Optimizing Noise Distributions for Differential Privacy

Differential privacy (DP) [1], [2] has become the de facto standard for privacy-preserving machine learning. It provides rigorous guarantees by ensuring that the presence or absence of any individual record in the training data cannot be inferred from the model's outputs. These guarantees are typically enforced through mechanisms that inject carefully calibrated noise into the learning process, masking sensitive information while preserving utility. In practice, such mechanisms are applied repeatedly, with the number of compositions (sequential uses, such as queries or training updates) depending on the application. While Gaussian and Laplace noise are widely used due to their simplicity and analytical tractability, they are not always optimal. For example, the Cactus distribution is optimal in the large-composition regime [3], where mechanisms are applied thousands of times, whereas the Staircase mechanism is optimal for pure ε-DP in the single-composition setting [4]. In this work, we introduce a unified framework for designing noise distributions tailored to a fixed number of compositions. Our framework recovers known optimal distributions in the large- and single-composition regimes as special cases, while also yielding new mechanisms that significantly improve the privacy–utility tradeoff in moderate-composition settings.

We introduce a unified optimization framework for designing both continuous and discrete noise distributions for (ε, δ)-DP. Given user-defined parameters—(i) the failure probability δ, (ii) the number of compositions, (iii) query sensitivity, and (iv) average distortion per query (e.g., mean squared error)—the framework outputs an optimized additive noise distribution that minimizes the privacy budget ε while preserving the same distortion level.

At the heart of our approach is the minimization of Rényi DP, a composition-friendly relaxation of DP. By tuning the Rényi parameter α, the framework adapts to a wide range of user-defined settings. We reduce the resulting optimization to a finite-dimensional convex program and solve it efficiently using a preconditioned gradient descent algorithm implemented in JAX.

Empirical results demonstrate that our optimized noise distributions improve the privacy–utility tradeoff over standard baselines (Gaussian, Laplace, Cactus, and Staircase) in regimes where no optimal distribution is known, while recovering the Staircase and Cactus mechanisms in the regimes where they are optimal. The gains are especially significant in the moderate composition regime (10–40 compositions), which aligns well with practical applications involving sensitive tabular data—particularly in healthcare. Experiments on real-world datasets such as Breast Cancer and Diabetes demonstrate up to a 10% reduction in mean squared error (MSE) when targeting small privacy budgets ($\varepsilon < 1$), as summarized in Table I.

TABLE I
MSE improvement (%) of our noise distribution over the best-performing baseline (Gaussian, Laplace, Cactus, or Staircase) for $\delta = 10^{-6}$ and 10 compositions (mean ± std over 20 seeds).

Dataset	$\varepsilon = 0.62$	0.69	0.78	0.84	0.97
Breast Cancer	8.28±0.19	9.14±0.19	9.63±0.23	8.61±0.17	10.34±0.20
Diabetes	8.11±0.19	9.06±0.21	9.43±0.16	8.48±0.17	10.06±0.17

References

[1] C. Dwork, K. Kenthapadi, F. McSherry, I. Mironov, and M. Naor, "Our data, ourselves: Privacy via distributed noise generation," in EUROCRYPT, S. Vaudenay, Ed., 2006, pp. 486–503.

[2] C. Dwork, F. McSherry, K. Nissim, and A. Smith, "Calibrating noise to sensitivity in private data analysis," in Proc. Theory of Cryptography (TCC), Berlin, Heidelberg, 2006, pp. 265–284.

[3] W. Alghamdi, S. Asoodeh, F. P. Calmon, O. Kosut, L. Sankar, and F. Wei, "Cactus mechanisms: Optimal differential privacy mechanisms in the large-composition regime," in 2022 IEEE International Symposium on Information Theory (ISIT), 2022, pp. 1838–1843.

[4] Q. Geng, P. Kairouz, S. Oh, and P. Viswanath, "The staircase mechanism in differential privacy," IEEE Journal of Selected Topics in Signal Processing, vol. 9, no. 7, pp. 1176–1184, 2015.