TIGHT END-TO-END ACCOUNTING FOR DIFFERENTIALLY PRIVATE SYNTHETIC DATA

Anonymous authors
Paper under double-blind review

ABSTRACT

Differentially private (DP) pipelines often combine heterogeneous mechanisms, yet their guarantees are best analyzed using different notions of privacy (e.g., bounded range, zCDP, ADP). When composing them, we have to analyze the guarantees through a reduced common representation that loses accounting precision. This issue is particularly relevant in DP synthetic data generation, where data preprocessing, generation algorithms, and `DP-SGD` could appear within the same pipeline. We develop novel tools for tight end-to-end accounting of such compositions using privacy loss distributions (PLD), f-DP, and Gaussian DP (GDP). In particular, we derive PLD and GDP analyses of bounded range mechanisms such as the Exponential mechanism. Building on these theoretical results, we introduce `AIM-GDP`, a GDP-native variant of `AIM` (a state-of-the-art DP synthetic data mechanism) that enables easy composability with complex pipelines via PLD, attack-aware risk calibration, and tight end-to-end auditing while consistently achieving better utility; and `MixCal`, an algorithm for calibrating adaptive mixed compositions using PLD accounting. Overall, our tools enable tighter end-to-end accounting and thus less noise at the same privacy level in DP synthetic data pipelines.

1 INTRODUCTION

A core strength of differential privacy (DP) (Dwork et al., 2006; 2014) is its composability: we can construct highly complex randomized algorithms by combining multiple simpler primitives. At the same time, in modern data-analytic workflows, the components are highly heterogeneous, e.g., including different data preprocessing algorithms, and using combinations of learning algorithms.

Despite the composability of DP, there is a bottleneck when analyzing the privacy guarantees of such heterogeneous workflows. DP primitives have been developed using disparate analytical approaches, e.g., zero-concentrated DP (zCDP) (Bun & Steinke, 2016; Dwork & Rothblum, 2016) for compositions of private selection and simple noise addition algorithms (Dong et al., 2020), bounded range (BR) (Durfee & Rogers, 2019) for private selection, Rényi DP (RDP) (Mironov, 2017) for data-dependent compositions (Papernot et al., 2018) and early analyses of `DP-SGD` (Abadi et al., 2016), and privacy loss distributions (PLDs) (Meiser & Mohammadi, 2018) for more recent analyses of `DP-SGD` (Koskela et al., 2020; Gopi et al., 2021; Doroshenko et al., 2022). Because of this, when composing multiple primitives analyzed using different approaches, we have to reduce the potentially rich representations to "lowest common denominators," incurring loss in accounting precision.

As an example, consider `AIM` (McKenna et al., 2022), a state-of-the-art algorithm for generating synthetic tabular data with DP. Internally, this algorithm is a fully adaptive composition (Rogers et al., 2016) of Exponential and Gaussian mechanisms (Dwork et al., 2014). The standard way to perform accounting for such compositions is using zCDP. This notion, although it is convenient and tightly characterizes the Exponential mechanism, does *not* tightly characterize the trade-off curve (Dong et al., 2022) of the Gaussian mechanism (Zhu et al., 2022; Riess et al., 2026). Because of this, the resulting privacy accounting is lossy both (1) within `AIM` itself, and (2) when `AIM` is further composed with `DP-SGD` for multimodal synthetic data generation (Hu et al., 2026). In contrast, Gaussian DP (GDP) (Dong et al., 2022) optimally characterizes the Gaussian mechanism, and would compose well with `DP-SGD` or other pipelines using PLD accounting, but there are no general analyses of the Exponential mechanism in terms of GDP that do not require a special structure (Gopi et al., 2022).

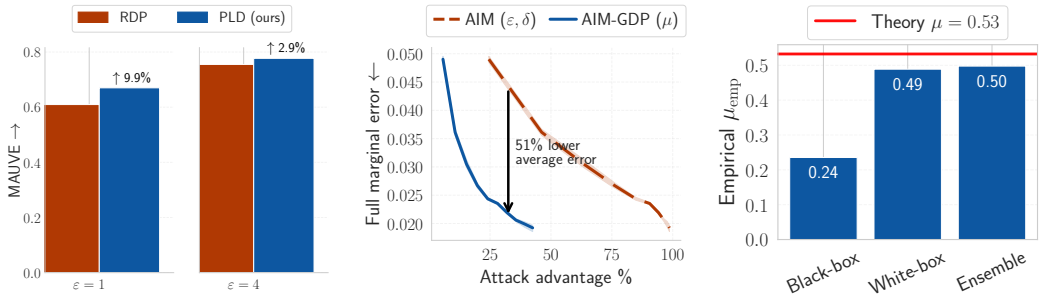
Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping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

Table 1: Novel theoretical results

	Mechanism	Analysis	Ref.
i	Bounded range	f-DP/GDP	Sec. 3
ii	Bounded range	PLD	Sec. 3
iii	Partition selection	PLD	App. B
iv	Poisson subsampling	PLD	App. B
v	Group privacy	Exponential	App. B

Table 2: Specific applications and algorithms/pipelines

Application	Algorithm/pipeline	Can use
Discrete synthesis	`AIM-GDP`	i
PLD calibration	`MixCal`	ii-iii
End-to-end synthesis	`AIM-GDP` + DP preproc. + `MixCal`	i-v
Multimodal synthesis	`AIM-GDP` + DP-SGD + `MixCal`	i-v



(a) Improved utility (b) Attack-aware calibration (c) First end-to-end audit

Figure 1: **Our methods provide improved utility and enable new capabilities in DP synthetic data workflows.** *Top:* Summary of theoretical results and algorithmic contributions. *Bottom:* (a) Our methods improve utility in synthetic text generation using LLMs. (b) Our methods enable to efficiently calibrate DP workflows to attack risk, with *the first application of attack-aware calibration to synthetic data.* (c) Our methods support tight end-to-end privacy audits, with *the first end-to-end audit of the synthetic data generator,* `AIM-GDP`, our version of `AIM`.

In this work, we improve the accounting precision by developing methods that are more suited for end-to-end accounting of realistic synthetic data workloads. Our core contributions are:

1. **PLD, f-DP, and GDP analyses of BR mechanisms for private selection.** Using these results, we show that GDP enables more precise analyses than zCDP in fully adaptive compositions where the majority of the composed mechanisms are Gaussian (as is the case in `AIM`). We provide additional results that enable the support of more workloads, e.g., involving private partition selection, data subsampling, and multiple user contributions in Appendix B (see summary in Table 1).
2. **`AIM-GDP`.** Building on the GDP analysis of the Exponential mechanism, we propose `AIM-GDP`, a variant of `AIM` that natively satisfies GDP. It provides comparable/better utility (Fig. 1a) with additional benefits. First, `AIM-GDP` enables the **first efficient attack-aware calibration for DP synthetic data** (Kulynych et al., 2024; 2025), enabling practitioners to set privacy budgets based on target risk levels (Fig. 1b). Second, it supports tight audits (Nasr et al., 2023; Annamalai et al., 2024), making it the **first end-to-end audited** implementation of `AIM` (Fig. 1c). Third, it is **easily composable** in complex pipelines, as GDP is compatible with tight PLD composition analyses.
3. **`MixCal`.** We introduce `MixCal`, a new algorithm for calibrating parameters in mixed-mechanism adaptive compositions that combines the advantages of zCDP or GDP for simple calibration of parameters across multiple components, and of PLD accounting for precision. Building on this, we extend `AIM-GDP` to more complex pipelines by combining it with DP preprocessing for end-to-end synthetic data generation and with `DP-SGD` for multimodal synthesis (see summary in Table 2).
4. We demonstrate the combinations of these algorithms in realistic synthetic data workflows, showing both performance improvements at the same privacy level and new capabilities. Upon acceptance, we will release open-source software implementing our methods.

2 Problem Statement

2.1 Motivation and Objective

Our main motivating setting is private tabular or multimodal synthetic data generation, even though, as we show next, our technical results are relevant to any mixed-mechanism compositions.

Discrete data. For synthetic tabular data, the best-performing DP mechanisms are marginal-based (Tao et al., 2022; Chen et al., 2025); which typically operate on discrete domains, and consist of a combination of selection steps (e.g., using the Exponential mechanism), and measurement steps (e.g., using the Gaussian mechanism). Some mechanisms (like `AIM`) further rely on *fully-adaptive* composition, privacy budget assigned to each Exponential and Gaussian mechanism is a data-dependent random variable.

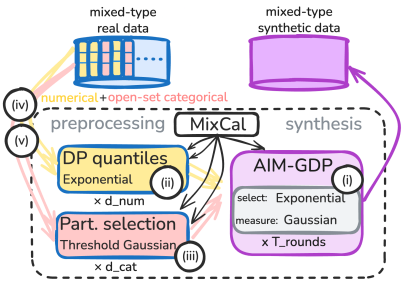


Figure 2: End-to-end synthesis.

End-to-end data. Beyond discrete data, real-world datasets often contain a mix of categorical attributes with potentially unknown domains (open-set attributes) and numerical attributes. To leverage marginal-based mechanisms, we must map our data to a discrete domain. For numerical attributes, data-dependent discretization, e.g., via DP quantiles (Kaplan et al., 2022), is a preferred approach (Ganev et al., 2025b). For open-set categorical attributes, private domain discovery (Gopi et al., 2020) poses an additional challenge, and, to the best of our knowledge, existing synthetic data implementations lack such support. Fig. 2 shows an example pipeline, which we discuss in detail in Section 4.

Multimodal data. Beyond tabular data, generating synthetic *unstructured* data usually involves using a `DP-SGD`-style mechanism (e.g., finetuning an LLM) (Ponomareva et al., 2026). ACTG, a state-of-the-art mechanism for this task consists of a synthetic tabular generation step plus *conditional* finetuning step (Hu et al., 2026). This work relies on RDP accounting due to the reliance on `AIM` that only provides a zCDP guarantee.

We provide a review of related mechanisms for data preprocessing and synthetization used in DP synthetic data pipelines in Table 6 in the Appendix.

Our goal. The objective of our work is to provide methods for tight end-to-end privacy analyses for such heterogeneous pipelines. We observe two key challenges: (1) we lack precise analyses of mixed-mechanism compositions involving the Exponential mechanism for both fully adaptive and adaptive settings, and Threshold Gaussian, and (2) it is unclear how to calibrate the parameters of the component mechanisms in such composed pipelines. In the remainder, we introduce several fundamental results, and new algorithms, to solve these challenges.

2.2 Technical Background

We provide a concise overview of the technical notions we use. For a more in-depth treatment, we refer the reader to Zhu et al. (2022), Gomez et al. (2026) or any of the cited works in this section. Throughout, P and Q are probability distributions on a measurable space $\mathbb{Y}$. In an abuse of notation, we also write $P(y)$ and $Q(y)$ to denote their densities with respect to a common σ-finite measure, e.g., $P+Q$. We write $(x)_+ \triangleq \max\{x, 0\}$ and use the conventions $\log(a/0) = +\infty$ and $\log(0/a) = -\infty$ for $a > 0$, and $e^{-\infty} = 0$.

We begin by defining two divergences: the hockeystick divergence of order $e^\varepsilon > 0$ and the Rényi divergence of order $\tau > 1$, respectively:

$$H_{e^\varepsilon}(P \parallel Q) \triangleq \sup_{E \subseteq \mathbb{Y}} (P(E) - e^\varepsilon Q(E)), \quad D_\tau(P \parallel Q) \triangleq \frac{1}{\tau - 1} \log \mathbb{E}_{y \sim Q} \left(\frac{P(y)}{Q(y)} \right)^\tau .$$

With these divergences, we can define two notions of differential privacy (DP). A mechanism $M(\cdot)$ satisfies (ε, δ)-DP if $H_{e^\varepsilon}(M(S) \parallel M(S')) \leq \delta$ for any *neighbouring dataset* pair $S \simeq S'$. A mechanism satisfies $(\tau, \varepsilon_{\text{RDP}})$-Rényi DP (RDP) if $D_\tau(M(S), M(S')) \leq \varepsilon_{\text{RDP}}$. Zero-concentrated DP (zCDP) (Dwork & Rothblum, 2016; Bun & Steinke, 2016) is a special case of RDP. A mechanism is ρ-zCDP if it is $(\tau, \tau \cdot \rho)$-RDP for any $\tau > 1$.

The function $\delta : \mathbb{R} \to [0, 1]$ such that a mechanism satisfies $(\varepsilon, \delta(\varepsilon))$-DP for any $\varepsilon \in \mathbb{R}$ is called a *privacy profile* (Balle et al., 2018), and is closely related to the operational formulation of DP known as f-DP (Dong et al., 2022). Let $T(P \parallel Q)$ denote the trade-off function which maps the Type I error (denoted by α) of the most powerful test distinguishing $H_0 \sim P$ and $H_1 \sim Q$ to its Type II error (denoted by β). We say a mechanism satisfies f-DP if $f = T(P \parallel Q)$ for some pair (P, Q) and $f \leq T(M(S) \parallel M(S'))$ for all $S \simeq S'$. An important special case of f-DP is Gaussian DP (GDP).

We say a mechanism is μ-GDP if it satisfies f_μ-DP, with $f_\mu(\alpha) = \Phi(\Phi^{-1}(1-\alpha) - \mu)$, where $\Phi(t)$ is the CDF of the standard normal distribution (Dong et al., 2022).

The celebrated Gaussian mechanism (Dwork et al., 2014), which adds Gaussian noise $\mathcal{N}(0, \sigma^2 \cdot I_d)$ to a vector query q with $\|q(S) - q(S')\|_2 \leq 1$ for any pair $S \simeq S'$, is μ-GDP with $\mu = 1/\sigma$. The Gaussian mechanism also satisfies ρ-zCDP with $\rho = \mu^2/2$.

One of the most fundamental properties about differential privacy is *composition*, where the outputs of one mechanism are used as the input of another. For example, the composition of M_1 and M_2 is $M_{\text{composed}} = M_1(D, M_2(D))$. We refer to M_1 and M_2 as *primitive* mechanisms. We distinguish between two types of composition: the first, which we call *adaptive composition* or simply *composition*, is when the number of primitive mechanisms and the choice of all (ε, δ) is fixed ahead of time. The second type, which we call *fully adaptive composition*, occurs when the number of primitive mechanisms is not known, and only a fixed budget (ε, δ) is known ahead of time (Rogers et al., 2016).

RDP admits a simple composition rule: if M_1 satisfies $(\tau, \varepsilon_{\text{RDP}})$-RDP and M_2 satisfies $(\tau, \varepsilon'_{\text{RDP}})$, then the composed mechanism satisfies $(\tau, \varepsilon_{\text{RDP}} + \varepsilon'_{\text{RDP}})$-RDP. For ρ-zCDP, the composition rule is particularly simple: sum the ρ of the components. This RDP composition result holds for *both* adaptive and fully-adaptive composition (Feldman & Zrnic, 2021). GDP composition also admits a simple form $\sqrt{\mu_1^2 + \mu_2^2}$, and this holds in fully-adaptive setting (Smith & Thakurta, 2022; Koskela et al., 2023). Tight composition results for (ε, δ)-DP and its generalizations (i.e. privacy profiles, f-DP) are not as simple, and there are different results for adaptive and fully-adaptive composition.

For adaptive composition, tight composition results can be derived via dominating pairs (Zhu et al., 2022). A pair of distributions P, Q is a *dominating pair* for a mechanism $M(\cdot)$ if for any $\varepsilon \in \mathbb{R}$, we have $H_{e^\varepsilon}(P \parallel Q) \geq \sup_{S \simeq S'} H_{e^\varepsilon}(M(S) \parallel M(S'))$. We say (P, Q) is a tightly dominating pair if the inequality is equality. Given a dominating pair (P, Q), composition becomes simple: if M_1 has dominating pair (P_1, Q_1) and M_2 has dominating pair (P_2, Q_2), then the composed mechanism has dominating pair $(P_1 \otimes P_2, Q_1 \otimes Q_2)$.

Privacy Loss Distributions (PLDs) and *Privacy Loss Random Variables* (PLRVs) (Sommer et al., 2019) make this analysis even simpler. Given a dominating pair (P, Q), define the random variable $\omega_{P/Q} \triangleq \log {}^{P(y)}/_{Q(y)}$ with $y \sim P$ as the privacy loss random variable, and, its distribution as the PLD. If M_1 has PLRV ω_{P_1/Q_1} and M_2 has PLRV ω_{P_2/Q_2}, then the composed mechanism has PLRV $\omega_{P_1/Q_1} + \omega_{P_2/Q_2}$. Because this addition is over independent random variables, the composed mechanism's PLD is a convolution of the individual PLDs. Given a PLD, we can compute the privacy profile as $\delta(\varepsilon) = \mathbb{E}[(1 - e^{\varepsilon - \omega_{P/Q}})_+]$. Thus, PLDs provide a recipe for tight composition of mechanisms: first, compute the PLD of each primitive, then convolve their PLDs and compute the composed profile or the trade-off function from the convolved PLD.

For fully-adaptive composition, the PLD approach generally does not provide valid bounds (Regehr et al., 2026; Tran et al., 2026). As far as we are aware, the tightest known bounds for general heterogeneous fully-adaptive composition of mechanisms are ρ-zCDP or μ-GDP (Smith & Thakurta, 2022; Koskela et al., 2023).

We conclude this section with one last privacy representation: *bounded range* (BR) (Durfee & Rogers, 2019). We say a mechanism $M : 2^{\mathbb{D}} \to \mathbb{Y}$ satisfies ν-bounded range (BR) if for any $S \simeq S'$ there exists $0 \leq t \leq \nu$ such that for all measurable $E \subseteq \mathbb{Y}$:

$$\log\left(\frac{\Pr[M(S) \in E]}{\Pr[M(S') \in E]}\right) \in [-t, -t + \nu]. \tag{1}$$

Bounded range is especially suitable to characterizing the standard Exponential mechanism. An *Exponential mechanism* with a discrete output space $\mathbb{Y}$ returns $\Pr[M(S) = y] \propto \exp\left(-\nu/2 \cdot q(S, y)\right)$, where $q : 2^{\mathbb{D}} \to \mathbb{R}$ is a query such that for any $S \simeq S'$ and $y \in \mathbb{Y}$, we have $|q(S, y) - q(S', y)| \leq 1$. The exponential mechanism satisfies ν-BR. Notably, it also satisfies zCDP with $\rho = 1/8 \cdot \nu^2$ (Durfee & Rogers, 2019), with a tighter expression recently given by Harrison & Manurangsi (2025).

3 New Building Blocks for Heteregeneous Accounting

In this section, we provide several new theoretical results that we later use as building blocks.

Novel analyses of Bounded Range mechanisms. We introduce a tight characterization of bounded-range mechanisms by deriving the its privacy profile, trade-off function, and PLD. From the trade-off function, we also derive a tight GDP bound. The latter is useful for analyzing *fully adaptive compositions* of mechanisms involving BR, such as the `AIM` algorithm for synthetic data generation.

Theorem 3.1. *Let $M(\cdot)$ satisfy ν-BR. The density $g_\nu : \mathbb{R} \to \mathbb{R}$ of its tightly dominating PLD is:*

$$g_\nu(x) = D_\nu \, e^{x/2} \, \mathbf{1}_{[-\nu,\nu]}(x), \quad \text{where } D_\nu = \tfrac{1}{2}\left(e^{\nu/2} - e^{-\nu/2}\right)^{-1}. \tag{2}$$

Moreover, $M(\cdot)$ satisfies f-DP and has privacy profile $\delta(\varepsilon)$:

$$f(\alpha) = \frac{1-\alpha}{\alpha e^{\nu} + 1 - \alpha}, \quad \delta(\varepsilon) = \begin{cases} 1 - e^{\varepsilon}, & \varepsilon < -\nu, \\ \frac{\left(e^{\nu/2} - e^{\varepsilon/2}\right)^2}{e^{\nu}-1}, & -\nu \le \varepsilon \le \nu \\ 0, & \varepsilon \ge \nu \end{cases} \tag{3}$$

As they are derived from the tightly dominating PLD, these two bounds are tight: every point (i.e. every $(\varepsilon, \delta(\varepsilon))$ or $(\alpha, f(\alpha))$ is achieved by some ν-BR mechanism.

Theorem 3.2. *Suppose that a mechanism $M(\cdot)$ satisfies ν-BR. Then, it also satisfies μ-GDP with:*

$$\mu = -2\Phi^{-1}\left(\tfrac{1}{e^{\nu/2}+1}\right). \tag{4}$$

We provide the proofs of both statements in Appendix A. There are various ways of comparing our new bounds to the previous tightest known zCDP bounds. We choose to compare the trade-off functions due to their interpretability We show the optimal bounded-range trade-off function, our derived GDP trade-off function, and the tightest known trade-off function implied by zCDP in Fig. 3. We observe that our optimal trade-off function provides a tighter bound than zCDP in all regimes, and our GDP bound provides a tighter bound than zCDP in some, but not all regimes.

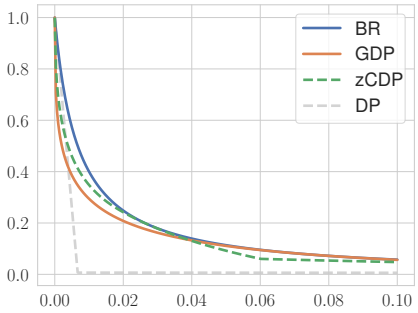


Figure 3: Trade-off curves of a BR mechanism (higher is more precise) with $\nu = 4$, given by our f-DP (Theorem 3.1) and GDP analyses (Theorem 3.2), and zCDP using the recent optimal conversion (Harrison & Manurangsi, 2025). To show the trade-off curves given by zCDP we use the conversion from RDP to f-DP (Riess et al., 2026), for orders > 1.

As we consider compositions of BR mechanisms, it is natural to ask if our analyses actually yield better results than zCDP. In practice, we care about compositions of mixed-mechanisms, often compositions of Gaussians and Exponential mechanisms. Fig. 4 illustrates that our GDP analysis is tighter for mixed-mechanism compositions containing BR mechanisms where the majority of the budget is spent in Gaussian mechanisms, which is the setting of `AIM`, as we leverage next; and the PLD analysis is tighter for non-fully adaptive compositions, which is the setting of preprocessing workloads.

Subsampling, group privacy, tight partition selection. We provide a number of additional analyses for mechanisms of interest in various realistic data generation pipelines. All of these results are deliberately designed to be compatible with open-source software PLD libraries such as `dp_accounting`. Due to space constraints, we defer details to Appendix B.

Poisson subsampling of arbitrary mechanisms. We provide a method for converting a PLD of an arbitrary mechanism to the PLD for the mechanism wrapped with Poisson subsampling (Abadi et al., 2016). Previously, only additive mechanisms like Gaussian and Laplace mechanisms were supported. By enabling Poisson subsampling for arbitrary mechanisms, it becomes possible scale to much larger datasets than used in this work. We leave this possibility as future work.

Group privacy for Exponential mechanisms on partitioned datasets. We demonstrate that partitioning a dataset and then running a ν-BR mechanism on each subset of the partition satisfies $k\nu$-bounded range for groups of k records. This enables a privacy analysis of mechanism such as DP quantiles (Kaplan et al., 2022) even when users can contribute multiple records.

Characterization of Threshold Gaussian and Threshold Laplace. The Threshold Gaussian and Laplace mechanisms (Gopi et al., 2020) are useful mechanisms for DP partition selection. We provide a tighter characterization of its privacy guarantees in terms of PLD and privacy profiles.

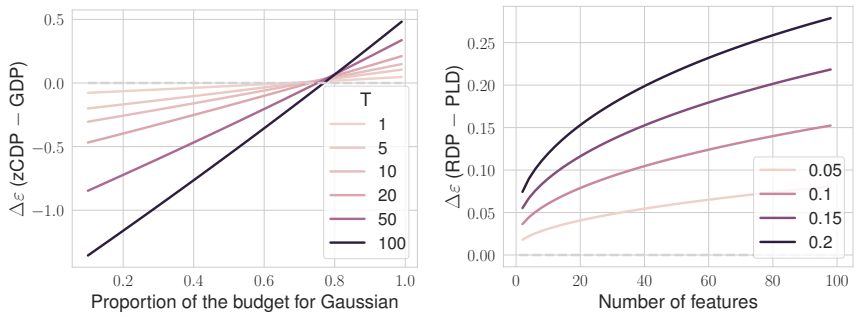


Figure 4: **Our methods enable tighter accounting in the regimes of interest.** On y axis, we show the ε decrease (> 0 means our methods are better). *Left:* Using GDP, as a function of the budget proportion of Gaussian mechanisms (x axis) and number of iterations (hue) in `AIM`-like workloads containing fully adaptive compositions of Exponential and Gaussian mechanisms. *Right:* Using PLD, as a function of the number of features (x axis) and the noise budget (hue) in data preprocessing-like workloads containing compositions of Exponential and Threshold Gaussian mechanisms.

4 Applications

Next, we use the results in Section 3 to develop better algorithms for heterogeneous data analysis.

Discrete synthetic data with `AIM-GDP`. The developments in Section 3 can be used to improve the privacy analysis of any mechanism that uses the exponential mechanism as a primitive. One such example is `AIM` (McKenna et al., 2022), a state-of-the-art mechanism for synthetic data generation that operates within the select-measure-generate paradigm. `AIM` selects the worst-approximated marginal using Exponential mechanism, "measures" the selected marginal query using Gaussian mechanism, and repeats these two steps until the budget is depleted. Afterwards, a synthetic dataset is generated based on the measured marginals about real data. `AIM` is designed to spend the majority of the budget on the marginal measurements (Gaussian mechanism) because the signal is divided among the histogram bins and is hence more sensitive to noise. Whereas the original algorithm uses zCDP, our new GDP analysis in Section 3 provides improvements in exactly the regime where the majority of the budget in a mixed-mechanism composition is spent in Gaussian mechanisms. As a result, a GDP analysis of `AIM` is tighter than the original zCDP analysis. We present algorithms for both the original `AIM`, and our version, `AIM-GDP`, in Appendix E.

End-to-end synthetic data with `AIM-GDP`, DP preprocessing, and `MixCal`. One downside of `AIM` is that it can only be run on datasets with closed-domain or discrete attributes. In contrast, real-world tabular datasets typically contain a mix of numerical attributes and open-set categorical attributes with unbounded vocabularies (Fig. 2). Applying `AIM` to real datasets thus requires preprocessing these attributes with DP. As this is done on top of `AIM` without referencing the privacy budget, preprocessing workloads can be thought of as primitives that we compose *non*-fully adaptively with `AIM`. For numerical attributes, we focus on using DP quantiles via recursive bisection (Kaplan et al., 2022) (see Algorithm 8 in the Appendix for reference). Because this algorithm is a composition of standard Exponential mechanisms, it directly benefits from our PLD characterization of BR. For open-set categorical attributes, existing implementations either perform data-dependent domain construction without accounting for it (Qian et al., 2023), or do not support it at all (OpenDP, 2021; Mahiou et al., 2025). We introduce the first principled integration of private partition selection, e.g., the Threshold Gaussian mechanism (Gopi et al., 2020) (see Algorithm 9 in the Appendix for reference), into synthetic data pipelines, using our new PLD characterization to tightly account for domain discovery. Concretely, given a tabular dataset with numerical and open-set categorical attributes, first, we apply DP Quantiles and the Threshold Gaussian mechanism to map these attributes to discrete ones, and then, apply `AIM-GDP` to generate synthetic data. The Gaussian PLD from `AIM-GDP`, the PLDs of private set partition via Threshold Gaussian, and of discretization via DP quantiles (we introduce both in Appendix B), are composable in a straightforward way using standard PLD accountants. Taken together, these results extend the usage of `AIM` to a larger class of tabular datasets. Moreover, in such

workflows, different components might measure the same queries. In Appendix B, we also describe how to enable the re-use of such query outputs.

Calibrating the components. In practice, however, there is a salient challenge when using PLD accounting in such workflows. We normally have a target guarantee, e.g., $(\varepsilon^\star, \delta^\star)$, or (α, β) for f-DP based attack-aware calibration (see Section 5). Therefore, we require a method to assign noise parameters (e.g. Gaussian parameter σ, Exponential Mechanism parameter ν) to the primitive mechanisms such that the composed PLD satisfies the target guarantee. A key challenge of this approach is that there is no clear way to decompose our target such that $\varepsilon_{\text{preprocess}} + \varepsilon_{\text{AIM}} = \varepsilon^\star$. This makes it difficult to "assign" more privacy budget to `AIM` and less to preprocessing, which is often desirable in practice.

To solve this problem, we introduce a "dummy" privacy budget, denoted as b. As a simple example, suppose we are handling a composition of Gaussians, Exponential Mechanisms, and `AIM-GDP`, all of which have a known PLD and ρ-zCDP guarantee. Starting from, e.g., the target $(\varepsilon^\star, \delta^\star)$, we use a conversion (e.g., Bun & Steinke, 2016) to zCDP to obtain the dummy privacy budget $b = \rho(\varepsilon^\star, \delta^\star)$. If we had a weight w_1 for preprocessing and weight w_2 for `AIM-GDP` such that $w_1 + w_2 = 1$, we calibrate the preprocessing noise parameters to satisfy $w_1 \cdot b$-zCDP budget and calibrate `AIM-GDP` to satisfy $w_2 \cdot b$-zCDP budget. We then use PLD accounting to analyze the composed mechanism. As our calibration itself is not precise, we would add more noise than necessary. We thus perform binary search over b: if the resulting ε is larger than the target ε, we decrease b. Otherwise, we increase it. This allows us to meaningfully assign more or less privacy budget to preprocessing vs `AIM-GDP`, while preserving the tighter analysis of PLD over zCDP. This algorithm can be generalized to use any "dummy" budget that is additive under composition, e.g., GDP, and to variable numbers of compositions. We call this algorithm `MixCal` and present it in Algorithm 1 in the Appendix.

Multimodal synthetic data with `AIM-GDP`, `DP-SGD`, and `MixCal`. Recent work in synthetic text data generation with LLMs used `AIM` as a building block. For instance, the ACTG mechanism (Hu et al., 2026) consists of several steps: (1) annotate input text data using a predefined tabular schema, (2) run `AIM` on the generated tabular data with a fraction of the budget, and (3) run `DP-SGD` to finetune a language model to predict the text given the tabular metadata (see Fig. 5). The privacy analysis is hence a composition of `AIM` with `DP-SGD`, and because `AIM` satisfies zCDP, Hu et al. (2026) used RDP instead of the tighter PLD accounting. By replacing `AIM` with `AIM-GDP`, we can analyze the algorithm end-to-end with PLD, and use `MixCal` to allocate the budget. Beyond precision, PLD accounting enables more capabilities, such as using random allocation (Feldman & Shenfeld, 2025) instead of Poisson subsampling in `DP-SGD`, which provides better utility in most regimes and is only analyzable using PLD, and subsampling large datasets using our new analyses in Appendix B.

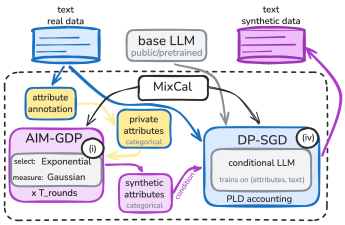


Figure 5: Multimodal synthesis.

Group privacy. In practice, users could contribute multiple items to a dataset. E.g., an author can contribute multiple papers to a repository like bioRxiv. In these cases, example-level privacy is not a sufficient safeguard. In Appendix B, we provide the group privacy analyses for the key primitives that enable valid analyses under multiple contributions for the mechanisms of interest in this paper.

5 Experimental Evaluation

We now evaluate the algorithms introduced in Section 4 across three synthetic data generation applications: discrete, end-to-end, and multimodal.

5.1 Discrete Synthetic Data with `AIM-GDP`

We evaluate `AIM-GDP` on four standard discrete datasets (with known categorical features; see Table 5 in the Appendix) across four privacy budgets, $\varepsilon \in \{0.5, 1, 2, 5\}$, $\delta = 10^{-5}$, with five runs per setting. We assess utility using the mean TV distance between the marginal distributions of real data and model projections, computed over 1/2/3-way and full workloads. We also demonstrate two additional capabilities enabled by `AIM-GDP`: attack-aware risk calibration, and tight end-to-end privacy auditing.

Table 3: **AIM-GDP consistently shows better utility.** We show 1/2/3-way and full marginal errors, number of measured marginals, and runtime, averaged across four datasets and four privacy budgets.

	Algorithm	1-way	2-way	3-way	Full	#Marginals	Time (s)
Discrete	AIM	0.0035	0.0224	0.0556	0.0451	40.33	**399.2**
	AIM-GDP (ours)	**0.0033**	**0.0216**	**0.0542**	**0.0439**	**40.68**	404.6
End-to-end	AIM	0.0038	0.0284	0.0763	0.0613	42.20	**426.3**
	AIM-GDP (ours)	**0.0036**	**0.0274**	**0.0746**	**0.0599**	**42.38**	444.8

Utility. AIM-GDP consistently outperforms the standard AIM analyzed with zCDP, reducing marginal error by 3% on average across datasets/privacy budgets. See top of Table 3 for a summary. In the Appendix, we present detailed comparisons (Fig. 8), and additional diagnostics over 25 runs on the Adult dataset with $\varepsilon = 1$, which show that AIM-GDP adds less noise per measurement step, enabling it to run for more iterations (Fig. 9), while selecting a very similar set of marginals (Fig. 10). Thus, out tighter accounting improves utility without substantially changing AIM's underlying behavior. These results provide evidence that **AIM-GDP can serve as a drop-in replacement for standard AIM**.

Attack-aware calibration. Choosing ε is challenging in practice because it does not directly quantify an operational privacy risk. In contrast, the μ-GDP guarantee of AIM-GDP, and the PLD representation of AIM-GDP after composition with data preprocessing, directly characterize (Kulynych et al., 2024) the full trade-off of any membership inference attacks (MIAs) (see, e.g., Shokri et al., 2017; Carlini et al., 2022), which is a standard way to quantify the operational privacy risk. This enables practitioners to specify an operational risk target, e.g., a maximum attack true positive rate (TPR) at a given low false positive rate (FPR), and calibrate AIM-GDP directly to it, rather than first selecting (ε, δ) and subsequently translating it to attack risk (Kulynych et al., 2024). See Appendix D for details. In Fig. 1b we show the results of calibrating AIM-GDP to attack advantage (Yeom et al., 2018), a standard risk measure that equals maximum TPR $-$ FPR of MIAs. As the plot shows, attack-aware calibration substantially improves the utility at the same level of risk (e.g., by 51% on Adult). To the best of our knowledge, this is the first such application to DP synthetic data.

Auditing. It is crucial to audit (see, e.g. Nasr et al., 2023; Annamalai et al., 2024; Ganev et al., 2025a) practical implementations of DP algorithms, as they often introduce software bugs that invalidate the theoretical guarantees. Our GDP analysis enables us to tightly audit the privacy guarantees of the AIM-GDP. To do so, we train 10,000 models on worst-case neighboring datasets and use the FPR/TPR of a MIA to statistically estimate a lower bound on the empirical privacy parameter μ_{emp}. See Appendix D for details. As shown in Fig. 1c, our white-box and ensemble attacks achieve $\mu_{\text{emp}} = 0.49$ and 0.50, respectively, closely approaching the theoretical bound of $\mu = 0.53$. To the best of our knowledge, this is the first tight end-to-end empirical audit of fully adaptive AIM, i.e., including its data-dependent marginal selection, as prior work restricted AIM to fixed marginal structures (e.g., all 1-way marginals) and did not establish comparable tightness (Ganev et al., 2026).

5.2 End-to-End Synthetic Data with Data Preprocessing

We now evaluate AIM-GDP in an end-to-end synthetic data pipeline, following the evaluation setup from the previous section, but privately preprocessing every column via either DP quantiles or Threshold Gaussian, as described in Section 4. We use MixCal with zCDP as "dummy" budget to calibrate the pipeline components, allocating 10% for preprocessing and 90% for synthetization.

As before, AIM-GDP outperforms AIM (with RDP accounting), reducing marginal error by over 2% across datasets (see bottom of Table 3 and Fig. 11 in the Appendix). Compared with the discrete setting in Section 5.1, errors decrease for predominantly categorical datasets but increase for predominantly numerical ones, showing that there are avenues for improvement in numerical preprocessing mechanisms.

Next, we illustrate how our end-to-end analyses can be used to support large-scale settings (e.g., more than 200K records) using the Poisson subsampling results (see Section 3). For this, we simulate a workload with 16 numerical attributes, each discretized into 16 bins using 4 iterations of DP Quantiles. The resulting pipeline composes 64 Exponential mechanisms. We compare our tight PLD accounting,

Table 4: **Our analyses enable to add substantially less noise in ACTG pipelines (combination of synthetization and DP-SGD).** The table shows DP-SGD noise multipliers σ_{SGD} at fixed iterations $T = 512, K = 8, \delta = 10^{-6}$ for varying values of ε. Poisson subsampling uses $q = K/T$.

ε	RDP (Poisson)	PLD (Poisson)	PLD (Random Alloc.)
1	2.34	**2.18**	2.25
2	1.40	1.32	**1.30**
4	0.97	0.92	**0.89**
8	0.74	0.71	**0.69**

which applies subsampling directly in PLD space, against an RDP baseline that first bounds the Exponential mechanisms via zCDP. Our analysis yields a lower ε across all sampling rates, improving over RDP by 5.6% without subsampling ($q = 1$) and by *up to 66.7%* at $q = 2^{-6}$, with the gains increasing as the subsampling probability decreases. We show the results in Fig. 12 in the Appendix.

5.3 Generating Synthetic Text Data with LLMs and Tabular Conditioning

To further demonstrate the benefits our methods, we replicate the attribute-conditioned text generation (ACTG) pipeline (Hu et al., 2026). This work has improved on existing DP synthetic text evaluations (Xie et al., 2024; Tan et al., 2025), and found that ACTG outperformed strong baselines such as vanilla DP-SGD, private evolution (Xie et al., 2024), and CTCL (Tan et al., 2025). ACTG assumes records that consist of structured tabular attributes and text. ACTG uses AIM to produce a synthetic dataset of structured attributes, then finetunes a language model with DP-SGD to generate text conditioned on these attributes. We reproduce this pipeline on the *bioRxiv* dataset using the Gemma 3 1B base model. ACTG has a reinforcement learning stage, but, as it is independent of the privacy analysis, we omit it in our demonstration. We compare the original AIM vs. AIM-GDP with end-to-end PLD accounting using MixCal. We used budget allocation weights that replicate the allocation used by Hu et al.. See further details in Appendix D.

We report our results in terms of the standard MAUVE (Pillutla et al., 2023) scores in Fig. 1a. We observe consistent improvement in using PLD with MixCal and AIM-GDP, over RDP accounting with AIM. This gain is due to a combination of using AIM-GDP over AIM and using lower noise parameters for DP-SGD when using PLD accounting compared to RDP accounting. We show the improvements in noise DP-SGD multipliers in Table 4, using Poisson subsampling and Random Allocation, which is a new capability enabled by PLD accounting.

6 Conclusions

We unified disparate DP primitives under a single analytical and algorithmic framework, providing both qualitative and quantitative advantages for DP synthetic data pipelines with heterogeneous components. To support this, we presented GDP and PLD analyses for bounded range mechanisms, alongside a series of other results such as the PLD analysis of private partition selection (e.g., Threshold Gaussian). These analyses provide a versatile toolbox for future DP mechanism design. Applying these theoretical results, we introduced MixCal and AIM-GDP to calibrate and tightly account for mixed-mechanism workloads. By eliminating analytical slack in heterogeneous composition, our framework yields immediate utility gains across tabular and multimodal use cases without requiring major mechanism modifications. Ultimately, this work establishes a principled foundation for building and tightly accounting for increasingly complex, multimodal DP mechanisms.

Limitations. Our GDP and PLD analyses of BR mechanisms do not provide tighter accounting than standard zCDP in *all* cases. E.g., as we show in Section 3, GDP is only beneficial in heterogeneous pipelines when the majority of the budget is used for the Gaussian mechanisms, which is the case for AIM, but not other fully adaptive workflows, e.g., interactive statistical queries. Our MixCal algorithm enables allocating the budget according to pre-specified weights, but we do not introduce a way for optimally choosing these weights. Finding optimal way to do so is an avenue for future work. Finally, although the GDP analysis used in AIM-GDP is more precise than the zCDP analysis, it is not as tight as it would have been had the algorithm used only homogeneous components internally. Future work could compare our version to possible variants of AIM that, e.g., only use Gaussian mechanisms, which would enable even tighter analyses.

AI USE STATEMENT

We have used AI-based systems to assist with the implementation of the code for the experiments, visualizations, and LaTeX source code, as well as with proof ideation and verification. All the final text and proofs are written by the authors.

REFERENCES

Martin Abadi, Andy Chu, Ian Goodfellow, H Brendan McMahan, Ilya Mironov, Kunal Talwar, and Li Zhang. Deep learning with differential privacy. In *ACM SIGSAC conference on computer and communications security*, 2016.

Meenatchi Sundaram Muthu Selva Annamalai, Georgi Ganev, and Emiliano De Cristofaro. "What do you want from theory alone?" Experimenting with Tight Auditing of Differentially Private Synthetic Data Generation. In *USENIX Security*, 2024.

Sergül Aydöre, William Brown, Michael Kearns, Krishnaram Kenthapadi, Luca Melis, Aaron Roth, and Ankit Siva. Differentially private query release through adaptive projection. In *International Conference on Machine Learning (ICML)*. 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 (NeurIPS)*, 2018.

Rina Foygel Barber and John C Duchi. Privacy and statistical risk: Formalisms and minimax bounds. *arXiv preprint arXiv:1412.4451*, 2014.

David Blackwell. Equivalent Comparisons of Experiments. *The Annals of Mathematical Statistics*, 24 (2):265 – 272, 1953. doi: 10.1214/aoms/1177729032. URL https://doi.org/10.1214/aoms/1177729032.

Mark Bun and Thomas Steinke. Concentrated differential privacy: Simplifications, extensions, and lower bounds. In *Theory of cryptography conference*, 2016.

Nicholas Carlini, Steve Chien, Milad Nasr, Shuang Song, Andreas Terzis, and Florian Tramer. Membership inference attacks from first principles. In *2022 IEEE Symposium on Security and Privacy (SP)*, 2022.

Kai Chen, Xiaochen Li, Chen Gong, Ryan McKenna, and Tianhao Wang. Benchmarking Differentially Private Tabular Data Synthesis:[Experiments & Analysis]. *PACMMOD*, 2025.

Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the ACM SIGKDD international conference on knowledge discovery and data mining*, pp. 785–794, 2016.

Giovanni Cherubin, Boris Köpf, Andrew Paverd, Shruti Tople, Lukas Wutschitz, and Santiago Zanella-Béguelin. Closed-form bounds for DP-SGD against record-level inference attacks. In *33rd USENIX Security Symposium (USENIX Security 24)*, pp. 4819–4836, 2024.

Damien Desfontaines, James Voss, Bryant Gipson, and Chinmoy Mandayam. Differentially private partition selection. *Proceedings on Privacy Enhancing Technologies (PoPETs)*, 2022(1), 2022.

Jinshuo Dong, David Durfee, and Ryan Rogers. Optimal differential privacy composition for exponential mechanisms. In *International Conference on Machine Learning*, pp. 2597–2606. PMLR, 2020.

Jinshuo Dong, Aaron Roth, and Weijie J Su. Gaussian differential privacy. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 2022.

Vadym Doroshenko, Badih Ghazi, Pritish Kamath, Ravi Kumar, and Pasin Manurangsi. Connect the dots: Tighter discrete approximations of privacy loss distributions. *Proceedings on Privacy Enhancing Technologies*, 2022.

David Durfee and Ryan M Rogers. Practical differentially private top-k selection with pay-what-you-get composition. *Advances in Neural Information Processing Systems*, 32, 2019.

Cynthia Dwork and Guy N Rothblum. Concentrated differential privacy. *arXiv preprint arXiv:1603.01887*, 2016.

Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. In *Proceedings of the Theory of Cryptography Conference*, 2006.

Cynthia Dwork, Aaron Roth, et al. The algorithmic foundations of differential privacy. *Foundations and Trends in Theoretical Computer Science*, 2014.

Vitaly Feldman and Moshe Shenfeld. Privacy amplification by random allocation. *arXiv preprint arXiv:2502.08202*, 2025.

Vitaly Feldman and Tijana Zrnic. Individual privacy accounting via a renyi filter. *Advances in Neural Information Processing Systems*, 34:28080–28091, 2021.

Georgi Ganev, Meenatchi Sundaram Muthu Selva Annamalai, and Emiliano De Cristofaro. The elusive pursuit of reproducing PATE-GAN: benchmarking, auditing, debugging. *TMLR*, 2025a.

Georgi Ganev, Meenatchi Sundaram Muthu Selva Annamalai, Sofiane Mahiou, and Emiliano De Cristofaro. The Importance of Being Discrete: Measuring the Impact of Discretization in End-to-End Differentially Private Synthetic Data. In *ACM CCS*, 2025b.

Georgi Ganev, Meenatchi Sundaram Muthu Selva Annamalai, and Bogdan Kulynych. Tight Auditing of Differential Privacy in MST and AIM. In *TPDP*, 2026.

Quan Geng and Pramod Viswanath. Optimal noise adding mechanisms for approximate differential privacy. *IEEE Transactions on Information Theory*, 2015.

Jennifer Gillenwater, Matthew Joseph, and Alex Kulesza. Differentially private quantiles. In *International Conference on Machine Learning (ICML)*. PMLR, 2021.

Juan Felipe Gomez, Bogdan Kulynych, Georgios Kaissis, Flavio du Pin Calmon, Jamie Hayes, Borja Balle, and Antti Honkela. Gaussian dp for reporting differential privacy guarantees in machine learning. In *2026 IEEE Conference on Secure and Trustworthy Machine Learning (SaTML)*. IEEE, 2026.

Sivakanth Gopi, Pankaj Gulhane, Janardhan Kulkarni, Judy Hanwen Shen, Milad Shokouhi, and Sergey Yekhanin. Differentially private set union. In Hal Daumé III and Aarti Singh (eds.), *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 3627–3636. PMLR, 13–18 Jul 2020. URL https://proceedings.mlr.press/v119/gopi20a.html.

Sivakanth Gopi, Yin Tat Lee, and Lukas Wutschitz. Numerical composition of differential privacy. *Advances in Neural Information Processing Systems*, 34:11631–11642, 2021.

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

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

Charlie Harrison and Pasin Manurangsi. Optimal partition selection with Rényi differential privacy. In *Symposium on Foundations of Responsible Computing (FORC)*, 2026.

Yuzheng Hu, Ryan McKenna, Da Yu, Shanshan Wu, Han Zhao, Zheng Xu, and Peter Kairouz. ACTG-ARL: Differentially Private Conditional Text Generation with RL-Boosted Control. In *ICML*, 2026.

Peter Kairouz, Sewoong Oh, and Pramod Viswanath. The composition theorem for differential privacy. In *International conference on machine learning*. PMLR, 2015.

Haim Kaplan, Shachar Schnapp, and Uri Stemmer. Differentially Private Approximate Quantiles. In *ICML*, 2022.

Antti Koskela, Joonas Jälkö, and Antti Honkela. Computing tight differential privacy guarantees using fft. In *International Conference on Artificial Intelligence and Statistics*, pp. 2560–2569. PMLR, 2020.

Antti Koskela, Marlon Tobaben, and Antti Honkela. Individual privacy accounting with gaussian differential privacy. In *International Conference on Learning Representations*, 2023.

Bogdan Kulynych, Yao-Yuan Yang, Yaodong Yu, Jaroslaw Blasiok, and Preetum Nakkiran. What you see is what you get: Principled deep learning via distributional generalization. *Advances in Neural Information Processing Systems (NeurIPS)*, 2022.

Bogdan Kulynych, Juan Felipe Gomez, Georgios Kaissis, Flavio du Pin Calmon, and Carmela Troncoso. Attack-aware noise calibration for differential privacy. *Advances in Neural Information Processing Systems (NeurIPS)*, 2024.

Bogdan Kulynych, Juan Gomez, Georgios Kaissis, Jamie Hayes, Borja Balle, Flavio Calmon, and Jean Raisaro. Unifying re-identification, attribute inference, and data reconstruction risks in differential privacy. *Advances in Neural Information Processing Systems*, 2025.

Terrance Liu, Giuseppe Vietri, and Zhiwei Steven Wu. Iterative methods for private synthetic data: Unifying framework and new methods. *Advances in Neural Information Processing Systems (NeurIPS)*, 34, 2021.

Sofiane Mahiou, Amir Dizche, Reza Nazari, Xinmin Wu, Ralph Abbey, Jorge Silva, and Georgi Ganev. dpmm: Differentially Private Marginal Models, a Library for Synthetic Tabular Data Generation. In *TPDP*, 2025.

Ryan McKenna, Daniel Sheldon, and Gerome Miklau. Graphical-model Based Estimation and Inference for Differential Privacy. In *ICML*, 2019.

Ryan McKenna, Siddhant Pradhan, Daniel R Sheldon, and Gerome Miklau. Relaxed marginal consistency for differentially private query answering. *Advances in Neural Information Processing Systems (NeurIPS)*, 2021.

Ryan McKenna, Brett Mullins, Daniel Sheldon, and Gerome Miklau. AIM: An adaptive and iterative mechanism for differentially private synthetic data. In *VLDB*, 2022.

Sebastian Meiser and Esfandiar Mohammadi. Tight on budget? tight bounds for r-fold approximate differential privacy. In *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*, pp. 247–264, 2018.

Ilya Mironov. Rényi differential privacy. In *Computer Security Foundations Symposium (CSF)*. IEEE, 2017.

Milad Nasr, Jamie Hayes, Thomas Steinke, Borja Balle, Florian Tramèr, Matthew Jagielski, Nicholas Carlini, and Andreas Terzis. Tight auditing of differentially private machine learning. In *32nd USENIX Security Symposium (USENIX Security 23)*, 2023.

OpenDP. SmartNoise SDK: Tools for Differential Privacy on Tabular Data. https://github.com/opendp/smartnoise-sdk, 2021.

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

Krishna Pillutla, Lang Liu, John Thickstun, Sean Welleck, Swabha Swayamdipta, Rowan Zellers, Sewoong Oh, Yejin Choi, and Zaid Harchaoui. Mauve scores for generative models: Theory and practice. *Journal of Machine Learning Research*, 24(356):1–92, 2023.

Natalia Ponomareva, Hussein Hazimeh, Alex Kurakin, Zheng Xu, Carson Denison, H Brendan McMahan, Sergei Vassilvitskii, Steve Chien, and Abhradeep Guha Thakurta. How to DP-fy ML: A practical guide to machine learning with differential privacy. *Journal of Artificial Intelligence Research*, 77:1113–1201, 2023.

Natalia Ponomareva, Zheng Xu, H Brendan McMahan, Peter Kairouz, Lucas Rosenblatt, Vincent Cohen-Addad, Cristóbal Guzmán, Ryan McKenna, Galen Andrew, Alex Bie, et al. How to DP-Fy Your Data: A Practical Guide to Generating Synthetic Data With Differential Privacy. *JAIR*, 2026.

Zhaozhi Qian, Rob Davis, and Mihaela Van Der Schaar. Synthcity: A Benchmark Framework for Diverse Use Cases of Tabular Synthetic Data. In *NeurIPS Datasets and Benchmarks Track*, 2023.

Matthew Regehr, Bingshan Hu, Ethan Leeman, Pasin Manurangsi, Pierre Tholoniat, and Mathias Lécuyer. Privacy filters are captured by residues: A characterization of free natural filters and the cost of adaptivity. *arXiv preprint arXiv:2602.15815*, 2026.

Anneliese Riess, Juan Felipe Gomez, Flavio du Pin Calmon, Julia Anne Schnabel, and Georgios Kaissis. Optimal conversion from Rényi differential privacy to f-differential privacy. In *International Conference on Machine Learning*. PMLR, 2026.

Ryan M Rogers, Aaron Roth, Jonathan Ullman, and Salil Vadhan. Privacy odometers and filters: Pay-as-you-go composition. *Advances in Neural Information Processing Systems*, 29, 2016.

Michael R Sampford. Some inequalities on mill's ratio and related functions. *The Annals of Mathematical Statistics*, 24(1):130–132, 1953.

Reza Shokri, Marco Stronati, Congzheng Song, and Vitaly Shmatikov. Membership inference attacks against machine learning models. In *2017 IEEE symposium on security and privacy (SP)*. IEEE, 2017.

Amanpreet Singh, Mike D'Arcy, Arman Cohan, Doug Downey, and Sergey Feldman. Scirepeval: A multi-format benchmark for scientific document representations. In *Proceedings of the 2023 conference on empirical methods in natural language processing*, pp. 5548–5566, 2023.

Adam Smith and Abhradeep Thakurta. Fully adaptive composition for gaussian differential privacy. *arXiv preprint arXiv:2210.17520*, 2022.

David M. Sommer, Sebastian Meiser, and Esfandiar Mohammadi. Privacy loss classes: The central limit theorem in differential privacy. *Proceedings on Privacy Enhancing Technologies*, 2019(2): 245–269, 2019. doi: 10.2478/popets-2019-0029.

Bowen Tan, Zheng Xu, Eric Xing, Zhiting Hu, and Shanshan Wu. Synthesizing privacy-preserving text data via finetuning without finetuning billion-scale llms. *arXiv preprint arXiv:2503.12347*, 2025.

Yuchao Tao, Ryan McKenna, Michael Hay, Ashwin Machanavajjhala, and Gerome Miklau. Benchmarking Differentially Private Synthetic Data Generation Algorithms. In *PPAI*, 2022.

Long Tran, Antti Koskela, Ossi Räisä, and Antti Honkela. f-differential privacy filters: Validity and approximate solutions. *arXiv preprint arXiv:2602.06756*, 2026.

Chulin Xie, Zinan Lin, Arturs Backurs, Sivakanth Gopi, Da Yu, Huseyin A Inan, Harsha Nori, Haotian Jiang, Huishuai Zhang, Yin Tat Lee, et al. Differentially private synthetic data via foundation model apis 2: Text. *arXiv preprint arXiv:2403.01749*, 2024.

Samuel Yeom, Irene Giacomelli, Matt Fredrikson, and Somesh Jha. Privacy risk in machine learning: Analyzing the connection to overfitting. In *Computer Security Foundations Symposium (CSF)*. IEEE, 2018.

Santiago Zanella-Béguelin, Lukas Wutschitz, Shruti Tople, Ahmed Salem, Victor Rühle, Andrew Paverd, Mohammad Naseri, Boris Köpf, and Daniel Jones. Bayesian Estimation of Differential Privacy. In *ICML*, 2023.

Zhikun Zhang, Tianhao Wang, Ninghui Li, Jean Honorio, Michael Backes, Shibo He, Jiming Chen, and Yang Zhang. PrivSyn: Differentially private data synthesis. In *30th USENIX Security Symposium (USENIX Security 21)*, 2021.

Yuqing Zhu, Jinshuo Dong, and Yu-Xiang Wang. Optimal accounting of differential privacy via characteristic function. In *International Conference on Artificial Intelligence and Statistics*. PMLR, 2022.

A Omitted Proofs

A.1 BR to f-DP, privacy profiles and PLD

The proof of the theorem rests on one result, which Dong et al. (2020) establish in the course of proving their reduction to generalized randomized response (their Lemma 4.1).

Lemma A.1 (Dong et al., 2020). *Suppose that a mechanism $M : 2^{\mathbb{D}} \to \mathbb{Y}$ satisfies ν-BR. Then, for any $S \simeq S'$ there exists $0 \le t \le \nu$ such that for any test $\phi : \mathbb{Y} \to \{0, 1\}$, we have:*

$$\beta_\phi \ge \max\{1 - e^t \alpha_\phi,\ e^{t-\nu}(1 - \alpha_\phi)\}. \tag{5}$$

Moreover for each $\alpha_\phi \in [0, 1]$, there exists the mechanism, such that the inequality turns to the equality.

Theorem 3.1. *Let $M(\cdot)$ satisfy ν-BR. The density $g_\nu : \mathbb{R} \to \mathbb{R}$ of its tightly dominating PLD is:*

$$g_\nu(x) = D_\nu\, e^{x/2}\, \mathbf{1}_{[-\nu,\nu]}(x), \quad \text{where } D_\nu = \tfrac{1}{2}\left(e^{\nu/2} - e^{-\nu/2}\right)^{-1}. \tag{2}$$

Moreover, $M(\cdot)$ satisfies f-DP and has privacy profile $\delta(\varepsilon)$:

$$f(\alpha) = \frac{1-\alpha}{\alpha e^\nu + 1 - \alpha}, \quad \delta(\varepsilon) = \begin{cases} 1 - e^\varepsilon, & \varepsilon < -\nu, \\ \frac{\left(e^{\nu/2} - e^{\varepsilon/2}\right)^2}{e^\nu - 1}, & -\nu \le \varepsilon \le \nu \\ 0, & \varepsilon \ge \nu \end{cases} \tag{3}$$

As they are derived from the tightly dominating PLD, these two bounds are tight: every point (i.e. every $(\varepsilon, \delta(\varepsilon))$ or $(\alpha, f(\alpha))$ is achieved by some ν-BR mechanism.

Proof. We prove it in the following steps:

1. Prove the form of the f-DP trade-off function and its tightness.
2. Derive the privacy profile from f-DP.
3. Derive the PLD from the privacy profile.

Step 1: The f-DP trade-off function and its tightness.

To show the result, for a given α it suffices to find the smallest Lemma A.1 bound for all $t \in [0, \nu]$:

$$f(\alpha) = \min_{t \in [0,\nu]} \max\{1 - e^t \alpha,\ e^{t-\nu}(1 - \alpha)\}. \tag{6}$$

As the first term is decreasing in t, and the second one is increasing, the optimal value must occur at their intersection. Equating the two terms and solving for t we get:

$$1 - e^t \alpha = e^{t-\nu}(1 - \alpha) \Rightarrow e^t = \frac{1}{\alpha + (1-\alpha)e^{-\nu}}. \tag{7}$$

Substituting e^t into Eq. (6), we have:

$$f(\alpha) = 1 - \frac{\alpha}{\alpha + (1-\alpha)e^{-\nu}} = \frac{(1-\alpha)e^{-\nu}}{\alpha + (1-\alpha)e^{-\nu}} = \frac{1-\alpha}{\alpha e^\nu + 1 - \alpha}. \tag{8}$$

The tightness follows from the fact that each point $(\alpha, f(\alpha)$ is attainable by Lemma A.1.

Step 2: From f-DP to privacy profile.

Let $a = e^\nu - 1$, so $f(\alpha) = \frac{1-\alpha}{1+a\alpha}$, and write

$$\delta(\varepsilon) = \sup_{\alpha \in [0,1]} g(\alpha), \text{ where } g(\alpha) = 1 - e^\varepsilon \alpha - f(\alpha).$$

Concavity. As $f'(\alpha) = -e^{\nu}/(1+a\alpha)^2$ and $f'' > 0$, the function g is concave. Its derivative $g'(\alpha) = e^{\nu}/(1+a\alpha)^2 - e^{\varepsilon}$ decreases from $g'(0) = e^{\nu} - e^{\varepsilon}$ to $g'(1) = e^{-\nu} - e^{\varepsilon}$.

Boundary cases.

If $\varepsilon \geq \nu$, then $g' \leq 0$ throughout, so $\delta = g(0) = 0$. If $\varepsilon \leq -\nu$, then $g' \geq 0$ throughout, so $\delta = g(1) = 1 - e^{\varepsilon}$.

Interior case: $|\varepsilon| < \nu$. In this case, $g'(\alpha^*) = 0$ at $1 + a\alpha^* = s := e^{(\nu-\varepsilon)/2} \in (1, e^{\nu})$. At this point $f(\alpha^*) = (e^{\nu} - s)/(as)$, and $e^{\varepsilon}s = e^{\nu}/s = e^{(\nu+\varepsilon)/2}$. Therefore:

$$\delta(\varepsilon) = \frac{a - e^{\varepsilon}(s-1) - (e^{\nu} - s)/s}{a} = \frac{e^{\nu} + e^{\varepsilon} - 2e^{(\nu+\varepsilon)/2}}{e^{\nu} - 1} = \frac{\left(e^{\nu/2} - e^{\varepsilon/2}\right)^2}{e^{\nu} - 1}.$$

At $\varepsilon = \pm\nu$ this equals 0 and $1 - e^{-\nu}$ respectively, which agrees with the boundary cases. The tightness follows from the tightness for f-DP tradeoff function.

Step 3: From privacy profile to PLD.

If $\delta(\varepsilon)$ is twice-differentiable profile function, then by Sommer et al. (2019), we have that $g = \delta'' - \delta'$ is the density of PLD.

On $(-\nu, \nu)$, differentiating $\delta_\nu = \left(e^{\nu} - 2e^{(\nu+\varepsilon)/2} + e^{\varepsilon}\right)/(e^{\nu} - 1)$ twice gives

$$\delta''_\nu - \delta'_\nu = \frac{\left(e^{\varepsilon} - \frac{1}{2}e^{(\nu+\varepsilon)/2}\right) - \left(e^{\varepsilon} - e^{(\nu+\varepsilon)/2}\right)}{e^{\nu} - 1} = \frac{e^{\nu/2}}{2\left(e^{\nu} - 1\right)} e^{\varepsilon/2} = D_\nu\, e^{\varepsilon/2}, \tag{9}$$

where $2\left(e^{\nu/2} - e^{-\nu/2}\right) = 2e^{-\nu/2}\left(e^{\nu} - 1\right)$ gives

$$D_\nu = \frac{1}{2\left(e^{\nu/2} - e^{-\nu/2}\right)} = \frac{e^{\nu/2}}{2\left(e^{\nu} - 1\right)}. \tag{10}$$

Observe that off $[-\nu, \nu]$, the density vanishes: $\delta_\nu \equiv 0$ on (ν, ∞), and on $(-\infty, -\nu)$ we have $\delta''_\nu - \delta'_\nu = -e^{\varepsilon} + e^{\varepsilon} = 0$.

□

Next, we provide several observations about these results.

Observation 1. The tight PLD for ν-BR mechanisms is not ν-BR (but 2ν-BR). As a result, there is no one single dominating PLD (or, equivalently, dominating pair) which is ν-BR. That contrasts with classes of other common mechanisms, like Laplace, Gaussian (incl. sampled with fixed probability).

Observation 2. Theorem 3 of Dong et al. (2020) provides the exact optimal $\delta_{\mathrm{OPT}}(\mathcal{M}^k_{\mathrm{BR}}, \varepsilon)$ for the homogeneous non-adaptive composition of k mechanisms that are ν-BR.

At $k = 1$, the two results coincide. For any $k > 1$, however, the PLD approach of the k-fold convolution of g_ν should be larger (i.e., worse guarantees) as the result from Dong et al. (2020) is optimal for the homogeneous composition of BR mechanisms.

At the same time, the PLD analysis has some advantages over the optimal expression. First, it enables simple composing with any mechanisms that also support PLD analyses, unlike the optimal expression for BR. Second, relatedly, it is easy to use in existing PLD accounting software (e.g., in dp_accounting library). The looseness is thus the price of that flexibility.

Observation 3. Empirical evaluations (see Fig. 6) show that, under homogeneous composition of BR mechanisms, for some number of compositions, zCDP provides a better (ε, δ)-DP bound. The explanation is that both are upper bounds and the PLD of BR is not Bounded Range itself, thus it might have larger zCDP.

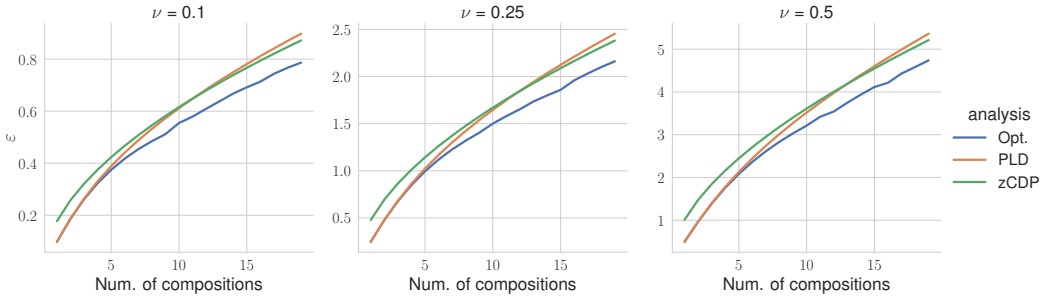


Figure 6: The DP ε at $\delta = 10^{-5}$ reported using the PLD (ours), zCDP (Harrison & Manurangsi, 2025), and optimal (Dong et al., 2020) analyses for the compositions of homogeneous BR mechanisms.

Observation 4. We observe an interesting form in terms of *total-variation (TV) privacy* (see, e.g., Geng & Viswanath, 2015; Barber & Duchi, 2014; Kulynych et al., 2022):

Definition A.1. An algorithm satisfies η-TV privacy if it satisfies $(0, \eta)$-DP.

TV is not a useful privacy notion on its own, but it provides simple forms of risk and generalization bounds (Cherubin et al., 2024; Kulynych et al., 2025). For the classical notion of $(\varepsilon, 0)$-DP, Kairouz et al. (2015) showed that an $(\varepsilon, 0)$-DP mechanism satisfies η-TV privacy with:

$$\eta = \frac{e^{\varepsilon} - 1}{e^{\varepsilon} + 1} = \tanh(\varepsilon/2). \tag{11}$$

The conversion from BR to TV privacy is as follows:

Corollary A.1. *Suppose that a mechanism $M(\cdot)$ satisfies ν-BR. Then, it satisfies η-TV privacy with:*

$$\eta = \frac{e^{\frac{\nu}{2}} - 1}{e^{\frac{\nu}{2}} + 1} = \tanh(\nu/4).$$

Curiously, this is exactly the conversion from ε-DP to TV privacy from Eq. (11) with a difference of a factor of two.

A.2 Relation to the Laplace Mechanism

Write Lap_ν for the Laplace mechanism of scale $1/\nu$ on a statistic of sensitivity 1, which is ν-DP, and let $\omega_\nu^{\mathrm{Lap}}$ denote its PLD. Let

$$\omega_\nu^{\mathrm{RR}} := \begin{cases} \nu, & \text{w.p. } \dfrac{e^\nu}{1+e^\nu}, \\ -\nu, & \text{w.p. } \dfrac{1}{1+e^\nu}, \end{cases} \tag{12}$$

be the PLD of randomized response at parameter ν, which tightly dominates the class of ν-DP mechanisms (Kairouz et al., 2015), and let ω_ν^{BR} be the PLD with density g_ν of Theorem 3.1. All three are supported in $[-\nu, \nu]$.

The Laplace PLD is well-known:

Lemma A.2 (Sommer et al., 2019, App. A.3)**.** *$\omega_\nu^{\mathrm{Lap}}$ is the measure on $[-\nu, \nu]$ consisting of two atoms,*

$$\Pr\left[\omega_\nu^{\mathrm{Lap}} = \nu\right] = \tfrac{1}{2}, \qquad \Pr\left[\omega_\nu^{\mathrm{Lap}} = -\nu\right] = \tfrac{1}{2}e^{-\nu}, \tag{13}$$

together with an absolutely continuous part on $(-\nu, \nu)$ of density

$$\tfrac{1}{4}\, e^{-\nu/2}\, e^{w/2}. \tag{14}$$

Proposition A.1. *For every $\nu > 0$,*

$$\omega_\nu^{\mathrm{Lap}} = \lambda_\nu\, \omega_\nu^{\mathrm{RR}} + (1 - \lambda_\nu)\, \omega_\nu^{\mathrm{BR}}, \qquad \lambda_\nu = \frac{1 + e^{-\nu}}{2}, \tag{15}$$

as an identity of distributions on $\mathbb{R}$.

Proof. By Lemma A.2 the atoms carry total mass $\frac{1}{2}\left(1+e^{-\nu}\right)=\lambda_\nu$, and after normalisation their masses are those of ω_ν^{RR}, since $\frac{1/2}{\lambda_\nu}=\frac{1}{1+e^{-\nu}}=\frac{e^\nu}{1+e^\nu}$. The continuous part carries the remaining mass,

$$\tfrac{1}{4}e^{-\nu/2}\int_{-\nu}^{\nu}e^{w/2}\mathrm{d}w=\tfrac{1}{2}\left(1-e^{-\nu}\right)=1-\lambda_\nu, \tag{16}$$

and normalising its density gives $\frac{\frac{1}{4}e^{-\nu/2}}{1-\lambda_\nu}e^{w/2}=\frac{e^{w/2}}{2\left(e^{\nu/2}-e^{-\nu/2}\right)}=D_\nu e^{w/2}$, i.e. exactly g_ν. □

Because $\delta_\omega(\varepsilon)=\mathbb{E}\left[(1-e^{\varepsilon-\omega})_+\right]$ is *linear* in the law of ω, (15) passes to privacy profiles.

Corollary A.2. *For every $\varepsilon\in\mathbb{R}$,*

$$\delta_{\mathrm{Lap}_\nu}(\varepsilon)=\lambda_\nu\,\delta_{\omega_\nu^{\mathrm{RR}}}(\varepsilon)+(1-\lambda_\nu)\,\delta_\nu(\varepsilon)\;\geq\;\delta_\nu(\varepsilon), \tag{17}$$

that is, $\omega_\nu^{\mathrm{BR}}\preceq\omega_\nu^{\mathrm{Lap}}$. In particular the Laplace PLD is itself a valid, but strictly loose, dominating PLD for the ν-BR class.

Proof. The equality is linearity of δ in the mixture. Since ω_ν^{RR} dominates every ν-DP mechanism (Kairouz et al., 2015), we have $\delta_{\omega_\nu^{\mathrm{RR}}}\geq\delta_\nu$ pointwise. Substituting this into the equality and using $\lambda_\nu+(1-\lambda_\nu)=1$ gives the inequality. □

A.3 BR TO GDP

In this section, we prove the GDP guarantee for BR mechanisms:

Theorem 3.2. *Suppose that a mechanism $M(\cdot)$ satisfies ν-BR. Then, it also satisfies μ-GDP with:*

$$\mu=-2\Phi^{-1}\left(\tfrac{1}{e^{\nu/2}+1}\right). \tag{4}$$

First, we make use of the properties of the *inverse Mills ratio*.

Proposition A.2 (Sampford, 1953)**.** *The inverse Mills ratio $\lambda(t)\triangleq\varphi(t)/\Phi(t)$, where $\varphi(t)$ is the density of the standard Normal distribution, is convex on $\mathbb{R}$. In particular, λ' is nondecreasing.*

We express the statement in Theorem 3.2 as a condition in terms of an expression $\log\left(\Phi(t/2)/\Phi(-t/2)\right)$, which has the following properties:

Lemma A.3. *Let $L(t)\triangleq\log\frac{\Phi(t/2)}{\Phi(-t/2)}$. Then, L is odd and differentiable. Its derivative L' is even, and nondecreasing on $[0,\infty)$.*

Proof. Oddness is immediate from $L(-t)=\log\frac{\Phi(-t/2)}{\Phi(t/2)}=-L(t)$. Differentiating, we get:

$$L'(t)=\frac{\mathrm{d}}{\mathrm{d}t}\log\Phi\left(\frac{t}{2}\right)-\frac{\mathrm{d}}{\mathrm{d}t}\log\Phi\left(-\frac{t}{2}\right)=\frac{\varphi(t/2)}{2\,\Phi(t/2)}+\frac{\varphi(-t/2)}{2\,\Phi(-t/2)}=\frac{1}{2}\left[\lambda\left(\frac{t}{2}\right)+\lambda\left(-\frac{t}{2}\right)\right],$$

using $\varphi(-s)=\varphi(s)$. The derivative L' is even as L is odd. Finally, for any $t\geq0$,

$$\frac{\mathrm{d}}{\mathrm{d}t}L'(t)=\frac{1}{4}\left[\lambda'\left(\frac{t}{2}\right)-\lambda'\left(-\frac{t}{2}\right)\right]\geq0,$$

as λ' is nondecreasing by Proposition A.2 and $t/2\geq-t/2$ for $t\geq0$. □

Finally, we can show the proof of Theorem 3.2.

Proof. To prove the statement, we need to show that we have, for any $\alpha\in[0,1]$:

$$\Phi(\Phi^{-1}(1-\alpha)-\mu)\leq\frac{1-\alpha}{\alpha e^\nu+1-\alpha}, \tag{18}$$

where μ is defined as in Eq. (4). To accomplish this, we will simplify this condition to a form that is more amenable to analysis. For this, first, let us write down ν in terms of μ by solving Eq. (4):

$$\Phi\left(-\frac{\mu}{2}\right) = \frac{1}{e^{\frac{\nu}{2}}+1}.$$

As $\Phi(t) = 1 - \Phi(-t)$, we have that:

$$\Phi\left(\frac{\mu}{2}\right) = 1 - \frac{1}{e^{\frac{\nu}{2}}+1} = \frac{e^{\frac{\nu}{2}}}{e^{\frac{\nu}{2}}+1}.$$

Dividing the two expressions, we have:

$$\frac{\Phi(\mu/2)}{\Phi(-\mu/2)} = e^{\frac{\nu}{2}}.$$

Using the definition of $L(t)$ in Lemma A.3, we thus have:

$$\nu = 2L\left(\mu\right). \tag{19}$$

Next, denoting by $z = \Phi^{-1}(1-\alpha)$, so that $\alpha = 1 - \Phi(z) = \Phi(-z)$, we have:

$$\Phi(z-\mu) \leq \frac{\Phi(z)}{\Phi(-z)e^{\nu} + \Phi(z)}. \tag{20}$$

Observing that the denominator in r.h.s. is always positive, we can multiply both sides:

$$\Phi(z-\mu)\Phi(-z)e^{\nu} + \Phi(z-\mu)\Phi(z) \leq \Phi(z).$$

Rewriting to group the $\Phi(z)$ terms:

$$\Phi(z-\mu)\Phi(-z)e^{\nu} \leq \Phi(z)\left(1 - \Phi(z-\mu)\right) = \Phi(z)\Phi(\mu - z).$$

Taking logarithms on both sides:

$$\log\Phi(z-\mu) + \log\Phi(-z) + \nu \leq \log\Phi(z) + \log\Phi(\mu-z)$$
$$\nu + \log\frac{\Phi(z-\mu)}{\Phi(\mu-z)} \leq \log\frac{\Phi(z)}{\Phi(-z)}.$$

Using the definition of $L(t)$ above and the form of ν in Eq. (19):

$$2L\left(\mu\right) \leq L(2z) - L(2z-2\mu).$$

Next, to derive a useful form, we write $z = \frac{\mu}{2} + \frac{u}{2}$ for some $u \in \mathbb{R}$:

$$2L\left(\mu\right) \leq L\left(u+\mu\right) - L(u-\mu).$$

As L is odd by Lemma A.3, we have $-L(u-\mu) = L(\mu-u)$. Therefore, the condition in Eq. (18) that we need to show is equivalent to the following:

$$D(u) \triangleq L(\mu+u) + L(\mu-u) - 2L(\mu) \geq 0 \qquad \forall u \in \mathbb{R}. \tag{21}$$

This form is easier to analyze. We start by observing that, as D is even, we have $D(u) = D(|u|)$ for any $u \in \mathbb{R}$, thus it suffices to consider $u \geq 0$. Moreover, observing that $D(0) = 0$, it suffices to show that D is nondecreasing on $[0, \infty)$. Its derivative is:

$$D'(u) = L'(\mu+u) - L'(\mu-u).$$

By Lemma A.3, L' is even and nondecreasing on $[0,\infty)$. This implies that $L'(a) \geq L'(b)$ whenever $|a| \geq |b|$. As $|\mu+u| \geq |\mu-u|$ for $\mu, u \geq 0$, we get $D'(u) \geq 0$, and the claim follows. □

B Additional Results

B.1 Poisson Subsampling for Arbitrary Privacy Loss Distributions

We present a novel method for converting a privacy guarantee for an arbitrary mechanism to a privacy guarantee for a Poisson subsampled version of that mechanism, which is compatible with modern privacy accounting libraries such as `dp_accounting`.

Our method operates by representing the privacy guarantee as a privacy profile. Although we could then convert the privacy profile to a PLD representation, as before, we found empirically that in this case the PLD approach is less numerically stable to apply the conversion to than the $\delta(\varepsilon)$ representation. Hence, we present the conversion in the $\delta(\varepsilon)$ representation.

We recall the hockey-stick divergence has a form $H_\alpha(P, Q) = \int (P(x) - \alpha Q(x))_+ dx$, and if a dominating pair P, Q for M satisfies $H_{e^\varepsilon}(P, Q) \leq \delta$ then M is (ε, δ)-DP. Finally, Theorem 11 of (Zhu et al., 2022) shows that if P, Q is a dominating pair for the unsampled mechanism in the remove direction, then (P', Q), where $P' = q \cdot P + (1-q) \cdot Q$, is a dominating pair for Poisson sampled mechanism with sampling probability q in the remove direction, and vice-versa in the add direction. The following theorem gives a conversion from hockey-stick divergences between P, Q to hockey-stick divergences between P', Q, which by the above facts suffices to convert a privacy profile for an unsampled mechanism to a privacy profile for its Poisson sampled counterpart.

Theorem B.1. *Let $P' = q \cdot P + (1-q) \cdot Q$. Then:*

$$H_{e^\varepsilon}(P', Q) = \begin{cases} q \cdot H_{\frac{e^\varepsilon + q - 1}{q}}(P, Q) & \text{if } e^\varepsilon > 1 - q, \\ 1 - e^\varepsilon & \text{otherwise.} \end{cases}$$

$$H_{e^\varepsilon}(Q, P') = \begin{cases} (1 - e^\varepsilon(1-q)) \cdot H_{\frac{e^\varepsilon \cdot q}{1 - e^\varepsilon(1-q)}}(Q, P) & \text{if } e^\varepsilon < \frac{1}{1-q}, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. We have:

$$H_{e^\varepsilon}(P', Q) = \int (P'(y) - e^\varepsilon Q(y))_+ dy = \int (q \cdot P(y) + (1 - q - e^\varepsilon) Q(y))_+ dy.$$

If $e^\varepsilon \leq 1 - q$ then $1 - q - e^\varepsilon \geq 0$ and:

$$\begin{aligned} &\int (q \cdot P(y) + (1 - q - e^\varepsilon) Q(y))_+ dy \\ =&\int q \cdot P(y) + (1 - q - e^\varepsilon) Q(y) dy \\ =&q \int P(y) dy + (1 - q - e^\varepsilon) \int Q(y) dy \\ =&q + 1 - q - e^\varepsilon = 1 - e^\varepsilon. \end{aligned}$$

Otherwise:

$$\begin{aligned} H_{e^\varepsilon}(P', Q) &= \int (q \cdot P(y) - (e^\varepsilon + q - 1) Q(y))_+ dy \\ &= q \int (P(y) - \frac{e^\varepsilon + q - 1}{q} Q(y))_+ dy = q \cdot H_{\frac{e^\varepsilon + q - 1}{q}}(P, Q), \end{aligned}$$

Similarly we have:

$$H_{e^\varepsilon}(Q, P') = \int (Q(y) - e^\varepsilon P'(y))_+ dy = \int ((1 - e^\varepsilon(1-q)) \cdot Q(y) - e^\varepsilon q \cdot P(y))_+ dy.$$

If $e^\varepsilon < \frac{1}{1-q}$ then:

$$\begin{aligned} &\int ((1 - e^\varepsilon(1-q)) \cdot Q(y) - e^\varepsilon q \cdot P(y))_+ dy. \\ =&(1 - e^\varepsilon(1-q)) \int (Q(y) - \frac{e^\varepsilon q}{1 - e^\varepsilon(1-q)} \cdot P(y))_+ dy \\ =&(1 - e^\varepsilon(1-q)) H_{\frac{e^\varepsilon q}{1 - e^\varepsilon(1-q)}}(Q, P), \end{aligned}$$

Otherwise $1 - e^{\varepsilon}(1-q) \leq 0$ and so:

$$\int ((1 - e^{\varepsilon}(1-q)) \cdot Q(y) - e^{\varepsilon} q \cdot P(y))_+ dy = \int 0 \, dy = 0.$$

□

B.2 Exact Characterization for Private Partition Selection

In this section, we present characterizations of mechanisms such as the Threshold Gaussian and Threshold Laplace (Gopi et al., 2020) for private partition selection. It rests on the following result:

Lemma B.1. *Let M be a mechanism with the following property: There exists P, Q such that for any inputs $S \simeq S'$ where S contains an example not in S', there is an event E occurring with probability at least $1 - \tau$ such that:*

$$T(M(S)|E, M(S'))(\alpha) \geq T(P, Q)(\alpha)$$

for all $\alpha \in [0, 1]$, where $A|E$ denotes the random variable A conditioned on event E. Then $P \times \mathsf{Bern}(\tau), Q \times \mathsf{Bern}(0)$ is a dominating pair of M, where $\times$ denotes a product distribution.

Proof. We use the following fact, a corollary of Blackwell's informativeness theorem (Blackwell, 1953): P, Q is a dominating pair for M if and only if for any $S \simeq S'$, there is a randomized function f such that $f(P) \equiv M(S), f(Q) \equiv M(S')$ where $\equiv$ denotes distributional equivalence.

Fix a $S \simeq S'$, and let f be the randomized function such that $M(S)|E \equiv f(P), M(S') \equiv f(Q)$. By the preceding fact, it suffices to define a randomized function f' such that $f'(P \times \mathsf{Bern}(\tau)) \equiv M(S), f'(Q \times \mathsf{Bern}(0)) \equiv M(S')$. Our function is, $f'(x, 0) = f(x)$, and $f'(x, 1) =$ a sample from $M(S)|\bar{E}$ w.p. $\Pr[\bar{E}]/\tau$, and otherwise $f'(x, 1) = f(x)$.

We have:

$$\begin{aligned} f'(P \times \mathsf{Bern}(\tau)) &\equiv (1-\tau) \cdot f(P) + \tau \cdot \left(\frac{\Pr[\bar{E}]}{\tau} \cdot M(S)|\bar{E} + \left(1 - \frac{\Pr[\bar{E}]}{\tau}\right) f(P) \right) \\ &\equiv (1 - \tau + \tau - \Pr[\bar{E}]) \cdot f(P) + \Pr[\bar{E}] \cdot M(S)|\bar{E} \\ &\equiv \Pr[E] \cdot M(S)|E + \Pr[\bar{E}] \cdot M(S)|\bar{E} \\ &\equiv M(S). \end{aligned}$$

and more straightforwardly, $f'(Q \times \mathsf{Bern}(0)) \equiv f(Q) \equiv M(S')$, completing the proof. □

The proof of Gopi et al. (2020, Theorem 5.1) exactly specifies an event E for their mechanism (namely, E is the event that for $S \simeq S'$ and $S \supset S'$, $M(S)$ is in the support of $M(S')$), which is what implies that their policy Gaussian (aka threshold Gaussian) algorithm is a composition of a Gaussian mechanism and a mechanism with dominating pair $\mathsf{Bern}(\tau), \mathsf{Bern}(0)$.

B.3 Group Privacy Guarantees

In real-world settings, individual users often contribute multiple records to the dataset. Practitioners typically handle this by enforcing hard contribution bounding, randomly truncating each user's history to at most one record ($K = 1$). While conceptually simple, contribution bounding introduces a *truncation bias*.

An alternative is to retain up to $K > 1$ records per user and evaluate the resulting user-level guarantee via group privacy. A user with K records contributes up to $K\times$ more signal to linear queries, equivalently the sensitivity increases by K, requiring $K\times$ more noise to preserve the same signal-to-noise ratio/privacy guarantee. Below, we formalize how group privacy plumbs through our unified framework across each primitive class.

- *Gaussian Mechanism:* If a user contributes at most K records, the global L_2 sensitivity of a marginal scales by at most K. Under Gaussian Differential Privacy (GDP), the guarantee scales strictly as $\mu \to K\mu$. The dominating pair remains Gaussian ($G_{K\mu}$), retaining exact hypothesis-testing duality.

- *Exponential Mechanism:* When each record can alter the selection score by at most Δs, K records alter the score by at most $K\Delta s$, scaling the effective privacy parameter $\varepsilon \to K\varepsilon$.
- *Threshold Gaussian:* A user with K records can alter the true count of any partition by at most K. The continuous Gaussian count measurement scales sensitivity by K. For partition selection at cutoff τ, the failure event $(0, \delta_{\text{thresh}})$ corresponds to the probability that an empty partition containing only the user exceeds the threshold:

$$\delta_{\text{thresh}}(K) = 1 - \Phi\left(\frac{\tau - K}{\sigma}\right).$$

  Operationally, this means that items with true count $< K$ cannot be returned by the mechanism.

Because every mechanism in the taxonomy admits an exact representation parameterized by K, our framework allows us to output precise privacy profiles for multiple group sizes. With this approach, we can calibrate our mechanism to any group size K, then post-hoc report the privacy profiles for any other K' (noting that if there are Gaussian Thresholding mechanisms, we must have $K' \leq K$).

One could alternatively consider taking the record-level privacy guarantee of the composed mechanism, and trying to convert this in a black-box manner to a group privacy guarantee. For example, any f-DP mechanism satisfies f'-DP with respect to add-or-remove-up-to-k where $f'(x) = 1 - (1-f)^{\circ k}(x)$, $\circ k$ denotes self-composition of a function k times. However, black-box approaches to group privacy are generally loose. Especially for mechanisms like the exponential mechanism, a black-box privacy guarantee might be pointwise tight everywhere, as in at each FPR-FNR pair there is an instantiation of the mechanism and two inputs S, S' that achieve that FPR-FNR pair, but not "globally" tight in the sense that a single mechanism and pair of inputs achieves all the FPR-FNR pairs simultaneously. A black-box group privacy analysis might try to combine information different FPR-FNR points that are not "globally" tight, thus giving an upper bound that no instantiation of the mechanism and inputs can realize, whereas a "white-box" group privacy analysis looking directly at $M(S), M(S')$ for $S \simeq S'$ would avoid this issue entirely.

B.4 Group Privacy of Hierarchical Exponential Mechanisms

In this section, we give a group privacy guarantee for Algorithm 8. That is, we give a privacy guarantee under "add-or-remove-up-to-k" adjacency $\simeq_{AR\leq k}$, i.e. $S \simeq_{AR\leq k} S'$ if $S \supseteq S'$ and S contains at most k elements not in S (or vice-versa).

Theorem B.2. *Any ν-BR mechanism with respect to the add-or-remove-one adjacency is $k\nu$-BR with respect to $\simeq_{AR\leq k}$.*

Proof. Let $S \simeq_{AR\leq k} S'$, assuming without loss of generality $|S \setminus S'| = k$, and let $S_0 = S, S_1, S_2, \ldots, S_k = S'$ be a sequence of datasets where $|S_i - S_{i+1}| = 1$. Since the exponential mechanism is ν-BR, for any output set E and any i we have $\ln\left(\frac{\Pr[M(S_i)\in E]}{\Pr[M(S_{i+1})\in E]}\right) \in [-t_i, -t_i + \nu]$ for some $t_0, t_1, \ldots, t_{k-1}$. In turn, we have:

$$\begin{aligned}\ln\left(\frac{\Pr[M(S)\in E]}{\Pr[M(S')\in E]}\right) &= \ln\left(\frac{\Pr[M(S_0)\in E]}{\Pr[M(S_k)\in E]}\right) \\ &= \sum_{i=0}^{k-1} \ln\left(\frac{\Pr[M(S_i)\in E]}{\Pr[M(S_{i+1})\in E]}\right)\end{aligned}$$

is in $\left[-\sum_{i=0}^{k-1} t_i, -\sum_{i=0}^{k-1} t_i + k\nu\right]$. □

Theorem B.3. *Let $f : \mathbb{D} \to [m]$ be any function from the data universe to $[m]$. Let M be the following mechanism: Given the input S, partition S into $\{S_1, S_2, \ldots S_m\}$ where $S_i = \{x \in S : f(x) = i\}$, then return $(M_1(S_1), M_2(S_2), \ldots, M_m(S_m))$ where each M_i is an (independent) ν-BR mechanism (with respect to add-or-remove-one). Then M is $k\nu$-BR with respect to $\simeq_{AR\leq k}$.*

Proof. Let $S \simeq_{AR \leq k} S'$, assuming without loss of generality $|S \setminus S'| = k$. Let k_i be the number of elements in $S \setminus S'$ in S_i. By Theorem B.2, for some $t_1, t_2, \ldots t_m$ and any i, output o_i for M_i we have

$$\ln\left(\frac{M_i(S)(o_i)}{M_i(S')(o_i)}\right) \in [-t_i, -t_i + k_i\nu].$$

Since the mechanisms are run independently and the elements in S' are assigned to the same subset of the partition whether we receive input S or S', we have for any output $o = (o_1, o_2, \ldots o_m)$ for M:

$$\ln\left(\frac{M(S)(o)}{M(S')(o)}\right) = \sum_{i=1}^{m} \ln\left(\frac{M_i(S)(o_i)}{M_i(S')(o_i)}\right)$$

$$\in \left[-\sum_{i=1}^{m} t_i, -\sum_{i=1}^{m} t_i + \nu \sum_{i=1}^{m} k_i\right] = \left[-\sum_{i=1}^{m} t_i, -\sum_{i=1}^{m} t_i + k\nu\right].$$

□

Theorem B.4. *Algorithm 8 satisfies any privacy guarantee satisfied by the composition of a $k\varepsilon_1$-BR, $k\varepsilon_2$-BR, …, and $k\varepsilon_L$-BR mechanisms.*

Proof. We can assume the mechanism is run in a "top-down" fashion, i.e. we run all mechanisms at $\ell = i$ before running any mechanism at level $\ell = i - 1$. Viewing the concatenation of all mechanisms being run at $\ell = i$ as a single mechanism, Algorithm 8 is then just an adaptive composition of the per-level mechanisms. Each per-level mechanism is exactly a partition of the dataset and a run of the exponential mechanism on each subset of the partition, exactly as described/analyzed by Theorem B.3, giving the theorem. □

B.5 Measurement Recycling for Mixed Tabular Data

Beyond tight accounting, we can use an additional enhancement that allows us to skip the dedicated one-way marginal measurement step by reusing information measured during the preprocessing step of mixed tabular domains (open set discovery and quantile-based discretization). A naive implementation would treat preprocessing as disjoint from the application of the base mechanism by spending a fraction of the privacy budget to establish discrete domains, and a separate fraction of the privacy budget measuring one-way marginals. This is wasteful, as it pays twice for single-attribute information and diminishes the remaining budget available for capturing multivariate interactions. Rather than measuring one-way marginals anew, we eliminate this step entirely by recycling the private information obtained during preprocessing.

Under the hood, synthetic data mechanisms like `AIM` rely on estimation algorithms like PGM (McKenna et al., 2019). These attempt to estimate the discrete data distribution $\hat{p}$ given noisy linear measurements of its marginals of the form $y = Qp + \xi$, where p denotes the true data vector, Q is a linear query matrix, and ξ is zero-centered noise with standard deviation σ. Whereas closed-set categorical attributes directly yield complete noisy histograms and thus total dataset size estimates, open-set and numerical attributes provide only partial information that we treat as follows:

- For open-set categories with discovered frequent items $\widehat{\mathcal{V}} = \{k_1, \ldots, k_m\}$ and an unmeasured tail bucket at index 0, partition selection provides noisy counts $\tilde{y} \in \mathbb{R}^m$. We incorporate these as a partial linear measurement with query matrix $Q_{\text{open}} = [0_{m\times 1} \quad I_m]$, which depends only on the observed categories while leaving the unmeasured tail unconstrained.
- For continuous attributes handled via quantile discretization, the distribution over bins is approximately uniform by definition of quantiles. Given an estimate of the total dataset size $\tilde{N}$, we construct an identity query $Q_{\text{num}} = I_K$ with target $y_{\text{num}} = {}^{\tilde{N}}/_{K} \cdot \mathbf{1}_K$ and (heuristic) uncertainty $\sigma_{\text{num}} \propto {}^{1}/_{\sqrt{\rho}}$ calibrated to the ρ-zCDP quantile budget.

Accounting for precisely what was measured during preprocessing allows downstream models to initialize one-way marginals for free, preserving the entire synthesis budget for modeling higher-order dependencies.

When applied to 1D data, the improvement from measurement recycling was easily observable and matches expectations, as shown in Fig. 7. When we applied this to an end-to-end synthetic data

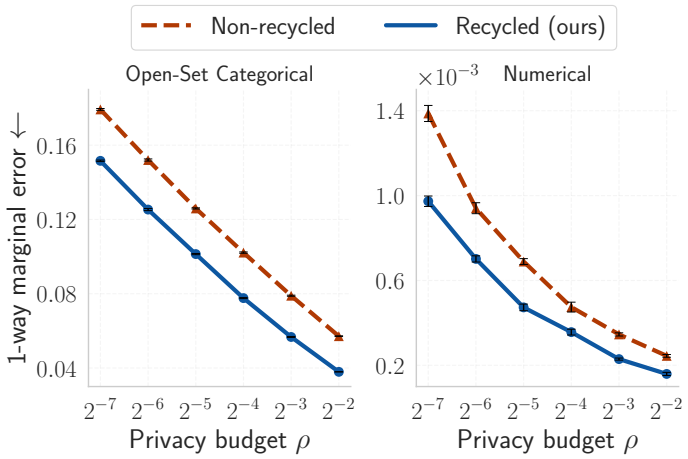


Figure 7: One-way marginal measurement recycling after preprocessing.

Algorithm 1 `MixCal`

Require: Target guarantee $(\varepsilon^*, \delta^*)$; mechanism family $M_i(\cdot)$ each parameterized by additive budget b; weights $w_1, \ldots, w_k$ with $\sum_i w_i = 1$

function PLDAccounting(b) ▷ privacy cost of the composition at budget b
  for $i = 1, \ldots, k$ **do**
    $b_i \leftarrow w_i\, b$ ▷ allocate the budget additively
    Calibrate $M_i(\cdot)$ to satisfy b_i privacy.
  return ε of $M_1 \circ \cdots \circ M_k$ at δ^* ▷ via a PLD accountant

Binary search for the largest b with PLDAccounting$(b) \leq \varepsilon^*$.
return Calibrated mechanism $M = M_1 \circ \cdots \circ M_k$.

mechanism like `AIM`, however, we observed no meaningful difference in performance between the recycling approach and the naive baseline that splits the budget between preprocessing and one-way marginal measurement. We therefore consider this a *negative* result. Future work, however, could investigate whether there exist settings where marginal recycling provides benefits.

C Omitted Details

C.1 `MixCal` Details

In `MixCal`, we introduce a budget parameter b that should represent an additive privacy notation like zCDP or GDP. Suppose that a budget split for a given component is b_i. Then, in `MixCal`, we calibrate the component to satisfy this privacy guarantee b_i.

A practically useful property of `MixCal` is that the correctness of the final guarantee only depends on the correctness of PLD accounting. Therefore, intermediate calibration steps to budget shares need only be approximate. For instance, when using GDP as the budget notion, we can calibrate a Gaussian mechanism exactly to its budget share $b_i = \mu_i$ by setting $\sigma = 1/\mu_i$, but when calibrating Threshold Gaussian with a relatively small probability of failure, e.g., $\tau < 10^{-5}$, using the results of Appendix B, we can ignore the failure term and treat the mechanism as Gaussian for the purposes of budget allocation in `MixCal`.

For instance, using a zCDP interpretation of b_i, we could use:

- Gaussian Mechanism: $\sigma = \sqrt{\frac{1}{2b_i}}$ (exact)
- Exponential Mechanism $\nu = \sqrt{8b_i}$ (pessimistic)

Table 5: Datasets used in our discrete/end-to-end synthetic data experiments. n is the number of training records, d the number of columns, and d_{num} and d_{cat} the numbers of numerical and categorical columns, respectively. The final column reports the categorical column with the largest domain and its domain size.

Dataset	n	d	d_{num}	d_{cat}	**Largest categorical domain**
Adult	32,561	11	0	11	native_country (42)
Fire	244,106	10	0	10	Station Area (46)
Page blocks	4,378	11	10	1	block (2)
Gas	22,191	11	10	1	TEY (2)

- Threshold Gaussian: $\sigma = \sqrt{\frac{1}{2b_i}}$ (approximate)
- DP-SGD noise multiplier: $\sigma = \sqrt{\frac{T}{2b_i}}$, where T is the number of iterations (approximate; assuming no amplification).

D Additional Experiment Details

D.1 Datasets

For our discrete and end-to-end synthetic data experiments, we use four standard datasets from the machine learning literature that have been extensively used in prior work on DP synthetic data (see, e.g., McKenna et al., 2022; Tao et al., 2022; Annamalai et al., 2024; Ganev et al., 2025b): Adult, Fire, Page blocks, and Gas. A summary is provided in Table 5. We use two versions of each dataset: 1) a fully discretized version, using uniform discretization into 16 bins, for our discrete experiments, and 2) the original, unprocessed version for our end-to-end experiments.

For our multimodal experiments, we use the bioRxiv dataset, a corpus of scientific abstracts ($n = 34k$, average number of tokens per sample around 300). We use the filtered dataset from (Hu et al., 2026), that is, we keep only the abstracts that appear after the knowledge cutoff date of Gemma-3 family models (Aug 2024). We perform train/validation/test split and obtain a train set of size $n = 28,846$.

D.2 Discrete Synthetic Data with `AIM-GDP`

Detailed results for all four (discrete) datasets are shown in Fig. 8, with additional diagnostic experiments on the Adult dataset presented in Figs. 9 and 10. We discuss these experiments in Section 5.1.

D.2.1 Attack-Aware Calibration

In the standard approach to calibrating DP mechanisms, we start with a target $(\varepsilon^\star, \delta^\star)$-DP guarantee for a small $\delta^\star$, usually smaller than the size of the dataset (see, e.g., Ponomareva et al., 2023). We can then evaluate the risk due to MIAs (Kulynych et al., 2024) or other attacks (Kulynych et al., 2025) by converting $(\varepsilon^\star, \delta^\star)$-DP guarantee to f-DP via a standard conversion (Dong et al., 2022):

$$f_{\varepsilon,\delta}(\alpha) = \max\{0, 1 - e^{\varepsilon}\alpha - \delta, e^{-\varepsilon}(1 - \alpha - \delta)\}. \tag{22}$$

Respectively, when we want to calibrate $\varepsilon^\star$ to a given level of risk in terms of f-DP, i.e., $(\alpha^\star, \beta^\star)$, we solve Eq. (22) for a given small $\delta^\star$. We refer to this as *standard risk calibration*.

In *attack-aware calibration* (Kulynych et al., 2024), we start with the f-DP target $(\alpha^\star, \beta^\star)$, and calibrate the parameters of the DP mechanism directly to satisfy the f-DP level. In the case of a GDP mechanism, we can then solve the direct f-DP expression to derive the μ parameter that satisfies the target $(\alpha^\star, \beta^\star)$. In the case of PLD analysis, we can can run search, e.g., bisection, for the mechanism parameters until the f-DP curve dervied from the PLD (Kulynych et al., 2024) satisfies the target requirement.

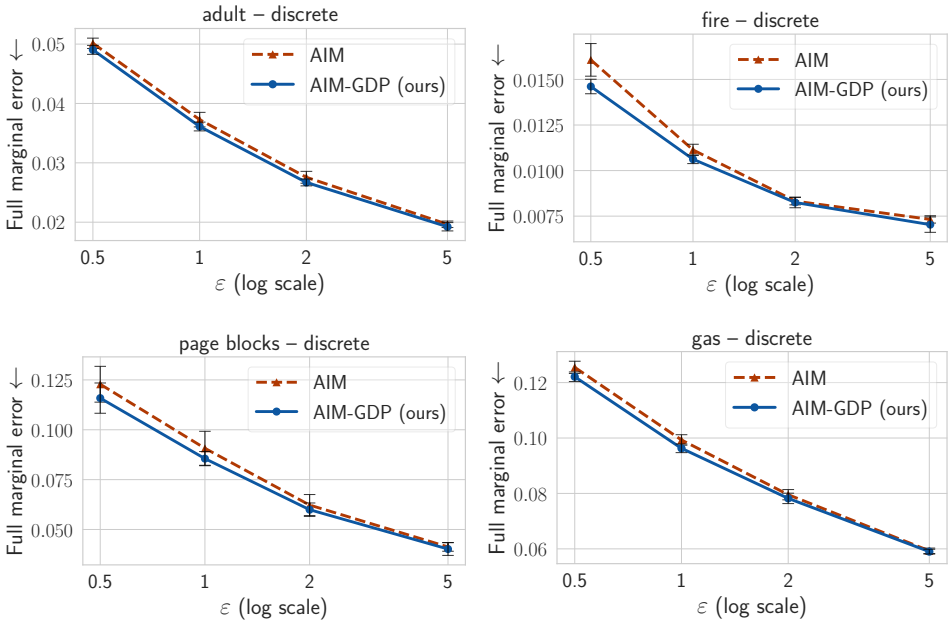


Figure 8: AIM-GDP vs. AIM across full marginal error across four datasets in a discrete pipeline. We report 95% bootstrap CI over the replications.

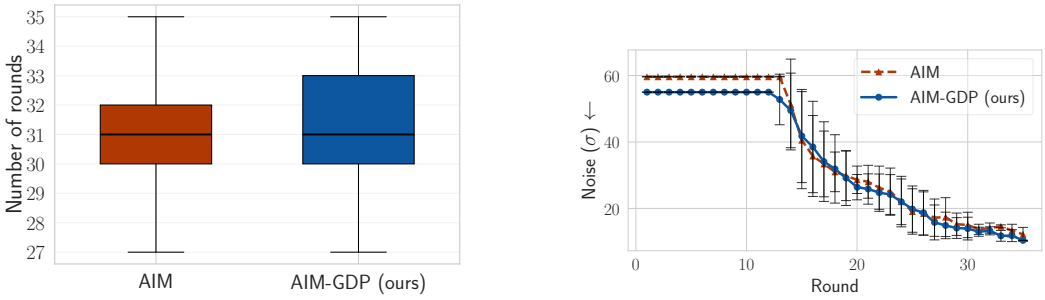


Figure 9: AIM-GDP vs. AIM across number of rounds and noise (σ) per round averaged across 25 runs on Adult in a discrete pipeline.

D.2.2 AUDITING

DP auditing empirically tests whether the privacy leakage of an implementation is consistent with its theoretical guarantee (Nasr et al., 2023; Annamalai et al., 2024). Following Ganev et al. (2026), we formulate auditing as a hypothesis-testing problem: a challenger repeatedly runs the mechanism on two neighboring datasets, and an adversary attempts to distinguish which dataset was used from the mechanism's output. The resulting FPR and TPR define an empirical point on the privacy trade-off curve. As AIM-GDP directly satisfies GDP, we convert this point into a statistical lower bound μ_{emp} on the privacy leakage and compare it directly with the theoretical μ.

We use the worst-case neighboring datasets $D_{\text{out}} = \{[0,0,0]\}^{10}$ and $D_{\text{in}} = D_{\text{out}} \cup \{[1,1,1]\}$, chosen to maximize the target record's influence on the measured marginals. We train 10,000 independent instances of AIM-GDP, equally split between D_{out} and D_{in}, and generate synthetic records from each model. Unlike prior audits of AIM, we run the full AIM-GDP model, including its adaptive/data-dependent marginal selection.

We consider black-box attacks using features computed from the released synthetic data, white-box attacks using the mechanism's internal measurements, and an ensemble using both. We split the 10,000 labeled runs into 4,000 training, 2,000 validation, and 4,000 test examples, train an XGBoost

classifier (Chen & Guestrin, 2016), and select the decision threshold on the validation set to maximize attack advantage ($\mathrm{TPR} - \mathrm{FPR}$). We evaluate the selected attack once on the held-out test set and use the resulting FPR/TPR with the Bayesian procedure (Zanella-Béguelin et al., 2023; Ganev et al., 2026) to obtain a 90%-posterior lower bound μ_{emp}. The resulting audit is discussed in Section 5.1.

D.3 End-to-End Synthetic Data with AIM-GDP, DP preprocessing, and MixCal

We show detailed results for all four (original/mixed-type) datasets in an end-to-end pipeline in Fig. 11. Additionally, we evaluate the pipeline under Poisson subsampling in simulation (Fig. 12). We discuss these in Section 5.2.

D.4 Generating Private Synthetic Data with LLMs using ACTG

We follow the conditional generation approach of Hu et al. (2026), which we summarize here. The goal is to generate a synthetic version of the *bioRxiv* dataset. First, Hu et al. fed each abstract in the *bioRxiv* dataset to Gemini-2.5-flash-lite to extract structured features. In particular, the following 8 features were extracted: primary-research-area, model-organism, experimental-approach, dominant-data-type, research-focus-scale, disease-mention, sample-size, research-goal. Each feature is categorical. See Figure 7 in Hu et al. (2026) for the allowed values of each category. The result is 8 categorical features for each abstract in the dataset.

These features are treated as sensitive, hence AIM is run on them to generate a list of synthetic features. The next step is to condition a DP-finetuned language model on these privatized features. Here, Gemma 3 1B is fine tuned with differential privacy on the feature/abstract pairs. We used a batch size of 448 and iterations $T = 512$ (8 epochs). We set clipping norm c to 1. We tune the learning rate $\eta \in \{1 \times 10^{-4}, 3 \times 10^{-4}, 1 \times 10^{-3}\}$, with 3×10^{-4} converging the fastest. We used a cosine decay learning rate schedule. We perform LoRA fine-tuning with a LoRA rank of 8, $\alpha = 16$ and dropout of 0.05. Training took 3 hours on 2 A100 GPUs, where data parallel was used to parallelize training. Since the 95% quantile of the abstracts in bioRxiv had 512 tokens, we used a context length of 512. We use the MAUVE scores (Pillutla et al., 2023) as the main evaluation metric for this demonstration.

We remark that we trained for a tenth as long as Hu et al. did, as they trained for 80 epochs whereas we trained for 8. Moreover, we used a smaller learning rate of 3×10^{-4} whereas they used 1×10^{-3}. While our evaluation loss values are within a tenth of (Hu et al., 2026), our MAUVE scores are smaller than (Hu et al., 2026) reported scores, likely due to this reason. Namely, we report a MAUVE score of 60% at $\varepsilon = 1$ for the RDP baseline, which corresponds to vanilla ACTG, where as (Hu et al., 2026) reports a little under 80%. Nevertheless, we are confident that, had we run for the full 80 epochs, we would still see improvement in our PLD approach since our approach allows for training at a smaller DP-SGD noise multiplier as reported in Table 4.

Once the model was fine tuned, we used nucleus decoding to sample conditioned on the features from AIM. We used temperature $T = 1.0$, top-p $= 0.95$, top-k $= 0$, and a batch size of 64. We sampled $N_{\text{syn}} = 5{,}000$ abstracts. Sampling took 3.5 hours on 2 A100 GPUs, where the parallelization was done by feeding half the features to condition on (2,500) to each GPU.

The MAUVE score from Fig. 1a was computed between the $5{,}000$ synthetic examples sampled from the fine tuned LLM and the $5{,}000$ real heldout examples from the test data from bioRxiv. MAUVE scores use an embedding model. Following (Hu et al., 2026), we used Spector2 (Singh et al., 2023), a specialized sentence embedding model finetuned on abstracts of scientific papers. (Hu et al., 2026) showed that this domain-specific embedding model improved the reliability of the MAUVE scores.

We repeated the ACTG training pipeline twice, once using Hu et al.'s original accounting approach, which used RDP accounting with AIM, and our proposed technique, which uses AIM-GDP and PLD accounting. The MixCal algorithm was used in the following manner. We discuss the PLD approach here. The RDP approach follows analogously. Suppose our target is for the composition of AIM-GDP and DP-SGD to satisfy $(\varepsilon^\star, \delta^\star)$-DP. Given a starting dummy budget $(\varepsilon_{\text{dummy}}, \delta_{\text{dummy}})$, we allocate 71% to AIM and 29% to DP-SGD following (Hu et al., 2026). By this, we mean we calibrate AIM-GDP to satisfy $(0.71 \cdot \varepsilon_{\text{dummy}}, 0.71 \cdot \delta_{\text{dummy}})$-DP, and DP-SGD to satisfy $(0.29 \cdot \varepsilon_{\text{dummy}}, 0.29 \cdot \delta_{\text{dummy}})$-DP. We then run PLD composition on the mechanism. As mentioned in the main body, the resulting PLD ε will larger than our target $\varepsilon^\star$, hence we run a binary search over the dummy value of $\varepsilon_{\text{dummy}}$ until the composed mechanisms under PLD composition satisfies our target $(\varepsilon^\star, \delta^\star)$.

Mechanism	Param	Base Events	Fully Adaptive	Accounting
Preprocessing & Domain Discovery				
DP Quantiles (Gillenwater et al., 2021)	ε	Exponential	No	ADP
DP Quantiles (Kaplan et al., 2022)	$\varepsilon_1, \ldots, \varepsilon_l$	Exponential	No	zCDP
Noisy Histogram	σ	Gaussian	No	GDP
Threshold Gaussian (Gopi et al., 2020)	σ, δ	Gaussian Failure	No	GDP
Optimal Partition (Desfontaines et al., 2022)	ε, δ	Laplace Failure	No	ADP
RDP Partition (Harrison & Manurangsi, 2026)	σ, δ	Gaussian Failure	No	RDP
Base Mechanism				
AIM (McKenna et al., 2022)	ρ	Exponential Gaussian	Yes	zCDP
AIM-GDP (Algorithm 5)	μ	Exponential Gaussian	Yes	GDP
GEM (Liu et al., 2021)	ρ	Exponential Gaussian	No	zCDP
RAP (Aydöre et al., 2021)	ρ	Exponential Gaussian	No	zCDP
PrivSyn (Zhang et al., 2021)	ρ	Exponential Gaussian	No	GDP
MST (McKenna et al., 2021)	ρ	Exponential Gaussian	No	zCDP
Language Model Fine-Tuning				
Poisson DP-SGD	σ	Sub. Gaussian	No	RDP, PLD
Random Allocation DP-SGD	σ	Sub. Gaussian	No	PLD
DP-BandMF	σ	Corr. Gaussian	No	Monte Carlo
DP-MF	σ	Gaussian	No	GDP

Table 6: Classification of differentially private synthetic data mechanisms across domain discovery, tabular synthesis, and language model fine-tuning pillars.

E Additional Figures

Algorithm 8 DP Quantile Discretization via Recursive Medians

Require: 1D histogram $\mathbf{c} = (c_0, \ldots, c_{K-1}) \in \mathbb{N}^K$, tree depth L, per-level privacy budgets $\{\epsilon_\ell\}_{\ell=1}^{L}$.
Ensure: Set of discretized bin cutpoints $\mathcal{E} \subset \{0, \ldots, K-1\}$.

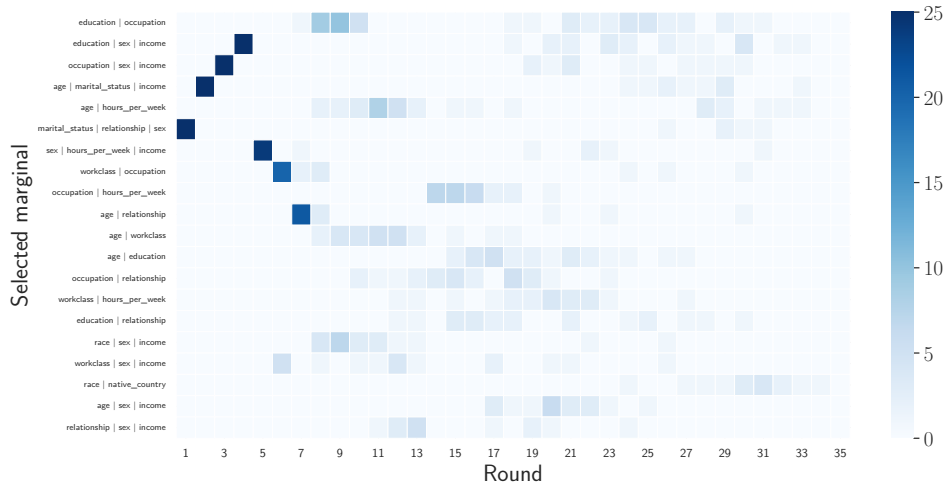
function RECURSIVEBISECTION(lo, hi, ℓ)
  if $\ell = 0$ **then**
    return $\emptyset$
  $r \leftarrow \frac{1}{2} \sum_{j=\text{lo}}^{\text{hi}-1} c_j$ ▷ Target median rank within current sub-interval
  for each candidate split $v \in \{\text{lo}, \ldots, \text{hi}-1\}$ **do**
    $L_v \leftarrow \sum_{j=\text{lo}}^{v-1} c_j, \quad R_v \leftarrow \sum_{j=\text{lo}}^{v} c_j$ ▷ Cumulative rank bounds of cell v
    $u(v) \leftarrow -\max(0,\, L_v - r,\, r - R_v)$ ▷ Distance to target median rank
  Sample $v^* \in \{\text{lo}, \ldots, \text{hi}-1\}$ w.p. $\propto \exp(\epsilon_\ell \cdot u(v))$ ▷ Exp. Mech. with $\Delta u = 1/2$
  $\mathcal{E}_{\text{left}} \leftarrow$ RECURSIVEBISECTION(lo, v^*, $\ell - 1$)
  $\mathcal{E}_{\text{right}} \leftarrow$ RECURSIVEBISECTION(v^*, hi, $\ell - 1$)
  return $\mathcal{E}_{\text{left}} \cup \{v^*\} \cup \mathcal{E}_{\text{right}}$

return RECURSIVEBISECTION$(0, K, L)$

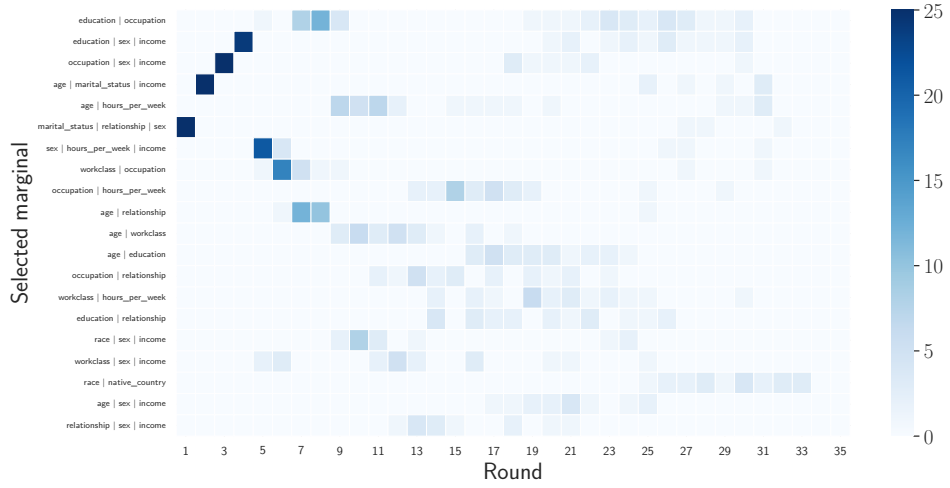
Algorithm 9 DP Open-Set Partition Selection via Threshold Gaussian

Require: Partition counts $\mathbf{y} \in \mathbb{N}^{|\mathcal{V}|}$ over active keys $\mathcal{V} = \{k : y_k \geq \tau_{\min}\}$, noise scale σ, target failure probability δ, minimum threshold $\tau_{\min}$.
Ensure: Discovered vocabulary $\widehat{\mathcal{V}}$, noisy partition counts $\tilde{\mathbf{y}}$, and closed-domain mapping $\mathcal{M}$.

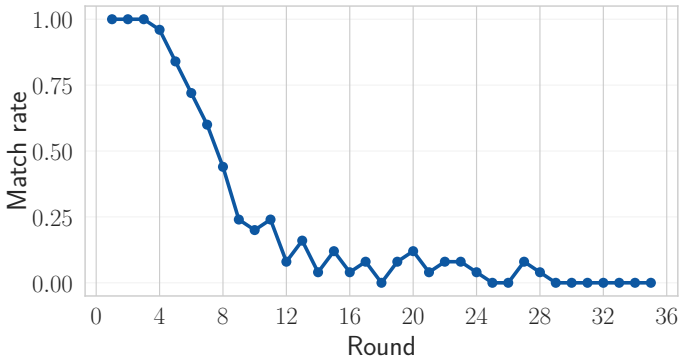
$T \leftarrow \tau_{\min} + \sigma\Phi^{-1}(1-\delta)$ ▷ Threshold cutoff bounding false positives to $\leq \delta$ per empty bin
$\widehat{\mathcal{V}} \leftarrow \emptyset, \quad \tilde{\mathbf{y}} \leftarrow \emptyset$
for each active partition key $k \in \mathcal{V}$ **do**
  Sample $\xi_k \sim \mathcal{N}(0, \sigma^2)$
  $\tilde{y}_k \leftarrow y_k + \xi_k$ ▷ Perturb count with Gaussian noise
  if $\tilde{y}_k \geq T$ **then**
    $\widehat{\mathcal{V}} \leftarrow \widehat{\mathcal{V}} \cup \{k\}$ ▷ Retain key in discovered vocabulary
    $\tilde{\mathbf{y}}[k] \leftarrow \tilde{y}_k$
Closed vocabulary domain $\mathcal{D} \leftarrow \widehat{\mathcal{V}} \cup \{\texttt{[OTHER]}\}$
Construct mapping $\mathcal{M}(k) = k$ if $k \in \widehat{\mathcal{V}}$ else `[OTHER]` ▷ Collapse unselected tail categories
return Discovered vocabulary $\widehat{\mathcal{V}}$, noisy counts $\tilde{\mathbf{y}}$, mapping $\mathcal{M}$



(a) AIM



(b) AIM-GDP (ours)



(c) Selected marginal match rate

Figure 10: AIM-GDP vs. AIM across the selected marginal per round averaged across 25 runs on Adult in a discrete pipeline.

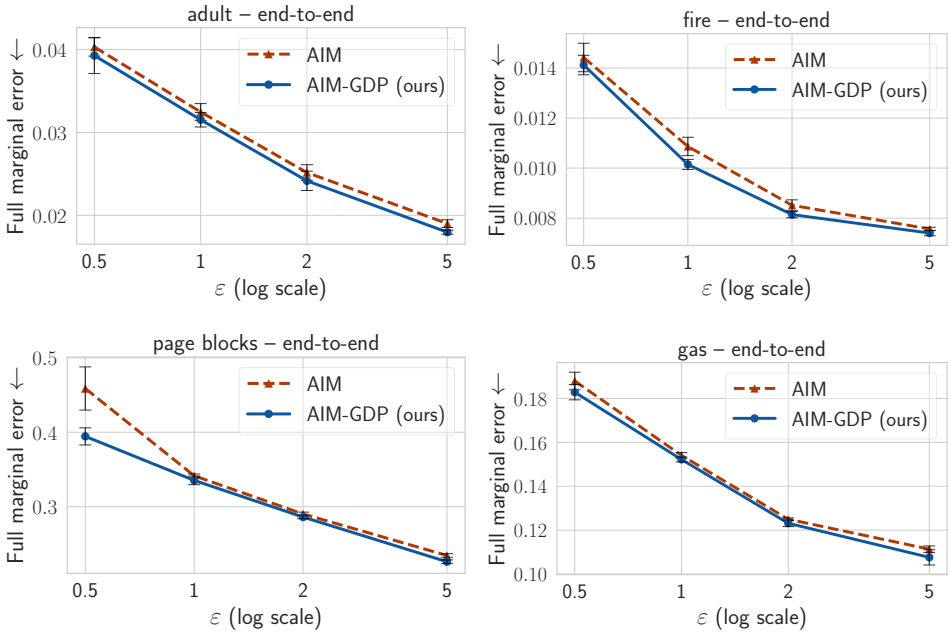


Figure 11: `AIM-GDP` vs. `AIM` across full marginal error across four datasets in an end-to-end pipeline. We report 95% bootstrap CI over the replications.

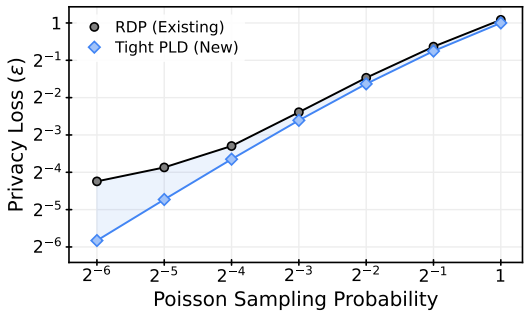


Figure 12: Privacy loss ε at $\delta = 10^{-6}$ across Poisson sampling probabilities $p = 2^{-k}$ ($k = 0, \ldots, 6$), with ρ calibrated to an un-amplified $\varepsilon = 1.0$ at $p = 1$.

Algorithm 2 AIM

Input: Dataset D, workload W, privacy parameter ρ (zCDP)
Output: Synthetic dataset $\hat{D}$
Hyper-Parameters: MAX-SIZE=80MB, $T = 16d$, $\alpha = 0.9$
$\sigma_0 = \sqrt{T/(2\,\alpha\,\rho)}$
$\rho_{\text{used}} = 0$
$t = 0$
Initialize $\hat{p}_t$ using Algorithm 3
$w_r = \sum_{s \in W} c_s \mid r \cap s \mid$
$\sigma_{t+1} \leftarrow \sigma_0$ $\quad \nu_{t+1} \leftarrow \sqrt{8(1-\alpha)\rho/T}$
while $\rho_{\text{used}} < \rho$ **do**
  $t = t + 1$
  $\rho_{\text{used}} \leftarrow \rho_{\text{used}} + \frac{1}{8}\nu_t^2 + \frac{1}{2\sigma_t^2}$
  $C_t = \left\{ r_t \in W_+ \mid \text{JT-SIZE}(r_1, \ldots, r_t) \leq \frac{\rho_{\text{used}}}{\rho} \cdot \text{MAX-SIZE} \right\}$
  select $r_t \in C_t$ using the exponential mechanism with:

$$q_r(D) = w_r\Big(\|M_r(D) - M_r(\hat{p}_{t-1})\|_1 - \sqrt{2/\pi}\cdot\sigma_t\cdot n_r \Big)$$

  measure marginal on r_t:

$$\tilde{y}_t = M_{r_t}(D) + \mathcal{N}(0, \sigma_t^2 \mathbb{I})$$

  estimate data distribution using PGM:

$$\hat{p}_t = \arg\min_{p \in S} \sum_{i=1}^{t} \frac{1}{\sigma_i} \|M_{r_i}(p) - \tilde{y}_i\|_2^2$$

  anneal ν_{t+1} and σ_{t+1} using Algorithm 4
generate synthetic data $\hat{D}$ from $\hat{p}_t$ using PGM
return $\hat{D}$

Algorithm 3 Initialize p_t (subroutine of Algorithm 2)

for $r \in \{r \in W_+ \mid |r| = 1\}$ **do**
  $t = t + 1 \quad \sigma_t \leftarrow \sigma_0 \quad r_t \leftarrow r$
  $\tilde{y}_t = M_r(D) + \mathcal{N}(0, \sigma_t^2 \mathbb{I})$
  $\rho_{\text{used}} \leftarrow \rho_{\text{used}} + \frac{1}{2\sigma_t^2}$
$\hat{p}_t = \arg\min_{p \in S} \sum_{i=1}^{t} \frac{1}{\sigma_i} \|M_{r_i}(p) - \tilde{y}_i\|_2^2$

Algorithm 4 Budget annealing (subroutine of Algorithm 2)

if $\|M_{r_t}(\hat{p}_t) - M_{r_t}(\hat{p}_{t-1})\|_1 \leq \sqrt{2/\pi} \cdot \sigma_t \cdot n_{r_t}$ **then**
  $\nu_{t+1} \leftarrow 2 \cdot \nu_t$
  $\sigma_{t+1} \leftarrow \sigma_t / 2$
else
  $\nu_{t+1} \leftarrow \nu_t$
  $\sigma_{t+1} \leftarrow \sigma_t$
if $(\rho - \rho_{\text{used}}) \leq 2\left(\frac{1}{2\sigma_{t+1}^2} + \frac{1}{8}\nu_{t+1}^2\right)$ **then**
  $\nu_{t+1} = \sqrt{8 \cdot (1-\alpha) \cdot (\rho - \rho_{\text{used}})}$
  $\sigma_{t+1} = \sqrt{1/(2 \cdot \alpha \cdot (\rho - \rho_{\text{used}}))}$

Algorithm 5 AIM-GDP

Input: Dataset D, workload W, privacy parameter μ^2 (GDP)
Output: Synthetic dataset $\hat{D}$
Hyper-Parameters: MAX-SIZE=80MB, $T = 16d$, $\alpha = 0.9$
$\sigma_0 = \sqrt{T/(\alpha\,\mu^2)}$
$\mu^2_{\text{used}} = 0$
$t = 0$
Initialize $\hat{p}_t$ using Algorithm 6
$w_r = \sum_{s \in W} c_s \mid r \cap s \mid$
$\sigma_{t+1} \leftarrow \sigma_0$ $\quad \nu_{t+1} \leftarrow b^{-1}\big((1-\alpha)\mu^2/T\big)$
while $\mu^2_{\text{used}} < \mu^2$ **do**
  $t = t + 1$
  $\mu^2_{\text{used}} \leftarrow \mu^2_{\text{used}} + b(\nu_t) + \frac{1}{\sigma_t^2}$
  $C_t = \left\{ r_t \in W_+ \mid \text{JT-SIZE}(r_1, \ldots, r_t) \leq \frac{\mu^2_{\text{used}}}{\mu^2} \cdot \text{MAX-SIZE} \right\}$
  select $r_t \in C_t$ using the exponential mechanism with:

$$q_r(D) = w_r\Big(\|M_r(D) - M_r(\hat{p}_{t-1})\|_1 - \sqrt{2/\pi}\cdot\sigma_t\cdot n_r \Big)$$

  measure marginal on r_t:

$$\tilde{y}_t = M_{r_t}(D) + \mathcal{N}(0, \sigma_t^2 \mathbb{I})$$

  estimate data distribution using PGM:

$$\hat{p}_t = \arg\min_{p \in S} \sum_{i=1}^{t} \frac{1}{\sigma_i} \|M_{r_i}(p) - \tilde{y}_i\|_2^2$$

  anneal ν_{t+1} and σ_{t+1} using Algorithm 7
generate synthetic data $\hat{D}$ from $\hat{p}_t$ using PGM
return $\hat{D}$

Algorithm 6 Initialize p_t (subroutine of Algorithm 5)

for $r \in \{r \in W_+ \mid |r| = 1\}$ **do**
  $t = t + 1 \quad \sigma_t \leftarrow \sigma_0 \quad r_t \leftarrow r$
  $\tilde{y}_t = M_r(D) + \mathcal{N}(0, \sigma_t^2 \mathbb{I})$
  $\mu^2_{\text{used}} \leftarrow \mu^2_{\text{used}} + \frac{1}{\sigma_t^2}$
$\hat{p}_t = \arg\min_{p \in S} \sum_{i=1}^{t} \frac{1}{\sigma_i} \|M_{r_i}(p) - \tilde{y}_i\|_2^2$

Algorithm 7 Budget annealing (subroutine of Algorithm 5)

if $\|M_{r_t}(\hat{p}_t) - M_{r_t}(\hat{p}_{t-1})\|_1 \leq \sqrt{2/\pi} \cdot \sigma_t \cdot n_{r_t}$ **then**
  $\nu_{t+1} \leftarrow 2 \cdot \nu_t$
  $\sigma_{t+1} \leftarrow \sigma_t / 2$
else
  $\nu_{t+1} \leftarrow \nu_t$
  $\sigma_{t+1} \leftarrow \sigma_t$
if $(\mu^2 - \mu^2_{\text{used}}) \leq 2\left(\frac{1}{\sigma_{t+1}^2} + b(\nu_{t+1})\right)$ **then**
  $\nu_{t+1} = b^{-1}\big((1-\alpha) \cdot (\mu^2 - \mu^2_{\text{used}})\big)$
  $\sigma_{t+1} = \sqrt{1/(\alpha \cdot (\mu^2 - \mu^2_{\text{used}}))}$

Figure 13: Left: the original AIM with zCDP accounting (orange). Right: AIM-GDP, the same algorithm with the budget tracked in GDP units μ^2, which compose additively (teal). The Gaussian mechanism with noise scale σ is exactly $1/\sigma$-GDP, and the exponential mechanism with parameter ν is a ν-BR mechanism with exact squared GDP budget $b(\nu) = \left(2\Phi^{-1}\left(\frac{1}{e^{\nu/2}+1}\right)\right)^2$ given by Theorem 3.2, with $b^{-1}(\mu^2) = 2\ln\frac{\Phi(\mu/2)}{\Phi(-\mu/2)}$ denoting its inverse used for calibrating the parameter to a given budget.