \mathbf{Y})}{p(\mathbf{Y})}, \tag{5}$$

where $\ell(\boldsymbol{\theta}; \mathbf{Y}) \triangleq \prod_{i=1}^n p(\mathbf{y}_i|\boldsymbol{\theta})$ denotes the likelihood function and $p(\mathbf{Y}) = \int \ell(\boldsymbol{\theta}; \mathbf{Y}) \pi(\boldsymbol{\theta}) d\boldsymbol{\theta}$ is the normalizing constant. Consider any pair of neighboring datasets $d(\mathbf{Y}, \mathbf{Y}') = 1$ with corresponding posterior distributions, $\pi(\boldsymbol{\theta}|\mathbf{Y})$ and $\pi(\boldsymbol{\theta}|\mathbf{Y}')$, which we refer to as **neighboring posteriors**. The privacy loss random variable $c(\boldsymbol{\theta})$ associated with releasing one posterior sample $\boldsymbol{\theta}$ is

$$c(\boldsymbol{\theta}) = \log \left(\frac{\pi(\boldsymbol{\theta}|\mathbf{Y})}{\pi(\boldsymbol{\theta}|\mathbf{Y}')} \right) = \log \left(\frac{\cancel{\pi(\boldsymbol{\theta})} \ell(\boldsymbol{\theta}; \mathbf{Y})}{p(\mathbf{Y})} \cdot \frac{p(\mathbf{Y}')}{\cancel{\pi(\boldsymbol{\theta})} \ell(\boldsymbol{\theta}; \mathbf{Y}')} \right) \text{ per Eq. (5)}$$

[1] $D_\infty(P\|Q) = \sup_{x \in \text{support } P} \log(P/Q)$ is the max divergence between P and Q when $\alpha = \infty$.

$$= \underbrace{\log\left(\frac{\ell(\boldsymbol{\theta};\mathbf{Y})}{\ell(\boldsymbol{\theta};\mathbf{Y}')}\right)}_{\triangleq\ c_1(\boldsymbol{\theta})} + \underbrace{\log\left(\frac{p(\mathbf{Y}')}{p(\mathbf{Y})}\right)}_{\triangleq\ c_2} = c_1(\boldsymbol{\theta}) + c_2. \tag{6}$$

Here, $c_1(\boldsymbol{\theta})$ is the pointwise log-likelihood ratio evaluated at $\boldsymbol{\theta}$, while c_2 represents the marginal log-likelihood ratio obtained after integrating out $\boldsymbol{\theta}$. $c_1(\boldsymbol{\theta})$ and c_2 are related; in particular, c_2 can be expressed as the logarithm of an exponential moment of $c_1(\boldsymbol{\theta})$ under the posterior $\pi(\boldsymbol{\theta}|\mathbf{Y})$. This relation is formalized in Lemma 2, with the proof is provided in Appendix A.1.

Lemma 2 (Marginal-Pointwise identity)**.** $c_2 = \log \mathbb{E}_{\boldsymbol{\theta}\sim\pi(\boldsymbol{\theta}|\mathbf{Y})}\left[\exp(-c_1(\boldsymbol{\theta}))\right]$. *Specifically, if there exist finite constants $a, b \in \mathbb{R}$ with $a < b$ such that $c_1(\boldsymbol{\theta}) \in [a, b]$ for $\forall \boldsymbol{\theta} \in \boldsymbol{\Theta}$, then $c_2 \in [-b, -a]$.*

3.2 Rényi Divergence Characterization

Building on the decomposition in Eq. (6), we establish an exact expression for the Rényi divergence of any order $\alpha > 1$ between neighboring posteriors in Theorem 3. The proof is provided in Appendix A.2.

Theorem 3 (Exact characterization of Rényi divergence between neighboring posteriors)**.** *Let $P = \pi(\boldsymbol{\theta}|\mathbf{Y})$ and $Q = \pi(\boldsymbol{\theta}|\mathbf{Y}')$ be a pair of neighboring posteriors. With $c_1(\boldsymbol{\theta})$ and c_2 defined in Eq.* (6)*, the Rényi divergence of order $\alpha > 1$ between P and Q is*

$$D_\alpha(P\|Q) = c_2 + (\alpha-1)^{-1}\log \mathbb{E}_{\boldsymbol{\theta}\sim P}\left[e^{(\alpha-1)c_1(\boldsymbol{\theta})}\right] \tag{7}$$

$$= \log\left[\left(\mathbb{E}_{\boldsymbol{\theta}\sim P}\left[e^{-c_1(\boldsymbol{\theta})}\right]\right)\cdot\left(\mathbb{E}_{\boldsymbol{\theta}\sim P}\left[e^{(\alpha-1)c_1(\boldsymbol{\theta})}\right]\right)^{\frac{1}{\alpha-1}}\right] \text{ per Lemma 2.} \tag{8}$$

Eq. (8) suggests that the Rényi divergence of order α – $D_\alpha(P\|Q)$ – between neighboring posteriors is fully determined by exponential moments $\mathbb{E}_{\boldsymbol{\theta}\sim P}[e^{tc_1(\boldsymbol{\theta})}]$, evaluated at $t = -1$ and $t = \alpha - 1$, of the log-likelihood ratio $c_1(\boldsymbol{\theta})$ under $P = \pi(\boldsymbol{\theta}|\mathbf{Y})$. This establishes Theorem 3 as the foundational result to build a unified framework for analyzing the inherent privacy guarantees of OPS.

Remark 4 (Reverse direction)**.** *Theorem 3 holds for any ordered pair of neighboring posteriors. Swapping the roles of $\mathbf{Y}$ and $\mathbf{Y}'$, let $\tilde{c}_1(\boldsymbol{\theta}) \triangleq \log\frac{\ell(\boldsymbol{\theta};\mathbf{Y}')}{\ell(\boldsymbol{\theta};\mathbf{Y})} = -c_1(\boldsymbol{\theta})$; then*

$$D_\alpha(Q\|P) = \log\left[\left(\mathbb{E}_{\boldsymbol{\theta}\sim Q}\left[e^{-\tilde{c}_1(\boldsymbol{\theta})}\right]\right)\cdot\left(\mathbb{E}_{\boldsymbol{\theta}\sim Q}\left[e^{(\alpha-1)\tilde{c}_1(\boldsymbol{\theta})}\right]\right)^{\frac{1}{\alpha-1}}\right], \tag{9}$$

with expectations taken under $Q = \pi(\boldsymbol{\theta}|\mathbf{Y}')$. Since $d(\mathbf{Y}, \mathbf{Y}') = 1 \Leftrightarrow d(\mathbf{Y}', \mathbf{Y}) = 1$, (α, ε)-RDP requires $\sup_{(\mathbf{Y},\mathbf{Y}'):d(\mathbf{Y},\mathbf{Y}')=1} \max\{D_\alpha(P\|Q), D_\alpha(Q\|P)\} \le \varepsilon$.

As Eqs. (7), (8) and (9) depend on the specific neighboring pair $\mathbf{Y}, \mathbf{Y}'$ under consideration, the task of deriving DP guarantees for OPS reduces to obtaining data-independent upper bounds for these exponential moments: of $c_1(\boldsymbol{\theta})$ under $\pi(\boldsymbol{\theta}|\mathbf{Y})$ and of $\tilde{c}_1(\boldsymbol{\theta})$ under $\pi(\boldsymbol{\theta}|\mathbf{Y}')$. To this end, we derive a series of results in Sections 4 to 6 under progressively weaker regularity conditions on $c_1(\boldsymbol{\theta})$. These results significantly expand the class of Bayesian models for which explicit upper bounds on D_α in both directions can be obtained, while also tightening the bounds available in the existing literature. Furthermore, the derived upper bounds on D_α can be interpreted under various DP notions, either through definitions or via well-established conversion results summarized in Proposition 1.

4 DP Guarantees of OPS with Uniformly Bounded Log-Likelihood Ratio

Under uniform boundedness of $c_1(\boldsymbol{\theta})$, we derive explicit DP guarantees for OPS in Corollaries 5 and 6 based on Theorem 3, and demonstrate their application to categorical likelihoods in Example 7.

Corollary 5 (Attainable RDP guarantees of OPS with uniformly bounded log-likelihood ratio)**.** *Suppose that $c_1(\boldsymbol{\theta})$ is uniformly bounded, that is, there exist finite constants $a, b \in \mathbb{R}$ with $a < b$ such that $c_1(\boldsymbol{\theta}) \in [a, b]$ for all $\boldsymbol{\theta} \in \boldsymbol{\Theta}$. Define $\Delta \triangleq b - a$ and $A \triangleq e^{-\Delta} \in (0, 1)$, $B \triangleq e^{(\alpha-1)\Delta} > 1$. Then, for every $\alpha > 1$,*

$$\max\{D_\alpha(P||Q), D_\alpha(Q||P)\} \le \varepsilon_{\mathrm{ub}}(\alpha, \Delta) \triangleq \log\left(\frac{(\alpha-1)(B-A)}{\alpha(B-1)}\right) + \frac{1}{\alpha-1}\log\left(\frac{B-A}{\alpha(1-A)}\right), \tag{10}$$

and releasing OPS satisfies $(\alpha, \varepsilon_{\mathrm{ub}}(\alpha, \Delta))$-RDP.

The full proof is in Appendix A.3. We provide the proof sketch below. First, we show that $D_\alpha(P||Q)$ is monotonically increasing with respect to the convex order of $c_1(\boldsymbol{\theta})$; Second, among all random variables supported on $[a, b]$, the endpoint-supported binary random variable X^* – defined by $\Pr(X^* = a) = 1 - p$ and $\Pr(X^* = b) = p$, with $\mathbb{E}(X^*) = \mathbb{E}[c_1(\boldsymbol{\theta})]$ – attains the maximal convex order. Evaluating $D_\alpha(P||Q)$ at X^* thus yields an analytical upper bound for a given pair of neighboring datasets, parameterized by p. Finally, by maximizing this neighboring pair-specific upper bound over $p \in [0, 1]$, we obtain the data-independent upper bound $\varepsilon_{\mathrm{ub}}(\alpha, \Delta)$.

The bound in Eq. (10) is sharp over all random variables supported on $[a, b]$, but may be conservative for a particular neighboring posterior pair. Since $\tilde{c}_1(\boldsymbol{\theta}) = -c_1(\boldsymbol{\theta}) \in [-b, -a]$ has the same range Δ, the same bound holds for $D_\alpha(Q\|P)$ by Remark 4. Corollary 5 reveals two insightful properties of $\varepsilon_{\mathrm{ub}}(\alpha, \Delta)$ in Eq. (10) for a fixed $\alpha > 1$. First, **shift-invariance** property: $\varepsilon_{\mathrm{ub}}(\alpha, \Delta)$ depends solely on the range $\Delta = b - a$, rather than the specific values of a and b, implying that a reparameterization that only shifts $c_1(\boldsymbol{\theta})$ does not affect the inherent DP guarantee offered by OPS. Second, **monotonic increase** of $\varepsilon_{\mathrm{ub}}(\alpha, \Delta)$ with Δ: larger values of Δ lead to weaker privacy guarantees, which aligns with the intuition that greater dissimilarity between the two neighboring posteriors implies weaker intrinsic privacy protection provided by OPS.

Corollary 6 below recovers the existing pure-DP guarantees for the symmetric case $a = -b$ (e.g., Thm. 2 in Dimitrakakis et al. [2014], Thm. 8 in Dimitrakakis et al. [2017], and Thm. 1 in Wang et al. [2015]) as a special case; the proof is in Appendix A.4.

Corollary 6 (Pure-DP guarantee)**.** *Under the conditions in Corollary 5, for every $\alpha > 1$,*

$$\max\{D_\alpha(P\|Q), D_\alpha(Q\|P)\} \le b - a = \Delta = \sup_{\alpha>1} \varepsilon_{\mathrm{ub}}(\alpha, \Delta), \tag{11}$$

implying OPS satisfies pure ε-DP with $\varepsilon = \Delta$.

Corollary 5 refines the pure-DP guarantee in Corollary 6 by providing an order-dependent Rényi divergence bound with two benefits. First, it addresses the attainability gap in the non-degenerate case of $a < b$. While Δ is a valid upper bound on $D_\alpha(P||Q)$, it is generally unattainable at finite α except $a = b$. Specifically, attaining $\Delta = b - a$ would require simultaneously that $(\alpha - 1)^{-1} \log \mathbb{E}_{\boldsymbol{\theta}\sim P}[e^{(\alpha-1)c_1(\boldsymbol{\theta})}] = b$ and $c_2 = -a$. However, per Lemma 2, c_2 decreases monotonically with $c_1(\boldsymbol{\theta})$: the condition $c_2 = -a$ implies $c_1(\boldsymbol{\theta}) = a$ almost surely, whereas the condition $(\alpha - 1)^{-1} \log \mathbb{E}_{\boldsymbol{\theta}\sim P}[e^{(\alpha-1)c_1(\boldsymbol{\theta})}] = b$ requires $c_1(\boldsymbol{\theta}) = b$ almost surely. These two conditions cannot hold simultaneously unless $b = a$. Corollary 5 resolves this attainability gap by explicitly accounting for the interdependence of $c_1(\boldsymbol{\theta})$ and c_2. Second, $\varepsilon_{\mathrm{ub}}(\alpha, \Delta)$ in Eq. (10) captures the full spectrum of Rényi divergence across all $\alpha > 1$, whereas Corollary 6 and prior works provide only the static, worst-case bound corresponding to the limit $\alpha \to \infty$. This order-dependent profile in Corollary 5 thus enables sharper RDP accounting for composition. For fixed Δ, expanding $\varepsilon_{\mathrm{ub}}(\alpha, \Delta)$ via Taylor expansion as $\alpha \to \infty$ gives $\varepsilon_{\mathrm{ub}}(\alpha, \Delta) = \Delta - \alpha^{-1} \log(\alpha(1 - e^{-\Delta})) + \mathcal{O}(\alpha^{-1})$, that is, $\varepsilon_{\mathrm{ub}}(\alpha, \Delta)$ converges to the pure-DP bound Δ at the rate of $\Theta(\log \alpha/\alpha)$ (see Appendix A.4 for the proof).

Example 7 (Categorical likelihood with arbitrary prior)**.** *Let $Y = \{y_i\}_{i=1}^n$ be a dataset with categorical observations $y_i \in \{1, \ldots, K\}$. Suppose $y_i|\boldsymbol{\theta} \overset{i.i.d}{\sim}$ Categorical($\boldsymbol{\theta}$), where $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_K)$ with $\sum_{j=1}^K \theta_j = 1$ and $\theta_j = \Pr(y_i = j)$. Let $s_j(Y) = \sum_{i=1}^n \mathbf{1}\{y_i = j\}$ denote the j-th category count.*

(1) Bounded neighboring (Y, Y'): there exist distinct indices $j \neq k$ for $j, k \in \{1, \ldots, K\}$ such that $s_j(Y) = s_j(Y') + 1, s_k(Y) = s_k(Y') - 1$, and $s_\ell(Y) = s_\ell(Y')$ for all $\ell \notin \{j, k\}$.

$$\begin{aligned} c_{1,\mathrm{bounded}}(\boldsymbol{\theta}) &= \log \frac{\ell(\boldsymbol{\theta}; Y)}{\ell(\boldsymbol{\theta}; Y')} = \sum_{\ell=1}^K \big(s_\ell(Y) - s_\ell(Y')\big) \log \theta_\ell = \log \frac{\theta_j}{\theta_k}, \\ D_{\alpha,\mathrm{bounded}}\big(P\|Q\big) &\le \max_{j\neq k} \Big\{ \log \mathbb{E}_{\boldsymbol{\theta}\sim P}[\theta_k/\theta_j] + (\alpha-1)^{-1} \log \mathbb{E}_{\boldsymbol{\theta}\sim P}\left[(\theta_j/\theta_k)^{\alpha-1}\right] \Big\}. \end{aligned} \tag{12}$$

(2) Unbounded neighboring (Y, Y'): there exists a unique index $j \in \{1, \ldots, K\}$ such that $s_j(Y) = s_j(Y') + 1$, and $s_\ell(Y) = s_\ell(Y')$ for all $\ell \neq j$.

$$\begin{aligned} c_{1,\mathrm{unbounded}}(\boldsymbol{\theta}) &= \log \frac{\ell(\boldsymbol{\theta}; Y)}{\ell(\boldsymbol{\theta}; Y')} = \sum_{\ell=1}^K \big(s_\ell(Y) - s_\ell(Y')\big) \log \theta_\ell = \log \theta_j, \\ D_{\alpha,\mathrm{unbounded}}(P\|Q) &\le \max_{j\in\{1,\ldots,K\}} \Big\{ \log \mathbb{E}_{\boldsymbol{\theta}\sim P}\big[\theta_j^{-1}\big] + (\alpha-1)^{-1} \log \mathbb{E}_{\boldsymbol{\theta}\sim P}\big[\theta_j^{\alpha-1}\big] \Big\}. \end{aligned} \tag{13}$$

Eqs. (12) and (13) are direct applications of Theorem 3; the proof is in Appendix A.7. Since $c_{1,\text{bounded}}(\boldsymbol{\theta})$ and $c_{1,\text{unbounded}}(\boldsymbol{\theta})$ diverge as $\theta_j \to 0$, we restrict the parameter space to obtain explicit DP guarantees. Specifically, suppose $\theta_j \in [\eta, 1-\eta]$ for some $\eta \in (0, 1/2)$ and all j; then $c_{1,\text{bounded}}(\boldsymbol{\theta}) = \log(\theta_j/\theta_k) \in [-\log\frac{1-\eta}{\eta}, \log\frac{1-\eta}{\eta}]$ and $c_{1,\text{unbounded}}(\boldsymbol{\theta}) = \log\theta_j \in [\log\eta, \log(1-\eta)]$, with ranges $\Delta_{\text{bounded}} = 2\log\frac{1-\eta}{\eta}$ and $\Delta_{\text{unbounded}} = \log\frac{1-\eta}{\eta}$, respectively. Hence the bounded neighboring dominates with $\Delta = 2\log\frac{1-\eta}{\eta}$. Some concrete examples of the inherent RDP guarantees for OPS when $\alpha = 2$ (where Eq. (10) reduces to $\varepsilon_{\text{ub}}(2, \Delta) = \log([1+\cosh(\Delta)]/2)$ with $\Delta = 2\log\frac{1-\eta}{\eta}$) and $\alpha = \infty$ are given below:

- If $\eta = 0.01$, Corollary 5 yields $(2, \sim 7.804)$-RDP, and Corollary 6 yields $(\infty, \sim 9.190)$-RDP.
- If $\eta = 0.1$, Corollary 5 yields $(2, \sim 3.033)$-RDP, and Corollary 6 yields $(\infty, \sim 4.394)$-RDP.
- If $\eta = 0.3$, Corollary 5 yields $(2, \sim 0.646)$-RDP, and Corollary 6 yields $(\infty, \sim 1.695)$-RDP.

We remark that Eqs. (12) and (13) can be evaluated exactly for specific prior choices to produce distribution-specific bounds. Examples include the Bernoulli-Beta pair (Appendix A.7.3) and Categorical-Dirichlet pair (Appendix A.7).

5 DP Guarantees of OPS with sub-Gaussian Log-Likelihood Ratio

In this section, we establish explicit DP guarantees for OPS under a weaker assumption than the uniform boundedness assumption in Section 4, that is, when $c_1(\boldsymbol{\theta})$ is sub-Gaussian. The main result (Corollary 8) applies to a broader class of Bayesian models than those covered in Corollary 5.

Corollary 8 (RDP guarantees of OPS with sub-Gaussian log-likelihood ratio)**.** *Suppose that there exists a variance proxy $\sigma^2_{\mathbf{Y},\mathbf{Y}'} \in (0, \infty)$ such that for $\forall t \in \mathbb{R}$, it holds that $\log \mathbb{E}_{\boldsymbol{\theta}\sim P}(e^{t(c_1(\boldsymbol{\theta}) - \mu_{\mathbf{Y},\mathbf{Y}'})}) \le t^2\sigma^2_{\mathbf{Y},\mathbf{Y}'}/2$ where $\mu_{\mathbf{Y},\mathbf{Y}'} = \mathbb{E}_{\boldsymbol{\theta}\sim P}[c_1(\boldsymbol{\theta})]$. Define*

$$\bar{\sigma}^2 \triangleq \sup_{(\mathbf{Y},\mathbf{Y}'):d(\mathbf{Y},\mathbf{Y}')=1} \sigma^2_{\mathbf{Y},\mathbf{Y}'}, \tag{14}$$

where the supremum is over all ordered neighboring pairs. Then, for every $\alpha > 1$,

$$\max\{D_\alpha(P||Q), D_\alpha(Q||P)\} \le \varepsilon_{\text{subGaus}}(\alpha, \bar{\sigma}^2) \triangleq \alpha\bar{\sigma}^2/2, \tag{15}$$

and releasing OPS satisfies $(\alpha, \varepsilon_{\text{subGaus}}(\alpha, \bar{\sigma}^2))$-RDP.

The proof is in Appendix A.5. Unlike the uniformly bounded case in Corollary 5, $\varepsilon_{\text{subGaus}}(\alpha, \bar{\sigma}^2)$ in Eq. (15) scales linearly with α and diverges as $\alpha \to \infty$. Therefore, the sub-Gaussian regime does not yield pure DP guarantees for OPS. Nevertheless, this linear dependence on α aligns naturally with other relaxed DP notions. Specifically, since Eq. (15) holds for all $\alpha \in (1, \infty)$, OPS satisfies ρ-zCDP with $\rho = \bar{\sigma}^2/2$, and $(\alpha\bar{\sigma}^2/2 + \frac{\log(1/\delta)}{\alpha-1}, \delta)$-DP for $\forall\ 0 < \delta < 1$ via Proposition 1. Furthermore, by Corollary 3 in Mironov [2017], the Gaussian mechanism with scale parameter σ^2 satisfies $(\alpha, \alpha/(2\sigma^2))$-RDP. This facilitates an intuitive interpretation of Corollary 8:

Remark: When $c_1(\boldsymbol{\theta})$ is sub-Gaussian, the inherent DP guarantees of OPS are conceptually equivalent to applying the Gaussian mechanism with scale parameter $\bar{\sigma}^{-2}$ defined in Eq. (14).

We apply Corollary 8 and Theorem 3 to generalized linear models (GLMs) with Gaussian priors (Section 5.1), and to models with Gaussian likelihoods (Section 5.2), including Gaussian priors on Gaussian means and Gaussian priors on Bayesian linear regression coefficients. Our results show that the choice of neighboring relation can affect the tail behavior of $c_1(\boldsymbol{\theta})$, leading to qualitatively different inherent privacy guarantees.

5.1 Generalized Linear Models with Gaussian Prior

Let each data record be $(\mathbf{x}_i, y_i)$, where $\mathbf{x}_i \in \mathbb{R}^d$ is the covariate vector and $y_i \in \mathcal{Y}$ is the response. Let $\mathbf{Y} = \{(\mathbf{x}_i, y_i)\}_{i=1}^n$ denote the dataset. Conditioned on $\mathbf{x}_i$ and the parameter $\boldsymbol{\beta} \in \mathbb{R}^d$, the response y_i follows an exponential family distribution: $p(y_i|\mathbf{x}_i, \boldsymbol{\beta}) = \exp(y_i\,\mathbf{x}_i^\top\boldsymbol{\beta} - \psi(\mathbf{x}_i^\top\boldsymbol{\beta}) + h(y_i))$, where $\psi(\cdot)$ is the log-partition function and $\exp(h(y_i))$ is the base measure (which may include a known dispersion parameter, such as σ^2 in Gaussian likelihood). We show that, under a bounded record-level

score condition, $c_1(\boldsymbol{\beta})$ for GLMs with a Gaussian prior on $\boldsymbol{\beta}$ is sub-Gaussian under the posterior $\pi(\cdot|\mathbf{Y})$. This yields explicit OPS privacy guarantees for both bounded and unbounded neighboring relations, as formalized in Proposition 9. The proof is in Appendix A.8.

Proposition 9 (GLMs with Gaussian prior)**.** *Suppose there exists a finite constant $L > 0$ such that for every $(\mathbf{x}_i, y_i)$ and every $\boldsymbol{\beta}$, $\|\nabla_{\boldsymbol{\beta}} \log p(y_i|\mathbf{x}_i, \boldsymbol{\beta})\| \leq L$, and impose a Gaussian prior on $\boldsymbol{\beta}$, i.e. $\pi(\boldsymbol{\beta}) = \mathcal{N}(0, \tau^2 \mathbf{I}_d)$. Then:*

(1) Bounded neighboring $(\mathbf{Y}, \mathbf{Y}')$: $c_1(\boldsymbol{\beta})$ is sub-Gaussian under $\pi(\boldsymbol{\beta}|\mathbf{Y})$ with variance proxy $\sigma^2_{\mathbf{Y},\mathbf{Y}'} \leq 4L^2\tau^2$ for every ordered neighboring pair and, per Corollary 8,

$$\max\{D_\alpha(P\|Q), D_\alpha(Q\|P)\} \leq \varepsilon_{\mathrm{GLM,bounded}}(\alpha) \triangleq \alpha\bar{\sigma}^2/2 \leq 2\alpha L^2\tau^2. \tag{16}$$

(2) Unbounded neighboring $(\mathbf{Y}, \mathbf{Y}')$: $c_1(\boldsymbol{\beta})$ is sub-Gaussian under $\pi(\boldsymbol{\beta}|\mathbf{Y})$ with variance proxy $\sigma^2_{\mathbf{Y},\mathbf{Y}'} \leq L^2\tau^2$ for every ordered neighboring pair and, per Corollary 8,

$$\max\{D_\alpha(P\|Q), D_\alpha(Q\|P)\} \leq \varepsilon_{\mathrm{GLM,unbounded}}(\alpha) \triangleq \alpha\bar{\sigma}^2/2 \leq \alpha L^2\tau^2/2. \tag{17}$$

Eqs. (16) *and* (17) *taken together, releasing OPS satisfies $(\alpha, 2\alpha L^2\tau^2)$-RDP.*

For GLMs with Gaussian priors, the bounded neighboring relation yields a larger worst-case Rényi divergence, specifically incurring a 4-fold increase relative to the unbounded case, similar to the observations in Example 7. Under bounded neighboring, the bound in Eq. (16) matches the strongly log-concave Lipschitz-perturbation analysis of Gopi et al. [2022]. However, our formulation explicitly treats both bounded and unbounded neighboring, yielding Eq. (17), and places the GLM results within a broader framework.

Proposition 9 shows how score bound L, prior variance τ^2, and neighboring relations jointly determine the inherent DP guarantee of OPS from GLMs with Gaussian prior. In particular, $\varepsilon_{\mathrm{GLM}}$ in Eqs. (16) and (17) scale quadratically with L, implying that records with large score norm can substantially weaken the intrinsic DP guarantee offered by OPS, consistent with the common practice of data clipping and normalization in many DP procedures. In addition, $\varepsilon_{\mathrm{GLM}}$ grows linearly with prior variance τ^2. A larger τ^2 corresponds to a less informative prior, and the posterior is more influenced by private data through the likelihood formulation, weakening the inherent DP guarantees of OPS.

$\varepsilon_{\mathrm{GLM}}$ in Eqs. (16) and (17) does not explicitly depend on the data size n because our proof controls the variance proxy $\sigma^2_{\mathbf{Y},\mathbf{Y}'}$ in Eq. (14) via a uniform score bound L and the curvature posterior induced by the Gaussian prior. Under the stated assumptions, the likelihood contributes nonnegative curvature but need not provide a strictly positive lower bound uniformly over all datasets and $\boldsymbol{\beta}$. The prior therefore supplies the uniform curvature lower bound $1/\tau^2$ used in our analysis. Deriving a sharper explicit n-dependent guarantee is possible but would require showing that the likelihood contributes a nontrivial curvature, such as by establishing an n-dependent lower bound on the Hessian of the negative log-posterior $H(\boldsymbol{\beta}) = \sum_{i=1}^n \psi''(\mathbf{x}_i^\top \boldsymbol{\beta})\mathbf{x}_i\mathbf{x}_i^\top + \tau^{-2} I_d$. However, the likelihood curvature $\psi''(\mathbf{x}_i^\top \boldsymbol{\beta})$ can be arbitrarily close to 0 in the worst case. For example, in logistic regression, $\psi''(\eta) = \sigma(\eta)(1 - \sigma(\eta))$ with $\sigma(t) = (1 + e^{-t})^{-1}$ vanishes as $|\mathbf{x}_i^\top \boldsymbol{\beta}| \to \infty$. The likelihood thus provides no guaranteed curvature in general, leaving the Gaussian prior as the only universal source of strong convexity $H(\boldsymbol{\beta}) \succeq \tau^{-2} I_d$. For instance, if $\sum_{i=1}^n \psi''(\mathbf{x}_i^\top \boldsymbol{\beta})\mathbf{x}_i\mathbf{x}_i^\top \succeq k(n)\mathbf{I}_d$ uniformly over all admissible datasets and $\boldsymbol{\beta}$ for some $k(n)$ increasing in n, then τ^{-2} can be replaced by $\tau^{-2} + k(n)$, yielding an n-dependent guarantee for a more restricted model class. We leave developing sharper, n-dependent DP guarantees under additional curvature or design assumptions for future investigation.

Example 10 (Logistic regression with Gaussian prior)**.** *Logistic regression is a special case of the GLM above with binary response $y_i \in \{0, 1\}$, log-partition function $\psi(t) = \log(1 + e^t)$ and $h(\cdot) = 0$. Suppose $\|\mathbf{x}_i\| \leq R < \infty$ for $\forall\, i$, then the condition in Proposition 9 holds with $L = R$.*

(1) Bounded neighboring $(\mathbf{Y}, \mathbf{Y}')$: WLOG, assume $(\mathbf{Y}, \mathbf{Y}')$ differ in the n-th observation.

$$\begin{aligned}\max\{D_\alpha(P\|Q), D_\alpha(Q\|P)\} &\leq \varepsilon_{\mathrm{logistic,bounded}}(\alpha) \\ &= \begin{cases} \alpha R^2\tau^2/2 & \text{if } \mathbf{x}_n = \mathbf{x}'_n \text{(and thus } y_n \neq y'_n\text{)} \\ 2\alpha R^2\tau^2 & \text{if } \mathbf{x}_n \neq \mathbf{x}'_n (y_n \neq y'_n \text{ or } y_n = y'_n) \end{cases}. \end{aligned} \tag{18}$$

(2) Unbounded neighboring $(\mathbf{Y}, \mathbf{Y}')$: Letting $L = R$ in Eq. (17) *yields*

$$\max\{D_\alpha(P\|Q), D_\alpha(Q\|P)\} \leq \varepsilon_{\mathrm{logistic,unbounded}}(\alpha) = \alpha R^2\tau^2/2. \tag{19}$$

Example 10 suggests that replacing a binary label y_n incurs the same worst-case privacy loss as removing an entire record $(\mathbf{x}_n, y_n)$, while replacing the covariate vector $\mathbf{x}_n$ results in a variance proxy 4-fold that of a record removal. It is worth noting that Theorem 16 in Geumlek et al. [2017] examines the same Bayesian logistic regression setting. However, its $(\alpha, 2\alpha R^2\tau^2)$-RDP guarantee for OPS is restricted to the bounded relation, whereas our results cover all neighboring scenarios; when $\mathbf{x}_n = \mathbf{x}'_n$, our bound $\alpha R^2\tau^2/2$ is also tighter by 75%, because label replacement cancels the log-partition terms so that $c_1(\boldsymbol{\beta})$ is R-Lipschitz rather than $2R$-Lipschitz.

5.2 Gaussian likelihood with Gaussian prior & linear regression with Gaussian prior

We analyze two foundational Bayesian models not previously studied for the inherent DP of OPS: (1) univariate Gaussian likelihood with a Gaussian prior for its mean and known variance (Proposition 11); (2) linear regression with a multivariate Gaussian prior for its coefficients and known variance (Proposition 12). Notably, both models exhibit fundamentally different privacy profiles depending on neighboring relation: c_1 is sub-Gaussian under bounded neighboring, but transitions to a non-sub-Gaussian regime under unbounded neighboring; moreover, the two directions of the Rényi divergence also differ in their range of finiteness.

Proposition 11 (Gaussian-Gaussian conjugate pair with known variance)**.** *Let $\mathbf{Y} = \{y_i\}_{i=1}^n$ be the dataset, where $y_i|\theta \overset{i.i.d}{\sim} \mathcal{N}(\theta, \sigma^2)$ with known variance σ^2. With prior $\pi(\theta) = \mathcal{N}(\mu_0, \tau^2)$, the posterior is $\theta|Y \sim \mathcal{N}(\mu_Y, v_Y^2)$ with $v_Y^2 = \sigma^2/\omega, \mu_Y = \mu_0 + \sum_{i=1}^n (y_i - \mu_0)/\omega$, and $\omega = n + \sigma^2/\tau^2$.*

(1) Bounded neighboring $(\mathbf{Y}, \mathbf{Y}')$: WLOG, assume $\mathbf{Y}, \mathbf{Y}'$ differ in their n-th record y_n and y'_n. Then $c_1(\theta)$ is Gaussian with variance $\sigma^2_{Y,Y'} = (y_n - y'_n)^2/(\sigma^2\omega)$. If all records lie in an interval of length $2k\sigma$ for some $k > 0$, then $\sigma^2_{Y,Y'} \le \bar{\sigma}^2 = 4k^2/\omega$, and per Corollary 8

$$\max\{D_\alpha(P\|Q), D_\alpha(Q\|P)\} \le \varepsilon_{\mathrm{GG,bounded}}(\alpha, \omega) \triangleq \alpha\bar{\sigma}^2/2 = 2k^2\alpha/\omega. \tag{20}$$

(2) Unbounded neighboring $(\mathbf{Y}, \mathbf{Y}')$: WLOG, assume $\mathbf{Y}'$ is obtained by removing the n-th record y_n in $\mathbf{Y}$. Then $c_1(\theta)$ is not sub-Gaussian; nevertheless, each term in the exact expression of the Rényi divergence $D_\alpha(P\|Q)$ and $D_\alpha(Q\|P)$ can be analytically derived using the moment-generating functions (MGFs) of the non-central χ^2 distribution with one degree of freedom. Specifically,

$$D_\alpha(P\|Q) = A(\alpha, \omega) + \frac{(\mu_Y - y_n)^2}{2\sigma^2} B(\alpha, \omega) \tag{21}$$

$$\text{where } A(\alpha,\omega) = -\frac{1}{2}\log\left(1-\omega^{-1}\right) - \frac{\log\left(1+\frac{\alpha-1}{\omega}\right)}{2(\alpha-1)}, \quad B(\alpha,\omega) = \left[\frac{1}{1-\frac{1}{\omega}} - \frac{1}{1+\frac{\alpha-1}{\omega}}\right],$$

$$D_\alpha(Q\|P) = \begin{cases} \widetilde{A}(\alpha,\omega) + \frac{(\mu_Y - y_n)^2}{2\sigma^2}\widetilde{B}(\alpha,\omega) & \text{if } 1 < \alpha < \omega, \\ \infty & \text{if } \alpha \ge \omega, \end{cases} \tag{22}$$

$$\text{where } \widetilde{A}(\alpha,\omega) = -\frac{1}{2}\log\frac{\omega}{\omega-1} - \frac{1}{2(\alpha-1)}\log\frac{\omega-\alpha}{\omega-1}, \quad \widetilde{B}(\alpha,\omega) = \frac{\alpha\omega}{(\omega-1)(\omega-\alpha)}.$$

If for all i, $|y_i - \mu_Y| \le k^*\sigma$ for some constant $0 < k^* < \infty$, then for $1 < \alpha < \omega$,

$$\begin{aligned}\max\{D_\alpha(P\|Q), D_\alpha(Q\|P)\} &\le \varepsilon_{\mathrm{GG,unbounded}}(\alpha,\omega) \\ &= \max\left\{A(\alpha,\omega) + \frac{k^{*2}}{2}B(\alpha,\omega),\ \widetilde{A}(\alpha,\omega) + \frac{k^{*2}}{2}\widetilde{B}(\alpha,\omega)\right\}.\end{aligned} \tag{23}$$

The proof is in Appendix A.9. To our knowledge, Proposition 11 gives the first quantitative characterization of how sample size n, prior variance τ^2, and neighboring relation jointly determine the inherent DP guarantees of OPS in the Gaussian-Gaussian model. There are several key takeaways:

- Prior regularization. In both neighboring relations, the influence of the prior is characterized by the precision term $\omega = n + \sigma^2/\tau^2$, where σ^2/τ^2 can be interpreted as the "effective sample size" contributed by the prior. A more informative prior (smaller τ^2) increases ω, leading to smaller values for both $\varepsilon_{\text{GG, bounded}}$ and $\varepsilon_{\text{GG, unbounded}}$, thereby tightening the privacy loss bounds.
- Bounded neighboring. The worst-case privacy loss of OPS in Eq. (20) is controlled by the maximum difference between possible replacements. For fixed α, k and σ^2, $\varepsilon_{\text{GG, bounded}}(\alpha,\omega) =$

$2k^2\alpha/\omega \to 0$ as $n \to \infty$ or $\tau^2 \to 0$. This indicates that OPS can achieve arbitrarily strong privacy through increasing the data size or decreasing the prior variance.

- Unbounded neighboring. $D_\alpha(P\|Q)$ in Eq. (21) decomposes into a structural scale-shift term $A(\alpha,\omega)$ and a data-dependent location-shift term $(\mu_Y - y_n)^2 B(\alpha,\omega)/(2\sigma^2)$. The lower bound of $D_\alpha(P\|Q)$ is attained when the removed record y_n coincides with the posterior mean μ_Y. In this case, while the removal does not shift the posterior mean, it reduces the posterior precision. This change in the "sharpness" of the distribution is a detectable signal of the record's presence. Thus, $D_\alpha(P\|Q)$ reduces to a strictly positive baseline $A(\alpha,\omega)$ – the minimum unavoidable privacy loss incurred by removal; likewise, $D_\alpha(Q\|P)$ reduces to $\widetilde{A}(\alpha,\omega) > 0$ for addition (Claim A.9).

  The two-sided upper bound is given in Eq. (23), where k^* represents the maximum deviation of a potentially removed record from μ_Y, scaled by σ. As $n \to \infty$, $\omega \to \infty$; A and $\widetilde{A}$ are $\mathcal{O}(\omega^{-2})$, B and $\widetilde{B}$ are $\mathcal{O}(\omega^{-1})$, so $\varepsilon_{\text{GG, unbounded}}(\alpha,\omega)$ vanishes and the admissible range $1 < \alpha < \omega$ expands. However, the two directions behave differently: $D_\alpha(P\|Q)$ (removal) is finite for all $\alpha > 1$, whereas $D_\alpha(Q\|P)$ (addition) diverges for $\alpha \geq \omega$, since removing a record inflates the posterior variance so that $\pi(\theta|\mathbf{Y}')$ has a heavier tail than $\pi(\theta|\mathbf{Y})$. Consequently, OPS does not satisfy pure DP under unbounded neighboring, and the neighboring relation affects not only the magnitude of the privacy loss but also the range of Rényi orders over which a finite guarantee exists.

Proposition 12 (Linear regression with Gaussian prior)**.** *Let each data record be $(\mathbf{x}_i, y_i)$, where $\mathbf{x}_i \in \mathbb{R}^d$ is the covariate vector, $y_i \in \mathbb{R}$ is the response, and $y_i|\mathbf{x}_i, \boldsymbol{\beta} \sim \mathcal{N}(\mathbf{x}_i^\top \boldsymbol{\beta}, \sigma^2)$ with known variance σ^2. Let $\mathbf{Y} = \{(\mathbf{x}_i, y_i)\}_{i=1}^n$ be the dataset, $\mathbf{x} \in \mathbb{R}^{n\times d}$ be the design matrix, $\mathbf{y} \in \mathbb{R}^n$ be the response vector. With Gaussian prior $\pi(\boldsymbol{\beta}) = \mathcal{N}(0, \tau^2\mathbf{I}_d)$, the posterior $\boldsymbol{\beta}|\mathbf{Y} \sim \mathcal{N}(\boldsymbol{\mu}_\mathbf{Y}, \Sigma_\mathbf{Y})$, where $\Sigma_\mathbf{Y}^{-1} = \sigma^{-2}\mathbf{x}^\top\mathbf{x} + \tau^{-2}\mathbf{I}_d$ and $\boldsymbol{\mu}_\mathbf{Y} = \Sigma_\mathbf{Y}\left(\sigma^{-2}\mathbf{x}^\top\mathbf{y}\right)$. Assume $\|\mathbf{x}_i\| \leq R < \infty$ for all i.*

(1) Bounded neighboring $(\mathbf{Y}, \mathbf{Y}')$ with $\mathbf{x}_n = \mathbf{x}_n'$: WLOG, suppose the n-th record $(\mathbf{x}_n, y_n)$ in $\mathbf{Y}$ is replaced by $(\mathbf{x}_n, y_n')$ in $\mathbf{Y}'$, where $y_n \neq y_n'$. Then $c_1(\boldsymbol{\beta})$ is Gaussian under $\pi(\boldsymbol{\beta}|\mathbf{Y})$ with variance $\sigma^2_{\mathbf{Y},\mathbf{Y}'} = (y_n - y_n')^2\mathbf{x}_n^\top\Sigma_\mathbf{Y}\mathbf{x}_n/\sigma^4$. If $|y_n - y_n'| \leq 2k\sigma$ for some $k > 0$, then $\sigma^2_{\mathbf{Y},\mathbf{Y}'} \leq \bar{\sigma}^2 = 4k^2\tau^2R^2/\sigma^2$, and per Corollary 8

$$\max\{D_\alpha(P\|Q), D_\alpha(Q\|P)\} \leq \varepsilon_{\text{LR,bounded}}(\alpha) \triangleq \alpha\bar{\sigma}^2/2 = 2\alpha k^2\tau^2R^2/\sigma^2. \tag{24}$$

(2) Unbounded neighboring $(\mathbf{Y}, \mathbf{Y}')$: WLOG, suppose $\mathbf{Y}'$ is obtained by removing the n-th record $(\mathbf{x}_n, y_n)$ in $\mathbf{Y}$. $c_1(\boldsymbol{\beta})$ is not sub-Gaussian under $\pi(\boldsymbol{\beta}|\mathbf{Y})$. If $x_n = 0$, then $P = Q$ and both divergences vanish. Otherwise, define $\omega_n = \sigma^2/(\mathbf{x}_n^\top\Sigma_\mathbf{Y}\mathbf{x}_n)$. With $A, B, \widetilde{A}, \widetilde{B}$ defined in Proposition 11,

$$D_\alpha(P\|Q) = A(\alpha,\omega_n) + \frac{(\mathbf{x}_n^\top\boldsymbol{\mu}_\mathbf{Y} - y_n)^2}{2\sigma^2}B(\alpha,\omega_n), \tag{25}$$

$$D_\alpha(Q\|P) = \begin{cases} \widetilde{A}(\alpha,\omega_n) + \frac{(\mathbf{x}_n^\top\boldsymbol{\mu}_\mathbf{Y} - y_n)^2}{2\sigma^2}\widetilde{B}(\alpha,\omega_n), & \text{if } 1 < \alpha < \omega_n, \\ \infty, & \text{if } \alpha \geq \omega_n. \end{cases} \tag{26}$$

Since $\omega_n \geq \omega_{\min} \triangleq 1 + \sigma^2/(\tau^2R^2)$ for every admissible pair, if for all i, $|\mathbf{x}_i^\top\boldsymbol{\mu}_\mathbf{Y} - y_i| \leq k^*\sigma$ for some $k^* > 0$, then for $1 < \alpha < \omega_{\min}$,

$$\max\{D_\alpha(P\|Q), D_\alpha(Q\|P)\} \leq \varepsilon_{\text{LR,unbounded}}(\alpha) = \varepsilon_{\text{GG,unbounded}}(\alpha,\omega_{\min}). \tag{27}$$

Proposition 12 extends the results in Proposition 11 to regression settings with covariates. In the bounded case, the privacy loss is determined by the difference between the replacement records and the posterior precision, and scales linearly with α. In the unbounded case, the data geometry determines the inherent DP guarantees of OPS through $\omega_n = \sigma^2/(\mathbf{x}_n^\top\Sigma_\mathbf{Y}\mathbf{x}_n)$. Specifically, $A(\alpha,\omega_n)$ is governed by the influence of the removed record on the estimate; high-influential points have low ω_n and may substantially reduce the posterior precision, creating a detectable signal even when the posterior mean is unchanged. As in Proposition 11, the reverse direction is finite only for $\alpha < \omega_n$; thus high-influence points, with small ω_n, both enlarge the privacy loss and shrink the range of Rényi orders admitting a finite guarantee.

6 Discussion

This work advances our understanding of the inherent privacy guarantees of Bayesian posterior sampling in a unified Rényi-divergence framework. Our results demonstrate that the intrinsic stochasticity of OPS provides meaningful privacy protection – shaped by the interplay among prior information, data size, model structure, and neighboring relationships – even without explicit noise injection. By offering explicit and interpretable privacy quantification across diverse Bayesian models, our analysis clarifies when and how OPS can serve as a reliable DP mechanism. Our results offer a principled basis for reducing external noise injection and thereby improving the utility of released statistical outputs while maintaining formal privacy protection. The work also offers insights on designing more transparent and effective private data analysis methods in sensitive applications.

There are several interesting directions for future work. For example, the sub-Gaussian regime examined in Section 5 can be further relaxed to sub-exponential log-likelihood ratios, thereby extending the framework to settings where both uniform boundedness and sub-Gaussianity fail.

Corollary 13 (RDP guarantees of OPS with sub-exponential log-likelihood ratio)**.** *Suppose that there exist* $\nu^2_{\mathbf{Y},\mathbf{Y}'} \in (0,\infty)$ *and* $b_{\mathbf{Y},\mathbf{Y}'} \in (0,\infty)$ *such that for* $\forall\ t \in \mathbb{R}$ *satisfying* $|t| < b^{-1}_{\mathbf{Y},\mathbf{Y}'}$*, it holds that* $\log \mathbb{E}_{\boldsymbol{\theta}\sim P}\big(e^{t(c_1(\boldsymbol{\theta})-\mu_{\mathbf{Y},\mathbf{Y}'})}\big) \le \frac{t^2\nu^2_{\mathbf{Y},\mathbf{Y}'}}{2(1-b_{\mathbf{Y},\mathbf{Y}'}|t|)}$ *where* $\mu_{\mathbf{Y},\mathbf{Y}'} = \mathbb{E}_{\boldsymbol{\theta}\sim P}\left[c_1(\boldsymbol{\theta})\right]$*. Define*

$$\varepsilon_{\mathrm{subExp}}(\alpha, b_{\mathbf{Y},\mathbf{Y}'}, \nu^2_{\mathbf{Y},\mathbf{Y}'}) \triangleq \frac{\nu^2_{\mathbf{Y},\mathbf{Y}'}}{2} \cdot \frac{\alpha(1-2b_{\mathbf{Y},\mathbf{Y}'}) + 2b_{\mathbf{Y},\mathbf{Y}'}}{(1-b_{\mathbf{Y},\mathbf{Y}'})\big(1-b_{\mathbf{Y},\mathbf{Y}'}(\alpha-1)\big)}. \tag{28}$$

When $b_{\mathbf{Y},\mathbf{Y}'} < \min\{(\alpha-1)^{-1}, 1\}$ *for all ordered neighboring pairs,*

$$\max\{D_\alpha(P||Q), D_\alpha(Q||P)\} \le \sup_{(\mathbf{Y},\mathbf{Y}'):d(\mathbf{Y},\mathbf{Y}')=1} \varepsilon_{\mathrm{subExp}}(\alpha, b_{\mathbf{Y},\mathbf{Y}'}, \nu^2_{\mathbf{Y},\mathbf{Y}'}), \tag{29}$$

and releasing OPS satisfies $(\alpha, \sup_{(\mathbf{Y},\mathbf{Y}')} \varepsilon_{\mathrm{subExp}}(\alpha, b_{\mathbf{Y},\mathbf{Y}'}, \nu^2_{\mathbf{Y},\mathbf{Y}'}))$*-RDP.*

The proof is in Appendix A.6. While $\varepsilon_{\mathrm{subExp}}(\alpha, b_{\mathbf{Y},\mathbf{Y}'}, \nu^2_{\mathbf{Y},\mathbf{Y}'})$ is tight for fixed sub-exponential parameters $(\nu^2_{\mathbf{Y},\mathbf{Y}'}, b_{\mathbf{Y},\mathbf{Y}'})$, defining these parameters for a specific model often introduces looseness. For example, in the unbounded case of the Gaussian-Gaussian pair (Proposition 11), the log-likelihood ratio satisfies the sub-exponential condition as shown in Appendix A.9, but the derived parameters rely on upper bounds and are thus sub-optimal. In contrast, directly exploiting the exact MGFs of a non-central χ^2 distribution and calculating $D_\alpha(P\|Q)$ via Theorem 3 yields sharper bounds. This illustrates the potential gap between general tail-based bounds and model-specific exact characterizations. Despite this, Corollary 13 remains a broadly applicable theoretical tool for analyzing inherent DP guarantees of OPS when neither uniform boundedness or sub-Gaussianity hold. Proposition 14 demonstrates this for linear regression under bounded-neighboring with covariate replacement, a case excluded from Proposition 12 because $c_1(\boldsymbol{\beta})$ contains a non-sub-Gaussian quadratic term.

Proposition 14 (Linear regression with Gaussian prior, bounded neighboring with $\mathbf{x}_n \neq \mathbf{x}'_n$)**.** *When* $\mathbf{x}_n \neq \mathbf{x}'_n$*,* $c_1(\boldsymbol{\beta})$ *contains both affine and quadratic terms in* $\boldsymbol{\beta}$ *and is generally not sub-Gaussian under* $\pi(\boldsymbol{\beta}|\mathbf{Y})$*. Suppose* $\|\mathbf{x}_i\| \le R < \infty$ *and* $|y_i| \le M_y < \infty$ *for* $\forall\ i$*, and assume the scale restriction* $\sigma^2 > 4\tau^2R^2$*, then* $c_1(\boldsymbol{\beta})$ *is sub-exponential with parameters*

$$b_{\mathbf{Y},\mathbf{Y}'} = \frac{4\tau^2R^2}{\sigma^2}, \quad \nu^2_{\mathbf{Y},\mathbf{Y}'} \le \frac{8M_y^2R^2\tau^2}{\sigma^4} + \frac{16\tau^4R^4}{\sigma^4} + \frac{8n^2\tau^6M_y^2R^6}{\sigma^8}, \tag{30}$$

which can be plugged into Corollary 13 for explicit RDP bounds.

The proof is in Appendix A.10.1. The parameters in Eq. (30) also illustrate the conservativeness that can arise when deriving model-specific sub-exponential bounds in Corollary 13. The Cauchy-Schwarz decomposition, and operator norm bounds can overestimate the scale parameter, narrowing the range of Rényi order α for which the analysis establishes a finite bound. In addition, bounding the posterior mean using only row-wise bounds introduces an n^2 term in the variance proxy, potentially inflating the privacy bound.

Developing tighter techniques for estimating $(\nu^2_{\mathbf{Y},\mathbf{Y}'}, b_{\mathbf{Y},\mathbf{Y}'})$, especially for complex models where exact MGFs are analytically intractable, is a promising direction for future work. Other future directions include leveraging OPS as a DP mechanism in applications such as privacy-preserving synthetic data generation and variational inference, with the goal of optimizing the trade-off between privacy guarantees and statistical utility.

Acknowledgments and Disclosure of Funding

We thank the anonymous reviewers for their helpful comments. The authors did not receive external funding for this work and have no competing interests to declare.

References

Cynthia Dwork, Krishnaram Kenthapadi, Frank McSherry, Ilya Mironov, and Moni Naor. Our data, ourselves: Privacy via distributed noise generation. In *Annual international conference on the theory and applications of cryptographic techniques*, pages 486–503. Springer, 2006a.

Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. In *Theory of cryptography conference*, pages 265–284. Springer, 2006b.

Ilya Mironov. Rényi differential privacy. In *2017 IEEE 30th Computer Security Foundations Symposium (CSF)*, pages 263–275. IEEE, 2017.

Mark Bun and Thomas Steinke. Concentrated differential privacy: Simplifications, extensions, and lower bounds. In *Theory of Cryptography Conference*, pages 635–658, 2016.

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

Kamalika Chaudhuri and Nina Mishra. When random sampling preserves privacy. In *Annual International Cryptology Conference*, pages 198–213. Springer, 2006.

Ninghui Li, Wahbeh Qardaji, and Dong Su. On sampling, anonymization, and differential privacy or, k-anonymization meets differential privacy. In *Proceedings of the 7th ACM Symposium on Information, Computer and Communications Security*, pages 32–33, 2012.

Latanya Sweeney. k-anonymity: A model for protecting privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 10(05):557–570, 2002.

Jairo Giraldo, Alvaro Cardenas, and Murat Kantarcioglu. Security and privacy trade-offs in cps by leveraging inherent differential privacy. In *2017 IEEE Conference on Control Technology and Applications (CCTA)*, pages 1313–1318. IEEE, 2017.

Zinan Lin, Vyas Sekar, and Giulia Fanti. On the privacy properties of gan-generated samples. In *International Conference on Artificial Intelligence and Statistics*, pages 1522–1530. PMLR, 2021.

Borja Balle, Gilles Barthe, and Marco Gaboardi. Privacy amplification by subsampling: Tight analyses via couplings and divergences. *Advances in neural information processing systems*, 31, 2018.

Thomas Steinke. Composition of differential privacy & privacy amplification by subsampling. *arXiv preprint arXiv:2210.00597*, 2022.

Xingyuan Zhao, Ruyu Zhou, and Fang Liu. Enhancing trade-offs in privacy, utility, and computational efficiency through multistage sampling technique (must). *IEEE Transactions on Privacy*, 2024.

Darakhshan J Mir. *Differential privacy: an exploration of the privacy-utility landscape*. Rutgers The State University of New Jersey, School of Graduate Studies, 2013.

Christos Dimitrakakis, Blaine Nelson, Aikaterini Mitrokotsa, and Benjamin IP Rubinstein. Robust and private bayesian inference. In *International Conference on Algorithmic Learning Theory*, pages 291–305. Springer, 2014.

Christos Dimitrakakis, Blaine Nelson, Zuhe Zhang, Aikaterini Mitrokotsa, and Benjamin IP Rubinstein. Differential privacy for bayesian inference through posterior sampling. *Journal of machine learning research*, 18(11):1–39, 2017.

Yu-Xiang Wang, Stephen Fienberg, and Alex Smola. Privacy for free: Posterior sampling and stochastic gradient monte carlo. In *International Conference on Machine Learning*, pages 2493–2502, 2015.

Xingyuan Zhao and Fang Liu. A new bound for privacy loss from bayesian posterior sampling. In *Proceedings of the Twelfth ACM Conference on Data and Application Security and Privacy*, pages 346–348, 2022.

Zuhe Zhang, Benjamin IP Rubinstein, and Christos Dimitrakakis. On the differential privacy of bayesian inference. In *Thirtieth AAAI Conference on Artificial Intelligence*, 2016.

James Foulds, Joseph Geumlek, Max Welling, and Kamalika Chaudhuri. On the theory and practice of privacy-preserving bayesian data analysis. pages 227–236, 2016.

Joseph Geumlek, Shuang Song, and Kamalika Chaudhuri. Renyi differential privacy mechanisms for posterior sampling. In *Advances in Neural Information Processing Systems*, pages 5289–5298, 2017.

Fangyuan Zhao, Xuebin Ren, Shusen Yang, Qing Han, Peng Zhao, and Xinyu Yang. Latent dirichlet allocation model training with differential privacy. *IEEE Transactions on Information Forensics and Security*, 16:1290–1305, 2020.

David M Blei, Andrew Y Ng, and Michael I Jordan. Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022, 2003.

Bai Li, Changyou Chen, Hao Liu, and Lawrence Carin. On connecting stochastic gradient mcmc and differential privacy. In *The 22nd International Conference on Artificial Intelligence and Statistics*, pages 557–566. PMLR, 2019.

Jack E Jewson, Sahra Ghalebikesabi, and Chris C Holmes. Differentially private statistical inference through $\backslash\beta$-divergence one posterior sampling. *Advances in Neural Information Processing Systems*, 36:76974–77001, 2023.

Kentaro Minami, Hiromi Arai, Issei Sato, and Hiroshi Nakagawa. Differential privacy without sensitivity. In *NIPS*, pages 956–964, 2016.

Yu-Xiang Wang. Per-instance differential privacy. *Journal of Privacy and Confidentiality*, 9(1), 2019.

Shenggang Hu, Louis Aslett, Hongsheng Dai, Murray Pollock, and Gareth O Roberts. Privacy guarantees in posterior sampling under contamination. *arXiv preprint arXiv:2403.07772*, 2024.

Daniel Kifer and Ashwin Machanavajjhala. No free lunch in data privacy. In *Proceedings of the 2011 ACM SIGMOD International Conference on Management of Data*, SIGMOD '11, page 193–204, New York, NY, USA, 2011. Association for Computing Machinery. ISBN 9781450306614. doi: 10.1145/1989323.1989345. URL `https://doi.org/10.1145/1989323.1989345`.

David Sommer, Sebastian Meiser, and Esfandiar Mohammadi. Privacy loss classes: The central limit theorem in differential privacy. *Cryptology ePrint Archive*, 2018.

Alfréd Rényi. On measures of entropy and information. In *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability, volume 1: contributions to the theory of statistics*, volume 4, pages 547–562. University of California Press, 1961.

Sivakanth Gopi, Yin Tat Lee, and Daogao Liu. Private convex optimization via exponential mechanism. In *Conference on Learning Theory*, pages 1948–1989. PMLR, 2022.

Moshe Shaked and J George Shanthikumar. *Stochastic orders*. Springer, 2007.

Michel Ledoux. *The concentration of measure phenomenon*. Number 89. American Mathematical Soc., 2001.

Appendix Contents

A.1 Proof of Lemma 2 **13**

A.2 Proof of Theorem 3 **14**

A.3 Proof of Corollary 5 **14**

A.4 Proof of Corollary 6 **16**

A.5 Proof of Corollary 8 **17**

A.6 Proof of Corollary 13 **18**

A.7 Proof of Example 7 for Categorical data **18**
A.7.1 bounded neighboring 19
A.7.2 unbounded neighboring 19
A.7.3 Proof of Bernoulli data 20

A.8 Proof of Proposition 9 for Generalized Linear Models **23**
A.8.1 bounded neighboring 23
A.8.2 unbounded neighboring 24
A.8.3 Proof of Example 10 for Logistic Regression 24

A.9 Proof of Proposition 11 for Gaussian-Gaussian conjugate pair **26**
A.9.1 bounded neighboring 26
A.9.2 unbounded neighboring 27

A.10 Proof of Propositions 12 and 14 for Linear Regression **32**
A.10.1 bounded neighboring 32
A.10.2 unbounded neighboring 36

A.1 Proof of Lemma 2

Proof. The posterior of parameter $\boldsymbol{\theta}$ given data $\mathbf{Y}$ is

$$\pi(\boldsymbol{\theta}|\mathbf{Y}) = \frac{\pi(\boldsymbol{\theta})\prod_{i=1}^{n} p(\mathbf{y}_i|\boldsymbol{\theta})}{\int \prod_{i=1}^{n} p(\mathbf{y}_i|\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}} \triangleq \frac{\pi(\boldsymbol{\theta})\ell(\boldsymbol{\theta};\mathbf{Y})}{\int \ell(\boldsymbol{\theta};\mathbf{Y})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}} = \frac{\pi(\boldsymbol{\theta})\ell(\boldsymbol{\theta};\mathbf{Y})}{p(\mathbf{Y})}. \tag{A.1}$$

For any pair of neighboring datasets $\mathbf{Y}$ and $\mathbf{Y}'$ where $d(\mathbf{Y},\mathbf{Y}')=1$, the privacy loss for releasing posterior sample $c(\boldsymbol{\theta})$ per Eq (2) can be decomposed as:

$$c(\boldsymbol{\theta}) = \log\left(\frac{\pi(\boldsymbol{\theta}|\mathbf{Y})}{\pi(\boldsymbol{\theta}|\mathbf{Y}')}\right) = \underbrace{\log\left(\frac{\ell(\boldsymbol{\theta};\mathbf{Y})}{\ell(\boldsymbol{\theta};\mathbf{Y}')}\right)}_{\triangleq\ c_1(\boldsymbol{\theta})} + \underbrace{\log\left(\frac{p(\mathbf{Y}')}{p(\mathbf{Y})}\right)}_{\triangleq\ c_2}. \tag{A.2}$$

Note that $p(\mathbf{Y}')$

$$= \int \ell(\boldsymbol{\theta};\mathbf{Y}')\pi(\boldsymbol{\theta})d\boldsymbol{\theta} = \int \frac{\ell(\boldsymbol{\theta};\mathbf{Y}')}{\ell(\boldsymbol{\theta};\mathbf{Y})}\ell(\boldsymbol{\theta};\mathbf{Y})\pi(\boldsymbol{\theta})d\boldsymbol{\theta} = \int \frac{\ell(\boldsymbol{\theta};\mathbf{Y}')}{\ell(\boldsymbol{\theta};\mathbf{Y})}\pi(\boldsymbol{\theta}|\mathbf{Y})p(\mathbf{Y})d\boldsymbol{\theta} \quad \text{per Eq. (A.1)}$$

$$= p(\mathbf{Y})\int \frac{\ell(\boldsymbol{\theta};\mathbf{Y}')}{\ell(\boldsymbol{\theta};\mathbf{Y})}\pi(\boldsymbol{\theta}|\mathbf{Y})d\boldsymbol{\theta}$$

$$\Rightarrow \frac{p(\mathbf{Y}')}{p(\mathbf{Y})} = \int \frac{\ell(\boldsymbol{\theta};\mathbf{Y}')}{\ell(\boldsymbol{\theta};\mathbf{Y})}\pi(\boldsymbol{\theta}|\mathbf{Y})d\boldsymbol{\theta} = \int e^{-c_1(\boldsymbol{\theta})}\pi(\boldsymbol{\theta}|\mathbf{Y})d\boldsymbol{\theta} = \mathbb{E}_{\boldsymbol{\theta}\sim\pi(\cdot|\mathbf{Y})}\left[e^{-c_1(\boldsymbol{\theta})}\right]. \tag{A.3}$$

Since $c_2 = \log(p(\mathbf{Y}')/p(\mathbf{Y}))$, we have

$$c_2 = \log\left(\frac{p(\mathbf{Y}')}{p(\mathbf{Y})}\right) = \log\left(\mathbb{E}_{\boldsymbol{\theta}\sim\pi(\cdot|\mathbf{Y})}\left[e^{-c_1(\boldsymbol{\theta})}\right]\right). \tag{A.4}$$

□

A.2 Proof of Theorem 3

Proof. Per the definition of Rényi divergence in Eq (4),

$$D_\alpha(P||Q) = \frac{1}{\alpha-1}\log \mathbb{E}_{x\sim Q}\left(\frac{P(x)}{Q(x)}\right)^\alpha = \frac{1}{\alpha-1}\log \mathbb{E}_{x\sim P}\left(\frac{P(x)}{Q(x)}\right)^{\alpha-1}. \tag{A.5}$$

For any neighboring datasets $\mathbf{Y}$ and $\mathbf{Y}'$, we can calculate the Rényi divergence between the corresponding neighboring posterior distributions $\pi(\boldsymbol{\theta}|\mathbf{Y})$ and $\pi(\boldsymbol{\theta}|\mathbf{Y}')$ as below

$$\begin{aligned} D_\alpha(\pi(\boldsymbol{\theta}|\mathbf{Y})\|\pi(\boldsymbol{\theta}|\mathbf{Y}')) &= \frac{1}{\alpha-1}\log \mathbb{E}_{\boldsymbol{\theta}\sim\pi(\cdot|\mathbf{Y})}\left[\left(\frac{\pi(\boldsymbol{\theta}|\mathbf{Y})}{\pi(\boldsymbol{\theta}|\mathbf{Y}')}\right)^{\alpha-1}\right] \\ &= \frac{1}{\alpha-1}\log \mathbb{E}_{\boldsymbol{\theta}\sim\pi(\cdot|\mathbf{Y})}\left[e^{(\alpha-1)c(\boldsymbol{\theta})}\right] \\ &= \frac{1}{\alpha-1}\log \mathbb{E}_{\boldsymbol{\theta}\sim\pi(\cdot|\mathbf{Y})}\left[e^{(\alpha-1)(c_1(\boldsymbol{\theta})+c_2)}\right] \quad \text{per Eq (A.2)} \\ &= c_2 + \frac{1}{\alpha-1}\log \mathbb{E}_{\boldsymbol{\theta}\sim\pi(\cdot|\mathbf{Y})}\left[e^{(\alpha-1)c_1(\boldsymbol{\theta})}\right] \end{aligned} \tag{A.6}$$

Plug Eq. (A.4) into Eq. (A.6),

$$D_\alpha(\pi(\boldsymbol{\theta}|\mathbf{Y})\|\pi(\boldsymbol{\theta}|\mathbf{Y}')) = \log \mathbb{E}_{\boldsymbol{\theta}\sim\pi(\cdot|\mathbf{Y})}\left[e^{-c_1(\boldsymbol{\theta})}\right] + \frac{1}{\alpha-1}\log \mathbb{E}_{\boldsymbol{\theta}\sim\pi(\cdot|\mathbf{Y})}\left[e^{(\alpha-1)c_1(\boldsymbol{\theta})}\right] \tag{A.7}$$

$$= \log\left[\left(\mathbb{E}_{\boldsymbol{\theta}\sim\pi(\cdot|\mathbf{Y})}\left[e^{-c_1(\boldsymbol{\theta})}\right]\right)\left(\mathbb{E}_{\boldsymbol{\theta}\sim\pi(\cdot|\mathbf{Y})}\left[e^{(\alpha-1)\cdot c_1(\boldsymbol{\theta})}\right]\right)^{\frac{1}{\alpha-1}}\right]. \tag{A.8}$$

□

A.3 Proof of Corollary 5

Proof. Assume $c_1(\boldsymbol{\theta})$ is uniformly bounded, i.e. there exist finite constants $a, b \in \mathbb{R}$ with $a < b$ such that $\forall \boldsymbol{\theta} \in \boldsymbol{\Theta}, c_1(\boldsymbol{\theta}) \in [a, b]$. Thus $-c_1(\boldsymbol{\theta}) \in [-b, -a]$ and per monotonicity.

$$c_2 = \log \mathbb{E}_{\boldsymbol{\theta}\sim\pi(\cdot|\mathbf{Y})}\left[e^{-c_1(\boldsymbol{\theta})}\right] \in [-b, -a] \tag{A.9}$$

$$\frac{1}{\alpha-1}\log \mathbb{E}_{\boldsymbol{\theta}\sim\pi(\cdot|\mathbf{Y})}\left[e^{(\alpha-1)c_1(\boldsymbol{\theta})}\right] \in [a, b]. \tag{A.10}$$

Taken together with the decomposition in Eq. (A.6), we obtain

$$D_\alpha(\pi(\boldsymbol{\theta}|\mathbf{Y})||\pi(\boldsymbol{\theta}|\mathbf{Y}')) = c_2 + \frac{1}{\alpha-1}\log \mathbb{E}_{\boldsymbol{\theta}\sim\pi(\cdot|\mathbf{Y})}\left[e^{(\alpha-1)\cdot c_1(\boldsymbol{\theta})}\right] \leq b - a. \tag{A.11}$$

Note that to attain the **equality** in Eq. (A.11) of the upper bound $b - a$ would require simultaneously $c_1(\boldsymbol{\theta}) = a$ (so that $c_2 = -a$) and $c_1(\boldsymbol{\theta}) = b$ (so that $\frac{1}{\alpha-1}\log \mathbb{E}_{\boldsymbol{\theta}\sim\pi(\cdot|\mathbf{Y})}\left[e^{(\alpha-1)c_1(\boldsymbol{\theta})}\right] = b$) almost surely, which is impossible unless $a = b$. Therefore, this bound is not tight and unattainable.

To obtain an attainable upper bound of $D_\alpha(\pi(\boldsymbol{\theta}|\mathbf{Y})||\pi(\boldsymbol{\theta}|\mathbf{Y}'))$, it suffices to maximize the two exponential moments of $c_1(\boldsymbol{\theta})$ involved in Eq. (A.11), subject to the constraint $c_1(\boldsymbol{\theta}) \in [a, b]$. This maximization can be achieved using convex-order arguments, as in [Shaked and Shanthikumar, 2007]. WLOG, let X be any random variable uniformly bounded with $[a, b]$, and X^* denote a binary random variable supported on $\{a, b\}$ as below, with the same mean as X:

$$X^* = \begin{cases} a, & \text{with probability } 1-p \\ b, & \text{with probability } p \end{cases}; \tag{A.12}$$

that is, $\mathbb{E}(X) = \mathbb{E}(X^*) = a(1-p) + bp$ for some $p \in [0, 1]$. For any convex function ϕ and any $x \in [a, b]$, convexity implies the chord bound:

$$\phi(x) \;\leq\; \lambda(x)\phi(b) + (1-\lambda(x))\phi(a) \quad \forall x \in [a, b] \tag{A.13}$$

Let $\lambda(x) = \frac{x-a}{b-a}$ and take expectation of $\phi(x)$ over its distribution, we have:

$$\mathbb{E}[\phi(X)] \leq p\phi(b) + (1-p)\phi(a) = \mathbb{E}[\phi(X^*)]. \tag{A.14}$$

Lemma A.1 (Theorem 3.A.7 in Shaked and Shanthikumar [2007]). *Let X and Y be two random variables with the same finite means. Then $X \leq_{\text{cx}} Y$ if and only if,*

$$\mathbb{E}(\phi(X)) \leq \mathbb{E}(\phi(Y)) \qquad \textit{for any convex function } \phi. \tag{A.15}$$

Hence, per the Lemma above, we have $X \leq_{\text{cx}} X^*$. Since X can be any random variable uniformly bounded with $[a, b]$, X^* is maximal in convex order among all such bounded random variables.

Also, since the exponential function e^x is convex, and the Rényi divergence functional D_α in Eq. (8) is fully determined by the exponential moments of $c_1(\boldsymbol{\theta})$, it follows that D_α is monotone with respect to the convex order. Setting $X = c_1(\boldsymbol{\theta})$, we therefore obtain $c_1(\boldsymbol{\theta}) \leq_{\text{cx}} X^*$, which implies that the $D_\alpha(\pi(\boldsymbol{\theta}|\mathbf{Y})||\pi(\boldsymbol{\theta}|\mathbf{Y}'))$ is maximized when $c_1(\boldsymbol{\theta})$ coincides with the endpoint-supported binary random variable X^* as defined in Eq. (A.12). Then per Eq. (A.7)

$$\begin{aligned}
& D_\alpha(\pi(\boldsymbol{\theta}|\mathbf{Y})||\pi(\boldsymbol{\theta}|\mathbf{Y}')) \\
=\ & \log \mathbb{E}_{\boldsymbol{\theta}\sim\pi(\cdot|\mathbf{Y})}\left[e^{-c_1(\boldsymbol{\theta})}\right] + \frac{1}{\alpha-1}\log \mathbb{E}_{\boldsymbol{\theta}\sim\pi(\cdot|\mathbf{Y})}\left[e^{(\alpha-1)c_1(\boldsymbol{\theta})}\right] \\
\leq\ & \log \mathbb{E}\left[e^{-X^*}\right] + \frac{1}{\alpha-1}\log \mathbb{E}\left[e^{(\alpha-1)X^*}\right] \quad \text{per convex order dominance} \\
=\ & \log\left((1-p)e^{-a} + pe^{-b}\right) + \frac{1}{\alpha-1}\log\Big((1-p)e^{(\alpha-1)a} + pe^{(\alpha-1)b}\Big) \\
=\ & \log\Big(1-p+pe^{-(b-a)}\Big) + \frac{1}{\alpha-1}\log\Big(1-p+pe^{(\alpha-1)(b-a)}\Big) \qquad \text{(A.16)} \\
\triangleq\ & D_\alpha(p) \qquad \text{(A.17)}
\end{aligned}$$

For a fixed pair of neighboring datasets $\mathbf{Y}$ and $\mathbf{Y}'$, in order to align the means of $c_1(\boldsymbol{\theta})$ and X^*, one should require,

$$\begin{aligned}
\mathbb{E}_{\boldsymbol{\theta}\sim\pi(\cdot|\mathbf{Y})}\big[c_1(\boldsymbol{\theta})\big] &= a(1-p) + bp \\
\Rightarrow \hat{p} &= \frac{\mathbb{E}_{\boldsymbol{\theta}\sim\pi(\cdot|\mathbf{Y})}\big[c_1(\boldsymbol{\theta})\big] - a}{b-a},
\end{aligned} \tag{A.18}$$

However, since $\mathbb{E}_{\boldsymbol{\theta}\sim\pi(\cdot|\mathbf{Y})}[c_1(\boldsymbol{\theta})]$ may take any value in $[a, b]$ across neighboring dataset pairs, the associated $\hat{p}$ ranges over $[0, 1]$. Thus, an upper bound on $D_\alpha(\pi(\boldsymbol{\theta}|\mathbf{Y})||\pi(\boldsymbol{\theta}|\mathbf{Y}'))$ that is valid over all neighboring datasets need to be obtained by maximizing $D_\alpha(p)$ in Eq. (A.16) over $p \in [0, 1]$.

Let $\Delta = b-a$ and define $A = e^{-\Delta} \in (0, 1)$, $B = e^{(\alpha-1)\Delta} > 1$. To maximize $D_\alpha(p)$ over $p \in [0, 1]$, take the first derivative on both sides with respect to p:

$$D'_\alpha(p) = \frac{dD_\alpha(p)}{dp} = \frac{A-1}{1-p+pA} + \frac{1}{\alpha-1}\cdot\frac{B-1}{1-p+pB}. \tag{A.19}$$

Setting $D'_\alpha(p) = 0$, we have

$$\begin{aligned}
p^* &= \arg\max_{p\in[0,1]} D_\alpha(p) \\
&= \frac{(B-1)-(\alpha-1)(1-A)}{\alpha(1-A)(B-1)} \\
&= \frac{\big(e^{(\alpha-1)\Delta}-1\big)-(\alpha-1)\big(1-e^{-\Delta}\big)}{\alpha\big(1-e^{-\Delta}\big)\big(e^{(\alpha-1)\Delta}-1\big)}.
\end{aligned} \tag{A.20}$$

Therefore,

$$\begin{aligned}
D_\alpha &\leq \log\left(1-p^*+p^*e^{-\Delta}\right) + \frac{1}{\alpha-1}\log\left(1-p^*+p^*e^{(\alpha-1)\Delta}\right) \\
&= \log\left(1 - \frac{(B-1)-(\alpha-1)(1-A)}{\alpha(1-A)(B-1)} + \frac{(B-1)-(\alpha-1)(1-A)}{\alpha(1-A)(B-1)}A\right) \\
&\quad + \frac{1}{\alpha-1}\log\left(1 - \frac{(B-1)-(\alpha-1)(1-A)}{\alpha(1-A)(B-1)} + \frac{(B-1)-(\alpha-1)(1-A)}{\alpha(1-A)(B-1)}B\right)
\end{aligned}$$

$$= \log\left(\frac{\cancel{\alpha(1-A)}(B-1) - \cancel{(1-A)}(B-1) + (\alpha-1)\cancel{(1-A)} - (\alpha-1)\cancel{(1-A)}A}{\cancel{\alpha(1-A)}(B-1)}\right)$$
$$+ \frac{1}{\alpha-1}\log\left(\frac{\alpha(1-A)\cancel{(B-1)} - \cancel{(B-1)} + B\cancel{(B-1)} - (\alpha-1)(1-A)\cancel{(B-1)}}{\alpha(1-A)\cancel{(B-1)}}\right)$$
$$= \log\left(\frac{(\alpha-1)(B-A)}{\alpha(B-1)}\right) + \frac{1}{\alpha-1}\log\left(\frac{B-A}{\alpha(1-A)}\right) \triangleq \varepsilon_{\mathrm{ub}}(\alpha,\Delta). \tag{A.21}$$

□

Reverse direction. If $c_1(\boldsymbol{\theta}) \in [a,b]$ for all $\boldsymbol{\theta}\in\Theta$, then $\tilde{c}_1(\boldsymbol{\theta}) = -c_1(\boldsymbol{\theta}) \in [-b,-a]$ for all $\boldsymbol{\theta}\in\Theta$, an interval of the same width $\Delta = b-a$. The argument above uses only the support width of the log-likelihood ratio and holds under an arbitrary posterior. Applying it to $\tilde{c}_1(\boldsymbol{\theta})$ under $Q = \pi(\boldsymbol{\theta}|\mathbf{Y}')$ via Eq. (9) yields $D_\alpha(Q\|P) \le \varepsilon_{\mathrm{ub}}(\alpha,\Delta)$. Hence $\max\{D_\alpha(P\|Q), D_\alpha(Q\|P)\} \le \varepsilon_{\mathrm{ub}}(\alpha,\Delta)$, and the $(\alpha, \varepsilon_{\mathrm{ub}}(\alpha,\Delta))$-RDP guarantee holds in both directions.

Claim A.2. *For $\forall\alpha > 1$, and for $\forall\Delta_2 \ge \Delta_1 \ge 0$, we have $0 \le \varepsilon_{\mathrm{ub}}(\alpha,\Delta_2) - \varepsilon_{\mathrm{ub}}(\alpha,\Delta_1) \le \Delta_2 - \Delta_1$.*

Proof. Take the first derivative of Eq. (A.16) with respect to Δ

$$\partial_\Delta D_\alpha(p,\Delta) = -\frac{pe^{-\Delta}}{1-p+pe^{-\Delta}} + \frac{pe^{(\alpha-1)\Delta}}{1-p+pe^{(\alpha-1)\Delta}} = h\left(e^{(\alpha-1)\Delta}\right) - h\left(e^{-\Delta}\right), \tag{A.22}$$

where $h(u) = \frac{pu}{1-p+pu}$ for $u > 0$. Since

$$\frac{\partial h(u)}{\partial u} = \frac{p(1-p+pu) - p^2u}{(1-p+pu)^2} = \frac{p(1-p)}{(1-p+pu)^2} \ge 0, \tag{A.23}$$

$h(u)$ is non-decreasing in u. Per $\alpha > 1$ and $\Delta \ge 0$, $e^{(\alpha-1)\Delta} \ge e^{-\Delta}$, then $\partial_\Delta D_\alpha(p,\Delta) \ge 0$ and the equality holds only in the degenerate cases $p = 0, p = 1$, or $\Delta = 0$.

Since the pointwise maximum preserves order, for any $\Delta_2 \ge \Delta_1$ and $\forall p \in [0,1], \forall\alpha > 1$:

$$0 \le D_\alpha(p,\Delta_2) - D_\alpha(p,\Delta_1) \le \Delta_2 - \Delta_1$$
$$\text{Taking the maximum over } p \text{ on both sides}$$
$$\Rightarrow 0 \le \varepsilon_{\mathrm{ub}}(\alpha,\Delta_2) - \varepsilon_{\mathrm{ub}}(\alpha,\Delta_1) \le \Delta_2 - \Delta_1. \tag{A.24}$$

We can conclude that $\varepsilon_{\mathrm{ub}}(\alpha,\Delta)$ is 1-Lipschitz in Δ. In other words, $\varepsilon_{\mathrm{ub}}(\alpha,\Delta)$ grows at most linearly with slope at most 1.

□

A.4 Proof of Corollary 6

Claim A.3. *For fixed Δ, as $\alpha\to\infty$, $\varepsilon_{\mathrm{ub}}(\alpha,\Delta) = \Delta - \alpha^{-1}\log\left(\alpha(1-e^{-\Delta})\right) + \mathcal{O}\left(\alpha^{-1}\right)$.*

Proof. $A = e^{-\Delta} \in (0,1)$. As $\alpha\to\infty$, $B = e^{(\alpha-1)\Delta} \to \infty$. Thus

$$\frac{B-A}{B-1} = 1 + \frac{1-A}{B-1} = 1 + \mathcal{O}(B^{-1}) = 1 + \mathcal{O}(e^{-(\alpha-1)\Delta}) \tag{A.25}$$

Plug Eq. (A.25) into the first term in Eq. (A.21)

$$\log\left(\frac{(\alpha-1)(B-A)}{\alpha(B-1)}\right) = \log\left(\frac{\alpha-1}{\alpha}\right) + \log\left(\frac{B-A}{B-1}\right)$$
$$= \log\left(1-\alpha^{-1}\right) + \log\left(1+\mathcal{O}(e^{-(\alpha-1)\Delta})\right)$$
$$= -\alpha^{-1} + \mathcal{O}(\alpha^{-2}) + \mathcal{O}\left(e^{-(\alpha-1)\Delta}\right) \quad \text{as } \alpha\to\infty \tag{A.26}$$

Since $B - A = e^{(\alpha-1)\Delta} - e^{-\Delta} = e^{-\Delta}\left(e^{\alpha\Delta} - 1\right)$,

$$\frac{1}{\alpha-1}\log\left(\frac{B-A}{\alpha(1-A)}\right) = \frac{1}{\alpha-1}\Big[\log(B-A) - \log(\alpha(1-A))\Big]$$

$$
\begin{aligned}
&= \frac{1}{\alpha-1}\Big[-\Delta + \log\left(e^{\alpha\Delta}-1\right) - \log(\alpha(1-A))\Big] \\
&= \frac{1}{\alpha-1}\Big[-\Delta + \alpha\Delta + \log\left(1-e^{-\alpha\Delta}\right) - \log(\alpha(1-A))\Big] \\
&= \Delta - \frac{1}{\alpha-1}\log(\alpha(1-A)) + \frac{\log\left(1-e^{-\alpha\Delta}\right)}{\alpha-1} \qquad \text{(A.27)}
\end{aligned}
$$

Plug Eqs. (A.27) and (A.26) in Eq. (A.21), then

$$
\begin{aligned}
\varepsilon_{\mathrm{ub}}(\alpha,\Delta) &= \frac{1}{\alpha-1}\log\left(\frac{B-A}{\alpha(1-A)}\right) + \log\left(\frac{(\alpha-1)(B-A)}{\alpha(B-1)}\right) \\
&= \Delta - \frac{1}{\alpha-1}\log(\alpha(1-A)) + \frac{\log\left(1-e^{-\alpha\Delta}\right)}{\alpha-1} - \alpha^{-1} + \mathcal{O}(\alpha^{-2}) + \underbrace{\mathcal{O}\left(e^{-(\alpha-1)\Delta}\right)}_{\text{dominated by } \alpha^{-2}} \\
&\quad \text{since } \frac{1}{\alpha-1} = \frac{1}{\alpha} + \mathcal{O}(\alpha^{-2}) \text{ and } \log\left(1-e^{-\alpha\Delta}\right) \approx e^{-\alpha\Delta} \quad \text{for sufficiently large } \alpha \\
&= \Delta - \left(\frac{1}{\alpha} + \mathcal{O}(\alpha^{-2})\right)\cdot\log(\alpha(1-A)) + \mathcal{O}\left(\frac{e^{-\alpha\Delta}}{\alpha}\right) - \alpha^{-1} + \mathcal{O}(\alpha^{-2}) \\
&\quad \text{since } \frac{e^{-\alpha\Delta}}{\alpha} \le \frac{1}{\alpha^2} \le \frac{\log\alpha}{\alpha^2} \le \frac{1}{\alpha} \le \frac{\log\alpha}{\alpha} \quad \text{for sufficiently large } \alpha \\
&= \Delta - \frac{1}{\alpha}\log(\alpha(1-e^{-\Delta})) + \mathcal{O}\left(\alpha^{-1}\right) \qquad \text{(A.28)}
\end{aligned}
$$

Similar to Appendix A.3, $\tilde{c}_1(\boldsymbol{\theta}) \in [-b,-a]$ under Q, we also have $D_\alpha(Q\|P) \le b-a = \Delta$ for every $\alpha > 1$. Hence $\max\{D_\infty(P\|Q), D_\infty(Q\|P)\} \le \Delta$, and the pure Δ-DP conclusion of Corollary 5 is two-sided. □

A.5 Proof of Corollary 8

Proof. Step 1: a fixed ordered pair. Fix an ordered neighboring pair $(\mathbf{Y},\mathbf{Y}')$, and let $P = \pi(\cdot|\mathbf{Y})$, $Q = \pi(\cdot|\mathbf{Y}')$. Suppose $c_1(\boldsymbol{\theta}) = \log\{\ell(\boldsymbol{\theta};\mathbf{Y})/\ell(\boldsymbol{\theta};\mathbf{Y}')\}$ is sub-Gaussian under P, i.e., there exists $\sigma^2_{\mathbf{Y},\mathbf{Y}'} \in (0,\infty)$ such that for all $t \in \mathbb{R}$,

$$
\log \mathbb{E}_{\boldsymbol{\theta}\sim P}\Big[e^{t(c_1(\boldsymbol{\theta})-\mu_{\mathbf{Y},\mathbf{Y}'})}\Big] \le \frac{t^2\sigma^2_{\mathbf{Y},\mathbf{Y}'}}{2}, \qquad \mu_{\mathbf{Y},\mathbf{Y}'} \triangleq \mathbb{E}_{\boldsymbol{\theta}\sim P}[c_1(\boldsymbol{\theta})]. \qquad \text{(A.29)}
$$

Setting $t = \alpha-1$ and $t=-1$ in Eq. (A.29), respectively, bounds the two exponential moments in Eq. (A.7):

$$
\begin{aligned}
\log \mathbb{E}_{\boldsymbol{\theta}\sim P}\big[e^{(\alpha-1)c_1(\boldsymbol{\theta})}\big] &\le (\alpha-1)\mu_{\mathbf{Y},\mathbf{Y}'} + \frac{(\alpha-1)^2\sigma^2_{\mathbf{Y},\mathbf{Y}'}}{2}, \\
c_2 = \log \mathbb{E}_{\boldsymbol{\theta}\sim P}\big[e^{-c_1(\boldsymbol{\theta})}\big] &\le -\mu_{\mathbf{Y},\mathbf{Y}'} + \frac{\sigma^2_{\mathbf{Y},\mathbf{Y}'}}{2}.
\end{aligned}
$$

Plugging both into Eq. (A.6),

$$
\begin{aligned}
D_\alpha(P\|Q) &= c_2 + \frac{1}{\alpha-1}\log \mathbb{E}_{\boldsymbol{\theta}\sim P}\big[e^{(\alpha-1)c_1(\boldsymbol{\theta})}\big] \\
&\le -\mu_{\mathbf{Y},\mathbf{Y}'} + \frac{\sigma^2_{\mathbf{Y},\mathbf{Y}'}}{2} + \mu_{\mathbf{Y},\mathbf{Y}'} + \frac{(\alpha-1)\sigma^2_{\mathbf{Y},\mathbf{Y}'}}{2} = \frac{\alpha}{2}\,\sigma^2_{\mathbf{Y},\mathbf{Y}'}. \qquad \text{(A.30)}
\end{aligned}
$$

Step 2: uniform, two-sided bound. Applying Step 1 to the reversed pair $(\mathbf{Y}',\mathbf{Y})$ gives $D_\alpha(Q\|P) \le \frac{\alpha}{2}\sigma^2_{\mathbf{Y}',\mathbf{Y}}$, where, by definition, $\sigma^2_{\mathbf{Y}',\mathbf{Y}}$ is the variance proxy of $\tilde{c}_1(\boldsymbol{\theta}) = -c_1(\boldsymbol{\theta})$ under Q. Since $(\mathbf{Y}',\mathbf{Y})$ is an ordered neighboring pair whenever $(\mathbf{Y},\mathbf{Y}')$ is, both $\sigma^2_{\mathbf{Y},\mathbf{Y}'}$ and $\sigma^2_{\mathbf{Y}',\mathbf{Y}}$ are bounded by

$$
\bar{\sigma}^2 \triangleq \sup_{(\mathbf{Y},\mathbf{Y}'):\, d(\mathbf{Y},\mathbf{Y}')=1} \sigma^2_{\mathbf{Y},\mathbf{Y}'}, \qquad \text{(A.31)}
$$

where the supremum is over all ordered neighboring pairs in the admissible data domain. Hence, by Eq. (A.30),

$$\sup_{(\mathbf{Y},\mathbf{Y}'):\, d(\mathbf{Y},\mathbf{Y}')=1} \max\{D_\alpha(P\|Q), D_\alpha(Q\|P)\} \le \frac{\alpha\bar{\sigma}^2}{2} \triangleq \varepsilon_{\text{subGaus}}(\alpha, \bar{\sigma}^2), \tag{A.32}$$

which completes the proof. □

A.6 Proof of Corollary 13

Proof. If assume $c_1(\boldsymbol{\theta})$ is sub-Exponential under $\pi(\cdot|\mathbf{Y})$, i.e. there exist $\nu^2_{\mathbf{Y},\mathbf{Y}'} \in (0,\infty)$ and $b_{\mathbf{Y},\mathbf{Y}'} \in (0,\infty)$ such that for all $t \in \mathbb{R}$ satisfying $|t| < 1/b_{\mathbf{Y},\mathbf{Y}'}$, it holds that

$$\log \mathbb{E}_{\boldsymbol{\theta}\sim P}\Big(e^{t(c_1(\boldsymbol{\theta})-\mu_{\mathbf{Y},\mathbf{Y}'})}\Big) \le \frac{t^2\nu^2_{\mathbf{Y},\mathbf{Y}'}}{2(1-b_{\mathbf{Y},\mathbf{Y}'}|t|)}; \tag{A.33}$$

Then, by setting t as -1 and $\alpha-1$ in Eq. (A.33) respectively, the required exponential moments of $c_1(\boldsymbol{\theta})$ under $\pi(\cdot|\mathbf{Y})$ in Eq. (A.7) can be derived as:

when $t = \alpha - 1$, we have:

$$\log \mathbb{E}_{\boldsymbol{\theta}\sim\pi(\cdot|\mathbf{Y})}\left(e^{(\alpha-1)\cdot c_1(\boldsymbol{\theta})}\right) \le (\alpha-1)\mu_{\mathbf{Y},\mathbf{Y}'} + \frac{(\alpha-1)^2\nu^2_{\mathbf{Y},\mathbf{Y}'}}{2(1-b_{\mathbf{Y},\mathbf{Y}'}(\alpha-1))}, \tag{A.34}$$

when $t = -1$, we have:

$$c_2 = \log \mathbb{E}_{\boldsymbol{\theta}\sim\pi(\cdot|\mathbf{Y})}[e^{-c_1(\boldsymbol{\theta})}] \le -\mu_{\mathbf{Y},\mathbf{Y}'} + \frac{\nu^2_{\mathbf{Y},\mathbf{Y}'}}{2(1-b_{\mathbf{Y},\mathbf{Y}'})}. \tag{A.35}$$

Plugging the above two into Eq. (A.6), under $b_{\mathbf{Y},\mathbf{Y}'} < 1$ and $b_{\mathbf{Y},\mathbf{Y}'}(\alpha-1) < 1$,

$$\begin{aligned}
D_\alpha(\pi(\boldsymbol{\theta}|\mathbf{Y})\|\pi(\boldsymbol{\theta}|\mathbf{Y}')) &= c_2 + \frac{1}{\alpha-1}\log \mathbb{E}_{\boldsymbol{\theta}\sim\pi(\cdot|\mathbf{Y})}\left[e^{(\alpha-1)\cdot c_1(\boldsymbol{\theta})}\right] \\
&\le -\mu_{\mathbf{Y},\mathbf{Y}'} + \frac{\nu^2_{\mathbf{Y},\mathbf{Y}'}}{2(1-b_{\mathbf{Y},\mathbf{Y}'})} + \frac{1}{\alpha-1}\Big[(\alpha-1)\mu_{\mathbf{Y},\mathbf{Y}'} + \frac{(\alpha-1)^2\nu^2_{\mathbf{Y},\mathbf{Y}'}}{2(1-b_{\mathbf{Y},\mathbf{Y}'}(\alpha-1))}\Big] \\
&= \frac{\nu^2_{\mathbf{Y},\mathbf{Y}'}}{2(1-b_{\mathbf{Y},\mathbf{Y}'})} + \frac{(\alpha-1)\nu^2_{\mathbf{Y},\mathbf{Y}'}}{2(1-b_{\mathbf{Y},\mathbf{Y}'}(\alpha-1))} \\
&= \frac{\nu^2_{\mathbf{Y},\mathbf{Y}'}}{2}\cdot\frac{1-b_{\mathbf{Y},\mathbf{Y}'}(\alpha-1)+(\alpha-1)(1-b_{\mathbf{Y},\mathbf{Y}'})}{(1-b_{\mathbf{Y},\mathbf{Y}'})(1-b_{\mathbf{Y},\mathbf{Y}'}(\alpha-1))} \\
&= \frac{\nu^2_{\mathbf{Y},\mathbf{Y}'}}{2}\frac{\alpha(1-2b_{\mathbf{Y},\mathbf{Y}'})+2b_{\mathbf{Y},\mathbf{Y}'}}{(1-b_{\mathbf{Y},\mathbf{Y}'})\big(1-b_{\mathbf{Y},\mathbf{Y}'}(\alpha-1)\big)} = \varepsilon_{\text{subExp}}(\alpha, b_{\mathbf{Y},\mathbf{Y}'}, \nu^2_{\mathbf{Y},\mathbf{Y}'}).
\end{aligned} \tag{A.36}$$

Applying the above to the reversed pair $(\mathbf{Y}', \mathbf{Y})$, i.e., to $\tilde{c}_1(\boldsymbol{\theta}) = -c_1(\boldsymbol{\theta})$ under $\pi(\cdot|\mathbf{Y}')$, gives $D_\alpha(\pi(\boldsymbol{\theta}|\mathbf{Y}')\|\pi(\boldsymbol{\theta}|\mathbf{Y})) \le \varepsilon_{\text{subExp}}(\alpha, b_{\mathbf{Y}',\mathbf{Y}}, \nu^2_{\mathbf{Y}',\mathbf{Y}})$, provided $b_{\mathbf{Y}',\mathbf{Y}} < \min\{(\alpha-1)^{-1}, 1\}$. Since $(\mathbf{Y}', \mathbf{Y})$ is an ordered neighboring pair whenever $(\mathbf{Y}, \mathbf{Y}')$ is, if $b_{\mathbf{Y},\mathbf{Y}'} < \min\{(\alpha-1)^{-1}, 1\}$ for all ordered neighboring pairs, then

$$\max\{D_\alpha(P\|Q), D_\alpha(Q\|P)\} \le \sup_{(\mathbf{Y},\mathbf{Y}'):\, d(\mathbf{Y},\mathbf{Y}')=1} \varepsilon_{\text{subExp}}(\alpha, b_{\mathbf{Y},\mathbf{Y}'}, \nu^2_{\mathbf{Y},\mathbf{Y}'}) \tag{A.37}$$

uniformly over all neighboring pairs, which completes the proof. □

A.7 Proof of Example 7 for Categorical data

Proof. Let $y_i \in \{1, \ldots, K\}$ denote a categorical outcome with K categories. Let $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_K)$, where $\sum_{j=1}^K \theta_j = 1$ and $\theta_j = \Pr(y_i = j)$. Assume $y_i|\boldsymbol{\theta} \sim \text{Categorical}(\boldsymbol{\theta})$ independently for $i = 1, \ldots, n$, and the likelihood of the dataset $Y = \{y_1, \ldots, y_n\}$ is $\ell(\boldsymbol{\theta}; \mathbf{Y}) = \prod_{j=1}^K \theta_j^{s_j(Y)}$, where $s_j(Y) = \sum_{i=1}^n \mathbf{1}\{y_i = j\}$ is the category counts.

A.7.1 bounded neighboring

WLOG, two neighboring datasets Y and Y' are both of size n and differ in exactly one observation. Specifically, there exist distinct indices $j \neq k$ and $j, k \in \{1, \ldots, K\}$ such that $s_j(Y) = s_j(Y') + 1$, $s_k(Y) = s_k(Y') - 1$, and $s_\ell(Y) = s_\ell(Y')$ for all $\ell \notin \{j, k\}$.

$$c_1(\boldsymbol{\theta}) = \log \frac{\ell(\boldsymbol{\theta}; Y)}{\ell(\boldsymbol{\theta}; Y')} = \sum_{\ell=1}^{K} \big(s_\ell(Y) - s_\ell(Y')\big) \log \theta_\ell = \log \frac{\theta_j}{\theta_k} \tag{A.38}$$

$$\Rightarrow D^{(jk)}_{\alpha,\text{bounded}}\big(\pi(\cdot|Y) \,\|\, \pi(\cdot|Y')\big) = \log \mathbb{E}_{\theta\sim\pi(\cdot|Y)}\left[\frac{\theta_k}{\theta_j}\right] + \frac{1}{\alpha - 1}\left(\log \mathbb{E}_{\theta\sim\pi(\cdot|Y)}\left[\left(\frac{\theta_j}{\theta_k}\right)^{\alpha-1}\right]\right). \tag{A.39}$$

Finally, to obtain an RDP guarantee that holds for all neighboring datasets under bounded neighboring, one takes the maximum over all ordered pairs (j, k) corresponding to a single-record change:

$$D_{\alpha,\text{bounded}}\big(\pi(\cdot|Y) \,\|\, \pi(\cdot|Y')\big) = \max_{j,k\in\{1,\ldots,K\}, j\neq k} \Big\{D^{(jk)}_{\alpha,\text{bounded}}\big(\pi(\cdot|Y) \,\|\, \pi(\cdot|Y')\big)\Big\}, \tag{A.40}$$

Since the maximum ranges over all ordered category pairs (j, k) and all datasets Y in the domain, the reverse ordering (Y', Y), which is itself a bounded-neighboring pair with j and k exchanged and posterior $\pi(\cdot|Y')$, is already included.

Dirichlet prior. Assume a Dirichlet prior: $\pi(\boldsymbol{\theta}) = \text{Dir}(\alpha_1, \ldots, \alpha_K)$. Then $\boldsymbol{\theta}|Y \sim \text{Dir}(\beta_1, \ldots, \beta_K)$, where $\beta_j = \alpha_j + s_j(Y)$ for $j \in \{1, \ldots, K\}$. We can calculate the moment below:

$$\mathbb{E}_{\boldsymbol{\theta}\sim\pi(\cdot|Y)}\left[\frac{\theta_k}{\theta_j}\right] = \frac{\beta_k}{\beta_j - 1} = \frac{\alpha_k + s_k(Y)}{\alpha_j + s_j(Y) - 1}, \text{ if } \beta_j > 1 \tag{A.41}$$

$$\mathbb{E}_{\boldsymbol{\theta}\sim\pi(\cdot|Y)}\left[\left(\frac{\theta_j}{\theta_k}\right)^{\alpha-1}\right] = \frac{\Gamma(\beta_j + \alpha - 1)}{\Gamma(\beta_j)} \cdot \frac{\Gamma(\beta_k - \alpha + 1)}{\Gamma(\beta_k)}, \text{ if } \beta_k > \alpha - 1 \tag{A.42}$$

Therefore, plug Eqns. (A.42)(A.41) into Eq (A.39), we can derive the closed-form expression for Rényi Divergence:

$$D^{(jk)}_{\alpha,\text{bounded}}\big(\pi(\cdot|Y) \,\|\, \pi(\cdot|Y')\big) = \log \frac{\beta_k}{\beta_j - 1} + \frac{1}{\alpha - 1}\left[\log\left(\frac{\Gamma(\beta_j - 1 + \alpha)}{\Gamma(\beta_j)} \cdot \frac{\Gamma(\beta_k + 1 - \alpha)}{\Gamma(\beta_k)}\right)\right].$$

Finally, with $\alpha < 1 + \min_k \beta_k$,

$$D_{\alpha,\text{bounded}}\big(\pi(\cdot|Y) \,\|\, \pi(\cdot|Y')\big) = \max_{j,k\in\{1,\ldots,K\}, j\neq k} \Big\{D^{(jk)}_{\alpha,\text{bounded}}\big(\pi(\cdot|Y) \,\|\, \pi(\cdot|Y')\big)\Big\}. \tag{A.43}$$

Special case. If $\exists\, \eta \in (0, 1/2)$ such that $\forall j \in \{1, \ldots, K\}, \theta_j \in [\eta, 1-\eta]$, then $|c_1(\boldsymbol{\theta})| \leq \log(\frac{1-\eta}{\eta})$.

A.7.2 unbounded neighboring

WLOG, let dataset Y has size n and dataset Y' is of size $n - 1$ by removing the n-th record $y_n \in \{1, \ldots, K\}$. Specifically, there exist one index $j \in \{1, \ldots, K\}$ such that $s_j(Y) = s_j(Y') + 1$, and $s_\ell(Y) = s_\ell(Y')$ for all $\ell \neq j$.

$$c_1(\boldsymbol{\theta}) = \log \frac{\ell(\boldsymbol{\theta}; Y)}{\ell(\boldsymbol{\theta}; Y')} = \sum_{\ell=1}^{K} \big(s_\ell(Y) - s_\ell(Y')\big) \log \theta_\ell = \log \theta_j \tag{A.44}$$

$$\Rightarrow D^{(j)}_{\alpha,\text{unbounded}}\big(\pi(\cdot|Y) \,\|\, \pi(\cdot|Y')\big) = \log \mathbb{E}_{\theta\sim\pi(\cdot|Y)}\big[\theta_j^{-1}\big] + \frac{1}{\alpha - 1}\left(\log \mathbb{E}_{\theta\sim\pi(\cdot|Y)}\big[\theta_j^{\alpha-1}\big]\right). \tag{A.45}$$

For the reverse direction, $\tilde{c}_1(\boldsymbol{\theta}) = -c_1(\boldsymbol{\theta}) = -\log \theta_j$, and by Eq. (9),

$$D^{(j)}_{\alpha,\text{unbounded}}\big(\pi(\cdot|Y') \,\|\, \pi(\cdot|Y)\big) = \log \mathbb{E}_{\boldsymbol{\theta}\sim\pi(\cdot|Y')}[\theta_j] + \frac{1}{\alpha - 1} \log \mathbb{E}_{\boldsymbol{\theta}\sim\pi(\cdot|Y')}\left[\theta_j^{-(\alpha-1)}\right]. \tag{A.46}$$

Finally, to obtain an RDP guarantee that holds for all neighboring datasets under unbounded neighboring, one takes the maximum over all j :

$$D_{\alpha,\text{unbounded}}\big(\pi(\cdot|Y)\,\|\,\pi(\cdot|Y')\big) = \max_{j\in\{1,\dots,K\}} \max\Big\{D^{(j)}_{\alpha,\text{unbounded}}\big(\pi(\cdot|Y)\,\|\,\pi(\cdot|Y')\big), \\ D^{(j)}_{\alpha,\text{unbounded}}\big(\pi(\cdot|Y')\,\|\,\pi(\cdot|Y)\big)\Big\},$$

Dirichlet prior. Assume a Dirichlet prior: $\pi(\boldsymbol{\theta}) = \mathrm{Dir}(\alpha_1,\dots,\alpha_K)$. Then $\boldsymbol{\theta}|Y \sim \mathrm{Dir}(\beta_1,\dots,\beta_K)$, where $\beta_j = \alpha_j + s_j(Y)$ for $j \in \{1,\dots,K\}$. And the marginal $\theta_j|Y \sim \text{Beta}(\beta_j, \sum_{\ell=1}^K \beta_\ell - \beta_j)$, based on which, we can calculate the moments below:

$$\mathbb{E}_{\boldsymbol{\theta}\sim\pi(\cdot|Y)}\left[\theta_j^{-1}\right] = \frac{\sum_{\ell=1}^K \beta_\ell - 1}{\beta_j - 1} \text{ if } \beta_j > 1 \tag{A.47}$$

$$\mathbb{E}_{\boldsymbol{\theta}\sim\pi(\cdot|Y)}\left[\theta_j^{\alpha-1}\right] = \frac{\Gamma(\beta_j+\alpha-1)}{\Gamma(\beta_j)} \cdot \frac{\Gamma(\sum_{\ell=1}^K \beta_\ell)}{\Gamma(\sum_{\ell=1}^K \beta_\ell + \alpha - 1)}. \tag{A.48}$$

Therefore, plug Eqns. (A.48)(A.47) into Eq (A.45), we can derive the closed-form expression for Rényi Divergence:

$$D^{(j)}_{\alpha,\text{unbounded}}\big(\pi(\cdot|Y)\,\|\,\pi(\cdot|Y')\big) = \log\frac{\sum_{\ell=1}^K \beta_\ell - 1}{\beta_j - 1} + \frac{1}{\alpha-1}\log\left[\frac{\Gamma(\beta_j+\alpha-1)}{\Gamma(\beta_j)} \cdot \frac{\Gamma(\sum_{\ell=1}^K \beta_\ell)}{\Gamma(\sum_{\ell=1}^K \beta_\ell + \alpha - 1)}\right]. \tag{A.49}$$

Since $s_j(Y) \geq 1$, $\beta_j \geq \alpha_j + 1 > 1$, so this direction is finite for every $\alpha > 1$. For the reverse direction, $\theta_j|Y' \sim \text{Beta}(\beta_j - 1, \sum_{\ell=1}^K \beta_\ell - \beta_j)$, when $1 < \alpha < \beta_j$

$$D^{(j)}_{\alpha,\text{unbounded}}\big(\pi(\cdot|Y')\,\|\,\pi(\cdot|Y)\big) = \log\frac{\beta_j - 1}{\sum_{\ell=1}^K \beta_\ell - 1} + \frac{1}{\alpha-1}\log\left[\frac{\Gamma(\beta_j-\alpha)}{\Gamma(\beta_j-1)} \cdot \frac{\Gamma(\sum_{\ell=1}^K \beta_\ell - 1)}{\Gamma(\sum_{\ell=1}^K \beta_\ell - \alpha)}\right], \tag{A.50}$$

but $= \infty$ for $\alpha \geq \beta_j$. Since $\beta_j \geq \alpha_j + 1$ with equality when $s_j(Y) = 1$, the two-sided divergence is finite uniformly over all unbounded neighboring pairs if and only if $\alpha < 1 + \min_j \alpha_j$; hence OPS with a Dirichlet prior does not satisfy pure DP under unbounded neighboring.

Special Case. If $\exists\, \eta \in (0, 1/2)$ such that $\forall j \in \{1,\dots,K\}, \theta_j \in [\eta, 1-\eta]$, then $|c_1(\boldsymbol{\theta})| \leq \log(\frac{1}{\eta})$. In the reverse direction, $\tilde{c}_1(\boldsymbol{\theta}) = -\log\theta_j$ has a range of the same width, so Corollary 5 applies in both directions.

□

A.7.3 Proof of Bernoulli data

Proof. Assume observed dataset $Y = \{y_i\}_{i=1}^n$, where $y_i|\theta \overset{\text{i.i.d}}{\sim} \text{Bernoulli}(\theta)$ with $\theta = \Pr(y_i = 1)$ and $1 - \theta = \Pr(y_i = 0)$.

A.7.3.1 bounded neighboring

WLOG, suppose neighboring datasets Y, Y', both of size n, differ in their n-th record y_n and y'_n. Then the log-likelihood ratio can be derived as

$$\ell(\theta;Y) = \theta^{y_n + \sum_{i=1}^{n-1} y_i}(1-\theta)^{n - y_n - \sum_{i=1}^{n-1} y_i}$$
$$\ell(\theta;Y') = \theta^{y'_n + \sum_{i=1}^{n-1} y_i}(1-\theta)^{n - y'_n - \sum_{i=1}^{n-1} y_i}$$
$$\Rightarrow c_1(\theta) = \log\frac{\ell(\theta;Y)}{\ell(\theta;Y')} = (y_n - y'_n)\log\left(\frac{\theta}{1-\theta}\right) = \begin{cases}\log\left(\frac{\theta}{1-\theta}\right) & \text{if } y_n - y'_n = 1 \\ -\log\left(\frac{\theta}{1-\theta}\right) & \text{if } y_n - y'_n = -1\end{cases} \tag{A.51}$$

If $y_n - y'_n = 1$ with $\mathbb{E}_{\boldsymbol{\theta}\sim\pi(\cdot|\mathbf{Y})}\left[\frac{1-\theta}{\theta}\right] < \infty$ and $\mathbb{E}_{\theta\sim\pi(\cdot|Y)}\left[\left(\frac{\theta}{1-\theta}\right)^{\alpha-1}\right] < \infty$:

$$D^{(+)}_{\alpha,\text{bounded}}\big(\pi(\cdot|Y)\,\|\,\pi(\cdot|Y')\big) = \log \mathbb{E}_{\theta\sim\pi(\cdot|Y)}\left[\frac{1-\theta}{\theta}\right] + \frac{1}{\alpha-1}\log \mathbb{E}_{\theta\sim\pi(\cdot|Y)}\left[\left(\frac{\theta}{1-\theta}\right)^{\alpha-1}\right].$$

If $y_n - y'_n = -1$ with $\mathbb{E}_{\theta\sim\pi(\cdot|Y)}\left[\frac{\theta}{1-\theta}\right] < \infty$ and $\mathbb{E}_{\theta\sim\pi(\cdot|Y)}\left[\left(\frac{1-\theta}{\theta}\right)^{\alpha-1}\right] < \infty$:

$$D^{(-)}_{\alpha,\text{bounded}}\big(\pi(\cdot|Y)\,\|\,\pi(\cdot|Y')\big) = \log \mathbb{E}_{\theta\sim\pi(\cdot|Y)}\left[\frac{\theta}{1-\theta}\right] + \frac{1}{\alpha-1}\log \mathbb{E}_{\theta\sim\pi(\cdot|Y)}\left[\left(\frac{1-\theta}{\theta}\right)^{\alpha-1}\right].$$

Without additional structure on the posterior, the worst-case Rényi divergence over bounded neighboring is given by

$$D_{\alpha,\text{bounded}}\big(\pi(\cdot|Y)\,\|\,\pi(\cdot|Y')\big) \le \max\{D^{(+)}_{\alpha,\text{bounded}},\, D^{(-)}_{\alpha,\text{bounded}}\}$$

The reverse ordering (Y', Y) is itself a bounded-neighboring pair with $y'_n - y_n = \mp 1$ and posterior $\pi(\cdot|Y')$, i.e., it is of the form $D^{(\mp)}_{\alpha,\text{bounded}}$ evaluated at the dataset Y'. Hence taking the maximum over $D^{(+)}$ and $D^{(-)}$ and over all datasets in the domain already covers the reverse direction.

Special case. If $\theta \in [\eta, 1-\eta]$ for some constant $\eta \in (0, 1/2)$. Since $\log(\frac{\theta}{1-\theta})$ is monotonically increasing, $|c_1(\theta)| \le \log(\frac{1-\eta}{\eta})$ uniformly.

Beta prior. If $\pi(\theta) = \text{Beta}(a, b)$, and define $s = \sum_{i=1}^n y_i$, then $\pi(\theta|Y) = \text{Beta}(A, B)$, where $A = a + s$, $B = b + n - s$.
If $y_n - y'_n = 1$ and $A > 1, B > \alpha - 1$

$$\mathbb{E}_{\theta\sim\pi(\cdot|Y)}\left[\frac{1-\theta}{\theta}\right] = \frac{B}{A-1}$$

$$\mathbb{E}_{\theta\sim\pi(\cdot|Y)}\left[\left(\frac{\theta}{1-\theta}\right)^{\alpha-1}\right] = \frac{\text{Beta}(A-1+\alpha, B+1-\alpha)}{\text{Beta}(A,B)} = \frac{\Gamma(A+\alpha-1)}{\Gamma(A)}\cdot\frac{\Gamma(B-(\alpha-1))}{\Gamma(B)}$$

$$\Rightarrow D^{(+)}_{\alpha,\text{bounded}}\big(\pi(\cdot|Y)\,\|\,\pi(\cdot|Y')\big) = \log\left[\frac{B}{A-1}\right] + \frac{1}{\alpha-1}\log\left[\frac{\Gamma(A+\alpha-1)}{\Gamma(A)}\cdot\frac{\Gamma(B-(\alpha-1))}{\Gamma(B)}\right]. \tag{A.52}$$

If $y_n - y'_n = -1$, and $B > 1, A > \alpha - 1$

$$\mathbb{E}_{\theta\sim\pi(\cdot|Y)}\left[\frac{\theta}{1-\theta}\right] = \frac{A}{B-1}$$

$$\mathbb{E}_{\theta\sim\pi(\cdot|Y)}\left[\left(\frac{1-\theta}{\theta}\right)^{\alpha-1}\right] = \frac{\text{Beta}(A+1-\alpha, B-1+\alpha)}{\text{Beta}(A,B)} = \frac{\Gamma(A-(\alpha-1))}{\Gamma(A)}\cdot\frac{\Gamma(B+(\alpha-1))}{\Gamma(B)}$$

$$\Rightarrow D^{(-)}_{\alpha,\text{bounded}}\big(\pi(\cdot|Y)\,\|\,\pi(\cdot|Y')\big) = \log\left[\frac{A}{B-1}\right] + \frac{1}{\alpha-1}\log\left[\frac{\Gamma(A-(\alpha-1))}{\Gamma(A)}\cdot\frac{\Gamma(B+(\alpha-1))}{\Gamma(B)}\right] \tag{A.53}$$

Remark A.4. *To ensure finite Beta-moment,* $\alpha < \alpha_{\max} = 1 + \min\{a+s,\ b+n-s\}$.

A.7.3.2 unbounded neighboring

WLOG, suppose neighboring datasets Y, Y', where Y is of size n and Y' is of size $n-1$ by removing the n-th record y_n. Then similarly, the log-likelihood ratio can be derived as

$$\ell(\theta; Y) = \theta^{y_n + \sum_{i=1}^{n-1} y_i}(1-\theta)^{n - y_n - \sum_{i=1}^{n-1} y_i}$$
$$\ell(\theta; Y') = \theta^{\sum_{i=1}^{n-1} y_i}(1-\theta)^{n-1-\sum_{i=1}^{n-1} y_i}$$

$$\Rightarrow c_1(\theta) = \log \frac{\ell(\theta; Y)}{\ell(\theta; Y')} = y_n \log \theta + (1 - y_n) \log(1 - \theta) = \begin{cases} \log \theta & \text{if } y_n = 1 \\ \log(1 - \theta) & \text{if } y_n = 0 \end{cases} \tag{A.54}$$

If $y_n = 1$,

$$D^{(1)}_{\alpha,\text{unbounded}}\big(\pi(\cdot|Y)\|\pi(\cdot|Y')\big) = \log \mathbb{E}_{\boldsymbol{\theta}\sim\pi(\cdot|Y)}\left[\theta^{-1}\right] + \frac{1}{\alpha - 1} \log \mathbb{E}_{\theta\sim\pi(\cdot|Y)}\left[\theta^{\alpha-1}\right].$$

If $y_n = 0$,

$$D^{(0)}_{\alpha,\text{unbounded}}\big(\pi(\cdot|Y) \,\|\, \pi(\cdot|Y')\big) = \log \mathbb{E}_{\boldsymbol{\theta}\sim\pi(\cdot|Y)}\left[(1-\theta)^{-1}\right] + \frac{1}{\alpha - 1} \log \mathbb{E}_{\theta\sim\pi(\cdot|Y)}\left[(1-\theta)^{\alpha-1}\right].$$

For the reverse direction, $\tilde{c}_1(\theta) = -c_1(\theta)$ and the expectations are taken under $\pi(\cdot|Y')$:

$$D^{(1)}_{\alpha,\text{unbounded}}\big(\pi(\cdot|Y')\|\pi(\cdot|Y)\big) = \log \mathbb{E}_{\theta\sim\pi(\cdot|Y')}\left[\theta\right] + \frac{1}{\alpha - 1} \log \mathbb{E}_{\theta\sim\pi(\cdot|Y')}\left[\theta^{-(\alpha-1)}\right],$$

$$D^{(0)}_{\alpha,\text{unbounded}}\big(\pi(\cdot|Y')\|\pi(\cdot|Y)\big) = \log \mathbb{E}_{\theta\sim\pi(\cdot|Y')}\left[1-\theta\right] + \frac{1}{\alpha - 1} \log \mathbb{E}_{\theta\sim\pi(\cdot|Y')}\left[(1-\theta)^{-(\alpha-1)}\right].$$

Without additional structure on the posterior, the worst-case Rényi divergence over unbounded neighboring is given by

$$D_{\alpha,\text{unbounded}} \le \max_{k\in\{0,1\}} \max\left\{D^{(k)}_{\alpha,\text{unbounded}}\big(\pi(\cdot|Y)\|\pi(\cdot|Y')\big),\, D^{(k)}_{\alpha,\text{unbounded}}\big(\pi(\cdot|Y')\|\pi(\cdot|Y)\big)\right\}$$

Special case: If $\theta \in [\eta, 1-\eta]$ for some $\eta \in (0, 1/2)$. $|c_1(\theta)| \le \max\{|\log \eta|, |\log(1-\eta)|\} = -\log \eta$ uniformly. In the reverse direction, $\tilde{c}_1(\theta) = -c_1(\theta)$ satisfies the same bound, so Corollary 5 applies in both directions.

Beta prior. If $\pi(\theta) = \text{Beta}(a, b)$, and define $s = \sum_{i=1}^n y_i$, then $\pi(\theta|Y) = \text{Beta}(A, B)$, where $A = a + s$, $B = b + n - s$. Then
If $y_n = 1$ and $A > 1$,

$$\mathbb{E}_{\theta\sim\pi(\cdot|Y)}\left[\theta^{-1}\right] = \frac{A + B - 1}{A - 1}; \tag{A.55}$$

$$\mathbb{E}_{\theta\sim\pi(\cdot|Y)}\left[\theta^{\alpha-1}\right] = \frac{\text{Beta}(A - 1 + \alpha, B)}{\text{Beta}(A, B)} = \frac{\Gamma(A + \alpha - 1)\Gamma(A + B)}{\Gamma(A)\,\Gamma(A + B + \alpha - 1)} \tag{A.56}$$

$$\Rightarrow D^{(1)}_{\alpha,\text{unbounded}}\big(\pi(\cdot|Y)\|\pi(\cdot|Y')\big) = \log\left(\frac{A + B - 1}{A - 1}\right) + \frac{1}{\alpha - 1} \log\left(\frac{\Gamma(A + \alpha - 1)\Gamma(A + B)}{\Gamma(A)\,\Gamma(A + B + \alpha - 1)}\right). \tag{A.57}$$

If $y_n = 0$, and $B > 1$,

$$\mathbb{E}_{\theta\sim\pi(\cdot|Y)}\left[(1-\theta)^{-1}\right] = \frac{A + B - 1}{B - 1}; \tag{A.58}$$

$$\mathbb{E}_{\theta\sim\pi(\cdot|Y)}\left[(1-\theta)^{\alpha-1}\right] = \frac{\text{Beta}(A, B - 1 + \alpha)}{\text{Beta}(A, B)} = \frac{\Gamma(A + B)\Gamma(B + \alpha - 1)}{\Gamma(A + B + \alpha - 1)\Gamma(B)} \tag{A.59}$$

$$\Rightarrow D^{(0)}_{\alpha,\text{unbounded}}\big(\pi(\cdot|Y)\|\pi(\cdot|Y')\big) = \log\left(\frac{A + B - 1}{B - 1}\right) + \frac{1}{\alpha - 1} \log\left(\frac{\Gamma(B + \alpha - 1)\Gamma(A + B)}{\Gamma(B)\,\Gamma(A + B + \alpha - 1)}\right). \tag{A.60}$$

For the reverse direction, if $y_n = 1$, then $\pi(\theta|Y') = \text{Beta}(A - 1, B)$ and if $1 < \alpha < A$,

$$D^{(1)}_{\alpha,\text{unbounded}}\big(\pi(\cdot|Y')\|\pi(\cdot|Y)\big) = \log\left(\frac{A - 1}{A + B - 1}\right) + \frac{1}{\alpha - 1} \log\left(\frac{\Gamma(A - \alpha)\Gamma(A + B - 1)}{\Gamma(A - 1)\,\Gamma(A + B - \alpha)}\right); \tag{A.61}$$

if $y_n = 0$, then $\pi(\theta|Y') = \text{Beta}(A, B - 1)$ and if $1 < \alpha < B$.

$$D^{(0)}_{\alpha,\text{unbounded}}\big(\pi(\cdot|Y')\|\pi(\cdot|Y)\big) = \log\left(\frac{B - 1}{A + B - 1}\right) + \frac{1}{\alpha - 1} \log\left(\frac{\Gamma(B - \alpha)\Gamma(A + B - 1)}{\Gamma(B - 1)\,\Gamma(A + B - \alpha)}\right). \tag{A.62}$$

Both equal ∞ when α exceeds the stated range.

Remark A.5. *The forward direction under unbounded neighboring imposes no upper bound on the Rényi order α, but the reverse direction is finite only for $\alpha < A$ (if $y_n = 1$) or $\alpha < B$ (if $y_n = 0$). Since $A = a + s \geq a + 1$ when $y_n = 1$ and $B = b + n - s \geq b + 1$ when $y_n = 0$, with equality attainable, the two-sided Rényi divergence is finite uniformly over all unbounded neighboring pairs if and only if $\alpha < 1 + \min\{a, b\}$. In particular, OPS with a Beta prior does not satisfy pure DP under unbounded neighboring.*

□

A.8 Proof of Proposition 9 for Generalized Linear Models

Proof. We consider generalized linear models (GLMs) with canonical link. Each record is $(\mathbf{x}_i, y_i)$, where $\mathbf{x}_i \in \mathbb{R}^d$ and $y_i \in \mathcal{Y} \subseteq \mathbb{R}$. The dataset is $\mathbf{Y} = \{(\mathbf{x}_i, y_i)\}_{i=1}^n$, and the model parameter is $\boldsymbol{\beta} \in \mathbb{R}^d$. Conditioned on $\mathbf{x}_i$ and $\boldsymbol{\beta}$, the response y_i follows an exponential family distribution as below:

$$p(y_i|\mathbf{x}_i, \boldsymbol{\beta}) = \exp\big(y_i\, \mathbf{x}_i^\top \boldsymbol{\beta} - \psi(\mathbf{x}_i^\top \boldsymbol{\beta}) + h(y_i)\big)\,, \tag{A.63}$$

where $\psi(\cdot)$ is the log-partition function and $h(\cdot)$ does not depend on $\boldsymbol{\beta}$. Then the log-likelihood is

$$\log p(y_i|\mathbf{x}_i, \boldsymbol{\beta}) = y_i \mathbf{x}_i^\top \boldsymbol{\beta} - \psi(\mathbf{x}_i^\top \boldsymbol{\beta}) + h(y_i) \tag{A.64}$$

$$\Rightarrow \ell(\boldsymbol{\beta}|\mathbf{Y}) = \sum_{i=1}^n \log p(y_i|\mathbf{x}_i, \boldsymbol{\beta}) = \sum_{i=1}^n \big(y_i\, \mathbf{x}_i^\top \boldsymbol{\beta} - \psi(\mathbf{x}_i^\top \boldsymbol{\beta})\big) + \sum_{i=1}^n h(y_i). \tag{A.65}$$

Suppose there exists a finite constant $L > 0$ such that for every $(\mathbf{x}_i, y_i)$ and every $\boldsymbol{\beta}$, $\|\nabla_{\boldsymbol{\beta}} \log p(y_i|\mathbf{x}_i, \boldsymbol{\beta})\| \leq L$. If imposing a Gaussian prior on $\boldsymbol{\beta}$, i.e. $\pi(\boldsymbol{\beta}) = \mathcal{N}(0, \tau^2 \mathbf{I}_d)$, then the posterior distribution can be derived as

$$\pi(\boldsymbol{\beta}|\mathbf{Y}) \propto \exp\left(\sum_{i=1}^n \big(y_i\, \mathbf{x}_i^\top \boldsymbol{\beta} - \psi(\mathbf{x}_i^\top \boldsymbol{\beta})\big) - \frac{\|\boldsymbol{\beta}\|_2^2}{2\tau^2}\right). \tag{A.66}$$

A.8.1 bounded neighboring

Let $\mathbf{Y}$ and $\mathbf{Y}'$ differ in exactly one record, i.e. $\mathbf{z}_n = (\mathbf{x}_n, y_n)$ in $\mathbf{Y}$ is replaced by $\mathbf{z}'_n = (\mathbf{x}'_n, y'_n)$ in $\mathbf{Y}'$, then

$$\begin{aligned} c_1(\boldsymbol{\beta}) = \log \frac{\ell(\boldsymbol{\beta}; \mathbf{Y})}{\ell(\boldsymbol{\beta}; \mathbf{Y}')} &= \log p(y_n|\mathbf{x}_n, \boldsymbol{\beta}) - \log p(y'_n|\mathbf{x}'_n, \boldsymbol{\beta}) \\ &= y_n \mathbf{x}_n^\top \boldsymbol{\beta} - \psi(\mathbf{x}_n^\top \boldsymbol{\beta}) + h(y_n) - \big(y'_n \mathbf{x}_n'^\top \boldsymbol{\beta} - \psi(\mathbf{x}_n'^\top \boldsymbol{\beta}) + h(y'_n)\big) \\ &= y_n \mathbf{x}_n^\top \boldsymbol{\beta} - y'_n \mathbf{x}_n'^\top \boldsymbol{\beta} - \big(\psi(\mathbf{x}_n^\top \boldsymbol{\beta}) - \psi(\mathbf{x}_n'^\top \boldsymbol{\beta})\big) + h(y_n) - h(y'_n) \end{aligned} \tag{A.67}$$

First, we prove $c_1(\boldsymbol{\beta})$ is Lipschitz.

$$\nabla_{\boldsymbol{\beta}} c_1(\boldsymbol{\beta}) = y_n \mathbf{x}_n - y'_n \mathbf{x}'_n - \big(\psi'(\mathbf{x}_n^\top \boldsymbol{\beta}) \mathbf{x}_n - \psi'(\mathbf{x}_n'^\top \boldsymbol{\beta}) \mathbf{x}'_n\big) \tag{A.68}$$

$$\begin{aligned} \Rightarrow \|\nabla_{\boldsymbol{\beta}} c_1(\boldsymbol{\beta})\| &= \Big\| y_n \mathbf{x}_n - y'_n \mathbf{x}'_n - \big(\psi'(\mathbf{x}_n^\top \boldsymbol{\beta}) \mathbf{x}_n - \psi'(\mathbf{x}_n'^\top \boldsymbol{\beta}) \mathbf{x}'_n\big) \Big\| \\ &= \Big\| \big(y_n - \psi'(\mathbf{x}_n^\top \boldsymbol{\beta})\big) \mathbf{x}_n - \big(y'_n - \psi'(\mathbf{x}_n'^\top \boldsymbol{\beta})\big) \mathbf{x}'_n \Big\| \\ &\leq |y_n - \psi'(\mathbf{x}_n^\top \boldsymbol{\beta})| \|\mathbf{x}_n\| + |y'_n - \psi'(\mathbf{x}_n'^\top \boldsymbol{\beta})| \|\mathbf{x}'_n\| \leq 2L \end{aligned} \tag{A.69}$$

Hence, $c_1(\boldsymbol{\beta})$ is Lipschitz continuous with constant $2L$. Then, we continue to prove the log-concavity of the posterior $\pi(\boldsymbol{\beta}|\mathbf{Y})$.

$$\log \pi(\boldsymbol{\beta}|\mathbf{Y}) = \sum_{i=1}^n \big(y_i\, \mathbf{x}_i^\top \boldsymbol{\beta} - \psi(\mathbf{x}_i^\top \boldsymbol{\beta})\big) - \frac{\|\boldsymbol{\beta}\|_2^2}{2\tau^2} + \text{constant} \tag{A.70}$$

$$\Rightarrow -\nabla_{\boldsymbol{\beta}}^2 \log \pi(\boldsymbol{\beta}|\mathbf{Y}) = \sum_{i=1}^n \psi''(\mathbf{x}_i^\top \boldsymbol{\beta}) \mathbf{x}_i \mathbf{x}_i^\top + \frac{1}{\tau^2} \mathbf{I}_d \succeq \frac{1}{\tau^2} \mathbf{I}_d \quad \text{per convexity of } \psi(\cdot). \tag{A.71}$$

So $\pi(\boldsymbol{\beta}|\mathbf{Y})$ is m-strongly log-concave with $m = \tau^{-2}$.

Lemma A.6 (Log-Sobolev Inequality [Ledoux, 2001]). *Let π be a probability measure on $\mathbb{R}^d$ with density $\pi(x) \propto e^{-U(x)}$, where U is twice continuously differentiable and m-strongly convex, i.e., $\nabla^2 U(x) \succeq m I_d$ for all x. Then π satisfies a logarithmic Sobolev inequality (LSI) with constant $1/m$:*

$$\mathrm{Ent}_\pi(g^2) \le \frac{2}{m}\,\mathbb{E}_\pi\left[\|\nabla g\|^2\right], \quad \textit{for all smooth } g.$$

Lemma A.7 (Herbst's Theorem [Ledoux, 2001]). *Let π be a probability measure satisfying a LSI with constant C_{LSI}. Then for any L-Lipschitz function f,*

$$\log \mathbb{E}_\pi\left[\exp\big(t(f - \mathbb{E}_\pi f)\big)\right] \le \frac{C_{\mathrm{LSI}}\, L^2\, t^2}{2}, \quad \forall t \in \mathbb{R}.$$

In particular, f is sub-Gaussian under π with variance proxy $C_{\mathrm{LSI}} L^2$.

Per above lemmas, $c_1(\boldsymbol{\beta})$ is sub-Gaussian under $\pi(\cdot|\mathbf{Y})$ with

$$\sigma^2_{\mathbf{Y},\mathbf{Y}'} \le \frac{(2L)^2}{m} = 4L^2\tau^2. \tag{A.72}$$

The same argument applies to the reversed pair $(\mathbf{Y}', \mathbf{Y})$: $\tilde{c}_1(\boldsymbol{\beta}) = -c_1(\boldsymbol{\beta})$ has the same Lipschitz constant $2L$, and $\pi(\boldsymbol{\beta}|\mathbf{Y}')$ is also τ^{-2}-strongly log-concave since $\mathbf{Y}'$ also has n records. Hence $\sigma^2_{\mathbf{Y}',\mathbf{Y}} \le 4L^2\tau^2$, and $\bar{\sigma}^2 \le 4L^2\tau^2$. Plug into Corollary 8

$$\max\{D_\alpha(\pi(\boldsymbol{\beta}|\mathbf{Y})\|\pi(\boldsymbol{\beta}|\mathbf{Y}')),\, D_\alpha(\pi(\boldsymbol{\beta}|\mathbf{Y}')\|\pi(\boldsymbol{\beta}|\mathbf{Y}))\} \le \frac{\alpha\bar{\sigma}^2}{2} \le 2\alpha L^2\tau^2 \tag{A.73}$$

A.8.2 unbounded neighboring

Let $\mathbf{Y}$ has size n and $\mathbf{Y}'$ is of size $n-1$ by removing the n-th record $\mathbf{z}_n = (\mathbf{x}_n, y_n)$ in $\mathbf{Y}$. Then

$$c_1(\boldsymbol{\beta}) = \log\frac{\ell(\boldsymbol{\beta};\mathbf{Y})}{\ell(\boldsymbol{\beta};\mathbf{Y}')} = \log p(y_n|\mathbf{x}_n,\boldsymbol{\beta}) = y_n\mathbf{x}_n^\top\boldsymbol{\beta} - \psi(\mathbf{x}_n^\top\boldsymbol{\beta}) + h(y_n). \tag{A.74}$$

Similarly, we prove $c_1(\boldsymbol{\beta})$ is Lipschitz.

$$\begin{aligned} \nabla_{\boldsymbol{\beta}} c_1(\boldsymbol{\beta}) &= y_n\mathbf{x}_n - \psi'(\mathbf{x}_n^\top\boldsymbol{\beta})\mathbf{x}_n \\ \Rightarrow \|\nabla_{\boldsymbol{\beta}} c_1(\boldsymbol{\beta})\| &= \left\| \left(y_n - \psi'(\mathbf{x}_n^\top\boldsymbol{\beta})\right)\mathbf{x}_n \right\| \le L. \end{aligned} \tag{A.75}$$

Combined with log-concavity, $c_1(\boldsymbol{\beta})$ is sub-Gaussian under $\pi(\cdot|\mathbf{Y})$ with

$$\sigma^2_{\mathbf{Y},\mathbf{Y}'} \le \frac{L^2}{m} \le L^2\tau^2. \tag{A.76}$$

For the reversed pair $(\mathbf{Y}', \mathbf{Y})$, $\tilde{c}_1(\boldsymbol{\beta}) = -\log p(y_n|\mathbf{x}_n,\boldsymbol{\beta})$ has the same gradient norm bound L, and

$$-\nabla^2_{\boldsymbol{\beta}} \log\pi(\boldsymbol{\beta}|\mathbf{Y}') = \sum_{i=1}^{n-1} \psi''(\mathbf{x}_i^\top\boldsymbol{\beta})\mathbf{x}_i\mathbf{x}_i^\top + \frac{1}{\tau^2}\mathbf{I}_d \succeq \frac{1}{\tau^2}\mathbf{I}_d, \tag{A.77}$$

where the empty sum equals 0 when $n = 1$. Hence $\pi(\boldsymbol{\beta}|\mathbf{Y}')$ is also τ^{-2}-strongly log-concave, $\sigma^2_{\mathbf{Y}',\mathbf{Y}} \le L^2\tau^2$, and $\bar{\sigma}^2 \le L^2\tau^2$. Plug into Corollary 8

$$\max\{D_\alpha(\pi(\boldsymbol{\beta}|\mathbf{Y})\|\pi(\boldsymbol{\beta}|\mathbf{Y}')),\, D_\alpha(\pi(\boldsymbol{\beta}|\mathbf{Y}')\|\pi(\boldsymbol{\beta}|\mathbf{Y}))\} \le \frac{\alpha\bar{\sigma}^2}{2} \le \frac{\alpha L^2\tau^2}{2}. \tag{A.78}$$

□

A.8.3 Proof of Example 10 for Logistic Regression

A.8.3.1 bounded neighboring

Proof. For logistic regression we have $\psi(t) = \log(1 + e^t)$, and $h(\cdot) = 0$, plug into Eq. (A.67):

$$c_1(\boldsymbol{\beta}) = y_n\mathbf{x}_n^\top\boldsymbol{\beta} - y_n'\mathbf{x}_n'^\top\boldsymbol{\beta} - \left(\log(1 + e^{\mathbf{x}_n^\top\boldsymbol{\beta}}) - \log(1 + e^{\mathbf{x}_n'^\top\boldsymbol{\beta}})\right) \tag{A.79}$$

First, we prove $c_1(\boldsymbol{\beta})$ is Lipschitz. Recall $\|\mathbf{x}\| \leq R$ and let $\sigma(t) = \frac{1}{1+e^{-t}}$, given $|\sigma(\cdot)| < 1$ and $\frac{d\log(1+e^t)}{dt} = \sigma(t)$, we can have

$$\nabla_{\boldsymbol{\beta}} c_1(\boldsymbol{\beta}) = y_n\mathbf{x}_n - y'_n\mathbf{x}'_n - \Big(\sigma(\mathbf{x}_n^\top\boldsymbol{\beta})\mathbf{x}_n - \sigma(\mathbf{x}_n'^\top\boldsymbol{\beta})\mathbf{x}'_n\Big)$$
$$= \begin{cases} (y_n - y'_n)\mathbf{x}_n & \text{if } \mathbf{x}_n = \mathbf{x}'_n \text{ and } y_n \neq y'_n \\ y_n\mathbf{x}_n - y_n\mathbf{x}'_n - \Big(\sigma(\mathbf{x}_n^\top\boldsymbol{\beta})\mathbf{x}_n - \sigma(\mathbf{x}_n'^\top\boldsymbol{\beta})\mathbf{x}'_n\Big) & \text{if } \mathbf{x}_n \neq \mathbf{x}'_n \text{ and } y_n = y'_n \\ y_n\mathbf{x}_n - y'_n\mathbf{x}'_n - \Big(\sigma(\mathbf{x}_n^\top\boldsymbol{\beta})\mathbf{x}_n - \sigma(\mathbf{x}_n'^\top\boldsymbol{\beta})\mathbf{x}'_n\Big) & \text{if } \mathbf{x}_n \neq \mathbf{x}'_n \text{ and } y_n \neq y'_n \\ 0 & \text{if } \mathbf{x}_n = \mathbf{x}'_n \text{ and } y_n = y'_n \end{cases}.$$

Therefore, when $\mathbf{x}_n = \mathbf{x}'_n$, the worst case is indeed flipping the label, then $\|\nabla_{\boldsymbol{\beta}} c_1(\boldsymbol{\beta})\| = \|(y_n - y'_n)\mathbf{x}_n\| \leq \|\mathbf{x}_n\| \leq R$. Otherwise,

$$\begin{aligned} \|\nabla_{\boldsymbol{\beta}} c_1(\boldsymbol{\beta})\| &\leq \|y_n\mathbf{x}_n - y'_n\mathbf{x}'_n - \sigma(\mathbf{x}_n^\top\boldsymbol{\beta})\mathbf{x}_n + \sigma(\mathbf{x}_n'^\top\boldsymbol{\beta})\mathbf{x}'_n\| \\ &\leq \|(y_n - \sigma(\mathbf{x}_n^\top\boldsymbol{\beta}))\mathbf{x}_n - (y'_n - \sigma(\mathbf{x}_n'^\top\boldsymbol{\beta}))\mathbf{x}'_n\| \\ &\leq \|(y_n - \sigma(\mathbf{x}_n^\top\boldsymbol{\beta}))\mathbf{x}_n\| + \|(y'_n - \sigma(\mathbf{x}_n'^\top\boldsymbol{\beta}))\mathbf{x}'_n\| \\ &\leq \underbrace{|y_n - \sigma(\mathbf{x}_n^\top\boldsymbol{\beta})|}_{\leq 1}\cdot\|\mathbf{x}_n\| + \underbrace{|y'_n - \sigma(\mathbf{x}_n'^\top\boldsymbol{\beta})|}_{\leq 1}\cdot\|\mathbf{x}'_n\| \leq 2R. \end{aligned} \tag{A.80}$$

Now we have proved that $c_1(\boldsymbol{\beta})$ is Lipschitz continuous with constant $L = R$ or $L = 2R$. Then, we similarly prove the log-concavity of the posterior $\pi(\cdot|\mathbf{Y})$.

$$\log\pi(\boldsymbol{\beta}|\mathbf{Y}) = \sum_{i=1}^n \big(y_i\,\mathbf{x}_i^\top\boldsymbol{\beta} - \log(1 + e^{\mathbf{x}_i^\top\boldsymbol{\beta}})\big) - \frac{\|\boldsymbol{\beta}\|_2^2}{2\tau^2} + \text{constant}$$
$$\Rightarrow -\nabla^2_{\boldsymbol{\beta}}\log\pi(\boldsymbol{\beta}|\mathbf{Y}) = \sum_{i=1}^n \sigma(\mathbf{x}_i^\top\boldsymbol{\beta})(1 - \sigma(\mathbf{x}_i^\top\boldsymbol{\beta}))\mathbf{x}_i\mathbf{x}_i^\top + \frac{1}{\tau^2}\mathbf{I}_d \succeq \frac{1}{\tau^2}\mathbf{I}_d \text{ for } \forall\boldsymbol{\beta}\,. \tag{A.81}$$

So $\pi(\boldsymbol{\beta}|\mathbf{Y})$ is m-strongly log-concave with $m = \tau^{-2}$. Therefore,

$$\sigma^2_{\mathbf{Y},\mathbf{Y}'} \leq \frac{L^2}{m} \leq \begin{cases} R^2\tau^2 & \text{if } \mathbf{x}_n = \mathbf{x}'_n \\ (2R)^2\tau^2 = 4R^2\tau^2 & \text{if } \mathbf{x}_n \neq \mathbf{x}'_n \end{cases}. \tag{A.82}$$

For the reversed pair $(\mathbf{Y}', \mathbf{Y})$, the Lipschitz constants R (if $\mathbf{x}_n = \mathbf{x}'_n$) and $2R$ (if $\mathbf{x}_n \neq \mathbf{x}'_n$) are symmetric in the two records, and $\pi(\boldsymbol{\beta}|\mathbf{Y}')$ is also τ^{-2}-strongly log-concave, so the same bounds hold for $\sigma^2_{\mathbf{Y}',\mathbf{Y}}$. Plug into Corollary 8

$$\max\{D_\alpha(\pi(\boldsymbol{\beta}|\mathbf{Y})\|\pi(\boldsymbol{\beta}|\mathbf{Y}')),\, D_\alpha(\pi(\boldsymbol{\beta}|\mathbf{Y}')\|\pi(\boldsymbol{\beta}|\mathbf{Y}))\} \leq \frac{\alpha\bar{\sigma}^2}{2} \leq \begin{cases} \frac{1}{2}\alpha R^2\tau^2 & \text{if } \mathbf{x}_n = \mathbf{x}'_n \\ 2\alpha R^2\tau^2 & \text{if } \mathbf{x}_n \neq \mathbf{x}'_n \end{cases}. \tag{A.83}$$

where $\bar{\sigma}^2$ is taken over the corresponding class of neighboring pairs.

For comparison, Theorem 16 in [Geumlek et al., 2017] gives $(\alpha, 2\alpha R^2\tau^2)$-RDP for the replacement case; our bound improves to $\alpha R^2\tau^2/2$ when $\mathbf{x}_n = \mathbf{x}'_n$. □

A.8.3.2 unbounded neighboring

Proof. Now we have $\psi(t) = \log(1 + e^t)$, and $h(\cdot) = 0$. Similarly, we prove $c_1(\boldsymbol{\beta})$ is Lipschitz first as what we did in bounded neighboring. Recall $\|\mathbf{x}\| \leq R$ and let $\sigma(t) = \frac{1}{1+e^{-t}}$, given $y \in \{0, 1\}, |\sigma(t)| < 1$ for $\forall t \in \mathbb{R}$ and $\frac{d\log(1+e^t)}{dt} = \sigma(t)$, we have

$$\begin{aligned} &\nabla_{\boldsymbol{\beta}} c_1(\boldsymbol{\beta}) = y_n\mathbf{x}_n - \sigma(\mathbf{x}_n^\top\boldsymbol{\beta})\mathbf{x}_n \quad \text{per Eq. (A.74)} \\ \Rightarrow\ &\|\nabla_{\boldsymbol{\beta}} c_1(\boldsymbol{\beta})\| = \|(y_n - \sigma(\mathbf{x}_n^\top\boldsymbol{\beta}))\mathbf{x}_n\| \leq \underbrace{|y_n - \sigma(\mathbf{x}_n^\top\boldsymbol{\beta})|}_{\leq 1}\|\mathbf{x}_n\| \leq R \end{aligned} \tag{A.84}$$

Now we have proved that $c_1(\boldsymbol{\beta})$ is Lipschitz continuous with constant $L = R$. Given $\pi(\boldsymbol{\beta}|\mathbf{Y})$ is m-strongly log-concave with $m = \tau^{-2}$ and the Lemmas, $c_1(\boldsymbol{\beta})$ is sub-Gaussian under $\pi(\cdot|\mathbf{Y})$ with

$$\sigma^2_{\mathbf{Y},\mathbf{Y}'} \leq \frac{L^2}{m} \leq R^2\tau^2 \tag{A.85}$$

For the reversed pair $(\mathbf{Y}', \mathbf{Y})$, $\|\nabla_{\boldsymbol{\beta}}\tilde{c}_1(\boldsymbol{\beta})\| = |y_n - \sigma(\mathbf{x}_n^\top\boldsymbol{\beta})|\,\|\mathbf{x}_n\| \leq R$, and $\pi(\boldsymbol{\beta}|\mathbf{Y}')$ is τ^{-2}-strongly log-concave by the argument for general GLMs above. Hence $\sigma^2_{\mathbf{Y}',\mathbf{Y}} \leq R^2\tau^2$ and $\bar{\sigma}^2 \leq R^2\tau^2$. Plug into Corollary 8

$$\max\{D_\alpha(\pi(\boldsymbol{\beta}|\mathbf{Y})\|\pi(\boldsymbol{\beta}|\mathbf{Y}')),\, D_\alpha(\pi(\boldsymbol{\beta}|\mathbf{Y}')\|\pi(\boldsymbol{\beta}|\mathbf{Y}))\} \leq \frac{\alpha\bar{\sigma}^2}{2} \leq \frac{\alpha R^2\tau^2}{2} \tag{A.86}$$

□

A.9 Proof of Proposition 11 for Gaussian-Gaussian conjugate pair

Proof. Assume observed dataset $Y = \{y_i\}_{i=1}^n$, where $y_i|\theta \overset{\text{i.i.d}}{\sim} \mathcal{N}(\theta, \sigma^2)$ and σ^2 is known. If imposing a Gaussian prior on θ, i.e. $\pi(\theta) = \mathcal{N}(\mu_0, \tau^2)$, then per conjugacy, the posterior $\theta|Y \sim \mathcal{N}(\mu_Y, v_Y^2)$, where

$$v_Y^2 = \left(\frac{n}{\sigma^2} + \frac{1}{\tau^2}\right)^{-1}, \qquad \mu_Y = v_Y^2\left(\frac{1}{\sigma^2}\sum_{i=1}^n y_i + \frac{\mu_0}{\tau^2}\right). \tag{A.87}$$

A.9.1 bounded neighboring

WLOG, suppose neighboring datasets Y, Y', both of size n, differ in their n-th record y_n and y_n'. Then the log-likelihood ratio $c_1(\theta)$ can be derived as

$$\ell(\theta; Y) = (2\pi\sigma^2)^{-n/2}\exp\left(-\frac{(y_n-\theta)^2}{2\sigma^2} - \sum_{i=1}^{n-1}\frac{(y_i-\theta)^2}{2\sigma^2}\right); \tag{A.88}$$

$$\ell(\theta; Y') = (2\pi\sigma^2)^{-n/2}\exp\left(-\frac{(y_n'-\theta)^2}{2\sigma^2} - \sum_{i=1}^{n-1}\frac{(y_i-\theta)^2}{2\sigma^2}\right) \tag{A.89}$$

$$\begin{aligned}\Rightarrow\ c_1(\theta) &= \log\frac{\ell(\theta;Y)}{\ell(\theta;Y')} = \frac{(y_n'-\theta)^2 - (y_n-\theta)^2}{2\sigma^2} \\ &= \frac{(y_n'^2 - y_n^2) - 2\theta(y_n' - y_n)}{2\sigma^2} \triangleq K + \frac{\Delta y}{\sigma^2}\theta,\end{aligned} \tag{A.90}$$

where $K = \frac{(y_n'^2 - y_n^2)}{2\sigma^2}$ and $\Delta y = y_n - y_n'$. Per the affine invariance of Gaussian distributions

$$c_1(\theta) \sim \mathcal{N}\left(K + \frac{\Delta y}{\sigma^2}\mu_Y, \left(\frac{\Delta y}{\sigma^2}\right)^2 v_Y^2\right). \tag{A.91}$$

Since the moment generating function (MGF) for $\mathcal{N}(\mu, \sigma^2)$ is $\exp(t\mu + \sigma^2t^2/2)$,

$$\begin{aligned}&\log\mathbb{E}_{\theta\sim\pi(\theta|Y)}\left(e^{t(c_1(\theta) - \mathbb{E}_{\theta\sim\pi(\theta|Y)}[c_1(\theta)])}\right) \\ =\ &\frac{t^2\left(\Delta y/\sigma^2\right)^2 v_Y^2}{2} = \frac{t^2}{2}\left(\frac{(y_n - y_n')^2}{\sigma^4}\left(\frac{n}{\sigma^2} + \frac{1}{\tau^2}\right)^{-1}\right) \triangleq \frac{t^2\sigma^2_{Y,Y'}}{2}.\end{aligned} \tag{A.92}$$

Then we conclude that $c_1(\theta)$ is sub-Gaussian with $\sigma^2_{Y,Y'} = \frac{(y_n - y_n')^2}{\sigma^4}\left(\frac{n}{\sigma^2} + \frac{1}{\tau^2}\right)^{-1}$.

To obtain $\bar{\sigma}^2 = \sup_{(Y,Y'):d(Y,Y')=1}\sigma^2_{Y,Y'}$, we need to assume a bounded data domain, i.e. $\exists\, c \in \mathbb{R}, k > 0$ such that $y \in [c - k\sigma,\, c + k\sigma]$ for every record, then $|y_n - y_n'| \leq 2k\sigma$, then

$$\sigma^2_{Y,Y'} \leq \frac{(2k\sigma)^2}{\sigma^4}\left(\frac{n}{\sigma^2} + \frac{1}{\tau^2}\right)^{-1} = \frac{4k^2}{n + \frac{\sigma^2}{\tau^2}} = \bar{\sigma}^2. \tag{A.93}$$

For the reversed pair (Y', Y), $\widetilde{c}_1(\theta) = -c_1(\theta)$ is Gaussian under $\pi(\theta|Y')$ with variance $(y_n - y'_n)^2 v_{Y'}^2/\sigma^4$; since Y' also has n records, $v_{Y'}^2 = v_Y^2$, so the same bound holds for $\sigma_{Y',Y}^2$. Plug Eq. (A.93) into Corollary 8,

$$\max\{D_\alpha(\pi(\theta|Y)\|\pi(\theta|Y')), D_\alpha(\pi(\theta|Y')\|\pi(\theta|Y))\} \le \frac{\alpha\bar{\sigma}^2}{2} \le \frac{2k^2\alpha}{n + \frac{\sigma^2}{\tau^2}}. \tag{A.94}$$

A.9.2 unbounded neighboring

Forward direction Let Y has size n and Y' is of size $n-1$ by removing the n-th record y_n in Y. Then

$$c_1(\theta) = \log \frac{\ell(\theta; Y)}{\ell(\theta; Y')} = \frac{-(y_n - \theta)^2}{2\sigma^2} - \frac{1}{2}\log(2\pi\sigma^2) \triangleq \frac{-Z^2}{2\sigma^2} - \frac{1}{2}\log(2\pi\sigma^2). \tag{A.95}$$

where $Z = \theta - y_n$. Under the posterior $\theta|Y \sim \mathcal{N}(\mu_Y, v_Y^2)$, then $Z|Y \sim \mathcal{N}(\mu_Y - y_n, v_Y^2)$, and Z^2/v_Y^2 follows a noncentral chi-square distribution with degree of freedom 1, that is

$$\frac{Z^2}{v_Y^2} \sim \chi_k^2(\lambda), \quad \text{with } k = 1, \lambda = \frac{(\mu_Y - y_n)^2}{v_Y^2}. \tag{A.96}$$

We note that $c_1(\theta)$ in Eq. (A.95) doesn't fall into the sub-Gaussian regime because $c_1(\theta)$ is a negative quadratic function of a Gaussian random variable, so its exponential moment $\mathbb{E}_P[e^{tc_1(\theta)}]$ is finite only for $|t|$ in a bounded neighborhood of zero (due to the noncentral χ^2 tail of Z^2), rather than for all $t \in \mathbb{R}$ as required by the sub-Gaussian condition. So there is no finite $\sigma_{Y,Y'}^2$.

But we can still leverage properties of χ^2 distribution to explicitly calculate each term in Eq. (7). First, the expectation of Z^2 is

$$\mathbb{E}_{\theta\sim\pi(\theta|Y)}\left[Z^2\right] = v_Y^2(k+\lambda) = v_Y^2\left(1 + \frac{(\mu_Y - y_n)^2}{v_Y^2}\right) = v_Y^2 + (\mu_Y - y_n)^2. \tag{A.97}$$

Then, recall the moment generating function of $\chi_k^2(\lambda)$ is $e^{\lambda t/(1-2t)}(1-2t)^{-k/2}$ for $t < 1/2$, hence

$$\begin{aligned}
&\mathbb{E}_{\theta\sim\pi(\theta|Y)}\left[e^{tZ^2/v_Y^2}\right] = \exp\left(\frac{(\mu_Y - y_n)^2}{v_Y^2}\cdot\frac{t}{1-2t}\right)(1-2t)^{-1/2} \\
\Rightarrow\ &\mathbb{E}_{\theta\sim\pi(\theta|Y)}\left[e^{tZ^2}\right] = \exp\left(\frac{(\mu_Y - y_n)^2}{v_Y^2}\cdot\frac{v_Y^2 t}{1-2v_Y^2 t}\right)\left(1-2tv_Y^2\right)^{-1/2} \\
&\qquad = (1-2v_Y^2 t)^{-1/2}\exp\left(\frac{(\mu_Y - y_n)^2 t}{1-2v_Y^2 t}\right), \qquad t < \frac{1}{2v_Y^2}.
\end{aligned} \tag{A.98}$$

Now we have two approaches to explicitly upper bound $D_\alpha(\pi(\theta|Y)\|\pi(\theta|Y'))$.

1) Derive sub-exponential parameters b_Y, ν_Y for $c_1(\theta)$ and apply Corollary 13. Since the derivation of b_Y, ν_Y is not tight, the resulting RDP bound is generally not tight;

2) Directly calculate each term in Eq. (A.6) using the closed-form moment generating function in Eq. (A.98). This approach yields an exact expression.

Method 1: sub-exponential parameters.[2] Our target is to identify b_Y, ν_Y^2 such that

$$\log \mathbb{E}_{\theta\sim P}\left(e^{t(c_1(\theta) - \mathbb{E}_P[c_1(\theta)])}\right) \le \frac{t^2\nu_Y^2}{2(1-b_Y|t|)} \tag{A.99}$$

for $|t| < 1/b_Y$, with $c_1(\theta)$ defined in Eq. (A.95).

$$\begin{aligned}
c_1(\theta) - \mathbb{E}_P\left[c_1(\theta)\right] &= \frac{-Z^2}{2\sigma^2} - \frac{1}{2}\log(2\pi\sigma^2) - \mathbb{E}_P\left(\frac{-Z^2}{2\sigma^2} - \frac{1}{2}\log(2\pi\sigma^2)\right) \\
&= -\frac{1}{2\sigma^2}\left(Z^2 - \mathbb{E}_P(Z^2)\right)
\end{aligned} \tag{A.100}$$

[2]For brevity, we write b_Y, ν_Y for $b_{Y,Y'}, \nu_{Y,Y'}$, and only the forward direction is illustrated here.

So we only need to examine $(Z^2 - \mathbb{E}_P(Z^2))$:

$$\begin{aligned}
&\log \mathbb{E}_P\left[e^{t(Z^2-\mathbb{E}_P(Z^2))}\right] \\
=\ &\log \mathbb{E}_P\left[e^{t(Z^2-v_Y^2-(\mu_Y-y_n)^2)}\right] \qquad \text{per Eq. (A.97)} \\
=\ &\log \mathbb{E}_P\left[e^{tZ^2}\right] - \left(v_Y^2 + (\mu_Y - y_n)^2\right)t \\
=\ &-\frac{1}{2}\log(1-2v_Y^2 t) - v_Y^2 t + (\mu_Y - y_n)^2 t\left(\frac{1}{1-2v_Y^2 t} - 1\right) \qquad \text{per Eq. (A.98)} \\
=\ &\frac{1}{2}\left[-\log(1-x) - x\right] + (\mu_Y - y_n)^2 t\left(\frac{1}{1-x} - 1\right) \qquad \text{let } x = 2v_Y^2 t. \qquad \text{(A.101)}
\end{aligned}$$

Since Eq. (A.98) requires $t < 1/(2v_Y^2)$, $x = 2v_Y^2 t < 1$. To ensure Eq. (A.101)$\leq \frac{t^2 \nu_Y'^2}{2(1-b_Y'|t|)}$ for some b_Y', ν_Y' when $|t| < 1/b_Y'$, we need some inequalities below for $|x| < 1$

$$-\log(1-x) - x \leq \frac{x^2}{2(1-|x|)}, \qquad \frac{1}{1-x} - 1 = \frac{x}{1-x} \leq \frac{|x|}{1-|x|}. \qquad \text{(A.102)}$$

Plug the inequalities above into Eq. (A.101):

$$\begin{aligned}
\text{Eq. (A.101)} &\leq \frac{1}{2}\cdot\frac{x^2}{2(1-|x|)} + (\mu_Y - y_n)^2|t|\cdot\frac{|x|}{1-|x|} \\
&= \frac{1}{2}\cdot\frac{4v_Y^4 t^2}{2(1-2v_Y^2|t|)} + (\mu_Y - y_n)^2|t|\cdot\frac{2v_Y^2|t|}{1-2v_Y^2|t|} \\
&= \frac{v_Y^4 t^2}{1-2v_Y^2|t|} + \frac{2v_Y^2(\mu_Y - y_n)^2 t^2}{1-2v_Y^2|t|} \\
&= \frac{2\left(v_Y^2 + 2(\mu_Y - y_n)^2\right)v_Y^2 t^2}{2(1-2v_Y^2|t|)} \triangleq \frac{t^2\nu_Y'^2}{2(1-b_Y'|t|)}. \qquad \text{(A.103)}
\end{aligned}$$

Therefore $b_Y' = 2v_Y^2$ and $\nu_Y'^2 = 2\left(v_Y^2 + 2(\mu_Y - y_n)^2\right)v_Y^2$.

Recall that $b_Y', \nu_Y'^2$ are for $(Z^2 - \mathbb{E}_P(Z^2))$, while $c_1(\theta) - \mathbb{E}_P\left[c_1(\theta)\right] = -\frac{1}{2\sigma^2}\left(Z^2 - \mathbb{E}_P(Z^2)\right)$, hence $b_Y = b_Y'/2\sigma^2$ and $\nu_Y^2 = \nu_Y'^2/4\sigma^4$. Plug in Corollary 13:

$$\begin{aligned}
D_\alpha(P||Q) &\leq \frac{\nu_Y^2}{2}\frac{\alpha(1-2b_Y) + 2b_Y}{(1-b_Y)\big(1-b_Y(\alpha-1)\big)} \\
&= \frac{v_Y^2\left(v_Y^2 + 2(\mu_Y - y_n)^2\right)}{4\sigma^4}\frac{\alpha\Big(1-\frac{2v_Y^2}{\sigma^2}\Big) + \frac{2v_Y^2}{\sigma^2}}{\Big(1-\frac{v_Y^2}{\sigma^2}\Big)\Big(1-\frac{v_Y^2}{\sigma^2}(\alpha-1)\Big)} \\
&= \frac{v_Y^2\left(v_Y^2 + 2(\mu_Y - y_n)^2\right)}{4\sigma^2}\cdot\frac{\alpha\sigma^2 - 2v_Y^2(\alpha-1)}{(\sigma^2 - v_Y^2)\left(\sigma^2 - v_Y^2(\alpha-1)\right)}, \qquad \text{(A.104)}
\end{aligned}$$

where v_Y^2 and μ_Y are defined in Eq. (A.87).

Method 2: exact moment calculation For the first term in Eq. (A.6), per the $c_1(\theta)$ defined in Eq. (A.95), and the moment generating function of Z^2 derived in Eq. (A.98)

$$\begin{aligned}
-c_1(\theta) &= \frac{Z^2}{2\sigma^2} + \frac{1}{2}\log(2\pi\sigma^2) \\
\mathbb{E}_P[e^{-c_1(\theta)}] &= \mathbb{E}_P\left[\exp\left(\frac{Z^2}{2\sigma^2} + \frac{1}{2}\log(2\pi\sigma^2)\right)\right] \\
&= (2\pi\sigma^2)^{\frac{1}{2}}\mathbb{E}_P\left[\exp\left(\frac{Z^2}{2\sigma^2}\right)\right] \\
&= (2\pi\sigma^2)^{\frac{1}{2}}\left(1 - 2v_Y^2\frac{1}{2\sigma^2}\right)^{-\frac{1}{2}}\exp\left(\frac{(\mu_Y - y_n)^2\frac{1}{2\sigma^2}}{1-2v_Y^2\frac{1}{2\sigma^2}}\right) \text{ plug } t = \frac{1}{2\sigma^2} \text{ in Eq. (A.98)}
\end{aligned}$$

$$= (2\pi\sigma^2)^{\frac{1}{2}} \left(1 - \frac{v_Y^2}{\sigma^2}\right)^{-\frac{1}{2}} \exp\left(\frac{(\mu_Y - y_n)^2}{2(\sigma^2 - v_Y^2)}\right). \tag{A.105}$$

This requires $t = \frac{1}{2\sigma^2} < \frac{1}{2v_Y^2}$, equivalently $v_Y^2 < \sigma^2$, recall $v_Y^2 = \left(\frac{n}{\sigma^2} + \frac{1}{\tau^2}\right)^{-1}$, this holds for $n \geq 1$ in the Gaussian-Gaussian conjugate model. Then per Lemma 2

$$c_2 = \log \mathbb{E}_P[e^{-c_1(\theta)}] = \frac{1}{2}\log(2\pi\sigma^2) - \frac{1}{2}\log\left(1 - \frac{v_Y^2}{\sigma^2}\right) + \frac{(\mu_Y - y_n)^2}{2(\sigma^2 - v_Y^2)}. \tag{A.106}$$

Next, for the second term in Eq. (A.6), we also use the moment generating function for Z^2 in Eq. (A.98) to exactly calculate the $(\alpha - 1)$ moment for $c_1(\theta)$:

$$\begin{aligned}
(\alpha-1)c_1(\theta) &= -\frac{\alpha-1}{2\sigma^2}Z^2 - \frac{\alpha-1}{2}\log(2\pi\sigma^2), \\
\mathbb{E}_P[e^{(\alpha-1)c_1(\theta)}] &= \mathbb{E}_P\left[\exp\left(-\frac{\alpha-1}{2\sigma^2}Z^2 - \frac{\alpha-1}{2}\log(2\pi\sigma^2)\right)\right] \\
&= (2\pi\sigma^2)^{-\frac{\alpha-1}{2}}\mathbb{E}_P\left[\exp\left(-\frac{\alpha-1}{2\sigma^2}Z^2\right)\right] \text{ plug } t = \frac{-(\alpha-1)}{2\sigma^2} \text{ in Eq. (A.98)} \\
&= (2\pi\sigma^2)^{-\frac{\alpha-1}{2}}\left(1 - 2v_Y^2\left(\frac{1-\alpha}{2\sigma^2}\right)\right)^{-\frac{1}{2}} \exp\left(\frac{(\mu_Y - y_n)^2\left(\frac{1-\alpha}{2\sigma^2}\right)}{1 - 2v_Y^2\left(\frac{1-\alpha}{2\sigma^2}\right)}\right) \\
&= (2\pi\sigma^2)^{-\frac{\alpha-1}{2}}\left(1 + v_Y^2\left(\frac{\alpha-1}{\sigma^2}\right)\right)^{-\frac{1}{2}} \exp\left(\frac{(\mu_Y - y_n)^2(1-\alpha)}{2(\sigma^2 + (\alpha-1)v_Y^2)}\right).
\end{aligned} \tag{A.107}$$

This requires $t = \frac{-(\alpha-1)}{2\sigma^2} < \frac{1}{2v_Y^2}$. Since $\alpha > 1$, $\frac{-(\alpha-1)}{2\sigma^2} < 0$, so this requirement always holds. Next, the 2nd term in Eq. (A.6) can be calculated as

$$\begin{aligned}
&\frac{1}{\alpha-1}\log\mathbb{E}_P\left[e^{(\alpha-1)c_1(\theta)}\right] \\
=& -\frac{1}{2}\log(2\pi\sigma^2) - \frac{1}{2(\alpha-1)}\log\left(1 + v_Y^2\left(\frac{\alpha-1}{\sigma^2}\right)\right) - \frac{(\mu_Y - y_n)^2}{2(\sigma^2 + (\alpha-1)v_Y^2)}.
\end{aligned} \tag{A.108}$$

Plug c_2 in Eq. (A.106) and $\frac{1}{\alpha-1}\log\mathbb{E}_P\left[e^{(\alpha-1)\cdot c_1(\theta)}\right]$ in Eq. (A.108) into Eq. (A.6), we have

$$\begin{aligned}
&D_\alpha(\pi(\theta|Y)||\pi(\theta|Y')) \\
=& c_2 + \frac{1}{\alpha-1}\log\mathbb{E}_P\left[e^{(\alpha-1)\cdot c_1(\theta)}\right] \\
=& \cancel{\frac{1}{2}\log(2\pi\sigma^2)} - \frac{1}{2}\log\left(1 - \frac{v_Y^2}{\sigma^2}\right) + \frac{(\mu_Y - y_n)^2}{2(\sigma^2 - v_Y^2)} \\
&\cancel{-\frac{1}{2}\log(2\pi\sigma^2)} - \frac{1}{2(\alpha-1)}\log\left(1 + v_Y^2\left(\frac{\alpha-1}{\sigma^2}\right)\right) - \frac{(\mu_Y - y_n)^2}{2(\sigma^2 + (\alpha-1)v_Y^2)} \\
=& -\frac{1}{2}\log\left(1 - \frac{v_Y^2}{\sigma^2}\right) - \frac{1}{2(\alpha-1)}\log\left(1 + v_Y^2\left(\frac{\alpha-1}{\sigma^2}\right)\right) + \frac{(\mu_Y - y_n)^2}{2(\sigma^2 - v_Y^2)} - \frac{(\mu_Y - y_n)^2}{2(\sigma^2 + (\alpha-1)v_Y^2)}.
\end{aligned}$$

Recall

$$v_Y^2 = \left(\frac{n}{\sigma^2} + \frac{1}{\tau^2}\right)^{-1}, \qquad \mu_Y = v_Y^2\left(\frac{1}{\sigma^2}\sum_{i=1}^n y_i + \frac{\mu_0}{\tau^2}\right). \tag{A.109}$$

Then define $\omega = n + \frac{\sigma^2}{\tau^2}$ and assume $y_n \in [\mu_Y - k^*\sigma, \mu_Y + k^*\sigma]$.

$$\begin{aligned}
&D_\alpha(\pi(\theta|Y)||\pi(\theta|Y')) \\
=& -\frac{1}{2}\log\left(1 - \frac{v_Y^2}{\sigma^2}\right) - \frac{1}{2(\alpha-1)}\log\left(1 + v_Y^2\left(\frac{\alpha-1}{\sigma^2}\right)\right) + \frac{(\mu_Y - y_n)^2}{2(\sigma^2 - v_Y^2)} - \frac{(\mu_Y - y_n)^2}{2(\sigma^2 + (\alpha-1)v_Y^2)} \\
=& -\frac{1}{2}\log\left(1 - \omega^{-1}\right) - \frac{1}{2(\alpha-1)}\log\left(1 + \frac{\alpha-1}{\omega}\right) + \frac{(\mu_Y - y_n)^2}{2\sigma^2}\left[\frac{1}{1-\omega^{-1}} - \frac{1}{1 + \frac{\alpha-1}{\omega}}\right]
\end{aligned}$$

$$\triangleq A(\alpha,\omega)+\frac{(\mu_Y-y_n)^2}{2\sigma^2}B(\alpha,\omega)$$
$$\le A(\alpha,\omega)+\frac{k^{*2}}{2}B(\alpha,\omega) \tag{A.110}$$

Since $\mu_Y=\big(\frac{\sigma^2}{\tau^2}\mu_0+\sum_{i=1}^n y_i\big)/\omega$ is a convex combination of μ_0 and $\{y_i\}$, if all records and μ_0 lie in $[c-M,c+M]$, then $|\mu_Y-y_n|\le 2M$, i.e., $k^*=2M/\sigma$ is a data-independent choice.

Claim A.8. *For all $\alpha>1$ and $\omega>1$, we have $A(\alpha,\omega)>0$ and $B(\alpha,\omega)>0$. Furthermore, holding α fixed, both $A(\alpha,\omega)$ and $B(\alpha,\omega)$ decrease monotonically as ω increases.*

Proof. Assume $n\ge 1$, then $\omega=n+\sigma^2/\tau^2>1$. Let $x=\omega^{-1}$ and $t=\alpha-1$, then $x\in(0,1)$ and $t>0$, and

$$A(\alpha,\omega)=-\frac{1}{2}\log(1-x)-\frac{1}{2t}\log(1+xt)\triangleq A(x). \tag{A.111}$$

Thus to prove $A(\alpha,\omega)>0$ is equivalent to prove

$$-\log(1-x)>\frac{1}{t}\log(1+xt). \tag{A.112}$$

Since $-\log(1-x)>x$ for $x\in(0,1)$, and $\log(1+xt)<xt\Rightarrow t^{-1}\log(1+xt)<x$,

$$-\log(1-x)>x>\frac{1}{t}\log(1+xt), \tag{A.113}$$

which completes the proof of strictly positivity of $A(\alpha,\omega)$. Then, take derivative of x on $A(x)$:

$$A'(x)=\frac{1}{2(1-x)}-\frac{1}{2(1+xt)}>0, \tag{A.114}$$

per $x\in(0,1)$ and $t>0$. Thus,

$$\frac{\partial A}{\partial\omega}=A'(x)\frac{dx}{d\omega}=-\omega^{-2}A'(x)<0. \tag{A.115}$$

Furthermore, take derivative of x on $B(x)=\frac{1}{1-x}-\frac{1}{1+xt}$

$$B'(x)=\frac{1}{(1-x)^2}+\frac{t}{(1+xt)^2}>0\quad\Rightarrow\frac{\partial B}{\partial\omega}<0 \tag{A.116}$$

thus, both $A(\alpha,\omega)$ and $B(\alpha,\omega)$ decrease monotonically as ω increases. □

Reverse direction We next analyze the reverse Rényi divergence $D_\alpha(\pi(\theta|Y')\|\pi(\theta|Y))$, where Y has size n and Y' is of size $n-1$, obtained by removing the n-th record y_n from Y. Define

$$\widetilde{c}_1(\theta)\triangleq\log\frac{\ell(\theta;Y')}{\ell(\theta;Y)}=-\log p(y_n|\theta)=\frac{(y_n-\theta)^2}{2\sigma^2}+\frac{1}{2}\log(2\pi\sigma^2)\triangleq\frac{Z'^2}{2\sigma^2}+\frac{1}{2}\log(2\pi\sigma^2)$$
$$=-c_1(\theta), \tag{A.117}$$

where $Z'=\theta-y_n$. Under the posterior $\theta|Y'\sim\mathcal{N}(\mu_{Y'},v_{Y'}^2)$, $Z'|Y'\sim\mathcal{N}(\mu_{Y'}-y_n,v_{Y'}^2)$, and $Z'^2/v_{Y'}^2$ follows a noncentral chi-square distribution with one degree of freedom:

$$\frac{Z'^2}{v_{Y'}^2}\sim\chi_1^2(\lambda'),\qquad\lambda'=\frac{(\mu_{Y'}-y_n)^2}{v_{Y'}^2}. \tag{A.118}$$

Recall that the MGF of $\chi_k^2(\lambda)$ is $\exp\big(\frac{\lambda t}{1-2t}\big)(1-2t)^{-k/2}$ for $t<\frac{1}{2}$. Therefore,

$$\mathbb{E}_{\theta\sim\pi(\theta|Y')}\big[e^{tZ'^2}\big]=\big(1-2v_{Y'}^2t\big)^{-1/2}\exp\left(\frac{(\mu_{Y'}-y_n)^2t}{1-2v_{Y'}^2t}\right),\qquad t<\frac{1}{2v_{Y'}^2}. \tag{A.119}$$

We now directly calculate the two terms in Eq. (A.6) using the closed-form MGF in Eq. (A.119). For the first term, per Eq. (A.117), $-\widetilde{c}_1(\theta)=-\frac{Z'^2}{2\sigma^2}-\frac{1}{2}\log(2\pi\sigma^2)$, and hence

$$\mathbb{E}_{\theta\sim\pi(\theta|Y')}\big[e^{-\widetilde{c}_1(\theta)}\big]=(2\pi\sigma^2)^{-1/2}\,\mathbb{E}_{\theta\sim\pi(\theta|Y')}\Big[\exp\Big(-\frac{Z'^2}{2\sigma^2}\Big)\Big]$$

$$= (2\pi\sigma^2)^{-1/2}\Big(1+\frac{v_{Y'}^2}{\sigma^2}\Big)^{-1/2}\exp\Big(-\frac{(\mu_{Y'}-y_n)^2}{2(\sigma^2+v_{Y'}^2)}\Big), \tag{A.120}$$

where we plug $t=-1/(2\sigma^2)$ into Eq. (A.119). Since $t<0$, the MGF condition always holds. Therefore, per Lemma 2,

$$\widetilde{c}_2 = \log \mathbb{E}_{\theta\sim\pi(\theta|Y')}\big[e^{-\widetilde{c}_1(\theta)}\big] = -\frac{1}{2}\log(2\pi\sigma^2) - \frac{1}{2}\log\Big(1+\frac{v_{Y'}^2}{\sigma^2}\Big) - \frac{(\mu_{Y'}-y_n)^2}{2(\sigma^2+v_{Y'}^2)}. \tag{A.121}$$

Next, for the second term in Eq. (A.6), $(\alpha-1)\widetilde{c}_1(\theta) = \frac{\alpha-1}{2\sigma^2}Z'^2 + \frac{\alpha-1}{2}\log(2\pi\sigma^2)$. Thus,

$$\begin{aligned}\mathbb{E}_{\theta\sim\pi(\theta|Y')}\big[e^{(\alpha-1)\widetilde{c}_1(\theta)}\big] &= (2\pi\sigma^2)^{(\alpha-1)/2}\,\mathbb{E}_{\theta\sim\pi(\theta|Y')}\Big[\exp\Big(\frac{\alpha-1}{2\sigma^2}Z'^2\Big)\Big]\\ &= (2\pi\sigma^2)^{(\alpha-1)/2}\Big(1-\frac{(\alpha-1)v_{Y'}^2}{\sigma^2}\Big)^{-1/2}\exp\Big(\frac{(\alpha-1)(\mu_{Y'}-y_n)^2}{2(\sigma^2-(\alpha-1)v_{Y'}^2)}\Big),\end{aligned} \tag{A.122}$$

where we plug $t=\frac{\alpha-1}{2\sigma^2}$ into Eq. (A.119). This requires

$$\frac{\alpha-1}{2\sigma^2} < \frac{1}{2v_{Y'}^2}, \quad \text{or equivalently} \quad (\alpha-1)v_{Y'}^2 < \sigma^2. \tag{A.123}$$

Consequently,

$$\begin{aligned}\frac{1}{\alpha-1}\log\mathbb{E}_{\theta\sim\pi(\theta|Y')}\big[e^{(\alpha-1)\widetilde{c}_1(\theta)}\big] =& \frac{1}{2}\log(2\pi\sigma^2) - \frac{1}{2(\alpha-1)}\log\Big(1-\frac{(\alpha-1)v_{Y'}^2}{\sigma^2}\Big)\\ &+\frac{(\mu_{Y'}-y_n)^2}{2(\sigma^2-(\alpha-1)v_{Y'}^2)}.\end{aligned} \tag{A.124}$$

Plugging Eqs. (A.121) and (A.124) into Eq. (A.6), the $\frac{1}{2}\log(2\pi\sigma^2)$ terms cancel, and we obtain

$$\begin{aligned}D_\alpha(\pi(\theta|Y')\|\pi(\theta|Y)) =& -\frac{1}{2}\log\Big(1+\frac{v_{Y'}^2}{\sigma^2}\Big) - \frac{1}{2(\alpha-1)}\log\Big(1-\frac{(\alpha-1)v_{Y'}^2}{\sigma^2}\Big)\\ &+\frac{(\mu_{Y'}-y_n)^2}{2(\sigma^2-(\alpha-1)v_{Y'}^2)} - \frac{(\mu_{Y'}-y_n)^2}{2(\sigma^2+v_{Y'}^2)}.\end{aligned} \tag{A.125}$$

Recall that $v_{Y'}^2 = \big(\frac{n-1}{\sigma^2}+\frac{1}{\tau^2}\big)^{-1} = \frac{\sigma^2}{\omega-1}$ with $\omega = n+\frac{\sigma^2}{\tau^2}$. Then the finiteness condition in Eq. (A.123) becomes

$$\frac{\alpha-1}{\omega-1} < 1 \quad\Longleftrightarrow\quad 1<\alpha<\omega. \tag{A.126}$$

For $1<\alpha<\omega$, Eq. (A.125) becomes

$$\begin{aligned}D_\alpha(\pi(\theta|Y')\|\pi(\theta|Y)) =& -\frac{1}{2}\log\Big(\frac{\omega}{\omega-1}\Big) - \frac{1}{2(\alpha-1)}\log\Big(\frac{\omega-\alpha}{\omega-1}\Big)\\ &+\frac{(\mu_{Y'}-y_n)^2}{2\sigma^2}\Big[\frac{1}{1-\frac{\alpha-1}{\omega-1}} - \frac{1}{1+\frac{1}{\omega-1}}\Big].\end{aligned} \tag{A.127}$$

The posterior means satisfy

$$\omega\mu_Y = (\omega-1)\mu_{Y'} + y_n \;\Rightarrow\; \mu_{Y'}-y_n = \frac{\omega}{\omega-1}(\mu_Y - y_n). \tag{A.128}$$

Therefore, for $1<\alpha<\omega$,

$$\begin{aligned}D_\alpha(\pi(\theta|Y')\|\pi(\theta|Y)) &= -\frac{1}{2}\log\Big(\frac{\omega}{\omega-1}\Big) - \frac{1}{2(\alpha-1)}\log\Big(\frac{\omega-\alpha}{\omega-1}\Big) + \frac{(\mu_Y-y_n)^2}{2\sigma^2}\frac{\alpha\omega}{(\omega-1)(\omega-\alpha)}\\ &\triangleq \widetilde{A}(\alpha,\omega) + \frac{(\mu_Y-y_n)^2}{2\sigma^2}\widetilde{B}(\alpha,\omega),\end{aligned} \tag{A.129}$$

where

$$\widetilde{A}(\alpha,\omega) \triangleq -\frac{1}{2}\log\Big(\frac{\omega}{\omega-1}\Big) - \frac{1}{2(\alpha-1)}\log\Big(\frac{\omega-\alpha}{\omega-1}\Big),$$

$$\widetilde{B}(\alpha,\omega) \triangleq \frac{\alpha\omega}{(\omega-1)(\omega-\alpha)}. \tag{A.130}$$

For comparison, the corresponding quantities for the forward direction in Eq. (A.110) are

$$A(\alpha,\omega) = -\frac{1}{2}\log\Big(\frac{\omega-1}{\omega}\Big) - \frac{1}{2(\alpha-1)}\log\Big(1+\frac{\alpha-1}{\omega}\Big),$$
$$B(\alpha,\omega) = \frac{\alpha\omega}{(\omega-1)(\omega+\alpha-1)}.$$

If $|\mu_Y - y_n| \le k^*\sigma$, then for every $1<\alpha<\omega$,

$$D_\alpha(\pi(\theta|Y')\|\pi(\theta|Y)) \le \widetilde{A}(\alpha,\omega) + \frac{k^{*2}}{2}\widetilde{B}(\alpha,\omega). \tag{A.131}$$

For $\alpha \ge \omega$, the MGF in Eq. (A.122) is not finite, and therefore

$$D_\alpha(\pi(\theta|Y')\|\pi(\theta|Y)) = \infty, \quad \alpha \ge \omega; \qquad \text{in particular, } D_\infty(\pi(\theta|Y')\|\pi(\theta|Y)) = \infty. \tag{A.132}$$

Claim A.9. *For all $1<\alpha<\omega$, $\widetilde{A}(\alpha,\omega) > 0$ and $\widetilde{B}(\alpha,\omega) > 0$. Holding α fixed, both decrease monotonically in ω on (α,∞), and $\widetilde{A}(\alpha,\omega) \to \infty$ as $\omega \downarrow \alpha$.*

Proof. Let $u = (\omega-1)^{-1}$ and $t = \alpha - 1 > 0$. The condition $1<\alpha<\omega$ is equivalent to $0<tu<1$, and $\omega/(\omega-1) = 1+u$, $(\omega-\alpha)/(\omega-1) = 1-tu$. Hence $\widetilde{A} = -\frac{1}{2}\log(1+u) - \frac{1}{2t}\log(1-tu)$ and $\widetilde{B} = \frac{\alpha u(1+u)}{1-tu}$. Since $-\log(1-y) > y$ for $y \in (0,1)$ and $\log(1+u) < u$ for $u>0$, we have $-\frac{1}{t}\log(1-tu) > u > \log(1+u)$, so $\widetilde{A} > 0$; $\widetilde{B} > 0$ is immediate. Further, $\frac{\partial \widetilde{A}}{\partial u} = \frac{1}{2}\big(\frac{1}{1-tu} - \frac{1}{1+u}\big) > 0$, and $\widetilde{B}$ is increasing in u because its numerator increases and its positive denominator decreases in u. Since u is decreasing in ω, both $\widetilde{A}$ and $\widetilde{B}$ decrease in ω. Finally, $tu \uparrow 1$ as $\omega \downarrow \alpha$, so $-\log(1-tu) \to \infty$. □

Two-sided guarantee. Combining Eqs. (A.110) and (A.131), if $|\mu_Y - y_n| \le k^*\sigma$, then for every $1<\alpha<\omega$,

$$\begin{aligned}&\max\{D_\alpha(\pi(\theta|Y)\|\pi(\theta|Y')), D_\alpha(\pi(\theta|Y')\|\pi(\theta|Y))\}\\ \le\ &\max\Big\{A(\alpha,\omega) + \frac{k^{*2}}{2}B(\alpha,\omega),\ \widetilde{A}(\alpha,\omega) + \frac{k^{*2}}{2}\widetilde{B}(\alpha,\omega)\Big\},\end{aligned} \tag{A.133}$$

so OPS satisfies RDP of this order only for $1<\alpha<\omega$, and does not satisfy pure DP under unbounded neighboring. Note that $\widetilde{B}(\alpha,\omega) > B(\alpha,\omega)$ since $\omega-\alpha < \omega+\alpha-1$. As $\omega = n + \sigma^2/\tau^2$ depends on n, a guarantee holding uniformly over all $n \ge 1$ is obtained by evaluating at $\omega_{\min} = 1 + \sigma^2/\tau^2$, and is finite only for $1<\alpha<1+\sigma^2/\tau^2$. □

A.10 Proof of Propositions 12 and 14 for Linear Regression

Proof. Let $\mathbf{x}_i \in \mathbb{R}^d$ be covariates and $y_i \in \mathbb{R}$ be response for the i-th individual. Given the dataset $\mathbf{Y} = \{(\mathbf{x}_i, y_i)\}_{i=1}^n$, let $\mathbf{x} \in \mathbb{R}^{n\times d}$ be the design matrix, and $\mathbf{y} \in \mathbb{R}^n$ be the response vector. The response follows a Gaussian distribution: $y_i|\mathbf{x}_i, \boldsymbol{\beta} \sim \mathcal{N}(\mathbf{x}_i^\top\boldsymbol{\beta}, \sigma^2)$. Impose a Gaussian prior on regression coefficient $\boldsymbol{\beta}$, i.e. $\pi(\boldsymbol{\beta}) = \mathcal{N}(0, \tau^2\mathbf{I}_d)$, then the posterior is $\boldsymbol{\beta}|\mathbf{Y} = \mathcal{N}(\boldsymbol{\mu}_\mathbf{Y}, \Sigma_\mathbf{Y})$, where $\Sigma_\mathbf{Y} = \Big(\frac{1}{\sigma^2}\mathbf{x}^\top\mathbf{x} + \frac{1}{\tau^2}\mathbf{I}_d\Big)^{-1}$, and $\boldsymbol{\mu}_\mathbf{Y} = \Sigma_\mathbf{Y}\big(\frac{1}{\sigma^2}\mathbf{x}^\top\mathbf{y}\big)$.

A.10.1 bounded neighboring

Let $\mathbf{Y}$ and $\mathbf{Y}'$ differ in exactly one record, i.e. $\mathbf{z}_n = (\mathbf{x}_n, y_n)$ in $\mathbf{Y}$ is replaced by $\mathbf{z}'_n = (\mathbf{x}'_n, y'_n)$ in $\mathbf{Y}'$, then

$$\begin{aligned}c_1(\boldsymbol{\beta}) = \log\frac{\ell(\boldsymbol{\beta};\mathbf{Y})}{\ell(\boldsymbol{\beta};\mathbf{Y}')} &= \log p(y_n|\mathbf{x}_n,\boldsymbol{\beta}) - \log p(y'_n|\mathbf{x}'_n,\boldsymbol{\beta})\\ &= -\frac{(y_n - \mathbf{x}_n^\top\boldsymbol{\beta})^2 - (y'_n - \mathbf{x}_n'^\top\boldsymbol{\beta})^2}{2\sigma^2}.\end{aligned} \tag{A.134}$$

Case 1: If $\mathbf{x}_n = \mathbf{x}'_n$, then

$$c_1(\boldsymbol{\beta}) = -\frac{(y_n^2 - y_n'^2) - 2(y_n - y'_n)\mathbf{x}_n^\top \boldsymbol{\beta}}{2\sigma^2}. \tag{A.135}$$

Since $\boldsymbol{\beta}|\mathbf{Y} = \mathcal{N}(\boldsymbol{\mu}_\mathbf{Y}, \Sigma_\mathbf{Y})$ and per the linear transformation of Gaussian random variable, we have $\mathbf{x}_n^\top \boldsymbol{\beta}|\mathbf{Y} \sim \mathcal{N}(\mathbf{x}_n^\top \boldsymbol{\mu}_\mathbf{Y}, \mathbf{x}_n^\top \Sigma_\mathbf{Y} \mathbf{x}_n)$. Therefore, $c_1(\boldsymbol{\beta})$ under $P = \pi(\cdot|\mathbf{Y})$ is Gaussian with variance

$$\sigma^2_{\mathbf{Y},\mathbf{Y}'} = \mathrm{Var}_{\boldsymbol{\beta}\sim\pi(\cdot|\mathbf{Y})}[c_1(\boldsymbol{\beta})] = \frac{(y_n - y'_n)^2}{\sigma^4}\, \mathbf{x}_n^\top \Sigma_\mathbf{Y} \mathbf{x}_n. \tag{A.136}$$

Since $\mathbf{x}^\top \mathbf{x}$ is positive semi-definite, $\Sigma_\mathbf{Y} = \left(\frac{1}{\sigma^2}\mathbf{x}^\top\mathbf{x} + \frac{1}{\tau^2}\mathbf{I}_d\right)^{-1} \preceq \left(\frac{1}{\tau^2}\mathbf{I}_d\right)^{-1} = \tau^2 \mathbf{I}_d$. If additionally assume $|y_n - y'_n| \le 2k\sigma$, then

$$\begin{aligned}\sigma^2_{\mathbf{Y},\mathbf{Y}'} &= \frac{(y_n - y'_n)^2}{\sigma^4}\, \mathbf{x}_n^\top \Sigma_\mathbf{Y} \mathbf{x}_n \\ &\le \frac{4k^2\sigma^2}{\sigma^4} \mathbf{x}_n^\top (\tau^2 \mathbf{I}_d) \mathbf{x}_n = \frac{4k^2}{\sigma^2}\tau^2 \|\mathbf{x}_n\|^2 \le \frac{4k^2\tau^2 R^2}{\sigma^2} \triangleq \bar{\sigma}^2.\end{aligned} \tag{A.137}$$

For the reversed pair $(\mathbf{Y}', \mathbf{Y})$, $\tilde{c}_1(\boldsymbol{\beta}) = -c_1(\boldsymbol{\beta})$ is Gaussian under $\pi(\boldsymbol{\beta}|\mathbf{Y}')$ with variance $(y_n - y'_n)^2 \mathbf{x}_n^\top \Sigma_{\mathbf{Y}'} \mathbf{x}_n / \sigma^4$; since $\Sigma_{\mathbf{Y}'} \preceq \tau^2 \mathbf{I}_d$, the same bound holds for $\sigma^2_{\mathbf{Y}',\mathbf{Y}}$. Per Corollary 8,

$$\max\{D_\alpha(P\|Q), D_\alpha(Q\|P)\} \le \frac{\alpha\bar{\sigma}^2}{2} \le \frac{2\alpha k^2 \tau^2 R^2}{\sigma^2}. \tag{A.138}$$

Case 2: When $\mathbf{x}_n \neq \mathbf{x}'_n$, the log-likelihood ratio decomposes as $c_1(\boldsymbol{\beta}) = L(\boldsymbol{\beta}) + Q(\boldsymbol{\beta})$, where

$$L(\boldsymbol{\beta}) = -\frac{(y_n^2 - y_n'^2) - 2(y_n \mathbf{x}_n - y'_n \mathbf{x}'_n)^\top \boldsymbol{\beta}}{2\sigma^2}, \tag{A.139}$$

$$Q(\boldsymbol{\beta}) = -\frac{\boldsymbol{\beta}^\top A \boldsymbol{\beta}}{2\sigma^2}, \quad A \triangleq \mathbf{x}_n \mathbf{x}_n^\top - \mathbf{x}'_n \mathbf{x}_n'^\top. \tag{A.140}$$

Since A may be indefinite, $Q(\boldsymbol{\beta})$ is a linear combination of non-central χ^2 random variables with potentially negative coefficients under $\boldsymbol{\beta}|\mathbf{Y} \sim \mathcal{N}(\boldsymbol{\mu}_\mathbf{Y}, \Sigma_\mathbf{Y})$, so $c_1(\boldsymbol{\beta})$ is generally not sub-Gaussian. We show that it satisfies a sub-exponential MGF bound under additional boundedness assumptions, by bounding the MGF of each part separately and combining via Cauchy-Schwarz. Let $\mu_L = \mathbb{E}[L(\boldsymbol{\beta})]$, $\mu_Q = \mathbb{E}[Q(\boldsymbol{\beta})]$, and $\mu^{c_1}_\mathbf{Y} = \mu_L + \mu_Q$[3].

Step 1: MGF of the linear part. Since $L(\boldsymbol{\beta})$ is affine in $\boldsymbol{\beta}$ and $\boldsymbol{\beta}|\mathbf{Y} \sim \mathcal{N}(\boldsymbol{\mu}_\mathbf{Y}, \Sigma_\mathbf{Y})$, $L(\boldsymbol{\beta})$ is Gaussian under the posterior. Using $\Sigma_\mathbf{Y} \preceq \tau^2 \mathbf{I}_d$, $|y_i| \le M_y$, and $\|\mathbf{x}_i\| \le R$:

$$\begin{aligned}\mathrm{Var}[L(\boldsymbol{\beta})] = \frac{(y_n \mathbf{x}_n - y'_n \mathbf{x}'_n)^\top \Sigma_\mathbf{Y} (y_n \mathbf{x}_n - y'_n \mathbf{x}'_n)}{\sigma^4} &\le \frac{\|y_n \mathbf{x}_n - y'_n \mathbf{x}'_n\|^2 \tau^2}{\sigma^4} \\ &\le \frac{4M_y^2 R^2 \tau^2}{\sigma^4} \triangleq \bar{\sigma}_L^2.\end{aligned} \tag{A.141}$$

Therefore $L(\boldsymbol{\beta})$ is sub-Gaussian with variance proxy $\bar{\sigma}_L^2$ and for all $t \in \mathbb{R}$:

$$\log \mathbb{E}\left[e^{t(L(\boldsymbol{\beta}) - \mu_L)}\right] \le \frac{t^2 \bar{\sigma}_L^2}{2}. \tag{A.142}$$

Step 2: Centered MGF of the quadratic part. We compute μ_Q exactly. For $\boldsymbol{\beta} \sim \mathcal{N}(\boldsymbol{\mu}_\mathbf{Y}, \Sigma_\mathbf{Y})$:

$$\mu_Q = -\frac{1}{2\sigma^2}\mathbb{E}[\boldsymbol{\beta}^\top A \boldsymbol{\beta}] = -\frac{1}{2\sigma^2}\left(\mathrm{tr}(A\Sigma_\mathbf{Y}) + \boldsymbol{\mu}_\mathbf{Y}^\top A \boldsymbol{\mu}_\mathbf{Y}\right). \tag{A.143}$$

Define the symmetrized matrix $\tilde{A} \triangleq \Sigma_\mathbf{Y}^{1/2} A \Sigma_\mathbf{Y}^{1/2}$. By Sylvester's determinant identity:

$$\det\left(\mathbf{I} + \frac{t}{\sigma^2}\Sigma_\mathbf{Y} A\right) = \det\left(\mathbf{I} + \frac{t}{\sigma^2}\Sigma_\mathbf{Y}^{1/2} A \Sigma_\mathbf{Y}^{1/2}\right) = \det\left(\mathbf{I} + \frac{t}{\sigma^2}\tilde{A}\right), \tag{A.144}$$

[3]For brevity, we write $\mu^{c_1}_\mathbf{Y}, \nu^2_\mathbf{Y}, b_\mathbf{Y}$ for $\mu_{\mathbf{Y},\mathbf{Y}'}, \nu^2_{\mathbf{Y},\mathbf{Y}'}, b_{\mathbf{Y},\mathbf{Y}'}$ in Corollary 13.

so the positivity condition $\mathbf{I} + \frac{t}{\sigma^2}\tilde{A} \succ 0$ controls both the determinant and the existence of the MGF, while the inverse term is retained in the nonsymmetric form $(\mathbf{I} + \frac{t}{\sigma^2}\Sigma_{\mathbf{Y}} A)^{-1}$, which is algebraically equivalent. Since $\tilde{A}$ is symmetric:

$$\|\tilde{A}\|_{\mathrm{op}} = \|\Sigma_{\mathbf{Y}}^{1/2} A \Sigma_{\mathbf{Y}}^{1/2}\|_{\mathrm{op}} \le \|\Sigma_{\mathbf{Y}}\|_{\mathrm{op}} \|A\|_{\mathrm{op}} \le \tau^2(\|\mathbf{x}_n\|^2 + \|\mathbf{x}_n'\|^2) \le 2\tau^2 R^2, \tag{A.145}$$

so the positivity condition holds for $|t| < b_Q^{-1}$ where $b_Q \triangleq 2\tau^2 R^2/\sigma^2$. The raw MGF of $Q(\boldsymbol{\beta})$ is:

$$\mathbb{E}\left[e^{tQ(\boldsymbol{\beta})}\right] = \frac{\exp\left(-\frac{t}{2\sigma^2}\boldsymbol{\mu}_{\mathbf{Y}}^\top A\left(\mathbf{I} + \frac{t}{\sigma^2}\Sigma_{\mathbf{Y}} A\right)^{-1}\boldsymbol{\mu}_{\mathbf{Y}}\right)}{\sqrt{\det\left(\mathbf{I} + \frac{t}{\sigma^2}\tilde{A}\right)}}. \tag{A.146}$$

Taking the logarithm of Eq. (A.146) and centering by μ_Q via Eq. (A.143):

$$\begin{aligned}\log \mathbb{E}\left[e^{t(Q(\boldsymbol{\beta})-\mu_Q)}\right] = &-\frac{1}{2}\log\det\left(\mathbf{I} + \frac{t}{\sigma^2}\tilde{A}\right) - \frac{t}{2\sigma^2}\boldsymbol{\mu}_{\mathbf{Y}}^\top A\left(\mathbf{I} + \frac{t}{\sigma^2}\Sigma_{\mathbf{Y}} A\right)^{-1}\boldsymbol{\mu}_{\mathbf{Y}} \\ &+ \frac{t}{2\sigma^2}\left(\mathrm{tr}(A\Sigma_{\mathbf{Y}}) + \boldsymbol{\mu}_{\mathbf{Y}}^\top A\boldsymbol{\mu}_{\mathbf{Y}}\right).\end{aligned} \tag{A.147}$$

We bound each term in Eq. (A.147) separately.

Term 1 (log-determinant). Applying the scalar inequality $-\log(1+x) \le -x + x^2/(1-|x|)$ (valid for $|x| < 1$) eigenvalue-by-eigenvalue to $\tilde{A}$, and using $\|t\tilde{A}/\sigma^2\|_{\mathrm{op}} \le b_Q|t| < 1$:

$$-\frac{1}{2}\log\det\left(\mathbf{I} + \frac{t}{\sigma^2}\tilde{A}\right) \le -\frac{t}{2\sigma^2}\mathrm{tr}(\tilde{A}) + \frac{t^2\,\mathrm{tr}(\tilde{A}^2)}{2\sigma^4(1 - b_Q|t|)}, \tag{A.148}$$

where $\mathrm{tr}(\tilde{A}) = \mathrm{tr}(\Sigma_{\mathbf{Y}} A)$.

Term 2 (quadratic numerator). Since $(\mathbf{I} + \frac{t}{\sigma^2}\Sigma_{\mathbf{Y}} A)^{-1}$ is not symmetric in general, we use operator norms directly. Note first that

$$\left\|\frac{t}{\sigma^2}\Sigma_{\mathbf{Y}} A\right\|_{\mathrm{op}} \le \frac{|t|}{\sigma^2}\|\Sigma_{\mathbf{Y}}\|_{\mathrm{op}}\|A\|_{\mathrm{op}} \le b_Q|t| < 1, \tag{A.149}$$

so the resolvent bound gives

$$\left\|\left(\mathbf{I} + \frac{t}{\sigma^2}\Sigma_{\mathbf{Y}} A\right)^{-1}\right\|_{\mathrm{op}} \le \frac{1}{1 - b_Q|t|}. \tag{A.150}$$

Writing $(\mathbf{I} + \frac{t}{\sigma^2}\Sigma_{\mathbf{Y}} A)^{-1} = \mathbf{I} - \frac{t}{\sigma^2}\Sigma_{\mathbf{Y}} A(\mathbf{I} + \frac{t}{\sigma^2}\Sigma_{\mathbf{Y}} A)^{-1}$ and applying Eq. (A.150):

$$\begin{aligned}&-\frac{t}{2\sigma^2}\boldsymbol{\mu}_{\mathbf{Y}}^\top A\left(\mathbf{I} + \frac{t}{\sigma^2}\Sigma_{\mathbf{Y}} A\right)^{-1}\boldsymbol{\mu}_{\mathbf{Y}} \\ =\ &-\frac{t}{2\sigma^2}\boldsymbol{\mu}_{\mathbf{Y}}^\top A\boldsymbol{\mu}_{\mathbf{Y}} + \frac{t^2}{2\sigma^4}\boldsymbol{\mu}_{\mathbf{Y}}^\top A\Sigma_{\mathbf{Y}} A\left(\mathbf{I} + \frac{t}{\sigma^2}\Sigma_{\mathbf{Y}} A\right)^{-1}\boldsymbol{\mu}_{\mathbf{Y}} \\ \le\ &-\frac{t}{2\sigma^2}\boldsymbol{\mu}_{\mathbf{Y}}^\top A\boldsymbol{\mu}_{\mathbf{Y}} + \frac{t^2\|\boldsymbol{\mu}_{\mathbf{Y}}\|^2\|A\|_{\mathrm{op}}^2\|\Sigma_{\mathbf{Y}}\|_{\mathrm{op}}}{2\sigma^4(1 - b_Q|t|)},\end{aligned} \tag{A.151}$$

where the last step uses $\left|\boldsymbol{\mu}_{\mathbf{Y}}^\top A\Sigma_{\mathbf{Y}} A(\mathbf{I} + \frac{t}{\sigma^2}\Sigma_{\mathbf{Y}} A)^{-1}\boldsymbol{\mu}_{\mathbf{Y}}\right| \le \|\boldsymbol{\mu}_{\mathbf{Y}}\|^2\|A\|_{\mathrm{op}}^2\|\Sigma_{\mathbf{Y}}\|_{\mathrm{op}}/(1 - b_Q|t|)$ via Eq. (A.150).

Term 3 (centering).

$$\frac{t}{2\sigma^2}\left(\mathrm{tr}(A\Sigma_{\mathbf{Y}}) + \boldsymbol{\mu}_{\mathbf{Y}}^\top A\boldsymbol{\mu}_{\mathbf{Y}}\right) = \frac{t}{2\sigma^2}\mathrm{tr}(\tilde{A}) + \frac{t}{2\sigma^2}\boldsymbol{\mu}_{\mathbf{Y}}^\top A\boldsymbol{\mu}_{\mathbf{Y}}. \tag{A.152}$$

Combining Terms 1–3, the linear terms $-\frac{t}{2\sigma^2}\mathrm{tr}(\tilde{A})$ and $-\frac{t}{2\sigma^2}\boldsymbol{\mu}_{\mathbf{Y}}^\top A\boldsymbol{\mu}_{\mathbf{Y}}$ cancel exactly with Eq. (A.152), giving:

$$\log\mathbb{E}\left[e^{t(Q(\boldsymbol{\beta})-\mu_Q)}\right] \le \frac{t^2}{2(1-b_Q|t|)}\left(\frac{\mathrm{tr}(\tilde{A}^2)}{\sigma^4}+\frac{\|\boldsymbol{\mu}_{\mathbf{Y}}\|^2\|A\|_{\mathrm{op}}^2\|\Sigma_{\mathbf{Y}}\|_{\mathrm{op}}}{\sigma^4}\right) \triangleq \frac{t^2\nu_Q^2}{2(1-b_Q|t|)}. \tag{A.153}$$

Bounding each term using $\Sigma_{\mathbf{Y}} \preceq \tau^2\mathbf{I}_d$, $\|A\|_F^2 \le 8R^4$, $\|A\|_{\mathrm{op}} \le 2R^2$, and $\|\boldsymbol{\mu}_{\mathbf{Y}}\| \le n\tau^2 M_y R/\sigma^2$:

$$\mathrm{tr}(\tilde{A}^2) = \mathrm{tr}(\Sigma_{\mathbf{Y}}A\Sigma_{\mathbf{Y}}A) \le \tau^4\|A\|_F^2 \le 8\tau^4R^4, \tag{A.154}$$

$$\|\boldsymbol{\mu}_{\mathbf{Y}}\|^2\|A\|_{\mathrm{op}}^2\|\Sigma_{\mathbf{Y}}\|_{\mathrm{op}} \le \frac{n^2\tau^4M_y^2R^2}{\sigma^4}\cdot 4R^4\cdot\tau^2 = \frac{4n^2\tau^6M_y^2R^6}{\sigma^4}, \tag{A.155}$$

so that

$$\nu_Q^2 \le \frac{8\tau^4R^4}{\sigma^4}+\frac{4n^2\tau^6M_y^2R^6}{\sigma^8}. \tag{A.156}$$

Step 3: Combining via Cauchy-Schwarz. Since $L(\boldsymbol{\beta})$ and $Q(\boldsymbol{\beta})$ both depend on $\boldsymbol{\beta}$, they are not independent. Applying Cauchy-Schwarz:

$$\begin{aligned}\log\mathbb{E}\left[e^{t(c_1(\boldsymbol{\beta})-\mu_{\mathbf{Y}}^{c_1})}\right] &\le \frac{1}{2}\log\mathbb{E}\left[e^{2t(L-\mu_L)}\right]+\frac{1}{2}\log\mathbb{E}\left[e^{2t(Q-\mu_Q)}\right]\\ &\le \frac{(2t)^2\bar{\sigma}_L^2}{4}+\frac{(2t)^2\nu_Q^2}{4(1-b_Q|2t|)}\\ &= t^2\bar{\sigma}_L^2+\frac{t^2\nu_Q^2}{1-2b_Q|t|}\\ &\le \frac{t^2(\bar{\sigma}_L^2+\nu_Q^2)}{1-2b_Q|t|} \triangleq \frac{t^2\nu_{\mathbf{Y}}^2/2}{1-b_{\mathbf{Y}}|t|},\end{aligned} \tag{A.157}$$

where

$$\nu_{\mathbf{Y}}^2 \triangleq 2(\bar{\sigma}_L^2+\nu_Q^2) \le \frac{8M_y^2R^2\tau^2}{\sigma^4}+\frac{16\tau^4R^4}{\sigma^4}+\frac{8n^2\tau^6M_y^2R^6}{\sigma^8}, \tag{A.158}$$

$$b_{\mathbf{Y}} \triangleq 2b_Q = \frac{4\tau^2R^2}{\sigma^2}. \tag{A.159}$$

This establishes that $c_1(\boldsymbol{\beta})$ satisfies a sub-exponential MGF bound under $\pi(\boldsymbol{\beta}|\mathbf{Y})$ with parameters $(\nu_{\mathbf{Y}}^2, b_{\mathbf{Y}})$ in the sense of Eq. (A.33).

For the reversed pair $(\mathbf{Y}', \mathbf{Y})$, $\tilde{c}_1(\boldsymbol{\beta}) = -c_1(\boldsymbol{\beta}) = -L(\boldsymbol{\beta}) - Q(\boldsymbol{\beta})$, i.e., (L, Q, A) is replaced by $(-L, -Q, -A)$. Steps 1–3 use only $\Sigma \preceq \tau^2\mathbf{I}_d$, $\|\mathbf{x}_i\| \le R$, $|y_i| \le M_y$, $\|\boldsymbol{\mu}\| \le n\tau^2M_yR/\sigma^2$, $\|-A\|_{\mathrm{op}} = \|A\|_{\mathrm{op}}$ and $\|-A\|_F = \|A\|_F$, all of which also hold under $\pi(\boldsymbol{\beta}|\mathbf{Y}')$ since $\mathbf{Y}'$ also has n records. Hence $\tilde{c}_1(\boldsymbol{\beta})$ satisfies the same sub-exponential bound with parameters $(\nu_{\mathbf{Y}}^2, b_{\mathbf{Y}})$ in Eqs. (A.158)–(A.159).

Step 4: Applying Corollary 13. Corollary 13 requires $b_{\mathbf{Y}} < \min\{(\alpha-1)^{-1}, 1\}$. The condition $b_{\mathbf{Y}} < 1$ is an additional scale restriction requiring:

$$\frac{4\tau^2R^2}{\sigma^2} < 1 \iff \sigma^2 > 4\tau^2R^2. \tag{A.160}$$

Under Eq. (A.160) and $\alpha < 1+\sigma^2/(4\tau^2R^2)$, substituting $(\nu_{\mathbf{Y}}^2, b_{\mathbf{Y}})$ into Corollary 13:

$$\max\{D_\alpha(P\|Q), D_\alpha(Q\|P)\} \le \varepsilon_{\text{LR, Case2}}(\alpha, b_{\mathbf{Y}}, \nu_{\mathbf{Y}}^2) \triangleq \frac{\nu_{\mathbf{Y}}^2}{2}\cdot\frac{\alpha(1-2b_{\mathbf{Y}})+2b_{\mathbf{Y}}}{(1-b_{\mathbf{Y}})\,(1-b_{\mathbf{Y}}(\alpha-1))}. \tag{A.161}$$

Remark A.10. *Three sources of looseness affect Eq.* (A.161)*. First, the Cauchy-Schwarz step doubles b_Q, restricting the valid range of α. Second, the operator norm bound $\|\tilde{A}\|_{\mathrm{op}} \le 2\tau^2R^2$ ignores data-dependent tightening of $\Sigma_{\mathbf{Y}}$ due to the likelihood. Third, the n^2-dependent term in $\nu_{\mathbf{Y}}^2$ arises from bounding $\|\boldsymbol{\mu}_{\mathbf{Y}}\|$ using only row-wise boundedness; under a direct bound $\|\boldsymbol{\mu}_{\mathbf{Y}}\| \le C$ for some constant C, this term becomes n-independent. Despite this looseness, Eq.* (A.161) *yields an explicit finite-order RDP bound for OPS in the bounded-neighboring linear-regression setting with covariate replacement, and illustrates how the sub-exponential framework of Corollary 13 extends beyond the sub-Gaussian cases treated in the main text.*

A.10.2 unbounded neighboring

Let $\mathbf{Y}$ has size n and $\mathbf{Y}'$ is of size $n-1$ by removing the n-th record $\mathbf{z}_n = (\mathbf{x}_n, y_n)$ in $\mathbf{Y}$. Then

$$c_1(\boldsymbol{\beta}) = \log \frac{\ell(\boldsymbol{\beta}; \mathbf{Y})}{\ell(\boldsymbol{\beta}; \mathbf{Y}')} = \log p(y_n|\mathbf{x}_n, \boldsymbol{\beta}) = -\frac{(y_n - \mathbf{x}_n^\top \boldsymbol{\beta})^2}{2\sigma^2} - \frac{1}{2}\log(2\pi\sigma^2).$$

This expression is structurally identical to Eq. (A.95) by setting $Z = y_n - \mathbf{x}_n^\top \boldsymbol{\beta}$, where

$$Z|\mathbf{Y} \sim \mathcal{N}\left(y_n - \mathbf{x}_n^\top \boldsymbol{\mu}_{\mathbf{Y}}, \mathbf{x}_n^\top \Sigma_{\mathbf{Y}} \mathbf{x}_n\right).$$

In Proposition 11, $\omega = \sigma^2/v_Y^2$, where v_Y^2 is the posterior variance of $\theta - y_n$; here the corresponding quantity is the posterior variance of $Z|\mathbf{Y}$. Thus, the corresponding ω_n here is $\omega_n = \frac{\sigma^2}{\mathbf{x}_n^\top \Sigma_{\mathbf{Y}} \mathbf{x}_n}$, plug in Eq. (A.110)

$$\begin{aligned} D_\alpha(P\|Q) &= \underbrace{-\frac{\log(1-\omega_n^{-1})}{2} - \frac{\log(1+\frac{\alpha-1}{\omega_n})}{2(\alpha-1)}}_{A(\alpha,\omega_n)} + \frac{(\mathbf{x}_n^\top \boldsymbol{\mu}_{\mathbf{Y}} - y_n)^2}{2\sigma^2} \underbrace{\left[\frac{1}{1-\omega_n^{-1}} - \frac{1}{1+\frac{\alpha-1}{\omega_n}}\right]}_{B(\alpha,\omega_n)} \\ &\le A(\alpha,\omega_n) + \frac{k^{*2}}{2} B(\alpha,\omega_n), \end{aligned} \tag{A.162}$$

if $|\mathbf{x}_n^\top \boldsymbol{\mu}_{\mathbf{Y}} - y_n| \le k^* \sigma$ for some $k^* > 0$.

Reverse direction. We now bound $D_\alpha(\pi(\boldsymbol{\beta}|\mathbf{Y}')\|\pi(\boldsymbol{\beta}|\mathbf{Y}))$. The posterior under $\mathbf{Y}'$ is $\boldsymbol{\beta}|\mathbf{Y}' \sim \mathcal{N}(\boldsymbol{\mu}_{\mathbf{Y}'}, \Sigma_{\mathbf{Y}'})$ with $\Sigma_{\mathbf{Y}'}^{-1} = \Sigma_{\mathbf{Y}}^{-1} - \sigma^{-2}\mathbf{x}_n\mathbf{x}_n^\top$. Let $v_n'^2 \triangleq \mathbf{x}_n^\top \Sigma_{\mathbf{Y}'} \mathbf{x}_n$ and

$$\begin{aligned} \tilde{c}_1(\boldsymbol{\beta}) &= \log \frac{\ell(\boldsymbol{\beta}; \mathbf{Y}')}{\ell(\boldsymbol{\beta}; \mathbf{Y})} = \frac{(y_n - \mathbf{x}_n^\top \boldsymbol{\beta})^2}{2\sigma^2} + \frac{1}{2}\log(2\pi\sigma^2), \\ Z' &\triangleq y_n - \mathbf{x}_n^\top \boldsymbol{\beta} \,|\, \mathbf{Y}' \sim \mathcal{N}(y_n - \mathbf{x}_n^\top \boldsymbol{\mu}_{\mathbf{Y}'},\, v_n'^2). \end{aligned} \tag{A.163}$$

This is structurally identical to Eq. (A.117) with $\mu_{Y'}$ replaced by $\mathbf{x}_n^\top \boldsymbol{\mu}_{\mathbf{Y}'}$ and $v_{Y'}^2$ replaced by $v_n'^2$. Hence, by Eq. (A.125), $D_\alpha(Q\|P) < \infty$ if and only if $(\alpha-1)v_n'^2 < \sigma^2$. By the Sherman–Morrison formula,

$$\begin{aligned} &\Sigma_{\mathbf{Y}} = \Sigma_{\mathbf{Y}'} - \frac{\Sigma_{\mathbf{Y}'}\mathbf{x}_n\mathbf{x}_n^\top\Sigma_{\mathbf{Y}'}}{\sigma^2 + v_n'^2} \;\Rightarrow\; \mathbf{x}_n^\top \Sigma_{\mathbf{Y}} \mathbf{x}_n = \frac{\sigma^2 v_n'^2}{\sigma^2 + v_n'^2} \;\Rightarrow\; \omega_n = 1 + \frac{\sigma^2}{v_n'^2}, \\ &\text{i.e. } \frac{v_n'^2}{\sigma^2} = \frac{1}{\omega_n - 1}. \end{aligned}$$

Moreover, $\boldsymbol{\mu}_{\mathbf{Y}} = \Sigma_{\mathbf{Y}}(\Sigma_{\mathbf{Y}'}^{-1}\boldsymbol{\mu}_{\mathbf{Y}'} + \sigma^{-2}\mathbf{x}_n y_n)$ and $\mathbf{x}_n^\top \Sigma_{\mathbf{Y}} \Sigma_{\mathbf{Y}'}^{-1} = \frac{\sigma^2}{\sigma^2 + v_n'^2}\mathbf{x}_n^\top$, so

$$\mathbf{x}_n^\top \boldsymbol{\mu}_{\mathbf{Y}} = \frac{\sigma^2}{\sigma^2 + v_n'^2}\mathbf{x}_n^\top \boldsymbol{\mu}_{\mathbf{Y}'} + \frac{v_n'^2}{\sigma^2 + v_n'^2} y_n \;\Rightarrow\; \mathbf{x}_n^\top \boldsymbol{\mu}_{\mathbf{Y}'} - y_n = \frac{\omega_n}{\omega_n - 1}(\mathbf{x}_n^\top \boldsymbol{\mu}_{\mathbf{Y}} - y_n). \tag{A.164}$$

These coincide with $v_{Y'}^2 = \sigma^2/(\omega - 1)$ and Eq. (A.128) with ω replaced by ω_n. Repeating the derivation of Eq. (A.129), when $1 < \alpha < \omega_n$

$$D_\alpha(Q\|P) = \widetilde{A}(\alpha,\omega_n) + \frac{(\mathbf{x}_n^\top \boldsymbol{\mu}_{\mathbf{Y}} - y_n)^2}{2\sigma^2}\widetilde{B}(\alpha,\omega_n) \le \widetilde{A}(\alpha,\omega_n) + \frac{k^{*2}}{2}\widetilde{B}(\alpha,\omega_n), \tag{A.165}$$

and $D_\alpha(Q\|P) = \infty$ for $\alpha \ge \omega_n$.

Two-sided, data-independent guarantee. Since $\Sigma_{\mathbf{Y}'} \preceq \tau^2 \mathbf{I}_d$ and $\|\mathbf{x}_n\| \le R$, $v_n'^2 \le \tau^2 R^2$, hence $\omega_n \ge \omega_{\min} \triangleq 1 + \sigma^2/(\tau^2 R^2)$ for every admissible neighboring pair, with equality, e.g., when $n = 1$ and $\|\mathbf{x}_1\| = R$. By the monotonicity of A, B and $\widetilde{A}, \widetilde{B}$ in ω (Claim A.9 and A.8), for every $1 < \alpha < \omega_{\min}$,

$$\begin{aligned} &\max\{D_\alpha(P\|Q), D_\alpha(Q\|P)\} \\ &\le \max\left\{A(\alpha,\omega_{\min}) + \frac{k^{*2}}{2}B(\alpha,\omega_{\min}),\; \widetilde{A}(\alpha,\omega_{\min}) + \frac{k^{*2}}{2}\widetilde{B}(\alpha,\omega_{\min})\right\}, \end{aligned} \tag{A.166}$$

whereas for $\alpha \ge \omega_{\min}$ there exist admissible neighboring pairs with $D_\alpha(Q\|P) = \infty$. In particular, OPS does not satisfy pure DP in this setting.

□

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The abstract and introduction state the paper's theoretical contributions

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: The paper discusses limitations in Section 6, including the dependence of the guarantees on model assumptions, tail conditions on the log-likelihood ratio, and the looseness of general sub-exponential bounds relative to model-specific exact calculations.

   Guidelines:

   - The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

Answer: [Yes]

Justification: The paper is theoretical. Each theoretical result states its assumptions, and proofs are provided either in the main text or in the appendix.

Guidelines:

- The answer [N/A] means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [N/A]

   Justification: The paper does not report empirical experiments. The main contributions are theoretical.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
     (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
     (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [N/A]

Justification: The paper does not use datasets or code to support experimental claims. The work is theoretical.

Guidelines:

- The answer [N/A] means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [N/A]

   Justification: The paper does not include experiments. The work is theoretical.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [N/A]

   Justification: The paper does not include experiments. The work is theoretical.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
   - The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
   - The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)

- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
- If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [N/A]

   Justification: The paper does not include experiments. The work is theoretical.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: The research is theoretical and concerns privacy guarantees for Bayesian posterior sampling.

   Guidelines:

   - The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [Yes]

    Justification: Potential positive impacts include improved understanding of privacy-preserving Bayesian posterior sampling. Such understanding can help design mechanisms that preserve more utility in released statistical outputs while exploiting the inherent privacy provided by posterior randomness. potential limitations include the risk of generalizing the guarantees outside the stated assumptions.

    Guidelines:

    - The answer [N/A] means that there is no societal impact of the work performed.

- If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

    Answer: [N/A]

    Justification: The paper does not release models, datasets, or tools with high misuse risk.

    Guidelines:

    - The answer [N/A] means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [N/A]

    Justification: The paper does not use existing datasets, code, models, or other external assets.Related Prior work is listed in the References Section.

    Guidelines:

    - The answer [N/A] means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.

- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [N/A]

    Justification: The paper does not introduce or release new datasets, codebases, models, benchmarks, or other assets.

    Guidelines:

    - The answer [N/A] means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A]

    Justification: The paper does not involve crowdsourcing.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A]

    Justification: The paper does not involve human-subject research.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.

- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

    Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

    Answer: [N/A]

    Justification: The core method development does not involve LLMs as an important, original, or non-standard component.

    Guidelines:

    - The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
    - Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.