CS839: Differential Privacy and Learning February 19, 2026

10: Zero-Concentrated Differential Privacy, Part 1

Instructor: Gavin Brown *Scribe:* Leitian Tao

Disclaimer: This document is intended as an informal supplement to in-class note-taking. It has not been given the level of scrutiny expected in polished lecture notes, let alone that reserved for peer-reviewed publications.

1 Introduction

This document develops the necessary privacy accounting tools required for modern machine learning applications. However, the modern machine learning paradigm of stochastic, gradient-based optimization—present unique challenges. These algorithms typically require thousands of iterations over very high-dimensional parameter spaces, distributed across multiple nodes.

Under such constraints, the (ε, δ)-differential privacy composition theorems may yield loose, pessimistic bounds that degrade utility. To resolve this, we transition to new frameworks for analyzing privacy guarantees that focus on understanding privacy loss as a random variable. By examining the concentration of this random variable via its moment generating function (MGF), we can sometimes achieve tighter composition. This motivates the study of zero-concentrated differential privacy (zCDP) and Rényi differential privacy (RDP).

Before looking at those, we will finish up some details on the Propose-Test-Release framework and mention some variations on the definition of dataset adjacency. This last isn't about optimization in particular, but it's important and we need to mention it somewhere.

2 Propose–Test–Release (PTR) and Data Stability

The Propose–Test–Release (PTR) paradigm provides a robust mechanism for safely answering queries that have high global sensitivity but low local sensitivity for most realistic datasets. The core strategy is to (i) propose candidate "base" algorithm, which usually will not itself be DP, (ii) privately test whether the input dataset resides in a "safe" neighborhood where the base algorithm is stable under local perturbations, and (iii) release/run the base algorithm if the test passes, or output a default safe symbol (e.g., $\perp$) otherwise.

2.1 Safe Sets and Mode Estimation

Consider a scenario where the dataset consists of binary vectors $x = (x_1, \ldots, x_n) \in \{0,1\}^n$, and the goal is to securely release the majority bit, $\text{mode}(x) \in \{0,1\}$. A naive base algorithm $A(x) = \text{mode}(x)$ is not differentially private, as shifting a single record can alter the mode when the counts of 0s and 1s are balanced.

To formalize the conditions under which $A(x)$ can be safely released, we define the concept of a safe set relative to a chosen privacy parameter.

Definition 2.1 (Safe Set)**.** Let $\sim$ denote the chosen adjacency relation on datasets. For a given algorithm A and privacy parameters ε, δ, the safe set is defined as:

$$\mathrm{SAFE}_{\varepsilon,\delta}^{A} \triangleq \left\{ x \in \mathcal{X}^n : \forall x' \sim x,\ A(x) \approx_{\varepsilon,\delta} A(x') \right\}.$$

For the mode estimation task, a dataset x is deeply within the safe set if the frequency of the majority bit substantially exceeds that of the minority bit. If the dataset is safely bounded away from the decision boundary, the PTR mechanism cleanly releases mode(x). If it is close to an unsafe dataset, it outputs $\perp$. In between these extremes, PTR will either fail or release the mode, but the decision of whether to do so will be differfentially private.

3 Different Notions of Adjacency

Differential privacy is evaluated relative to a specific adjacency relation. The choice of adjacency influences the resulting sensitivity bounds and should be tailored to the underlying problem.

Definition 3.1 (Swap Adjacency)**.** Two datasets $x = (x_1, \ldots, x_n)$ and $x' = (x'_1, \ldots, x'_n)$ in $\mathcal{X}^n$ are *swap-adjacent* (or replace-one adjacent) if they differ in exactly one coordinate. This assumes the dataset size n is fixed and public.

Definition 3.2 (Add/Remove Adjacency)**.** Viewing datasets as multisets, x and x' are *add/remove adjacent* if one can be constructed from the other by inserting or deleting a single record. This allows the dataset size n to remain private.

In class so far we've always worked with swap adjacency. This is usually (but not always) simpler from a theoretical standpoint. Practitioners usually prefer add/remove, because the number of data points is essentially never public knowledge.

4 Privacy Loss and the Privacy Loss Random Variable

4.1 The Privacy Loss Random Variable

We think of an adversary receiving an output y; they know that the dataset was either x or x'. Recall the *privacy loss*, which measures the evidence for the hypothesis that $y \sim A(x)$ instead of $A(x')$.

Definition 4.1 (Privacy Loss)**.** Let A be a randomized algorithm and let x, x' be adjacent datasets. For an observed output $y \in \mathrm{supp}(A)$, the *privacy loss* evaluated at y is given by

$$L_A^{x \to x'}(y) \triangleq \log \frac{\mathbb{P}[A(x) = y]}{\mathbb{P}[A(x') = y]}.$$

(For continuous output spaces, the probabilities are replaced by their respective probability density functions).

When we think of Y as a random variable sampled according to $Y \sim A(x)$, the privacy loss itself becomes a random quantity.

Definition 4.2 (Privacy Loss Random Variable (PLRV))**.** The *privacy loss random variable* (PLRV) for adjacent datasets $x \sim x'$ under mechanism A is the random variable $L_A^{x \to x'}(Y)$, where $Y \sim A(x)$.

This definition possesses a natural hypothesis testing interpretation. Consider distinguishing between the null hypothesis $H_0 : Y \sim A(x)$ and the alternative $H_1 : Y \sim A(x')$. The PLRV is precisely the log-likelihood ratio statistic used to evaluate the relative evidence in favor of H_0 over H_1.

4.2 Gaussian PLRV Behavior

To gain intuition for the behavior of the PLRV, we consider the canonical case of releasing a continuous value subject to Gaussian noise.

Claim 4.3 (Gaussian PLRV)**.** *Let $A(x) \sim \mathcal{N}(\mu, \sigma^2)$ and $A(x') \sim \mathcal{N}(\mu', \sigma^2)$. If we sample $Y \sim A(x)$ and define the signal-to-noise ratio $\rho \triangleq \frac{(\mu-\mu')^2}{2\sigma^2}$, then the resulting PLRV is normally distributed:*

$$L_A^{x \to x'}(Y) \sim \mathcal{N}(\rho,\, 2\rho).$$

Proof. Evaluating the privacy loss function at a fixed point y gives:

$$\begin{aligned} L_A^{x \to x'}(y) &= \log \frac{\frac{1}{\sqrt{2\pi}\sigma} \exp\Big(-\frac{(y-\mu)^2}{2\sigma^2}\Big)}{\frac{1}{\sqrt{2\pi}\sigma} \exp\Big(-\frac{(y-\mu')^2}{2\sigma^2}\Big)} = \frac{-(y-\mu)^2 + (y-\mu')^2}{2\sigma^2} \\ &= \frac{(\mu-\mu')(2y-\mu-\mu')}{2\sigma^2}. \end{aligned}$$

Observe that $L_A^{x \to x'}(y)$ is an affine function of y. Because Y is drawn from a normal distribution $\mathcal{N}(\mu, \sigma^2)$, and affine transformations preserve Gaussianity, the variable $L_A^{x \to x'}(Y)$ is also normally distributed. A calculation of the expectation and variance yields $\mathbb{E}[L] = \rho$ and $\mathrm{Var}(L) = 2\rho$. □

5 Composition and the MGF Lens

In iterative algorithms such as DP-GD, we release a sequence of updates. Under independent[1] composition, the overall privacy loss behaves gracefully.

Claim 5.1 (Additivity of Privacy Loss)**.** *Suppose an algorithm A comprises k sequentially executed, independent mechanisms, $A(x) = \big(A_1(x), \ldots, A_k(x)\big)$. The total PLRV is exactly the sum of the individual PLRVs:*

$$L_A^{x \to x'}(Y) = \sum_{j=1}^{k} L_{A_j}^{x \to x'}(Y_j).$$

If we assume the individual privacy losses L_j are bounded and identically distributed, the Central Limit Theorem (CLT) suggests that the sum $\sum_j L_j$ will asymptotically converge to a Gaussian. We could formalize this intuition with the Berry–Esseen theorem, but its $O(1/\sqrt{k})$ convergence rate is insufficient for the strict, small-δ tail bounds required for privacy guarantees.

Instead, we control the tails of the PLRV directly via its moment generating function (MGF), $M_X(\lambda) \triangleq \mathbb{E}[e^{\lambda X}]$. By applying an exponential Markov bound (Chernoff bound), for any $\lambda > 0$ and threshold t:

$$\mathbb{P}[X \geq t] = \mathbb{P}[e^{\lambda X} \geq e^{\lambda t}] \leq e^{-\lambda t} M_X(\lambda).$$

[1] In much of the class, we will assume that each algorithm is fixed in advance and runs independently of the previous algorithms. This greatly simplifies our math. However, this assumption is not required: DP composes *adaptively*. This is crucial for all of our applications.

If a mechanism's privacy loss exhibits bounded MGF growth—specifically, if it is *subgaussian*—we can derive incredibly tight tail bounds that persist through thousands of compositions.

6 Advanced Accounting: zCDP and RDP

To systematize MGF-based privacy accounting, recent literature standardizes metrics that encapsulate subgaussian PLRV behavior. We focus on Zero-Concentrated Differential Privacy (zCDP) and Rényi Differential Privacy (RDP).

6.1 Zero-Concentrated Differential Privacy (zCDP)

Definition 6.1 (ρ-zCDP)**.** A randomized algorithm A satisfies ρ-zero-concentrated differential privacy (ρ-zCDP) if for all adjacent datasets x, x' and all $\lambda \geq 0$, the corresponding PLRV $L \triangleq L_A^{x \to x'}(Y)$ satisfies:

$$\mathbb{E}[e^{\lambda L}] \leq \exp\big(\lambda(\lambda+1)\rho\big).$$

The zCDP framework is very useful in many contexts because it yields strong and simple composition bounds. Specifically, standard pure differential privacy maps smoothly into zCDP: any ε-DP mechanism is inherently $\frac{1}{2}\varepsilon^2$-zCDP. Furthermore, zCDP scales linearly; the independent composition of k mechanisms, each satisfying ρ-zCDP, is $(k\rho)$-zCDP.

Once an algorithm's entire trajectory has been accounted for in the zCDP space, we can cleanly translate the final ρ value back into the standard (ε, δ)-DP regime.

Theorem 6.2 (Converting zCDP to Standard DP)**.** *If an algorithm A is ρ-zCDP, then for any failure probability $\delta \in (0,1)$, A satisfies (ε', δ)-DP where*

$$\varepsilon' = \rho + \sqrt{2\rho \log(1/\delta)}.$$

6.2 Rényi Differential Privacy (RDP)

Rényi Differential Privacy generalizes these MGF bounds by measuring the privacy loss through the Rényi divergence.

Definition 6.3 (Rényi Divergence)**.** For two probability distributions P, Q on a common measurable space, and for a defined order $\alpha > 1$, the Rényi divergence of order α is defined as:

$$D_\alpha(P\|Q) \triangleq \frac{1}{\alpha-1} \log \mathbb{E}_{Z\sim Q}\left[\left(\frac{P(Z)}{Q(Z)}\right)^\alpha\right].$$

Definition 6.4 ((α,ε)-RDP)**.** An algorithm A satisfies (α,ε)-RDP if for all adjacent datasets x, x', the Rényi divergence between their output distributions is bounded by ε:

$$D_\alpha\big(A(x) \,\|\, A(x')\big) \leq \varepsilon.$$

This Rényi divergence definition is functionally equivalent to asserting bounds on the MGF of the PLRV. Specifically, bounding the exponentiated privacy loss by $\exp((\alpha-1)\varepsilon)$ is identical to bounding the Rényi divergence of order α. Because zCDP requires a parabolic bound across all $\lambda > 0$, it can be conceptually viewed as a requirement that a mechanism satisfies a structured family of RDP guarantees across all orders $\alpha > 1$.

7 Further Exercises

Remark 7.1 (Histogram Estimation)**.** As an exercise in contrasting local and global sensitivity, one should consider how to approximate the exact mode of a dataset using a Laplacian or Gaussian noisy histogram. Analyzing the accuracy constraints of the histogram approach versus the Propose–Test–Release approach provides crucial intuition for modern mechanism design.

Remark 7.2 (Continuous Domains)**.** The foundational concepts introduced here, particularly the construction of safe sets, naturally extend to real-valued data domains ($\mathcal{X} = \mathbb{R}$). One must verify how sensitivity metrics and noise calibration shift when local perturbations can take arbitrarily small, but continuous, forms.

References

[1] Vadhan, S. *The complexity of differential privacy.* Tutorials on the Foundations of Cryptography, 2017. https://salil.seas.harvard.edu/sites/g/files/omnuum4266/files/salil/files/manuscript_2017.pdf

[2] Dwork, C., McSherry, F., Nissim, K., & Smith, A. *Calibrating noise to sensitivity in private data analysis.* Theory of Cryptography Conference, 2006. https://journalprivacyconfidentiality.org/index.php/jpc/article/view/405

[3] Bun, M., & Steinke, T. *Concentrated Differential Privacy: Simplifications, Extensions, and Lower Bounds.* Theory of Cryptography Conference (TCC), 2016. https://arxiv.org/abs/1605.02065

[4] Mironov, I. *Rényi Differential Privacy.* 2017. https://arxiv.org/abs/1702.07476