Journal of Machine Learning Research 27 (2026) 1-44 Submitted 5/25; Revised 7/26; Published 8/26

Particle Filter for Bayesian Inference on Privatized Data

Yu-Wei Chen CHEN4357@PURDUE.EDU
Department of Statistics
Purdue University
West Lafayette, IN 47907, USA

Pranav Sanghi PSANGHI04@GMAIL.COM
Department of Computer Science
Purdue University
West Lafayette, IN 47907, USA

Jordan Awan JAA557@PITT.EDU
Department of Statistics
University of Pittsburgh
Pittsburgh, PA 15260, USA

Editor: Debdeep Pati

Abstract

Differential Privacy (DP) is a probabilistic framework that protects privacy while preserving data utility. To protect the privacy of the individuals in the data set, DP requires adding a precise amount of noise to a statistic of interest; however, this noise addition alters the resulting sampling distribution, making statistical inference challenging. One of the main DP goals in Bayesian analysis is to make statistical inference based on the private posterior distribution. While existing methods have strengths in specific conditions, they can be limited by poor mixing, strict assumptions, or low acceptance rates. We propose a novel particle filtering algorithm, which features (i) consistent estimates, (ii) Monte Carlo error estimates and asymptotic confidence intervals, (iii) computational efficiency, and (iv) accommodation to a wide variety of priors, models, and privacy mechanisms with minimal assumptions. We empirically evaluate our algorithm through a variety of simulation settings as well as an application to a 2021 Canadian census data set, demonstrating the efficacy and adaptability of the proposed sampler.

Keywords: differential privacy (DP), sequential Monte Carlo (SMC), approximate Bayesian computation (ABC), Markov chain Monte Carlo (MCMC)

1. Introduction

Differential Privacy (DP; Dwork et al., 2006) is a rigorous probabilistic framework designed to protect the privacy of individuals in a data set by adding noise to the queries made on the data. It has become widely adopted in academia (Vadhan, 2017 from Harvard OpenDP), industrial companies (Erlingsson et al., 2014, Abadi et al., 2016 from Google Research; Dwork, 2006, Ding et al., 2017 from Microsoft Research), and governments (Abowd et al., 2022 from U.S. Census Bureau). Such privacy-protection mechanisms inevitably come with many challenges in conducting valid statistical inference on privatized data. One of the major challenges is that, given a prior model $\theta \sim \pi_0(\cdot)$ for the target parameters, a data

model $x \sim f(\cdot \mid \theta)$ for the sensitive data, and a privacy mechanism $s_{dp} \mid x \sim m(\cdot \mid x)$, the marginal private posterior distribution $\pi(\theta \mid s_{dp}) \propto \pi_0(\theta) \int f(x \mid \theta) m(s_{dp} \mid x)\, dx$ is typically intractable, as it requires integrating over all possible databases x. The doubly intractable computation (Murray et al., 2006) cannot be directly solved by traditional Markov Chain Monte Carlo (MCMC) methods, such as the Metropolis-Hastings algorithm, but requires a more careful approach.

Under the Bayesian regime, MCMC methods have arisen to exactly sample from or at least to target sampling from the private posterior $\pi(\theta \mid s_{dp})$ (Bernstein and Sheldon, 2018 & 2019; Gong, 2022; Ju et al., 2022). (i) Bernstein and Sheldon (2018 & 2019) are limited to certain priors, data models, and mechanisms. (ii) Gong (2022) provides i.i.d. samples from the exact private posterior but can suffer from low acceptance rates. (iii) While Ju et al. (2022) gives a more generic missing-data sampling approach, it is optimized for conjugate prior distributions. In other words, along with the usual mixing issues in MCMC methods, it can be inefficient when non-conjugate prior distributions are used. Even though conjugate priors are analytically convenient, they are known to be non-robust to mis-specifications (Berger, 1985); for more robust Bayesian inference, Berger et al. (1994) recommend using heavier-tailed prior distributions than those arising from standard conjugate priors.

Particle filtering (PF), also known as Sequential Monte Carlo (SMC), is a powerful method of obtaining approximate samples from a posterior distribution, assuming the prior is proper and both the prior and data model can be easily sampled. The algorithm begins with a set of particles drawn from the prior distribution, which are then iteratively perturbed and filtered. Then, it evolves recursively through the following three steps: propagation (mutation), reweighting (correction), and resampling (selection). Through repeated iterations, the particle system gradually generates samples that approximate the target posterior distribution.

Under mild conditions—generally weaker than (geometric) ergodicity assumptions required in MCMC—PF exhibits favorable particle behavior, including strong convergence and stability (Chopin, 2004), while naturally accommodating parallel computation. Moreover, leveraging its inherent sequential construction at no additional computational cost, PF provides an estimator of the marginal likelihood (model evidence) via the product of incremental normalizing-constant ratios, thereby enabling Bayesian model comparison and selection. However, the preservation of these properties under differential privacy is not automatic and requires careful design of the privacy-aware procedure, along with theoretical analysis.

Our Contributions. In this paper, we develop a custom particle filtering algorithm to sample from the private posterior distribution $\pi(\theta, x \mid s_{dp})$. To tailor the algorithm to the DP setting, we introduce two customizations: (i) an artificial schedule on the privacy parameter ϵ, and (ii) a rejection step based on Gong (2022) that incorporates the privacy schedule. The schedule has a tempering effect on the intermediate target distribution, and the rejection step helps to efficiently target the private posterior distribution, avoiding the bias introduced by approximate Bayesian computing (ABC) approaches. Following the generic structure of particle filtering, our method

- is widely applicable and only requires the ability to sample from the prior and data model, and the ability to evaluate the densities of the prior and the privacy mechanism,

- uses a novel tempering of the privacy budget ϵ, in place of the usual ABC schedule,
- allows for parallelization across the particles in each cycle,
- results in consistent estimates, the central limit theorem (CLT), and accurate estimates of the Monte Carlo standard errors (MCSE), giving an asymptotic confidence interval (CI) for the true posterior quantities, and
- yields an estimate of the marginal likelihood for the sensitive data.

Importantly, these estimates are nontrivial to derive in the DP setting and may be compromised without the specific structural design of our proposed algorithm (see Section 3 for further elaboration). To evaluate our method, we apply it to location-scale normal, linear regression, and Gaussian mixture models to demonstrate its improved performance on computation efficiency compared to the data augmentation MCMC of Ju et al. (2022) and the rejection ABC algorithm of Gong (2022), which will be referred to hereafter as DP-DA-MCMC and DP-Reject-ABC, respectively. Additionally, we apply our method to 2021 Canadian Census data to perform a DP logistic regression via objective perturbation, showcasing its practical application.

Organization. The remainder of the paper is organized as follows. Section 2 reviews the definition of ϵ-differential privacy and key properties. We also introduce the structure and statistics of a typical particle filtering method (not specific to the privatized data setting). Section 3 first introduces the difference between our algorithm and a typical particle filtering scheme. Section 3.1 presents our main algorithm along with its assumptions and designs. In Section 3.2 and Section 3.3, we prove our estimator has strong consistency and asymptotic normality. In Section 3.4 and Section 3.5, we derive the asymptotic confidence interval for the true parameter and marginal likelihood of the sensitive data. In Section 4, we compare our method to the DP-DA-MCMC and DP-Reject-ABC in various simulation studies. In Section 5, we analyze the model parameters by applying our algorithm to a logistic regression via objective perturbation on the 2021 Canadian census data set. Section 6 discusses the implications and limitations of our work. Proof, technical details, and additional results are deferred to the appendices. The source code used to implement the simulation study and data analysis is available at `https://github.com/psanghi04/DP-PF`.

Related Work. Common strategies for statistical inference on privatized data can roughly be categorized into two major perspectives: asymptotic frequentist and Bayesian. The first employs traditional statistical asymptotics to approximate the sampling distribution of differentially private statistics, arguing that the privacy noise is negligible compared to sampling errors in large samples. However, these approximations can fall short in finite samples and require specific methods to account for privacy effects (Smith, 2011; Cai et al., 2021; Wang et al., 2018) and are tailored to specific privacy mechanisms (Gaboardi et al., 2016; Gaboardi and Rogers, 2017).

The second category emphasizes the marginal likelihood of model parameters, which involves high-dimensional integrals over unobserved confidential databases that are analytically tractable only in simple cases (Awan and Slavković, 2018, 2020). To address these challenges, various Markov Chain Monte Carlo (MCMC) methods have been proposed for specific privacy mechanisms and data models, including approaches for exponential random graph models, exponential family models, and generalized linear models (Karwa et al.,

2017; Bernstein and Sheldon, 2018, 2019; Schein et al., 2019; Kulkarni et al., 2021). Recent developments include Gong (2022), who showed that approximate Bayesian computation (ABC) can provide samples consistent with the marginal likelihood and Bayesian posterior for certain differentially private statistics. Ju et al. (2022) proposed a method for Bayesian inference that integrates existing samplers for non-private data while accounting for the privacy mechanism, although it relies on strong assumptions regarding geometric ergodicity and may be inefficient in some cases.

In addition, some approaches do not fit into the two categories. Awan and Wang (2025) developed a simulation-based finite-sample inference strategy for privatized data, providing accurate confidence sets at the cost of increased computation. Other frequentist approaches use different approximations, such as parametric bootstrapping (Ferrando et al., 2022; Wang et al., 2025a) and non-parametric bootstrapping (Wang et al., 2025b). There is also an approach based on Bayesian asymptotics, but this is currently limited to additive noise mechanisms and summation-based summaries without clamping (Awan et al., 2026).

2. Background

In this section, we review the necessary background and set the notation for the paper.

We let $x = (x_1, \ldots, x_n)$ be a database in $\mathcal{X}^n$ and under a data model $f(\cdot \mid \theta)$, where θ is the parameter of interest. In the DP framework, s_{dp} is denoted as the observed, privatized statistic, and $m_\epsilon(\cdot \mid x)$ is the density of a privacy mechanism conditional on confidential data x, where ϵ is the privacy parameter. In a Bayesian setting, we assume a prior $\pi_0(\cdot)$ on $\theta \in \Theta$ and the Bayesian model can be summarized as follows:

$$\begin{aligned} \theta &\sim \pi_0(\theta), \\ x \mid \theta &\sim f(x \mid \theta), \\ s_{dp} \mid x &\sim m_\epsilon(s_{dp} \mid x). \end{aligned}$$

To infer θ based on s_{dp}, we are interested in the private posterior

$$\pi(\theta \mid s_{dp}) \propto \pi_0(\theta) \int_{\mathcal{X}^n} f(x \mid \theta) m_\epsilon(s_{dp} \mid x)\, dx,$$

or more generally, in

$$\pi(\theta, x \mid s_{dp}) \propto \pi_0(\theta) f(x \mid \theta) m_\epsilon(s_{dp} \mid x).$$

2.1 Differential Privacy

Differential privacy (DP), introduced in Dwork et al. (2006), provides a framework to quantify the privacy risks associated with an algorithm and offers techniques to design privacy mechanisms that control privacy loss. To protect sensitive information in the data set, DP methods require adding a proper amount of randomness to the summary statistic with the goal of ensuring that adversaries cannot easily determine whether a specific individual was part of the data set. Meanwhile, the utility of the private output is ideally still informative for many population-level statistics.

Definition 1 (Privacy Mechanism) *Given a set $\mathcal{X}^n$, which is the collection of all possible databases, a* privacy mechanism $\mathcal{M}$ *is defined as a set of probability measures $\{M_x \mid x \in \mathcal{X}^n\}$ which take values on a common measurable space $\mathcal{Y}$.*

Definition 2 (ϵ-Differential Privacy: Dwork et al., 2006) *Given a set $\mathcal{X}^n$ and the Hamming distance $d_H(\cdot,\cdot)$ on $\mathcal{X}^n \times \mathcal{X}^n$, a privacy mechanism $\mathcal{M}$ satisfies ϵ-differential privacy (ϵ-DP), if for any two databases $x, x' \in \mathcal{X}^n$ such that $d_H(x, x') \leqslant 1$,*

$$\Pr[\mathcal{M}(x) \in S] \leqslant \exp(\epsilon) \Pr[\mathcal{M}(x') \in S]$$

for all measurable sets S.

The privacy parameter ϵ, also known as the privacy (loss) budget, captures the worst privacy loss between adjacent data sets. It is used for controlling the trade-off between data privacy and utility. A large ϵ value may preserve strong data utility but can cause severe privacy leakage. Thus, a typical range for ϵ is from 0.01 to 10 to ensure sufficient privacy protection while maintaining data utility. A common ϵ-DP mechanism is the Laplace mechanism, which adds noise drawn from a Laplace distribution to the summary. As $d_H(x, x') = 1$, it means that the two databases differ in exactly one record, and hence they are called adjacent data sets.

Definition 3 (Sensitivity) *let $f : \mathcal{X}^n \to \mathbb{R}$ be a statistic. The sensitivity of f is defined*

$$\Delta f = \max_{\substack{x,x' \in \mathcal{X}^n \\ d_H(x,x') \leqslant 1}} \|f(x) - f(x')\|_1,$$

where $\|\cdot\|_1$ is the ℓ_1-norm.

The sensitivity of a function quantifies how much the output of the function can change when a single individual's data in the input data set is altered. It is a critical concept in DP as it determines the amount of noise needed to be added to the function's output to achieve a desired level of privacy.

Proposition 4 (Laplace Mechanism: Dwork et al., 2006) *Let $M(D) = f(D) + L$, where $L = Lap\left(0, \frac{\Delta f}{\epsilon}\right)$ with density $f_L(x) = \frac{\epsilon}{2\Delta f} \exp(-\frac{\epsilon}{\Delta f}|x|)$. Then, $\mathcal{M}$ satisfies ϵ-DP.*

The following are two important properties in ϵ-DP:

- (Composition) If a mechanism $\mathcal{M}_1 : \mathcal{X}^n \to \mathcal{Y}$ satisfies ϵ_1-DP and another mechanism $\mathcal{M}_2 : \mathcal{X}^n \to \mathcal{Z}$ satisfies ϵ_2-DP, then the composed mechanism which jointly releases $(\mathcal{M}_1(D), \mathcal{M}_2(D))$ satisfies $(\epsilon_1 + \epsilon_2)$-DP.
- (Post-Processing: Dwork et al. (2014)) If $\mathcal{M} : \mathcal{X}^n \to \mathcal{Y}$ is an ϵ-DP mechanism and $g : \mathcal{Y} \to \mathcal{Z}$ is another mechanism, then $g \circ \mathcal{M} : \mathcal{X}^n \to \mathcal{Z}$ satisfies ϵ-DP.

Composition describes the privacy leakage as multiple data summaries are published. Post-processing is a property of differential privacy that ensures the privacy guarantees are preserved under any data-independent transformation of the output, such as our posterior analysis. We use the Laplace mechanisms and their properties in our examples in Section 4.

Remark 5 *We use ϵ-DP for the presentation of this paper primarily for simplicity. Our proposed PF scheme can be easily modified to any other privacy guarantee, especially those with a single privacy parameter, such as μ-GDP (Dong et al., 2022) or ρ-zCDP (Bun and Steinke, 2016). Mechanisms for other frameworks, such as (ϵ, δ)-DP, f-DP (Dong et al., 2022), or Rényi DP (Mironov, 2017), may need some special treatment to identify a useful filtering schedule in our sampler.*

2.2 Particle Filters

The structure of a particle filtering method can be broken into three main iterative operations: resampling, propagation, and reweighting. Suppose there are N particles in the particle filtering system and T iterations before completion. Iteration $t = 0$ is when we initialize the particle system by generating N i.i.d. samples from the prior π_0. Then, in iteration $t \geqslant 1$, we perform the following steps subsequently to approximate the intermediate target distribution π_t, with π_T being the ultimate target:

Step 1. Resampling. The resampling step acts as pruning, helping the system avoid particle degeneration. Although this action introduces variation from the resampling process, it ultimately offsets the risk of particle degeneracy. The simplest resampling method is to perform multinomial resampling, which is simply weighted sampling with replacement:

$$\tilde{y}^{(1,t)}, \ldots, \tilde{y}^{(N,t)} \overset{i.i.d.}{\sim} \text{Multinomial}\left(1, \left(\mathbf{w}^{(1,t-1)}, \ldots, \mathbf{w}^{(N,t-1)}\right)\right),$$

where $\mathbf{w}^{(j,t-1)}$'s are normalized weights that sum up to 1 and $\tilde{y}^{(j,t)}$'s are resampled particles but not yet perturbed. As $t = 1$, the new particles are resampled from the N i.i.d. samples generated from the prior π_0; otherwise, they are resampled from the independent weighted samples $\{(y, \mathbf{w})^{(j,t-1)}\}_{j=1}^{N}$ from $\pi_{t-1}(\cdot)$, the intermediate distribution obtained at the end of iteration $t-1$. Other resampling schemes, such as stratified resampling, residual resampling (Liu and Chen, 1998), and systematic resampling (Carpenter et al., 2000), can introduce less noise and have lower variance in estimates than multinomial resampling. While we focus on multinomial sampling for simplicity, there should be no complication in extending our methodology to these other resampling schemes, each of which possesses a central limit theorem result (Chopin, 2004). Additional discussion is provided in Section 6.

Step 2. Propagation. The propagation step drives exploration of the parameter space by producing new particles from a perturbation kernel $K_t(\tilde{y}^{(j,t)}, \cdot)$ for all j. A typical particle filtering propagation step ends with new particles approximating

$$\eta_t(\cdot) = \int \pi_{t-1}(\tilde{y}) K_t(\tilde{y}, \cdot) d\tilde{y}.$$

A multivariate Gaussian kernel is often used when targeting a continuous distribution, for it is easily evaluated and sampled. The covariance of a Gaussian kernel can be specified in advance or estimated by the sample covariance from the particles of the previous iteration. Beaumont et al. (2009) and Filippi et al. (2013) derive an optimal kernel variance (scale) with high average acceptance rates for parametric perturbation kernels, such as uniform kernels and Gaussian kernels.

Step 3. Reweighting. The reweighting step recalibrates the importance of each particle. It uses the proposal distribution $\eta_t(\cdot)$ to target the iteration-wise distribution of interest

Algorithm 1 Generic Particle Filter

Input: number of particles N, number of iterations T, prior π_0, perturbation kernel K_t $(t = 1, \ldots, T)$

Initialize by generating, from π_0, N particles $\{y^{(i,0)}\}_{i=1}^N$ with weights $w_0 = \frac{1}{N}$.

for t = 1,...,T **do**

Step 1. **Resample** from y_{t-1} with weights w_{t-1} to get candidate particles $\{\tilde{y}^{(i,t)}\}_{i=1}^N$.

Step 2. **Propagate** by a perturbation kernel $y^{(i,t)} \sim K_t(\tilde{y}^{(i,t)}, \cdot)$.

Step 3. **Reweight** $y^{(i,t)}$ with importance weights as in (1).

end for

Output: $\{(y, \mathbf{w})^{(i,T)}\}_{i=1}^N$: weighted samples from $\pi_T(\cdot)$.

$\pi_t(\cdot)$ by considering the importance weight

$$w_t(y) = \frac{\pi_t(y)}{\eta_t(y)}, \tag{1}$$

which comes from the concept of importance sampling. Importance sampling is a technique used to estimate properties of a particular distribution π while only having samples generated from another distribution η. In such cases, $\mathbb{E}_\pi[\varphi(Y)] = \int \varphi(y) \frac{\pi(y)}{\eta(y)} \eta(y)\, dy$ can be estimated by

$$\frac{1}{N} \sum_{i=1}^{N} \varphi(y_i) \frac{\pi(y_i)}{\eta(y_i)},$$

where $Y \sim \pi$ but can only sample from a tractable proposal distribution $\eta(y)$. The ratio of the target to the proposal π/η is called the importance weight.

Remark 6 *The weight calculated in the reweighting step (Line 11) of Algorithm 3 is only proportional to the "ideal" importance weight, that is, the fraction of the target posterior to the proposal. However, this poses no issue as normalization is performed in the next step (Line 12).*

Algorithm 1 describes how a particle filtering system works and evolves. Figure 1 shows that as iterations proceed, the particles in the system gravitate towards a better proposal η_t than the prior π_0. The colored (overlapping) area expands in t and hence more particles are considered effective. On the other hand, if π_0 is used to target π, it reduces to a typical importance sampling (that is, the leftmost graph). In this case, the effective area is small.

Example 1 (ABC SMC: Toni et al., 2009) *For Bayesian analysis, PF is commonly used in combination with ABC, which is called ABC SMC. ABC SMC, described in Algorithm 2, incorporates an additional rejection step to rule out bad particles when its threshold $d(s, \cdot) > \Delta$. The function $d(s, \cdot)$ measures the distance between a statistic and its simulated counterpart, accepting particles within a range of Δ. Smaller thresholds require more computation to generate effective particles but can result in a smaller approximation error.*

By considering the events of $\{d(s, \cdot) \leqslant \Delta_t\}$ over iterations, ABC SMC can adaptively approach its target by shrinking Δ_t in t. This is essentially the "schedule" in ABC SMC.

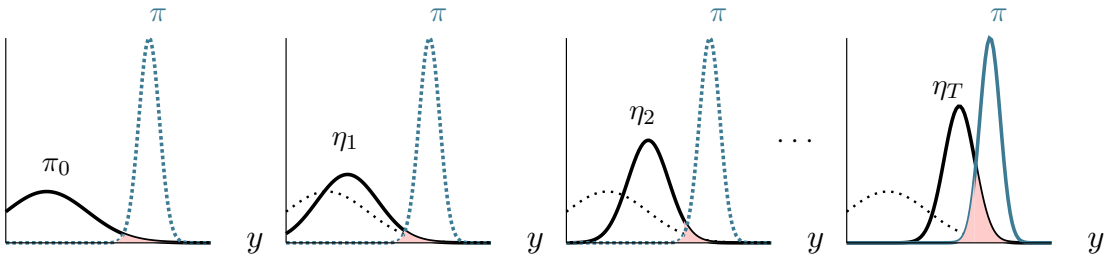

Figure 1: Diagram of a particle filter. Plots progress from left to right.

Algorithm 2 ABC SMC

Input: number of particles N, number of iterations T, prior π_0, perturbation kernel K_t, schedule Δ_t $(t = 1, \ldots, T)$

Initialize by generating, from π_0, N particles $\{y^{(i,0)}\}_{i=1}^N$ with weights $w_0 = \frac{1}{N}$

for t = 1,...,T **do**

Step 1. **Resample** from y_{t-1} with weights w_{t-1} to get candidate particles $\{\tilde{y}^{(i,t)}\}_{i=1}^N$.

Step 2. **Propagate** by a perturbation kernel $y^{(i,t)} \sim K_t(\tilde{y}^{(i,t)}, \cdot)$.

Step 2.5. **Reject** when $d > \Delta_t$ and go back to Step 2.

Step 3. **Reweight** $y^{(i,t)}$ with importance weights as in (1).

end for

Output: $\{(y, \mathbf{w})^{(i,T)}\}_{i=1}^N$: weighted samples from $\pi_T(\cdot)$.

More formally, the schedule targets $\pi(\theta \mid d(s, \cdot) \leqslant \Delta_t)$, where the tolerance Δ_t decreases with t until reaching the target ABC posterior distribution at the final step T. This process avoids rejecting too many particles in the first few iterations when the particle system is still close to prior knowledge.

3. Particle Filter for Differential Privacy

In this section, we present our particle filtering algorithm for DP, which generates weighted samples that approximately follow the private posterior distribution $\pi_\epsilon(\theta, x \mid s_{dp})$.

Our algorithm integrates the strengths of ABC SMC and DP-Reject-ABC. While ABC SMC effectively narrows down potential parameters, it introduces approximation errors as its final target distribution is an ABC posterior rather than the true posterior distribution. To avoid this issue, we incorporate techniques from DP-Reject-ABC, which produces i.i.d. samples from the exact posterior. To properly combine the two methods, we define the ϵ_t schedule as an increasing, positive sequence for the ϵ budget, with $\epsilon_T = \epsilon$. Additionally, the scale of the perturbation kernels can be set as a decreasing function of t, narrowing the exploration scope as the iterations proceed. The ϵ_t-schedule is analogous to the decreasing sequence Δ_t in ABC SMC, which fits the empirical evidence that smaller ϵ values result in a posterior that is closer to the prior (Ju et al., 2022).

We point out the key differences of our differentially private particle filtering algorithm compared to the previous two PF algorithms. Compared to Algorithm 2, *Steps 1. and 2.* are the same except that our algorithm requires specifying the ϵ_t-schedule.

Step 2.5. Rejection. The DP propagation step incorporates a rejection step (in Lines 7 to 9 in Algorithm 3) that helps target the exact intermediate private posterior $\pi_{\epsilon_t}(\cdot \mid s_{dp})$. Instead of function d in ABC, we have the probability of acceptance

$$r_t(x) = \frac{m_{\epsilon_t}(s_{dp} \mid x)}{\sup_x m_{\epsilon_t}(s_{dp} \mid x)},$$

where x is sampled from $f(\cdot \mid \theta)$. This is a direct and crucial consequence of the design of Algorithm 3, in which the mechanism density appears only in the propagation step but not in the reweighting or resampling steps. The tempering effect on the mechanism density m_{ϵ_t} enables the particle system to collect plausible particles in early iterations to avoid a high rejection rate and filter these through later iterations to meet the actual ϵ-DP guarantee.

Step 3. Reweighting. As usual, the reweighting step considers the importance weight. The numerator now becomes our target private posterior $\pi_{\epsilon_t}(\theta, x \mid s_{dp})$ and the denominator is the whole propagation density $h_t(\theta, x)$, which involves both perturbation and rejection. The properties of h_t are discussed in Appendix C.

3.1 Algorithm

In this section, we present the proposed algorithm, detailing its inputs, outputs, and assumptions, and discuss how its design underpins the theoretical guarantees established later.

For the input, we assume the following:

A.1 prior π_0 can be sampled and its density can be evaluated,

A.2 data model f can be sampled,

A.3 perturbation kernel K_t can be sampled, and its density can be evaluated for each t,

A.4 mechanism density $m_{\epsilon_t}(s_{dp}|\cdot)$ can be evaluated and has a maximum for each t,

A.5 $\mathbb{E}_{\pi_{\epsilon_T}(\theta,x|s_{dp})}[\varphi(\theta,x)]$, $c_{\eta_t}(s_{dp})$ and $c_{\epsilon_t}(s_{dp})$ are all finite.

For the output, $\{(\theta, x, \mathbf{w})^{(j,T)}\}_{j=1}^N$ are weighted samples that are approximately distributed as $\pi_\epsilon(\theta, x \mid s_{dp})$. More importantly, by taking the weighted average of the particles, we obtain the consistent estimator of $\mathbb{E}_{\pi_{\epsilon_T}(\theta,x|s_{dp})}[\varphi(\theta,x)]$ for a given function φ to be

$$\hat{\mathbb{E}}_T[\varphi] := \sum_{i=1}^N \mathbf{w}^{(i,T)} \varphi\left((\theta, x)^{(i,T)}\right). \tag{2}$$

The properties of (2) are studied in Section 3.2, where we also state additional technical conditions on the prior π_0 and on the perturbation kernel K_t to ensure that the weight function w_t is well defined, with $0 \leqslant w_t(\theta) < \infty$ for all $\theta \in \Theta_t$ at each t.

Algorithm 3 presents our differentially private particle filtering (DP-PF) approach for posterior inference. The for loop on Line 3 and the unnormalized weighting calculations on Line 11 can be processed in parallel across particles, substantially reducing computation time.

Algorithm 3 DP-PF

Input: privacy budget ϵ, sample size n, particles N, iterations T, prior π_0, privatized summary s_{dp}, sampler f, kernel K_t, schedule ϵ_t, mechanism densities m_{ϵ_t} $(t = 1, \ldots, T)$

1: **for** $t = 0, 1, \ldots, T$ **do**
2: Set ϵ_t and variance (scale) of K_t when $t > 0$
3: **for** $i = 1, .., N$ **do**
4: [**Initialization**] Sample $\theta^{(i,t)} \sim \pi_0$ if $t = 0$; else:
5: [**Resampling**] Sample $\tilde{\theta} \sim \{\theta^{(j,t-1)}\}_{j=1}^N$ using weights $\{\mathbf{w}^{(j,t-1)}\}_{j=1}^N$.
6: [**Propagation**] Sample $\theta \sim K_t(\tilde{\theta}, \cdot)$ until $\pi_0(\theta) > 0$.
7: Sample $x \sim f(\cdot \mid \theta)$, compute

$$r_t(x) = \frac{m_{\epsilon_t}(s_{dp} \mid x)}{\sup_x m_{\epsilon_t}(s_{dp} \mid x)}.$$

8: Sample $u \sim U(0, 1)$.
9: If $u > r_t(x)$, go to the propagation step; else: set $\theta^{(i,t)} \leftarrow \theta$ and $x^{(i,t)} \leftarrow x$.
10: **end for**
11: [**Reweighting**] Compute weights for all $i = 1, \ldots, N$:

$$\tilde{w}^{(i,t)} = \begin{cases} 1, & t = 0, \\ \tilde{w}^{(i,t)} \leftarrow \frac{\pi_0\left(\theta^{(i,t)}\right)}{\sum_{j=1}^N \mathbf{w}^{(j,t-1)} K_t(\theta^{(j,t-1)}, \theta^{(i,t)} \mid \pi_0(\theta^{(i,t)}) > 0, r_t(x^{(i,t)} > U)}, & \text{otherwise.} \end{cases}$$

12: Normalize the weights to get $\{\mathbf{w}^{(j,t)}\}$ s.t. $\sum_{j=1}^N \mathbf{w}^{(j,t)} = 1$.
13: **end for**

Output: $\{(\theta, x, \mathbf{w})^{(j,T)}\}_{j=1}^N$: weighted samples from $\pi_\epsilon(\theta, x \mid s_{dp})$.

Remark 7 *To decide what T to use, we note the following intuition: when rejection sampling maintains good acceptance and efficiency (for example, when ϵ and sample size n are small), we let T be small (close to 1); conversely, we let T be large ($\geqslant 10$).*

While many variations are possible, there is an inherent trade-off between computational cost and particle quality, especially in high dimensions. For instance, a DP importance sampling variant, relying only on perturbation and reweighting, can start with hundreds of thousands of particles efficiently, but without rejection, the weights concentrate quickly, leading to severe degeneracy and very small ESS. Conversely, incorporating ideas from Ju et al. (2022), which sequentially updates and stores synthetic data, may yield more stable but potentially poorly mixed particles at substantially greater computational cost.

DP-PF is designed not only to strike a balance between these two extremes, but also to exploit an advantage that alternative designs do not possess. The key structural feature is that the data and mechanism densities cancel in the weight update (Algorithm 3, Line 11), yielding a particularly simple form. As a result, in the reweighting step, the data model f need not be evaluated, and neither the model density nor the mechanism density appears.

This structure has two important consequences. First, the acceptance probability depends only on the proposal η_t (or kernel K_t); see Proposition 19. Second, the marginal

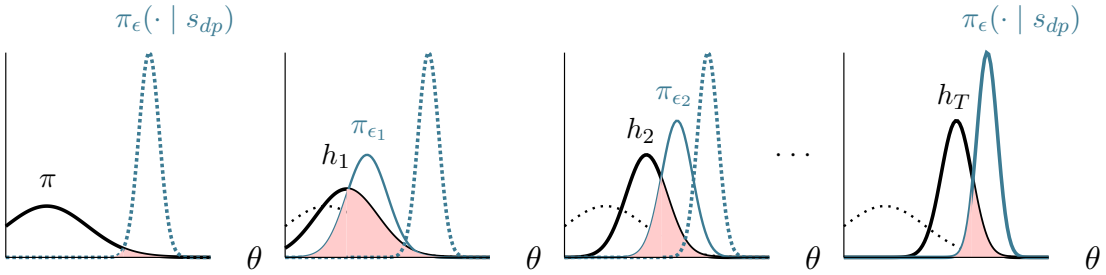

Figure 2: Diagram of DP-PF. Plots progress from left to right.

likelihood admits a clean characterization that the incremental evidence ratios depend only on the mechanism density; see Proposition 16. Moreover, since the mechanism density enters only through the rejection step and not the reweighting step, the algorithm integrates naturally into the classical SMC framework: the propagation step absorbs the rejection, while the weighting step remains unaffected. This cancellation is particularly valuable in high dimensions, as it avoids repeated evaluation of expensive data and DP mechanisms.

Unlike Figure 1, the colored area in Figure 2 increases as a consequence of two features of our proposed design: (i) the particle filtering procedure progressively shifts the proposals toward the private posterior, and (ii) the intermediate target distributions are flatter than the private posterior, thereby making the sampling more efficient at each step.

3.2 Consistency

In this section, we prove that our estimator (2) is consistent as the number of particles N tends to infinity.

Since our goal is to sample from and perform inference based on $\pi_{\epsilon_T}(\theta, x \mid s_{dp})$ (or simply its marginal $\pi_{\epsilon_T}(\theta \mid s_{dp})$), we assess the convergence of the weighted estimator (2) to its posterior expectation, componentwise:

$$\mathbb{E}_T[\varphi] := \mathbb{E}_{\pi_{\epsilon_T}(\theta,x|s_{dp})}[\varphi(\theta, x)]$$

for any measurable φ mapping from the product of the parameter space and the database space to d-dimensional real space such that $\mathbb{E}_T[\varphi]$ is finite. Throughout Section 3, moment conditions involving vector-valued φ are understood componentwise.

Theorem 8 (Strong Consistency) *Let $\varphi : \Theta_t \times \mathcal{X}^n \to \mathbb{R}^d$, such that $\mathbb{E}_{h_t}|w_t\varphi_j| < \infty$ for all $j = 1, \ldots, d$. Our estimator, defined in (2), converges almost surely:*

$$\hat{\mathbb{E}}_t[\varphi] = \sum_{i=1}^{N} \mathbf{w}^{(i,t)} \varphi\left((\theta, x)^{(i,t)}\right) \overset{a.s.}{\to} \mathbb{E}_t[\varphi]$$

as $N \to \infty$. In particular, for $t = T$, $\hat{\mathbb{E}}_T[\varphi] \overset{a.s.}{\to} \mathbb{E}_T[\varphi]$.

Theorem 8 plays a role analogous to ergodicity in DP-DA-MCMC. The almost sure convergence is with respect to the randomness of the PF while the DP summary is held

fixed. Although the DP-PF importance weight $\tilde{w}_t$ is only proportional to the ideal weight

$$w_t = \frac{\pi_{\epsilon_t}(\theta, x \mid s_{dp})}{h_t(\theta, x)},$$

the unknown proportionality constant is not an issue: it can be consistently estimated by setting $\varphi = 1$. Consequently, both the numerator and denominator of the self-normalized importance sampling estimator are consistent, and $\hat{\mathbb{E}}_t[\varphi]$ converges almost surely to $\mathbb{E}_t[\varphi]$ by the continuous mapping theorem. Theorem 8 therefore provides consistent estimates of posterior moments, such as $\varphi(\theta) = \theta^q$ for $q \in \mathbb{N}$, as well as posterior probabilities, obtained by taking $\varphi(\theta) = 1_{\{\theta \in S\}}$ for a measurable set $S \in \Theta_t$. While our primary interest is in functions of θ, functions of x may also be considered.

Remark 9 *$\mathbb{E}_{h_t}|w_t\varphi| < \infty$ is indeed a crucial but mild assumption in the following way.*

- *It ensures that $\mathbb{E}_t[\varphi]$ exists.*
- *It excludes the cases in which $w_t = \infty$ for any θ.*
- *It is mild because it is often reasonable to assume that the parameter space Θ_t and the weight w_t will be bounded in practice. In Euclidean space, if Θ_t is a compact set, for any continuous function φ defined on Θ_t,*

$$\mathbb{E}_{h_t}|w_t\varphi| \leqslant \sup_{\theta \in \Theta_t} |\varphi(\theta)| < \infty.$$

3.3 Central Limit Theorem

In addition to the consistency of $\hat{\mathbb{E}}_T[\varphi]$, it is also crucial to understand its stability. In this section, we provide the asymptotic distribution of the consistent estimator $\hat{\mathbb{E}}_T[\varphi]$, its asymptotic variance, and a $1 - \alpha$ confidence interval. We follow a standard central limit theorem (CLT) argument for particle filtering algorithms by Chopin (2004), but need to adjust for the parts that involve our DP components and rejection step; specifically, the propagation step. As in the previous section, this quantifies the randomness introduced by the PF system, not the randomness from the privacy mechanism or the data.

To make the CLT hold for the propagation step, we need to assume that

$$\mathbb{E}_{h_t}|w_t\varphi|^{2+\zeta} < \infty$$

for some $\zeta > 0$. It is worth noting that we only need to assume $\mathbb{E}_{h_t}|w_t\varphi| < \infty$ for consistency, but an additional finite absolute $2 + \zeta$ moment is required for the CLT. However, this is still fairly mild in a similar sense as discussed in Remark 9. In particular, the boundedness of Θ_t and the weights w_t implies that this condition holds.

Remark 10 *The assumption $\mathbb{E}_{h_t}|w_t\varphi|^{2+\zeta} < \infty$ with some $\zeta > 0$ ensures Lyapunov's condition for the CLT. It also implies $\mathbb{E}_{h_t}|w_t\varphi| < \infty$ by Jensen's inequality, and hence guarantees strong consistency.*

Theorem 11 (Central Limit Theorem) *Let $\varphi : \Theta_t \times \mathcal{X}^n \to \mathbb{R}^d$ satisfy the recursive measurability and moment conditions stated in Appendix C, and suppose there exists $\zeta > 0$ such that $\mathbb{E}_{h_t}|w_t\varphi_j|^{2+\zeta} < \infty$ for all $j = 1, \ldots, d$. Then,*

$$\sqrt{N}\left(\hat{\mathbb{E}}_t[\varphi] - \mathbb{E}_t[\varphi]\right) \xrightarrow{d} N(0, V_t(\varphi))$$

as $N \to \infty$.

Theorem 11 plays a role analogous to geometric ergodicity in DP-DA-MCMC. However, a relatively strong assumption has to be made for geometric ergodicity to hold for DP-DA-MCMC: there exists $0 < l \leqslant u < \infty$ such that $l \leqslant f(x \mid \theta) \leqslant u$ for all θ, x (Ju et al., 2022). In contrast, Theorem 11 has a relatively mild assumption. In addition, DP-DA-MCMC also requires ϵ-DP, while DP-PF allows for other privacy frameworks, such as μ-GDP and ρ-zCDP.

3.4 Monte Carlo Standard Errors and Asymptotic Confidence Intervals

The asymptotic normality established in the previous section characterizes the Monte Carlo error at the $\sqrt{N}$ rate. While this provides a basis for quantifying Monte Carlo standard errors (MCSE), we need to develop an estimator of the MCSE in order to construct confidence intervals (CI) for the true posterior quantities. To this end, we use the concept of effective sample size to account for the weighting of the particles.

Definition 12 (Effective Sample Size: Kong, 1992) *Suppose the target distribution is $\pi(x)$, the proposal distribution is $h(x)$, and the importance weight is $w(x) = \frac{\pi(x)}{h(x)}$. The quantity*

$$ESS_N := \frac{N}{1 + Var_h(w(X))}$$

is defined as the population effective sample size with respect to the distributions π and h.

Effective Sample Size (ESS) is a rough measure to evaluate the quality of importance weights in importance sampling. It indicates that N weighted samples from h create the same variation as ESS_N true samples from π do. A similar concept, relative numerical efficiency (RNE), is introduced by Geweke (1989), where $RNE \approx ESS/N$ but can occasionally exceed 1. Thus, for ordinary importance sampling, the numerical standard error of the estimate is the fraction of $(ESS)^{-1/2}$ or $(RNE \cdot N)^{-1/2}$ of the posterior standard deviation.

Lemma 13 (ESS estimate) *Let $w_i = \frac{\pi(X_i)}{h(X_i)}$ and $\widehat{ESS}_N = \frac{(\sum_{i=1}^N w_i)^2}{\sum_{i=1}^N w_i^2}$. Assume $(X_i)_{i \geqslant 1}$ has marginal distribution h and satisfies $\frac{1}{N}\sum_{i=1}^N w_i \xrightarrow{p} \mathbb{E}_h[w(X)]$ and $\frac{1}{N}\sum_{i=1}^N w_i^2 \xrightarrow{p} \mathbb{E}_h[w(X)^2]$, with $\mathbb{E}_h[w(X)^2] < \infty$. Then, as $N \to \infty$,*

$$\widehat{ESS}_N / N \xrightarrow{p} \frac{1}{1 + Var_h(w(X))}.$$

The moment assumptions require only a weak law of large numbers for $w(X)$ and $w(X)^2$, rather than independence of the samples. They hold for importance sampling and, more generally, for ergodic Markov chains in MCMC, where empirical averages converge to expectations under the stationary distribution. In particle filtering methods, although resampling induces dependence, standard consistency results ensure convergence of empirical averages under mild regularity conditions. Thus, independence is sufficient but not necessary for the consistency of $\widehat{ESS}_N$. Moreover, the definition of $\widehat{ESS}_N$ is invariant to normalization. Combined with Lemma 13, this provides a basis for quantifying the MCSE and constructing an asymptotic confidence interval.

Theorem 14 establishes a consistent estimator of the asymptotic variance $V_t(\varphi)$ and the resulting asymptotic confidence interval for $\mathbb{E}_t[\varphi]$. Unlike the standard approximation that omits the remainder term, our variance estimator retains this term to account for its potential contribution.

Theorem 14 (Asymptotic Confidence Interval) *Let $\varphi : \Theta_t \times \mathcal{X}^n \to \mathbb{R}$ satisfy the recursive measurability and moment conditions stated in Appendix C, and suppose there exists $\zeta > 0$ such that $\mathbb{E}_{h_t}|w_t\varphi|^{2+\zeta} < \infty$. Then,*

1. *a consistent estimate of $V_t(\varphi)$ is*

$$\hat{V}_t(\varphi) := \frac{\sum_{i=1}^N \mathbf{w}^{(i,t)} \left[\varphi\left((\theta,x)^{(i,t)}\right) - \hat{\mathbb{E}}_t[\varphi]\right]^2}{\widehat{ESS}_N/N} + \sum_{i=1}^N \mathbf{w}^{(i,t)} \left[\mathbf{w}^{(i,t)} - \frac{1}{N}\right] \left[\varphi\left((\theta,x)^{(i,t)}\right) - \hat{\mathbb{E}}_t[\varphi]\right]^2;$$

2. *an asymptotic $(1-\alpha)$ confidence interval for $\mathbb{E}_t[\varphi]$ is*

$$\hat{\mathbb{E}}_t[\varphi] \pm z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{V}_t(\varphi)}{N}},$$

where $z_{1-\frac{\alpha}{2}}$ is the $\left(1-\frac{\alpha}{2}\right)$ quantile of a standard normal distribution.

Remark 15 *In much of the literature, the second term of $V_t(\varphi)$, also known as the remainder term, is omitted under the assumption that it is negligible. However, Kong (1992) emphasize that the approximation may become inaccurate if the remainder term is not sufficiently small. Since we can estimate the remainder term by the second term in $\hat{V}_t(\varphi)$, we retain it in our estimator in Theorem 14 to address this potential issue.*

3.5 Marginal Likelihood of Privatized Data

In classical SMC methods for Bayesian inference, particles represent parameter values and are reweighted according to the likelihood of the observed data. This sequential reweighting implicitly accumulates the normalizing constants of the intermediate distributions, and their product yields the marginal likelihood of the observed data (Del Moral et al., 2006). Since our inference is based on the privatized statistic s_{dp}, the corresponding normalizing

constants define its marginal likelihood, providing a differentially private counterpart to the classical model evidence with a similar telescoping structure.

A key structural property of Algorithm 3 makes this construction clean. As noted in Section 3.1, the data and mechanism densities cancel in the weight update (Line 11), so m_{ϵ_t} enters the algorithm only through the propagation step (Line 7), not through reweighting. Consequently, ratios of mechanism densities can be evaluated as importance-sampling expectations under the intermediate private posteriors using the particle approximation already maintained by Algorithm 3, with no additional simulation.

For notational clarity, let

$$\kappa_t(\theta, x) \;=\; \pi_0(\theta)\, f(x \mid \theta)\, m_{\epsilon_t}(s_{dp} \mid x), \qquad c_{\epsilon_t}(s_{dp}) \;=\; \int_\Theta \int_{\mathcal{X}^n} \kappa_t(\theta, x)\, dx\, d\theta,$$

so that $\pi_{\epsilon_t}(\theta, x \mid s_{dp}) = \kappa_t(\theta, x)/c_{\epsilon_t}(s_{dp})$. We adopt the convention $m_{\epsilon_0}(s_{dp} \mid x) \equiv 1$ and $c_{\epsilon_0}(s_{dp}) = 1$ as a result, corresponding to the degenerate "no-release" baseline ($\epsilon_0 = 0$). Under this convention, $\pi_{\epsilon_0}(\theta, x \mid s_{dp}) = \pi_0(\theta) f(x \mid \theta)$ coincides with the prior predictive distribution, which is exactly the distribution sampled at the initialization step (Line 4) of Algorithm 3. We further define the incremental evidence ratio

$$\rho_t \;:=\; \frac{c_{\epsilon_t}(s_{dp})}{c_{\epsilon_{t-1}}(s_{dp})}, \qquad t = 1, \ldots, T,$$

so that the marginal likelihood factorizes as $c_{\epsilon_T}(s_{dp}) = \prod_{t=1}^{T} \rho_t$.

Theorem 16 (Marginal likelihood and its consistent estimation) *Let $\{\epsilon_t\}_{t=0}^{T}$ be the schedule of Algorithm 3 with $\epsilon_0 = 0$. For $t = 1, \ldots, T$, $\mathbb{E}_{h_t}\big[\, w_t m_{\epsilon_t}(s_{dp} \mid x)/m_{\epsilon_{t-1}}(s_{dp} \mid x) \,\big] < \infty$. Then,*

1. *Each incremental evidence ratio admits the representation*

$$\rho_t \;=\; \mathbb{E}_{\pi_{\epsilon_{t-1}}(\theta, x \mid s_{dp})}\left[\frac{m_{\epsilon_t}(s_{dp} \mid x)}{m_{\epsilon_{t-1}}(s_{dp} \mid x)} \right],$$

   so that $c_{\epsilon_T}(s_{dp}) = \prod_{t=1}^{T} \rho_t$.

2. *Each ρ_t is consistently estimated, as $N \to \infty$, by*

$$\hat{\rho}_t \;=\; \sum_{i=1}^{N} \mathbf{w}^{(i,t-1)}\, \frac{m_{\epsilon_t}(s_{dp} \mid x^{(i,t-1)})}{m_{\epsilon_{t-1}}(s_{dp} \mid x^{(i,t-1)})},$$

   where $\mathbf{w}^{(i,t)}$ and $x^{(i,t)}$ are the normalized weights and accepted synthetic data set produced at iteration t of Algorithm 3. Consequently, $\widehat{c_{\epsilon_T}(s_{dp})} = \prod_{t=1}^{T} \hat{\rho}_t$ is a consistent estimator of $c_{\epsilon_T}(s_{dp})$.

Theorem 16 is the differentially private analogue of the marginal-likelihood identity for tempered SMC samplers (Chopin, 2004; Del Moral et al., 2006). The consistency of $\hat{\rho}_t$ is an immediate application of Theorem 8 with $\varphi(\theta, x) = m_{\epsilon_t}(s_{dp} \mid x)/m_{\epsilon_{t-1}}(s_{dp} \mid x)$. The consistency of $\widehat{c_{\epsilon_T}(s_{dp})}$ then follows by the continuous mapping theorem. Note that the

integrability condition is mild. By Remark 9, when Θ_t is compact and w_t is bounded, it suffices that $m_{\epsilon_t}/m_{\epsilon_{t-1}}$ be bounded as a function of x. For the Laplace mechanism with $t > 1$,

$$\frac{m_{\epsilon_t}(s_{dp} \mid x)}{m_{\epsilon_{t-1}}(s_{dp} \mid x)} \leqslant \frac{\epsilon_t}{\epsilon_{t-1}} \exp\left(\frac{\epsilon_{t-1} - \epsilon_t}{\Delta f}|s_{dp} - f(x)|\right) \leqslant \frac{\epsilon_t}{\epsilon_{t-1}}$$

is uniformly bounded in x whenever $\epsilon_t \geqslant \epsilon_{t-1}$. Hence the condition is automatically satisfied under our DP-PF tempering schedule.

Algorithm 3 already produces $\mathbf{w}^{(i,t)}$ (Line 12) and $x^{(i,t)}$ (Line 7) at each iteration, so the only extra computation is the log-density difference $\log m_{\epsilon_t}(s_{dp} \mid x^{(i,t)}) - \log m_{\epsilon_{t-1}}(s_{dp} \mid x^{(i,t)})$. Given $\widehat{c_{\epsilon_T}(s_{dp})}$ for two competing models, the DP Bayes factor is

$$\mathrm{BF}_{12}^{\mathrm{DP}} := \frac{p_1(s_{dp})}{p_2(s_{dp})} \approx \frac{\widehat{c_{\epsilon_T}^{(1)}(s_{dp})}}{\widehat{c_{\epsilon_T}^{(2)}(s_{dp})}},$$

which is well-defined provided both models use the same privatized statistic and privacy mechanism; the prior on θ and the data model f may differ. Section 4 demonstrates this on a Gaussian mixture model selection problem.

4. Simulations

In this section, we investigate applications of our approach proposed in Section 3.1 to several common settings in DP Bayesian analysis, demonstrating the efficacy of the algorithm.

General setup. All resampling steps are conducted with multinomial resampling, and all perturbation kernels are Gaussian. Simulations were conducted on a computing cluster with multi-core compute nodes, 256 GB of memory, and a high-speed interconnect. We use 100 cores (1 node) per replicate per simulation with a maximum wall-time of 4 hours.

4.1 Consistency and Posterior Inference

We demonstrate the consistency of Algorithm 3 under a location-scale normal distribution, linear regression, and a Gaussian mixture model, using the Laplace mechanism and 1,000 replicates.

4.1.1 Location-scale Normal

Suppose the data model f is $y_1, \ldots, y_n \mid \mu, \sigma^2 \overset{iid}{\sim} N\left(\mu, \sigma^2\right)$ with prior distributions $\mu \sim N\left(m, \tau^2\right)$ and $\sigma^2 \sim \text{Inv. Gamma}\left(\alpha, \beta\right)$, where α is the shape parameter and β is the scale parameter. The private data $\tilde{y}_i$ is obtained by first clamping y_i to the interval $[-5, 5]$ and subsequently rescaling it to the interval $[-1, 1]$. Let $L_1, L_2 \overset{iid}{\sim} \text{Lap}(\frac{\Delta}{\epsilon})$, where $\Delta = 3$. Then, the privatized statistics s_{dp} are $(\sum \tilde{y}_i + L_1, \sum \tilde{y}_i^2 + L_2)$, which satisfies ϵ-DP.

Simulation setup. True parameters are set as $\mu^* = 1, \sigma^* = 1$, $\epsilon \in \{0.1, 0.5, 1, 2\}$ and $n \in \{100, 300, 500, 1000\}$. We generate a single value of s_{dp} for each pair of (ϵ, n). The number of particles N is set to achieve an average ESS of approximately 100 or higher. The hyperparameters $m = 0, \tau^2 = 4^2, \alpha = 1, \beta = 0.5$. We use an adaptive kernel with variance estimation and $T = 2$ ($\epsilon_t = (0.5\epsilon, \epsilon)$).

		DP-PF			DP-Reject-ABC		
ϵ	Sample Size	$\mathbb{E}(\mu \mid s_{dp})$	$\mathbb{E}(\sigma^2 \mid s_{dp})$	Sec./ESS†	$\mathbb{E}(\mu \mid s_{dp})$	$\mathbb{E}(\sigma^2 \mid s_{dp})$	Sec./ESS†
0.1	100	0.052 (0.0198)	4.268 (0.3584)	0.05	0.100 (0.0038)	2.284 (0.2292)	0.00
	300	1.014 (0.0146)	3.902 (0.1316)	0.05	1.722 (0.0020)	2.215 (0.0112)	0.00
	500	0.721 (0.0085)	2.275 (0.1667)	0.04	1.072 (0.0011)	1.062 (0.0033)	0.00
	1000	1.145 (0.0053)	1.585 (0.0404)	0.04	1.271 (0.0007)	1.788 (0.0037)	0.04
0.5	100	0.552 (0.0067)	1.580 (0.0558)	0.04	0.795 (0.0012)	0.669 (0.0022)	0.00
	300	1.123 (0.0036)	0.988 (0.0105)	0.03	1.171 (0.0005)	1.233 (0.0022)	0.03
	500	1.045 (0.0018)	0.691 (0.0041)	0.03	1.043 (0.0003)	0.942 (0.0013)	0.06
	1000	1.081 (0.0008)	0.785 (0.0027)	0.04	1.034 (0.0002)	1.329 (0.0008)	0.46
1.0	100	0.794 (0.0042)	0.925 (0.0268)	0.03	0.986 (0.0006)	0.579 (0.0014)	0.01
	300	1.126 (0.0015)	0.729 (0.0038)	0.03	1.096 (0.0003)	1.130 (0.0013)	0.09
	500	1.064 (0.0008)	0.626 (0.0022)	0.03	1.030 (0.0002)	1.010 (0.0008)	0.22
	1000	1.037 (0.0004)	0.865 (0.0022)	0.05	1.010 (0.0001)	1.218 (0.0004)	1.58
2.0	100	1.004 (0.0026)	0.632 (0.0126)	0.03	1.066 (0.0003)	0.564 (0.0010)	0.02
	300	1.099 (0.0006)	0.660 (0.0022)	0.03	1.062 (0.0002)	1.069 (0.0007)	0.29
	500	1.051 (0.0004)	0.713 (0.0019)	0.04	1.025 (0.0002)	1.038 (0.0014)	0.90
	1000	1.006 (0.0002)	1.030 (0.0014)	0.07	0.999 (0.0001)	1.150 (0.0002)	5.74

Table 1: Comparison of DP-PF and DP-Reject-ABC based on the averaged posterior means of μ and σ^2 over 1000 simulations for the location-scale normal model, and the seconds spent per ESS, with the values in parentheses representing MCSEs for each quantity. (†Parallelized over particles; 0.00 denotes runtimes < 0.005 seconds.)

Table 1 compares the performance of DP-PF with DP-Reject-ABC, an exact sampling method. Across various privacy levels and sample sizes, DP-PF achieves comparable estimates for both μ and σ. The computation time per effective sample size is also similar. This demonstrates that DP-PF provides accuracy on par with the exact method while maintaining computational efficiency. While DP-Reject-ABC is effective in this setting, in the next section, we see that it does not scale well when applied to more complex settings.

Figure 3 illustrates the 90% highest posterior density (HPD) intervals for the posterior distributions of μ and σ^2. We again compare DP-PF with DP-Reject-ABC, which serves as an exact posterior sampler. The privacy budget ϵ ranges from 0.1 to 2, and DP-PF uses $1,000$ particles. Figure 3 shows that the HPD lengths produced by DP-PF exhibit greater variability when ϵ is small. As ϵ increases, the HPD lengths quickly become stable and closely match those obtained by DP-Reject-ABC, particularly for $\epsilon \geqslant 0.5$.

4.1.2 Linear regression: Non-conjugate prior

Suppose the data model f is $y_i \mid x_i, \beta, \tau \sim N((1, x_i)\beta, \tau^{-1})$ and $x_i \mid \mu, \phi \sim N(\mu, \phi^{-1})$ with prior distributions $\beta \sim t_2(m, V, df)$, $\tau \sim \text{Weibull}(sc, sh)$, $\mu \sim t_p(\theta = 0_p, \Sigma = I_p, df = 2)$, and $\phi \sim$ folded $t_p(a = 0_p, B = I_p, df = 2)$. The private data $\tilde{x}_i$ and $\tilde{y}_i$ are obtained by clamping x_i and y_i to $[-5, 5]$ and subsequently rescaling them to $[-1, 1]$. Let $L_1, .., L_5 \overset{iid}{\sim} \text{Lap}(\frac{\Delta}{\epsilon})$, where $\Delta = 8$ and $p = 1$. Then, the privatized statistics s_{dp} are $(\sum \tilde{y}_i + L_1, \sum \tilde{x}_i\tilde{y}_i + L_2, \sum \tilde{y}_i^2 + L_3, \sum \tilde{x}_i + L_4, \sum \tilde{x}_i^2 + L_5)$, which satisfies ϵ-DP. In addition, the conjugate prior case is also implemented but deferred to Appendix A.

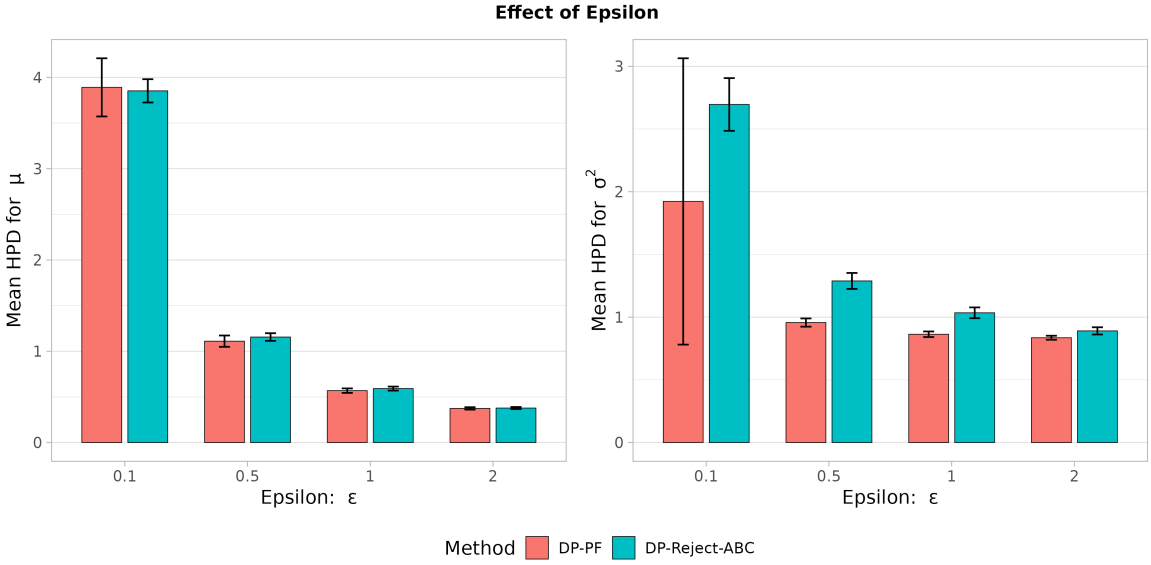

Figure 3: Comparison of the averaged lengths of the 90% from DP-PF HPDs with the MCSEs of the posterior distributions of μ and σ^2 given s_{dp}, under different values of ϵ.

		DP-PF		DP-DA-MCMC		DP-Reject-ABC	
ϵ	Sample Size	$\mathbb{E}(\beta_1 \mid s_{dp})$	Sec./ESS†	$\mathbb{E}(\beta_1 \mid s_{dp})$	Sec./ESS	$\mathbb{E}(\beta_1 \mid s_{dp})$	Sec./ESS†
0.5	300	1.973 (0.0128)	0.293	1.871 (0.0324)	629.809	2.083 (0.0031)	0.257
	500	1.772 (0.0079)	0.638	1.818 (0.0187)	1173.600	1.849 (0.0023)	7.079
1.0	300	1.983 (0.0085)	0.250	1.957 (0.0167)	370.535	2.116 (0.0021)	2.053
	500	1.889 (0.0070)	1.696	1.933 (0.0107)	857.429	1.934 (0.0016)	70.577
2.0	300	2.048 (0.0076)	0.634	2.075 (0.0113)	255.562	2.142 (0.0015)	23.235
	500	1.980 (0.0061)	7.481	1.995 (0.0071)	605.773	–	–

Table 2: Comparison of DP-PF, DP-DA-MCMC, and DP-Reject-ABC based on the averaged posterior mean of β_1 and seconds per effective sample size (Sec./ESS) over 1000 simulations for linear regression under a non-conjugate prior, with values in parentheses representing MCSEs for the posterior mean estimates. († Parallelized over particles.)

Simulation setup. True values of the parameters are set as $\beta^* = (0, 2)^\top$, $\tau^* = 1$, $\mu^* = 1$, and $\phi^* = 1$. The number of particles $N = 100$, $\epsilon \in \{0.5, 1, 2\}$ and $n \in \{300, 500\}$. The hyperparameters are set as $m = 0_2, V = I_2, df = 2, sc = 1.25, sh = 2, \theta = 0, \Sigma = 1, a = 0, B = 1$. We use a kernel with a pre-specified scale and $T = 10$ ($\epsilon_t = (0.1\epsilon, 0.2\epsilon, \ldots, \epsilon)$); results with different schedules are included in Section A.

ϵ	Sample Size	DP-PF 90% HPD	DP-DA-MCMC 90% HPD	DP-Reject-ABC 90% HPD
0.5	300	2.499 (0.0139)	4.485 (0.0625)	2.740 (0.0101)
	500	2.097 (0.0120)	3.009 (0.0274)	2.186 (0.0078)
1.0	300	1.935 (0.0093)	2.932 (0.0318)	1.864 (0.0063)
	500	1.435 (0.0072)	1.918 (0.0176)	1.362 (0.0042)
2.0	300	1.388 (0.0059)	1.899 (0.0189)	1.244 (0.0038)
	500	0.947 (0.0045)	1.164 (0.0084)	–

Table 3: Comparison of DP-PF, DP-DA-MCMC, and DP-Reject-ABC based on the averaged 90% HPD widths of β_1 over 1,000 simulations for linear regression under a non-conjugate prior, with values in parentheses representing MCSEs.

Table 2 demonstrates that DP-PF maintains accurate posterior mean estimation across a range of sample sizes and privacy budgets, even under a challenging heavy-tailed t prior with two degrees of freedom, whose variance is undefined. This prior is especially difficult for Monte Carlo-based DP methods because clamping may provide insufficient information to correct proposals far from the observed data. Since s_{dp} is fixed for each (n, ϵ) combination, posterior means may vary across settings when a "bad seed" generates an unrepresentative s_{dp}. Nevertheless, DP-PF remains computationally efficient. In contrast, DP-DA-MCMC has approximately twice the MCSEs and requires roughly 100 times more computation due to the additional Metropolis–Hastings step and its inability to parallelize. Although DP-Reject-ABC is theoretically exact, it failed to complete within 4 hours for $n = 500$ and $\epsilon = 2$ and is therefore omitted. Appendix A reports analogous results under conjugate priors, where DP-PF still outperforms DP-DA-MCMC under parallelization, although by a smaller margin. We also report the performance and computation time of DP-PF under an exponential ϵ_t schedule.

Table 3 shows that DP-PF produces 90% HPD widths that are consistently closer to those of DP-Reject-ABC than those produced by DP-DA-MCMC. This indicates that DP-PF more accurately captures posterior distributions in addition to providing accurate point estimates. The advantage is particularly pronounced when the privacy budget is moderate or large, where DP-PF closely tracks the benchmark while retaining its computational efficiency.

4.1.3 Gaussian mixture model

Suppose the data model f is $x_1, \ldots, x_n \mid \mu_1, \mu_2, \sigma_1^2, \sigma_2^2, p \overset{iid}{\sim} pN(\mu_1, \sigma_1^2) + (1-p)N(\mu_2, \sigma_2^2)$ with prior distributions $\mu_1, \mu_2 \overset{iid}{\sim} N(m, \tau^2)$ with the identifiability constraint $\mu_1 < \mu_2$ enforced by taking the minimum and maximum of two independent draws, $\sigma_k^2 \sim$ Inv. Gamma(α, β) for $k = 1, 2$, and $p \sim$ Beta(a, b). Let $L_1, \ldots, L_K \overset{iid}{\sim}$ Lap $\left(\frac{M}{n\epsilon}\right)$, where $K = \sum_{j=1}^{M} 2^j = 2^{M+1} - 2$ is the total number of histogram bins across all M layers. The privatized statistic s_{dp} is the M-layer hierarchical histogram, where layer j partitions the clamping region $[L, U]$ into 2^j equal bins and each layer receives an equal privacy budget of ϵ/M, giving a Laplace noise scale of $M/(n\epsilon)$ per layer. This satisfies ϵ-DP by the composition property.

ϵ	Layers	Sample Size	DP-PF				
			$\mathbb{E}(\mu_1 \mid s_{dp})$	$\mathbb{E}(\mu_2 \mid s_{dp})$	$\mathbb{E}(\sigma_1^2 \mid s_{dp})$	$\mathbb{E}(\sigma_2^2 \mid s_{dp})$	$\mathbb{E}(p \mid s_{dp})$
0.50	2	100	1.218 (0.114)	3.044 (0.089)	0.812 (0.102)	0.748 (0.103)	0.475 (0.044)
		200	0.961 (0.068)	2.927 (0.039)	0.510 (0.075)	0.349 (0.030)	0.419 (0.022)
		500	0.921 (0.027)	2.962 (0.015)	0.347 (0.021)	0.259 (0.011)	0.373 (0.009)
		1000	0.913 (0.016)	2.997 (0.012)	0.294 (0.012)	0.268 (0.007)	0.399 (0.006)
	3	100	1.264 (0.121)	2.967 (0.092)	0.835 (0.109)	0.724 (0.105)	0.482 (0.045)
		200	1.105 (0.049)	2.966 (0.035)	0.464 (0.067)	0.307 (0.037)	0.457 (0.022)
		500	0.944 (0.022)	2.995 (0.012)	0.419 (0.030)	0.272 (0.012)	0.389 (0.009)
		1000	0.894 (0.011)	2.959 (0.008)	0.287 (0.012)	0.277 (0.007)	0.386 (0.005)

Table 4: Average posterior estimates from DP-PF for the Gaussian Mixture Model parameters $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$, and p across 1,000 simulations, with the values in parentheses representing the MCSEs for each quantity.

ϵ	Layers	Sample Size	DP-DA-MCMC				
			$\mathbb{E}(\mu_1 \mid s_{dp})$	$\mathbb{E}(\mu_2 \mid s_{dp})$	$\mathbb{E}(\sigma_1^2 \mid s_{dp})$	$\mathbb{E}(\sigma_2^2 \mid s_{dp})$	$\mathbb{E}(p \mid s_{dp})$
0.5	2	100	0.999 (0.083)	3.019 (0.036)	1.196 (0.155)	0.348 (0.040)	0.577 (0.015)
		200	0.869 (0.097)	2.829 (0.023)	0.817 (0.100)	0.158 (0.005)	0.506 (0.011)
		500	0.607 (0.172)	2.915 (0.030)	1.053 (0.289)	0.156 (0.009)	0.485 (0.022)
		1000	0.491 (0.277)	2.954 (0.062)	1.138 (0.424)	0.183 (0.014)	0.532 (0.039)
	3	100	1.241 (0.026)	2.961 (0.011)	0.591 (0.053)	0.231 (0.014)	0.536 (0.010)
		200	1.256 (0.008)	2.951 (0.005)	0.431 (0.014)	0.147 (0.003)	0.507 (0.004)
		500	0.856 (0.039)	3.026 (0.006)	1.116 (0.105)	0.202 (0.005)	0.485 (0.010)
		1000	0.891 (0.002)	2.955 (0.002)	0.340 (0.012)	0.255 (0.002)	0.395 (0.002)

Table 5: Average posterior estimates from DP-DA-MCMC for the Gaussian Mixture Model parameters $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$, and p across 1,000 simulations, with the values in parentheses representing the MCSEs for each quantity.

Simulation setup. True parameters are set as $\mu_1^* = 1$, $\mu_2^* = 3$, $\sigma_1^{*2} = \sigma_2^{*2} = 0.25$, $p^* = 0.4$, with clamping bounds $L = 0$ and $U = 5$. We use $M \in \{2, 3\}, \epsilon \in \{0.5, 1, 2\}, n \in \{100, 200, 500, 1000\}$, and $N = 100$ particles. The hyperparameters are $m = 2.5$, $\tau^2 = 1.5^2$, $\alpha = 3$, $\beta = 0.5$, $a = b = 1$. We use an adaptive scaling of the perturbation kernel based on the covariance structure of particles in each iteration and $T = 10$ ($\epsilon_t = (0.1\epsilon, 0.2\epsilon, \ldots, \epsilon)$).

For the case of $M = 3$, both DP-PF and DP-DA-MCMC perform well, as the degrees of freedom of s_{dp} exceed the number of parameters. For the case of $M = 2$, this is not the case, and because of this, we expect the posterior distribution to be multi-modal. In this case, DP-PF still recovers sufficient information from s_{dp} and yields accurate posterior estimates for all parameters, whereas DP-DA-MCMC recovers only one location parameter well.

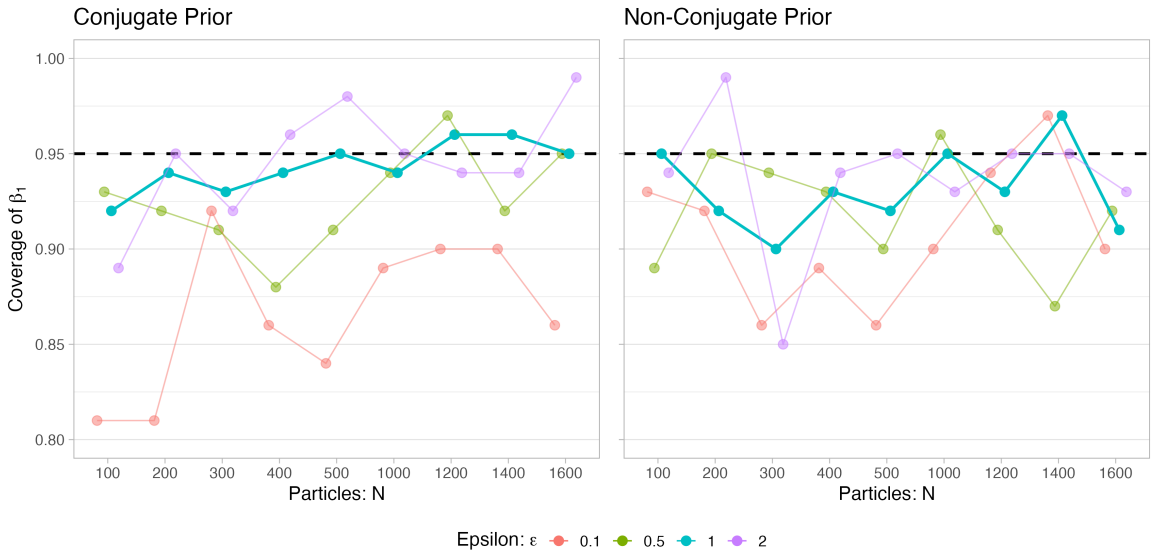

Figure 4: Comparison of conjugate and non-conjugate priors on the β_1 coverage in N for linear regression via DP-PF. The estimates are based on Theorem 14, which establishes 95% asymptotic confidence intervals for the posterior mean of β_1.

4.2 Asymptotic Confidence Interval Coverage

In this section, we demonstrate using the linear regression model under both conjugate and non-conjugate priors, introduced in Section 4.1 and Appendix A.2, that Algorithm 3 achieves the nominal 95% coverage of its asymptotic confidence interval. The values of s_{dp} are carefully chosen so that the private posterior mean of β_1 is known to be 0, which we use as the ground truth for assessing coverage.

Simulation setup. We set $L = -5, U = 5$ and the privatized statistics $s_{dp} = (0, 0, 50, 0, 50)$, in the same order as in Section 4.1.2. $n = 500$, $\epsilon \in \{0.1, 0.5, 1, 2\}$ and $N \in \{100, 200, 300, 500, 1000, 1200, 1400, 1600\}$. We use 100 independent simulations to assess whether the empirical coverage is consistent with the nominal 95% level.

Note that there are no true parameters in this setting, as the data are not generated from any model. Instead, we construct s_{dp} to have a convenient property: the posterior mean of β_1 is known to be 0 under the symmetric and well-behaved prior. This construction of s_{dp} allows us to evaluate coverage of the posterior mean, which would be difficult for a generic private summary statistic, as the posterior mean usually does not have a closed-form expression.

Figure 4 illustrates that the coverage of the 95% confidence interval (CI) for the posterior mean of β_1 generally approaches the nominal level as the number of particles N increases. For $\epsilon \geqslant 0.5$, the coverage stabilizes around 0.95 when $N > 1000$, indicating that DP-PF achieves reliable coverage with a sufficiently large number of particles. Moreover, as N increases, the coverage improves while the interval width decreases. Notably, smaller

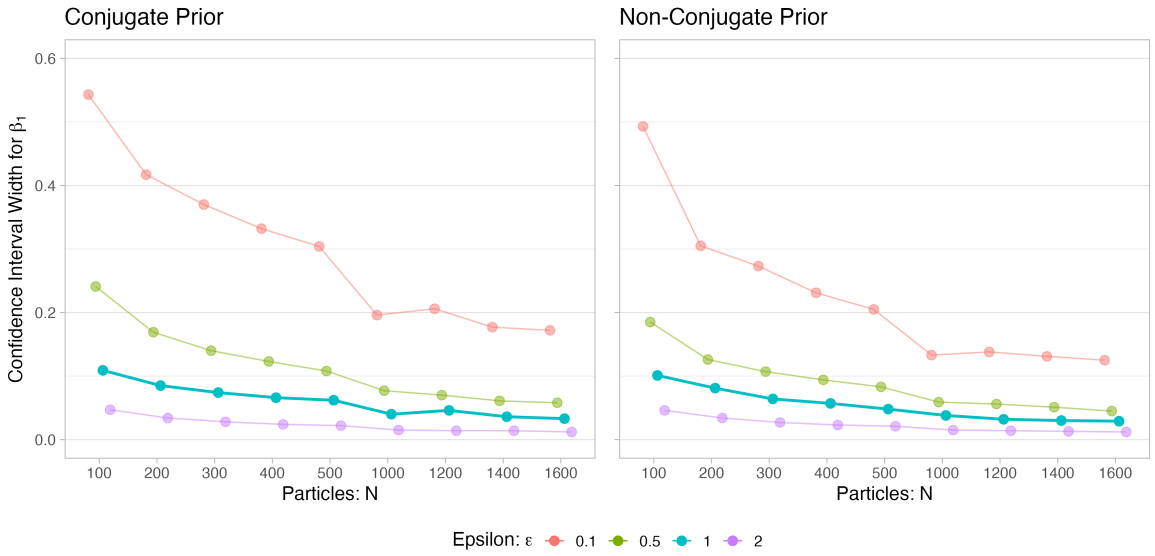

Figure 5: Comparison of conjugate and non-conjugate priors on the confidence interval width of β_1 in N for linear regression via DP-PF. The estimates are based on Theorem 14, which provides 95% asymptotic confidence intervals for the posterior mean of β_1.

values of ϵ require a larger N to achieve the same coverage and interval width as larger ϵ. This observation aligns with the findings for DP-DA-MCMC, which noted worse mixing of their chain for small ϵ. This, however, is not a major concern for our method, as smaller ϵ values allow for faster computations, making it feasible to use a larger N without significant computational overhead.

Figure 5 shows that the average CI width decreases as N increases. For $\epsilon \geqslant 0.5$, the CI width stabilizes below 0.05 when $N > 1000$, indicating that DP-PF achieves both precise and stable inference with a sufficient number of particles. This suggests that increasing the particle size beyond this point provides limited additional improvement in the precision of the resulting inference in this case.

4.3 Model Evidence

We consider a model comparison problem where three candidate Bayesian models are fit to the same privatized summary statistic s_{dp}, using 1,000 replicates. The goal is to assess which model best explains s_{dp} under a fixed privacy mechanism.

The three competing models are: (M1) a single Gaussian model (Section 4.1.3); (M2) a two-component Gaussian mixture model (Section 4.1.2); and (M3) a three-component

	M1 (Single Gaussian)	M2 (2-component)	M3 (3-component)
Mean log evidence	3.886 (0.0041)	4.820 (0.0048)	4.879 (0.0063)
$\hat{P}$(best model)	0.000 (0.0000)	0.397 (0.0154)	0.603 (0.0154)

Table 6: Model comparison via DP-PF marginal likelihood estimates over 1,000 replicates.

Gaussian mixture model:

$$x_i \mid \underline{\mu}, \underline{\sigma}^2, \underline{p} \overset{iid}{\sim} \sum_{k=1}^{3} p_k N(\mu_k, \sigma_k^2),$$

with ordered means $\mu_1 < \mu_2 < \mu_3$, priors $\mu_k \sim N(m, \tau^2)$, $\sigma_k^2 \sim$ Inv-Gamma(α, β), and $(p_1, p_2, p_3) \sim \text{Dir}(1, 1, 1)$. The true data-generating model is a 2-component Gaussian mixture (M2), and the privacy mechanism is the hierarchical histogram queries used in Section 4.1.3 with $M = 2$.

Simulation setup. All models are fit using the DP-PF algorithm with identical mechanisms, privacy budget, and hyperparameters as described above, with $n = 100$, $\epsilon = 0.5$, $N = 100$ particles, and $T = 10$ iterations (linear ϵ_t schedule). All three models use the same hyperparameters $m = 2.5$, $\tau^2 = 1.5^2$, $\alpha = 3$, $\beta = 0.5$ for mean and variance parameters; M2 additionally sets $a = b = 1$, while M3 uses a symmetric Dirichlet prior. Data are generated from a Gaussian mixture model with two components; that is, M2 is the true model. Results are summarized over 1,000 independent replications in Table 6, where $\hat{P}$(best model) denotes the proportion of replications in which each model achieves the highest log evidence.

Table 6 shows that M2 is selected as the best model in approximately 40% of the replications, while M3 is selected in about 60% of the time; M1 is never selected as the best model across all replications, hence its zero MCSE. Moreover, the log evidence of M2 and M3 is close relative to their MCSEs, indicating that neither model is strongly favored over the other.

5. Data Analysis

In this section, we demonstrate that our particle filtering algorithm can be applied to a non-additive privacy mechanism as well. In particular, we derive the acceptance probability for a logistic regression, using the objective perturbation mechanism (Chaudhuri et al., 2011) with the formulation given in Awan and Slavković (2021), which adds noise to the gradient of the log-likelihood before minimization, thereby creating non-additive noise. Moreover, since we assume that the independent variable is protected, we privatize its first two moments via the K-norm mechanism (Hardt and Talwar, 2010; Awan and Slavković, 2021), a multivariate additive-noise mechanism, which is similar to the setup used in Awan and Wang (2025).

We conduct an experiment with the 2021 Census Public Use Microdata Files (PUMF), which provide data on the characteristics of the Canadian population (Canada, 2022). We analyze the dependence between age and retirement status by inferring logistic regression under DP guarantees.

5.1 Experiment Setup

The PUMF data set contains 980,868 records of individuals, representing 2.5% of the Canadian population. Among the 144 variables in the data set, we chose three variables: the census metropolitan area or census agglomeration of current residence (named CMA), the age (named AGEGRP), and the private retirement income (named Retir). We extracted the records of AGEGRP and Retir belonging to the people in Victoria (according to the values in CMA). After pre-processing the records with unavailable or not applicable values, the sample size is 10,473; Retir is transformed into a binary variable with Retir = 1 for retirement and Retir = 0 otherwise, and AGEGRP is scaled into the range of $[0, 1]$. For the preliminary data analysis, we find the non-private regression coefficient between Retir and AGEGRP is $\hat{\beta} = (-3.825, 6.822)$, and the distribution of AGEGRP is near-uniform with softened edges.

The DP-PF algorithm is conducted with multinomial resampling and Gaussian perturbation kernels. We use a kernel with pre-specified scale and $T = 10$ ($\epsilon_t = (0.1\epsilon, 0.2\epsilon, \ldots, \epsilon)$ and $q_t = (0.02q, 0.02q, \ldots, 0.02q, q)$). $N = 200$. This additional q_t schedule creates a similar effect to the ϵ_t schedule, contributing to flatter intermediate target distributions and thereby increasing the acceptance probability r_t. The same computing environment is used as in Section 4.

5.2 Logistic Regression

The logistic regression model is assumed to be $P(Y = 1 \mid X) = \frac{1}{1+\exp\{-(\beta_0+\beta_1 X)\}}$ and $Z \overset{d}{=} \frac{X+1}{2} \sim \text{Beta}(a, b)$, with prior distributions $\beta_0, \beta_1 \overset{\text{iid}}{\sim} N(0, 16)$ and $a, b \overset{\text{iid}}{\sim} \text{Gamma}(6, 4)$. We set $\epsilon = 0.5$ and $q = 0.5$, and randomly draw $n = 1000$ samples from the Victoria data set. The $s_{dp} = (\beta_{0_{dp}}, \beta_{1_{dp}}, \sum z_i + K_1, \sum z_i^2 + K_2)$ we get from objective perturbation and K-norm mechanisms (Algorithms 4 and 5) is (-2.982634, 5.119215, 479.573100, 302.947785). We derive the acceptance probability of our algorithm on the objective perturbation for the regression coefficient and introduce the K-norm mechanism for the first two moments' sum of AGEGRP in Appendix B. We use the ℓ_∞-norm with $\Delta_\infty = 2$ for both mechanisms, allocating 90% of the privacy budget to the regression coefficients and 10% to the two moments. Our method gives the private posterior mean of $\beta_0 = -2.357$ and $\beta_1 = 3.816$ with an effective sample size of 150.787 and runtime of 3103.811 seconds.

Figure 6 compares the 95% most representative private curves with the 95% non-private posterior curves. The selected private curves correspond to the sampled β whose values are closest to the median, as measured by the Mahalanobis distance using the covariance matrix of β. Similarly, the selected non-private curves, obtained via the Bayesian Metropolis-Hastings algorithm, correspond to the 95% closest posterior samples to its median among the 200 thinned β samples, drawn from a total of 5,000 iterations after a 5,000-sample burn-in, with an acceptance rate of 20.72%. The blue curves represent the estimated logistic functions, while the black step lines indicate the observed proportions of retirement across age groups. The red dashed curve shows the logistic fit based on the original (non-private) data, and the horizontal dotted line represents the model determined by the prior means.

While the non-private posterior curves closely align with the red fitted curve, it is critical to note that the private estimation is based solely on differentially private (DP) summary statistics, without access to the original data. Despite some variability among

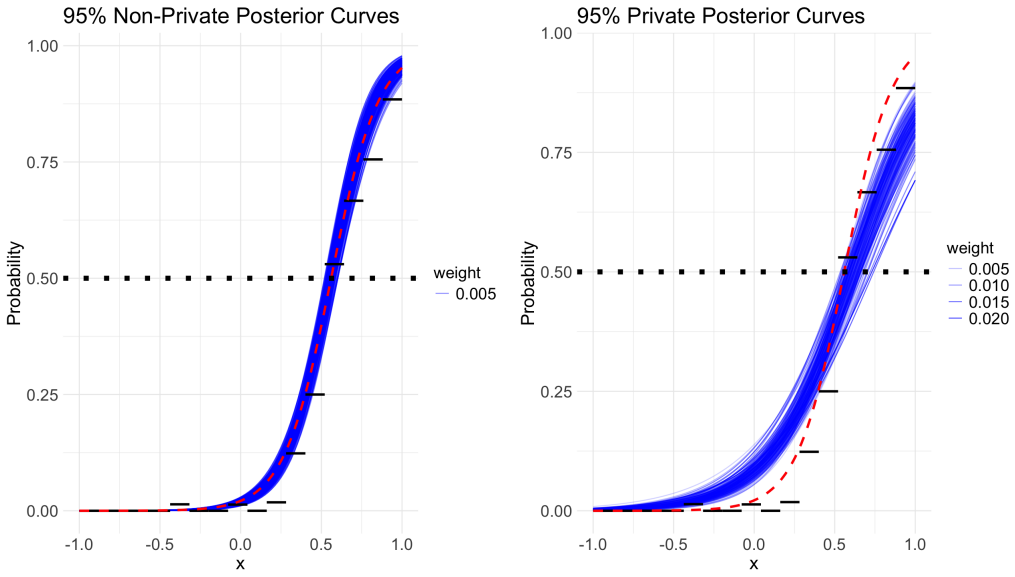

Figure 6: Comparison of 95% non-private and private posterior curves for logistic regression via Algorithms 4 and 5, which illustrates the result of Section 5.2. Black step lines: proportions of retirement from AGEGRP. Red dashed curve: fitted logistic curve. Horizontal dotted line: model given by the prior means.

the private curves, their alignment with the black step lines suggests that the privatized estimates capture the underlying data structure reasonably well. Although DP introduces randomness at the cost of information loss—causing our method to be more susceptible to regularization effects from the prior—the private logistic curves approximate the trend and shape of the non-private fit, maintaining a reasonable correspondence with the observed proportions and preserving key patterns in the data while ensuring privacy.

6. Discussion and Directions for Future Work

In this paper, we propose a novel differentially private particle filtering algorithm to generate approximate posterior samples based on confidential queries. Our sampler offers the significant advantage of producing consistent estimates for the parameter θ given s_{dp}. The tempering ϵ_t schedule enhances the efficiency of the sampler. Compared to MCMC methods, our sampler does not require strong assumptions for convergence or for (CLT) to hold.

Our method is conceptually related to existing approaches that treat the differentially private posterior as doubly intractable, but differs in assumptions, computation, and practical behavior. Compared with DP-DA-MCMC of Ju et al. (2022), which augments the confidential database and relies on a privacy-aware Metropolis-within-Gibbs sampler, DP-PF replaces Markov chain dynamics with sequential Monte Carlo, avoiding mixing issues, multimodality concerns, and the geometric ergodicity assumptions required for theoretical guarantees in DP-DA-MCMC, while accommodating non-conjugate priors more naturally. At the same time, because DP-DA-MCMC updates latent data sequentially, it may scale

more favorably with sample size n in some settings. Compared with DP-Reject-ABC of Gong (2022), which exploits additive-noise structure through rejection sampling and may suffer from low acceptance rates or limited applicability to other perturbation mechanisms, DP-PF combines a privacy-tempering schedule with sequential propagation and rejection correction to improve sampling efficiency while preserving the exactness of the target distribution. Empirically, our simulations demonstrate competitive or improved computational performance and accurate posterior inference, while the SMC framework additionally provides Monte Carlo error quantification and marginal likelihood estimation.

While our analysis in Section 3 focuses on multinomial resampling, the validity of our Monte Carlo estimates and confidence intervals extends to other resampling schemes as well. Extending the proof to stratified or systematic resampling is not straightforward, as their next-iteration particle systems are determined by a single uniform random draw, unlike multinomial and residual resampling, which generate independent particles. However, all these resampling schemes satisfy a central limit theorem, ensuring that the estimates and intervals remain valid under the appropriate asymptotic variances. Although different schemes may alter the form of the asymptotic variance, our Taylor expansion-based analysis still applies. Moreover, some schemes may achieve smaller asymptotic variances compared to multinomial resampling, depending on how they influence the resampled particle system. The detailed analysis is left for future work.

Our algorithm inevitably becomes more computationally demanding as the sample size and privacy budget ϵ increase, and its performance depends critically on the quality of the importance weights. In addition, the dimensions of both the parameter space θ and the privatized statistic s_{dp} introduce the curse of dimensionality common to importance sampling methods (Chopin and Papaspiliopoulos, 2020). Because the importance weights must be analytically tractable, some potentially efficient proposal distributions may be excluded. Another limitation is that each iteration generates synthetic data solely from θ, without leveraging information from previously generated data. Our empirical results also suggest that stable and reliable inference typically requires at least 50 observations per parameter, indicating that high-dimensional inference with our DP-PF remains fundamentally difficult. Future work may explore dimension-reduction techniques (Talwar et al., 2015) or efficient synthetic-data update approaches such as DP-DA-MCMC (Ju et al., 2022) to improve scalability in higher-dimensional settings.

An additional direction for future work is to improve the scalability and adaptivity of the DP-PF framework. A promising extension is adaptive tempering, where the increment of ϵ_t is determined dynamically based on particle degeneracy measures such as the effective sample size (ESS), following adaptive SMC ideas in Agapiou et al. (2017); Chopin and Papaspiliopoulos (2020). Such strategies may improve computational stability and efficiency by deciding the number of iterations of DP-PF and allocating its tempering scheme based on the evolving particle system, but they also introduce theoretical challenges because tempering schedules are data-dependent, complicating privacy accounting and the statistical analysis. As for high-dimensional problems, the synthetic data updates can further bottleneck parameter updates. While projecting updates onto a lower-dimensional feature space—by leveraging sufficient statistics or latent factors—offers a scalable and efficient solution, it is highly problem-specific and requires additional assumptions.

Acknowledgments

This work was supported in part by NSF grants SES-2150615 and SES-2610910. Part of this project was completed while Jordan Awan was a member of the Department of Statistics at Purdue University. The idea to manipulate the privacy budget in our sampler was inspired by preliminary work completed by a team consisting of Ruobin Gong, Nianqiao Ju, Vinayak Rao, and Jordan Awan, where a similar ϵ-schedule was used in an MCMC sampler; we are grateful for these prior discussions that led to this work.

Appendix A. More Results of Simulation

In this section, we present additional simulation results that could not be included in the main text but provide further support for our findings.

A.1 Location-Scale Normal: HPD in sample sizes

To examine the effect of sample size on the HPD intervals, we run DP-PF across a range of sample sizes. All other settings are identical to those used in Figure 3.

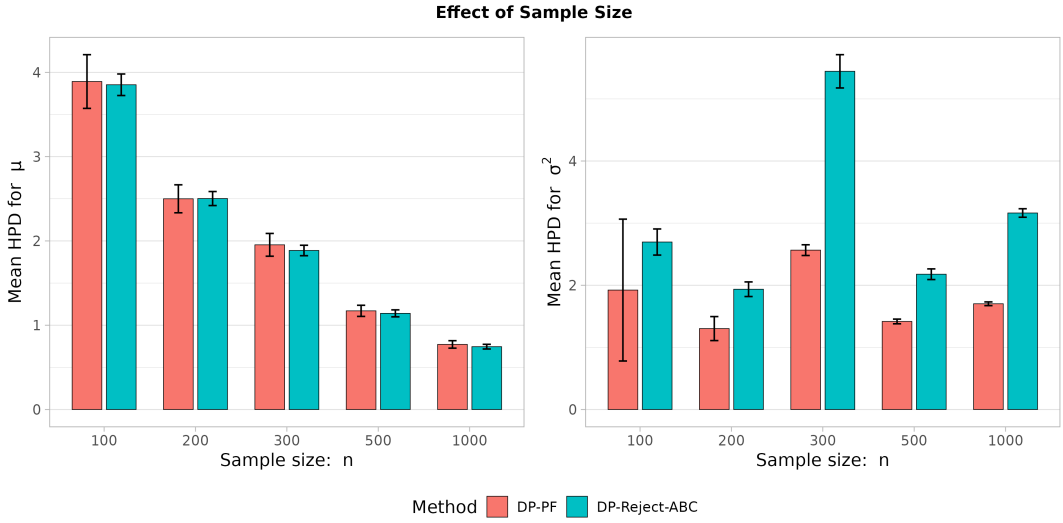

Figure 7: Comparison of the averaged lengths of the 90% HPDs from DP-PF with the MCSEs of the posterior distributions of μ and σ^2 given s_{dp}, under different values of sample sizes.

Figure 7 shows that the lengths of the 90% HPDs for μ obtained by DP-PF and DP-Reject-ABC are nearly indistinguishable relative to their MCSEs. In contrast, DP-PF appears to produce slightly shorter HPDs for σ^2 than DP-Reject-ABC. Aside from the case $n = 300$, where both methods exhibit an unexpected increase in HPD length, DP-PF appears to systematically underestimate the HPD length for σ^2. We expect this discrepancy to diminish with a larger number of particles, such as $N = 2000$ or higher.

A.2 Linear Regression: Conjugate prior

The data model f is the same as the non-conjugate prior but with conjugate prior assumptions $\beta \mid \tau \sim N_2(m, \tau^{-1}V^{-1}), \tau \sim \text{Gamma}\left(\frac{a}{2}, \frac{b}{2}\right), \mu \sim N(\theta, \Sigma)$, and $\phi \sim \chi^2(d)$. The privatized statistics s_{dp} are the same as in the non-conjugate prior setting.

Simulation setup. With bounds $L = -5, U = 5$, true values of the parameters are set as $\beta^* = (0, 2)^\top$, $\tau^* = 1$, $\mu^* = 1$, and $\phi^* = 1$. $\epsilon \in \{0.5, 1, 2\}$ and $n \in \{300, 500\}$. The hyper-parameters are set as $m = (0, 0)^\top$, $V = I_2$, $a = 2$, $b = 2$, $\theta = (0)^\top$, $\Sigma = 1$, and $d = 2$.

		DP-PF		DP-DA-MCMC		DP-Reject-ABC	
ϵ	Sample Size	$\mathbb{E}(\beta_1 \mid s_{dp})$	Sec./ESS	$\mathbb{E}(\beta_1 \mid s_{dp})$	Sec./ESS	$\mathbb{E}(\beta_1 \mid s_{dp})$	Sec./ESS
0.50	300	1.906 (0.0133)	0.296	1.679 (0.0223)	24.456	1.842 (0.0030)	0.150
	500	1.726 (0.0087)	0.868	1.618 (0.0109)	45.503	1.648 (0.0024)	4.959
1	300	1.900 (0.0092)	0.268	1.751 (0.0126)	17.024	1.895 (0.0020)	1.326
	500	1.845 (0.0083)	2.796	1.762 (0.0071)	36.384	1.791 (0.0015)	64.206
2	300	1.951 (0.0066)	0.591	1.896 (0.0095)	12.984	1.957 (0.0013)	17.053
	500	1.918 (0.0070)	11.657	1.876 (0.0050)	28.439	–	–

Table 7: Comparison of DP-PF, DP-DA-MCMC, and DP-Reject-ABC based on the averaged posterior mean of the regression coefficient β_1 and seconds per effective sample size (Sec./ESS) over 1000 simulations for linear regression under conjugate priors, with the values in parentheses representing the MCSEs for the posterior mean estimates. († Parallelized over particles.)

		DP-PF	DP-DA-MCMC	DP-Reject-ABC
ϵ	Sample Size	90% HPD	90% HPD	90% HPD
0.5	300	1.897 (0.0099)	2.825 (0.0386)	2.664 (0.0092)
	500	1.507 (0.0076)	1.927 (0.0243)	2.023 (0.0068)
1.0	300	1.635 (0.0073)	2.178 (0.0243)	1.905 (0.0060)
	500	1.192 (0.0054)	1.422 (0.0139)	1.401 (0.0044)
2.0	300	1.275 (0.0058)	1.672 (0.0173)	1.349 (0.0040)
	500	0.844 (0.0036)	1.038 (0.0092)	–

Table 8: Comparison of DP-PF, DP-DA-MCMC, and DP-Reject-ABC based on the averaged 90% HPD widths of β_1 over 1000 simulations for linear regression under conjugate priors, with values in parentheses representing MCSEs.

Table 7 along with Table 2 show that DP-PF consistently delivers strong accuracy across different sample sizes, privacy levels, and prior settings. Although DP-PF does not gain as much computational speed from conjugate priors as DP-DA-MCMC, it offers a key advantage that its implementation remains consistent and straightforward regardless of the prior. In contrast, while DP-DA-MCMC benefits from a simplified structure when conjugate priors are available, it becomes notably more complex to implement and slower to run in non-conjugate settings. This highlights a key strength of DP-PF that its simplicity and flexibility make it easier to apply across a wide range of models.

A.3 Linear Regression: on ϵ_t Schedule

To demonstrate the performance under different ϵ_t schedules, we implemented an exponential schedule in addition to the linear one using linear regression under the conjugate and non-conjugate priors stated in Section 4.1 and Appendix A.2. Under this schedule, the privacy budget increases rapidly during the early iterations and gradually approaches the

		DP-PF (linear ϵ_t schedule)			DP-PF (exponential ϵ_t schedule)		
ϵ	Sample Size	$\mathbb{E}(\beta_1 \mid s_{dp})$	$\text{Var}(\beta_1 \mid s_{dp})$	Sec./ESS†	$\mathbb{E}(\beta_1 \mid s_{dp})$	$\text{Var}(\beta_1 \mid s_{dp})$	Sec./ESS†
0.5	300	1.973 (0.0128)	0.523 (0.0133)	0.293	1.970 (0.0136)	0.554 (0.0167)	0.344
	500	1.772 (0.0079)	0.295 (0.0065)	0.638	1.778 (0.0084)	0.293 (0.0066)	1.010
1.0	300	1.983 (0.0085)	0.314 (0.0062)	0.250	1.987 (0.0094)	0.325 (0.0064)	0.349
	500	1.889 (0.0070)	0.178 (0.0029)	1.696	1.888 (0.0071)	0.182 (0.0056)	3.089
2.0	300	2.048 (0.0076)	0.187 (0.0051)	0.634	2.052 (0.0079)	0.185 (0.0073)	1.078
	500	1.980 (0.0061)	0.095 (0.0038)	7.481	1.980 (0.0064)	0.099 (0.0027)	14.021

Table 9: Comparison of DP-PF with linear and exponential ϵ_t schedules based on the averaged posterior means and variances of β_1 over 1000 simulations for a linear regression model under a non-conjugate prior, and the seconds spent per ESS, with the values in the parentheses representing MCSEs for the posterior mean and variance estimates. (†Parallelized over particles.)

		DP-PF (linear ϵ_t schedule)			DP-PF (exponential ϵ_t schedule)		
ϵ	Sample Size	$\mathbb{E}(\beta_1 \mid s_{dp})$	$\text{Var}(\beta_1 \mid s_{dp})$	Sec./ESS†	$\mathbb{E}(\beta_1 \mid s_{dp})$	$\text{Var}(\beta_1 \mid s_{dp})$	Sec./ESS†
0.5	300	1.906 (0.0133)	0.581 (0.0127)	0.296	1.920 (0.0145)	0.615 (0.0175)	0.421
	500	1.726 (0.0087)	0.373 (0.0067)	0.868	1.713 (0.0097)	0.373 (0.0077)	1.380
1.0	300	1.900 (0.0092)	0.310 (0.0073)	0.268	1.901 (0.0094)	0.306 (0.0062)	0.380
	500	1.845 (0.0083)	0.207 (0.0118)	2.796	1.841 (0.0074)	0.185 (0.0040)	4.202
2.0	300	1.951 (0.0066)	0.161 (0.0046)	0.591	1.952 (0.0062)	0.157 (0.0033)	0.990
	500	1.918 (0.0070)	0.104 (0.0034)	11.657	1.937 (0.0071)	0.111 (0.0047)	22.512

Table 10: Comparison of DP-PF with linear and exponential ϵ_t schedules based on the averaged posterior means and variances of β_1 over 1,000 simulations for a linear regression model under a conjugate prior, and the seconds spent per ESS, with the values in the parentheses representing MCSEs for the posterior mean and variance estimates. (†Parallelized over particles.)

total budget ϵ. Specifically,

$$\epsilon_t = \epsilon \left(1 - 0.9e^{-0.3(t-1)}\right), \qquad t = 1, \ldots, T-1,$$

and $\epsilon_T = \epsilon$.

From Tables 9 and 10, we observe that the posterior estimates and their MCSEs are similar under the two schedules. Although the linear schedule is somewhat more computationally efficient in terms of seconds per ESS, the exponential schedule achieves comparable overall performance.

Appendix B. Details of Real Data Analysis

For implementing Algorithm 3 with K-norm mechanisms, we need the following proposition.

Algorithm 4 Sampling $\propto \exp(-c\|V\|_\infty)$: Steinke and Ullman (2016)

Input: c and dimension d

1: Set U_j for $j = 1, \ldots, d$

2: Draw r from $\Gamma(\alpha = d + 1, \beta = c)$

3: Set $V = r \cdot (U_1, \ldots, U_d)^\top$

Output: V

Proposition 17 *If the K-norm mechanism is used as the privacy mechanism in Algorithm 3 for $T_z = (\sum z_i, \sum z_i^2)$, then the acceptance probability for the sample z is*

$$r_t(z) = \exp\left(\frac{-\epsilon_t}{\Delta}\|T_z - Z_{dp}\|_K\right)$$

at iteration t.

Proof The K-Norm mechanism (Hardt and Talwar, 2010) adds a random variable V to the summary statistics with density

$$f_v(v) = \frac{\exp(\frac{-\epsilon}{\Delta}\|v\|_K)}{\Gamma(d+1)\lambda(\frac{\Delta}{\epsilon}K)},$$

which satisfies ϵ-DP, where $\Gamma(\cdot)$ is the Gamma function and $\lambda(\cdot)$ is the Lebesgue volume function. Given fixed dimensions d and K, since both of the functions remain constant, the acceptance rate follows directly from calculating

$$r_t(z) = \frac{m_{\epsilon_t}(Z_{dp}|z)}{\sup_z m_{\epsilon_t}(Z_{dp}|z)} = \exp\left(\frac{-\epsilon_t}{\Delta}\|T_z - Z_{dp}\|_K - \frac{-\epsilon_t}{\Delta}\|0\|_K\right) = \exp\left(\frac{-\epsilon_t}{\Delta}\|T_z - Z_{dp}\|_K\right).$$

■

In general, sampling from a K-norm mechanism can be challenging as it requires a precise characterization of the sensitivity space and relies on rejection sampling to give an exact uniform sampling from the sensitivity space (Awan and Slavković, 2021). For the ℓ_∞-norm used in our data analysis, Algorithm 4 provides a simple way to sample from random vectors whose density is proportional to their ℓ_∞-norm.

To implement Algorithm 3 with the objective perturbation mechanism described in Algorithm 5 under a logistic model, we need the following proposition.

Proposition 18 *Under a logistic model, if the objective perturbation is used as the privacy mechanism in Algorithm 5 for the regression coefficient $\theta = (\beta_0, \beta_1, \ldots, \beta_{d-1})$, then the acceptance probability for the data set $D = (X, Y)$ is*

$$r_t(D) = \frac{\exp\left(\frac{-\epsilon_t}{\Delta_K}\|V(\theta_{dp}; D)\|_K\right)}{(\gamma_t + \lambda_1)(\gamma_t + \lambda_2)\cdots(\gamma_t + \lambda_d)}\gamma_t^d$$

at iteration t, where $\gamma_t = \frac{d/4}{\exp(\epsilon_t(1-q_t))-1}$ and $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_d$ are the eigenvalues of $\sum_{i=1}^n \left(\frac{\exp(\theta_{dp}^\top X_i)}{(1+\exp(\theta_{dp}^\top X_i))^2}\right) X_i X_i^\top$.

Algorithm 5 ℓ_∞ Objective Perturbation: Awan and Slavković (2021)

Input: $X \in \mathcal{X}^n, \varepsilon > 0$, a convex set $\Theta \subset \mathbb{R}^d$, a convex function $r : \Theta \to \mathbb{R}$, a convex loss $\hat{\mathcal{L}}(\theta; X) = \frac{1}{n}\sum_{i=1}^n \ell(\theta; x_i)$ defined on Θ such that $\nabla^2\ell(\theta; x)$ is continuous in θ and $x, \Delta > 0$ such that $\sup_{x,x'\in\mathcal{X}} \sup_{\theta\in\Theta} \|\nabla\ell(\theta; x) - \nabla\ell(\theta; x')\|_\infty \leqslant \Delta, \lambda > 0$ is an upper bound on the eigenvalues of $\nabla^2\ell(\theta; x)$ for all $\theta \in \Theta$ and $x \in \mathcal{X}$, and $q \in (0, 1)$.

1: Set $\gamma = \frac{\lambda}{\exp(\varepsilon(1-q))-1}$

2: sample V from density $\propto \exp(\frac{-\varepsilon q}{\Delta}\|V\|_\infty)$ from Algorithm 4

3: Compute $\theta_{dp} = \arg\min_\theta \frac{1}{n}\sum_{i=1}^n l(\theta, x_i) + \frac{1}{n}r(\theta) + \frac{\gamma}{2n}\theta^\top\theta + \frac{1}{n}V^\top\theta$.

Output: θ_{dp}

Proof For iteration t, let $\gamma_t = \frac{d/4}{\exp(\epsilon_t(1-q_t))-1}$. Consider

$$\theta_{dp} = \arg\min_\theta \hat{\mathcal{L}}(\theta; D) + \frac{\gamma_t}{2n}\theta^\top\theta + \frac{1}{n}V^\top\theta,$$

where V is drawn from $f(V) \propto \exp(\frac{-\epsilon_t}{\Delta_K}\|V\|_K)$ and $0 < q_t < 1$. Let $\theta_{dp} = a$, then we have the identity $\nabla_\theta\hat{\mathcal{L}}(a; X) + \frac{\gamma_t}{n}a + \frac{1}{n}V = 0$. Thus, we know

$$V(a; D) = -\left(\gamma_t a + n\nabla_\theta\hat{\mathcal{L}}(a; D)\right)$$
$$\nabla_\theta V(a; D) = -\left(\gamma_t I + n\nabla^2_\theta\hat{\mathcal{L}}(a; D)\right).$$

Then, the pdf of $\theta_{dp} = a$ given X can be expressed with change of variable

$$\frac{f_V(V(a; D))}{|\det \nabla V(a; D)|} = \frac{f_V\left(-(\gamma_t a + n\nabla_\theta\hat{\mathcal{L}}(a; D))\right)}{\det\left(\gamma_t I + n\nabla^2_\theta\hat{\mathcal{L}}(a; D)\right)}$$
$$= \frac{\frac{1}{\Gamma(d+1)\lambda(\frac{\Delta}{\epsilon_t}K)}\exp\left(\frac{-\epsilon_t}{\Delta}\|-\left(\gamma_t a + \sum_{i=1}^n(\frac{\exp(a^\top X_i)}{1+\exp(a^\top X_i)} - Y_i)X_i\right)\|_K\right)}{(\gamma_t + \lambda_1)(\gamma_t + \lambda_2)\cdots(\gamma_t + \lambda_d)}$$

because the Hessian of $\hat{\mathcal{L}}$ can be expressed as

$$\nabla^2_\theta\hat{\mathcal{L}}(a; D) = \frac{1}{n}\sum_{i=1}^n\left(\frac{\exp(a^\top X_i)}{(1+\exp(a^\top X_i))^2}\right)X_i X_i^\top,$$

where the eigenvalues of $n\nabla^2_\theta\hat{\mathcal{L}}(a; D)$ for the synthetic data D are denoted as $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_d$. Thus,

$$m_{\epsilon_t}(\theta_{dp} = a \mid D) = \frac{C_{\epsilon_t}\exp\left(\frac{-\epsilon_t}{\Delta}\|V(a; D)\|_K\right)}{(\gamma_t + \lambda_1)(\gamma_t + \lambda_2)\cdots(\gamma_t + \lambda_d)},$$

where $C_{\epsilon_t} = \frac{1}{\Gamma(d+1)\lambda(\frac{\Delta}{\epsilon_t}K)}$. Note that $\sup_X m_{\epsilon_t}(\theta_{dp} = a \mid D) = \frac{C_{\epsilon_t}}{\gamma_t^d}$. Therefore, the acceptance rate of our differentially private particle filtering algorithm at iteration t is

$$r_t(D) = \frac{m_{\epsilon_t}(\theta_{dp} = a \mid D)}{\sup_X m_{\epsilon_t}(\theta_{dp} = a \mid D)} = \frac{\exp\left(\frac{-\epsilon_t}{\Delta}\|V(a; D)\|_K\right)}{(\gamma_t + \lambda_1)(\gamma_t + \lambda_2)\cdots(\gamma_t + \lambda_d)}\gamma_t^d.$$

■

Appendix C. Proofs and Technical Details

Since our propagation involves perturbation and rejection, the whole propagation transition has density

$$h_t(\theta, x) = \frac{\eta_t(\theta) f(x \mid \theta) m_{\epsilon_t}(s_{dp} \mid x)}{\int\int \eta_t(\theta) f(x \mid \theta) m_{\epsilon_t}(s_{dp} \mid x)\ dx\ d\theta}$$

with its denominator denoted as $c_{\eta_t}(s_{dp})$. We prove that the accepted particles after the whole propagation are drawn from h_t in Proposition 19. In other words, the accepted particles are derived from conditioning on the successful events after perturbation. That is,

$$h_t(\cdot, x) = \int \pi_{\epsilon_{t-1}}(\tilde{\theta}) K_t(\tilde{\theta}, \cdot \mid A_t) d\tilde{\theta},$$

where $A_t = \{(\theta, x) : \pi_0(\theta) > 0, r_t(x) > U\}$ is the event of success. The importance weight becomes

$$w_t = \frac{\pi_{\epsilon_t}(\theta, x \mid s_{dp})}{h_t(\theta, x)},$$

where

$$\pi_{\epsilon_t}(\theta, x \mid s_{dp}) = \frac{\pi_0(\theta) f(x \mid \theta) m_{\epsilon_t}(s_{dp} \mid x)}{\int\int \pi_0(\theta) f(x \mid \theta) m_{\epsilon_t}(s_{dp} \mid x)\ dx\ d\theta}$$

with its denominator denoted as $c_{\epsilon_t}(s_{dp})$. Often, both of the normalizing constants $c_{\eta_t}(s_{dp})$ and $c_{\epsilon_t}(s_{dp})$ are intractable. Nevertheless, this can be addressed by using a consistent estimate of the fraction of the two normalizing constants (Lemma 21).

C.1 Properties of DP-PF (Algorithm 3)

In this section, we present two useful properties that are a direct consequence of the way Algorithm 3 is constructed.

Proposition 19 *The accepted samples $(\theta, x)^{(i,t)}$ in algorithm 3 follow distribution $h_t(\theta, x)$. Furthermore, let $Y^{(i,t)}$ be the trial times ($\geqslant 1$) until the first acceptance. Then,*

- *$Y^{(i,t)}$ is independent of the accepted samples $(\theta, x)^{(i,t)}$, and*
- *$Y^{(i,t)}$ follows Geom $\left(p = \frac{c_{\eta_t}(s_{dp})}{\sup_x m_{\epsilon_t}(s_{dp}|x)}\right)$.*

Proof Define $I_t = 1$, if $(\theta, x) \in A_t$; $I_t = 0$, otherwise. Since

$$P(I_t = 1 \mid \Theta_t = \theta, X = x) = \frac{m_{\epsilon_t}(s_{dp} \mid x)}{\sup_x m_{\epsilon_t}(s_{dp} \mid x)},$$

the probability of acceptance is

$$\begin{aligned} P(I_t = 1) &= \int\int P(I_t = 1 \mid \Theta_t = \theta, X = x)\ \eta(\theta) f(x \mid \theta)\ d\theta dx \\ &= \frac{1}{\sup_x m_{\epsilon_t}(s_{dp} \mid x)} \int\int \eta(\theta) f(x \mid \theta) m_{\epsilon_t}(s_{dp} \mid x)\ d\theta dx \\ &= \frac{c_{\eta_t}(s_{dp})}{\sup_x m_{\epsilon_t}(s_{dp} \mid x)}. \end{aligned}$$

The distribution of the accepted samples is

$$\begin{aligned} P(\Theta_t = \theta, X = x \mid I_t = 1) &= \frac{P(I_t = 1 \mid \Theta_t = \theta, X = x) P(\Theta_t = \theta, X = x)}{P(I_t = 1)} \\ &= \frac{\eta_t(\theta) f(x \mid \theta) m_{\epsilon_t}(s_{dp} \mid x)}{c_{\eta_t}(s_{dp})} \\ &= h_t(\theta, x). \end{aligned}$$

Thus, the density of the accepted samples is $h_t(\theta, x)$. Now, consider

$$\begin{aligned} & P(Y = s, \Theta_t = \theta, X = x) \\ =& \left(1 - \frac{c_{\eta_t}(s_{dp})}{\sup_x m_{\epsilon_t}(s_{dp} \mid x)}\right)^{s-1} \eta_t(\theta) f(x \mid \theta) \frac{m_{\epsilon_t}(s_{dp} \mid x)}{\sup_x m_{\epsilon_t}(s_{dp} \mid x)} \\ =& \left[\left(1 - \frac{c_{\eta_t}(s_{dp})}{\sup_x m_{\epsilon_t}(s_{dp} \mid x)}\right)^{s-1} \frac{c_{\eta_t}(s_{dp})}{\sup_x m_{\epsilon_t}(s_{dp} \mid x)}\right] \cdot \frac{\eta_t(\theta) f(x \mid \theta) m_{\epsilon_t}(s_{dp} \mid x)}{c_{\eta_t}(s_{dp})} \\ =& P(G = s) \cdot h_t(\theta, x). \end{aligned}$$

The conclusion follows. ■

Theorem 16 (Marginal likelihood and its consistent estimation) *Let $\{\epsilon_t\}_{t=0}^T$ be the schedule of Algorithm 3 with $\epsilon_0 = 0$. For $t = 1, \ldots, T$, $\mathbb{E}_{h_t}\big[\, w_t m_{\epsilon_t}(s_{dp} \mid x) / m_{\epsilon_{t-1}}(s_{dp} \mid x) \,\big] < \infty$. Then,*

1. *Each incremental evidence ratio admits the representation*

$$\rho_t \;=\; \mathbb{E}_{\pi_{\epsilon_{t-1}}(\theta, x \mid s_{dp})}\left[\frac{m_{\epsilon_t}(s_{dp} \mid x)}{m_{\epsilon_{t-1}}(s_{dp} \mid x)}\right],$$

   so that $c_{\epsilon_T}(s_{dp}) = \prod_{t=1}^T \rho_t$.

2. *Each ρ_t is consistently estimated, as $N \to \infty$, by*

$$\hat{\rho}_t \;=\; \sum_{i=1}^N \mathbf{w}^{(i,t-1)}\, \frac{m_{\epsilon_t}(s_{dp} \mid x^{(i,t-1)})}{m_{\epsilon_{t-1}}(s_{dp} \mid x^{(i,t-1)})},$$

   where $\mathbf{w}^{(i,t)}$ and $x^{(i,t)}$ are the normalized weights and accepted synthetic data set produced at iteration t of Algorithm 3. Consequently, $\widehat{c_{\epsilon_T}(s_{dp})} = \prod_{t=1}^T \hat{\rho}_t$ is a consistent estimator of $c_{\epsilon_T}(s_{dp})$.

Proof For the first result, observe that for any $t \geqslant 1$,

$$c_{\epsilon_t}(s_{dp}) = \int \kappa_t(\theta, x)\, d\theta dx = \int \kappa_{t-1}(\theta, x) \frac{\kappa_t(\theta, x)}{\kappa_{t-1}(\theta, x)}\, d\theta dx.$$

Dividing both sides by $c_{\epsilon_{t-1}}(s_{dp})$ yields

$$\begin{aligned} \rho_t \;=\; \frac{c_{\epsilon_t}(s_{dp})}{c_{\epsilon_{t-1}}(s_{dp})} &= \int \frac{\kappa_{t-1}(\theta, x)}{c_{\epsilon_{t-1}}(s_{dp})} \frac{\kappa_t(\theta, x)}{\kappa_{t-1}(\theta, x)}\, d\theta dx \\ &= \mathbb{E}_{\pi_{\epsilon_{t-1}}(\theta, x | s_{dp})} \left[\frac{\kappa_t(\theta, x)}{\kappa_{t-1}(\theta, x)} \right] \\ &= \mathbb{E}_{\pi_{\epsilon_{t-1}}(\theta, x | s_{dp})} \left[\frac{m_{\epsilon_t}(s_{dp} \mid x)}{m_{\epsilon_{t-1}}(s_{dp} \mid x)} \right], \end{aligned}$$

where the first identity follows from $\kappa_t(\theta, x) = \pi_{\epsilon_t}(\theta, x \mid s_{dp}) c_{\epsilon_t}(s_{dp})$, and the second identity

$$\frac{\kappa_t(\theta, x)}{\kappa_{t-1}(\theta, x)} = \frac{\pi_0(\theta) f(x \mid \theta) m_{\epsilon_t}(s_{dp} \mid x)}{\pi_0(\theta) f(x \mid \theta) m_{\epsilon_{t-1}}(s_{dp} \mid x)} = \frac{m_{\epsilon_t}(s_{dp} \mid x)}{m_{\epsilon_{t-1}}(s_{dp} \mid x)}$$

by the design of Algorithm 3. The factorization $c_{\epsilon_T}(s_{dp}) = \prod_{t=1}^{T} \rho_t$ follows by telescoping.

For the second result, define $\varphi_t(\theta, x) = m_{\epsilon_t}(s_{dp} \mid x) / m_{\epsilon_{t-1}}(s_{dp} \mid x)$. By the integrability assumption of the theorem, $\mathbb{E}_{h_t} |w_t \varphi_t| < \infty$, so Theorem 8 applies with $\varphi = \varphi_t$, giving

$$\hat{\rho}_t \;=\; \sum_{i=1}^{N} \mathbf{w}^{(i,t-1)}\, \varphi_t\Big(\theta^{(i,t-1)}, x^{(i,t-1)}\Big) \;\xrightarrow{a.s.}\; \mathbb{E}_{\pi_{\epsilon_{t-1}}(\theta, x | s_{dp})}[\varphi_t(\theta, x)] \;=\; \rho_t$$

as $N \to \infty$. Since $\widehat{c_{\epsilon_T}(s_{dp})} = \prod_{t=1}^{T} \hat{\rho}_t$ is a continuous function of $(\hat{\rho}_1, \ldots, \hat{\rho}_T)$, the continuous mapping theorem gives $\widehat{c_{\epsilon_T}(s_{dp})} \xrightarrow{a.s.} c_{\epsilon_T}(s_{dp})$. ■

C.2 On the Consistency and CLT of the DP-PF Estimator

In order to carefully define the measurability across iterations for the consistency and CLT proof. We use the following definition:

Definition 20 (Recursive Measurability Chopin, 2004) *$\mathcal{F}_t{}^p(d)$ is defined to be the set of measurable function $\varphi : \Theta_t \times \mathcal{X}^n \to \mathbb{R}^d$ such that for some p,*

$$\mathbb{E}_{h_t} |w_t \varphi|^p < \infty,$$

in which the function $\theta_{t-1} \mapsto \mathbb{E}_{K_t(\theta_{t-1}, \cdot | A_t)}[w_t(\cdot)\ \varphi(\cdot)]$ lies in $\mathcal{F}_{t-1}{}^p(1)$. For convenience, if $d = 1$, the (1) will be omitted, and if $p = 2^{\dagger}$, then there exists some $\zeta > 0$, such that $\varphi \in \mathcal{F}_t{}^{2+\zeta}(d)$.

From section 3, we know that after the whole propagation transition (*Step 2. & 2.5.*) in each t, we are able to sample from the proposal distribution $h_t(\theta, x)$. Thus, we use importance sampling to derive the importance weights $w_t = \frac{\pi_{\epsilon_t}(\theta, x | s_{dp})}{h_t(\theta, x)}$, such that

$$\mathbb{E}_t[\varphi] = \mathbb{E}_{h_t(\theta, x)}[w_t \varphi]. \tag{3}$$

While w_t satisfies (3), the normalizing constants of the two distributions are often intractable. Alternatively, we tackle

$$\tilde{w}_t = \frac{\pi_0(\theta)}{\eta_t(\theta)},$$

such that $\tilde{w}_t = w_t \gamma_t$ with $\gamma_t := \left(\frac{c_{\eta_t}(s_{dp})}{c_{\epsilon_t}(s_{dp})}\right)^{-1}$ being constant at each iteration.

In exchange for not being able to quantify the normalizing constants, we need the following lemma to set up for consistency.

Lemma 21 *Let $\varphi : \Theta_t \times \mathcal{X}^n \to \mathbb{R}^d$, such that $\mathbb{E}_{h_t}|w_t \varphi_j| < \infty$ for all $j = 1, \ldots, d$,*

$$\frac{1}{N} \sum_{i=1}^{N} \tilde{w}^{(i,t)} \varphi\left((\theta, x)^{(i,t)}\right) \overset{a.s.}{\to} \gamma_t \mathbb{E}_t[\varphi] \tag{4}$$

as $N \to \infty$. In the case $\varphi \equiv 1$, we obtain $\frac{1}{N} \sum_{i=1}^{N} \tilde{w}^{(i,t)} \overset{a.s.}{\to} \gamma_t$.

Proof The argument can be applied coordinate-wise, so we prove the result for $d = 1$. For any φ satisfying $\mathbb{E}_{h_t}|w_t \varphi| < \infty$,

$$\mathbb{E}_{h_t(\theta,x)}[\tilde{w}_t \varphi] = \gamma_t \mathbb{E}_t[\varphi].$$

Now, let $Y_i = \tilde{w}^{(i,t)} \varphi\left((\theta, x)^{(i,t)}\right)$ for all $i = 1, .., N$. Since $Y_1, \ldots, Y_N$ are i.i.d., the conclusion follows by the strong law of large numbers (See Section 2.4 in Durrett (2019)). ■

We use Lemma 21 to prove the strong consistency of our estimator $\mathbb{E}_t[\varphi]$.

Theorem 8 (Strong Consistency) *Let $\varphi : \Theta_t \times \mathcal{X}^n \to \mathbb{R}^d$, such that $\mathbb{E}_{h_t}|w_t \varphi_j| < \infty$ for all $j = 1, \ldots, d$. Our estimator, defined in (2), converges almost surely:*

$$\hat{\mathbb{E}}_t[\varphi] = \sum_{i=1}^{N} \mathbf{w}^{(i,t)} \varphi\left((\theta, x)^{(i,t)}\right) \overset{a.s.}{\to} \mathbb{E}_t[\varphi]$$

as $N \to \infty$. In particular, for $t = T$, $\hat{\mathbb{E}}_T[\varphi] \overset{a.s.}{\to} \mathbb{E}_T[\varphi]$.

Proof The argument can be applied coordinate-wise, so we prove the results for $d = 1$. Let $f(x) = \frac{1}{x}$, which is continuous on $x > 0$. Since $\frac{1}{N} \sum_{i=1}^{N} \tilde{w}^{(i,t)} \overset{a.s.}{\to} \gamma_t$,

$$f\left(\frac{1}{N} \sum_{i=1}^{N} \tilde{w}^{(i,t)}\right) = \frac{1}{\frac{1}{N} \sum_{i=1}^{N} \tilde{w}^{(i,t)}} \overset{a.s.}{\to} \frac{1}{\gamma_t}.$$

by the continuous mapping theorem (Theorem 3.2.10 in Durrett (2019)). With (4), the conclusion follows by the continuous mapping theorem for a bivariate function. ■

Another crucial result is the central limit theorem for our estimator $\mathbb{E}_t[\varphi]$. It requires a higher-order integrability assumption.

Theorem 11 (Central Limit Theorem) *Let $\varphi : \Theta_t \times \mathcal{X}^n \to \mathbb{R}^d$ satisfy the recursive measurability and moment conditions stated in Appendix C, and suppose there exists $\zeta > 0$ such that $\mathbb{E}_{h_t}|w_t\varphi_j|^{2+\zeta} < \infty$ for all $j = 1, \ldots, d$. Then,*

$$\sqrt{N}\left(\hat{\mathbb{E}}_t[\varphi] - \mathbb{E}_t[\varphi]\right) \xrightarrow{d} N(0, V_t(\varphi))$$

as $N \to \infty$.

Proof In addition to recursive measurability, we also assume the unit function $\theta_t \mapsto 1$ belongs to ${\mathcal{F}_T}^{2^\dagger}$. This, for example, ensures $\frac{1}{N}\sum_{i=1}^N \tilde{w}^{(i,t)} \xrightarrow{a.s.} \gamma_t$ and sets the ground for CLT.

Define the following quantities for the asymptotic variances in the CLT. Let $\varphi : \Theta_0 \times \mathcal{X}^n \to \mathbb{R}^d$, and $\widetilde{V}_0(\varphi) = \mathrm{Var}_{\pi_0}(\varphi)$, which is a typical asymptotic variance of a CLT where the sample mean converges to its population mean of π_0. Then, by induction, $\forall \varphi : \Theta_t \times \mathcal{X}^n \to \mathbb{R}^d$, let

$$\begin{aligned}
\breve{V}_t(\varphi) &:= \widetilde{V}_{t-1}[\mathbb{E}_{K_t(\cdot,\cdot|A_t)}(\varphi)] + \mathbb{E}_{\pi_{\epsilon_{t-1}}}[\mathrm{Var}_{K_t(\cdot,\cdot|A_t)}(\varphi)], && t > 0,\ (\textit{propagation}) \\
V_t(\varphi) &:= \breve{V}_t\left[w_t\left(\varphi - \mathbb{E}_{\pi_{\epsilon_t}}(\varphi)\right)\right], && t > 0,\ (\textit{reweighting}) \\
\widetilde{V}_t(\varphi) &:= V_t(\varphi) + U_t(\varphi), && t > 0.\ (\textit{resampling})
\end{aligned} \tag{5}$$

As stated by Chopin (2004), the CLT proof of Theorem 11 is done by mathematical induction with three lemmas, each of which states a CLT for each step of the particle filtering. Essentially, the lemmas quantify the randomness incurred from doing resampling, propagating, and re-weighting. The inductive hypothesis is as follows, for any given $t > 0$, we assume that $\forall \varphi \in {\mathcal{F}_{t-1}}^{2^\dagger}(d)$,

$$\sqrt{N}\left[\frac{1}{N}\sum_{j=1}^N \varphi\left((\tilde{\theta}, x)^{(j,t-1)}\right) - \mathbb{E}_{\pi_{\epsilon_{t-1}}}(\varphi)\right] \xrightarrow{d} N(0, \widetilde{V}_{t-1}(\varphi)). \tag{6}$$

Lemma 22 (propagation: Chopin, 2004) *Let $\psi : \Theta_t \to \mathbb{R}^d$, such that the function $\nu : \theta_{t-1} \mapsto \mathbb{E}_{K_t(\theta_{t-1},\cdot|A_t)}[\psi(\cdot) - \mathbb{E}_{h_t}(\psi)]$ lies in ${\mathcal{F}_{t-1}}^{2^\dagger}$ and there exists $\zeta > 0$ such that $\mathbb{E}_{h_t}|\psi|^{2+\zeta} < \infty$, then*

$$\sqrt{N}\left[\frac{1}{N}\sum_{j=1}^N \psi((\theta, x)^{(j,t)}) - \mathbb{E}_{h_t}(\psi)\right] \xrightarrow{d} N(0, \breve{V}_{t-1}(\psi)).$$

under the inductive hypothesis (6)*.*

Lemma 22 quantifies the MCSE from propagation. The asymptotic variance accounts for randomness from the previous target posterior $\pi_{\epsilon_{t-1}}$ and the propagation process. To adapt the proof to our DP-PF algorithm, we redefine the function ν slightly differently from its original proof, where ν represents a $(2+\delta)$-th central moment of ψ with respect to its perturbation kernel. In our DP setting, after perturbation, particles are selected based on the success event $A_t = \{\pi_0(\cdot) > 0, r_t(x) > U\}$. Consequently, the propagation distribution becomes $K_t(\theta_{t-1}, \cdot \mid A_t)$, incorporating both the kernel perturbation and particle selection to target the posteriors.

Lemma 23 (reweighting: Chopin, 2004) *Let $\varphi \in {\mathcal{F}_t}^{2^\dagger}(d)$ and the function $\theta_t \mapsto 1$ lies in ${\mathcal{F}_t}^{2^\dagger}$, then*

$$\sqrt{N}\left(\hat{\mathbb{E}}_t[\varphi] - \mathbb{E}_t[\varphi]\right) \xrightarrow{d} N(0, V_t(\varphi)),$$

under the inductive hypothesis (6)*.*

Lemma 23 quantifies the MCSE from reweighting. From (5), we can see that the asymptotic variance of reweighting is a rescale from that of propagation. At $t = T$, Algorithm 3 stops at this step.

Lemma 24 (resampling: Chopin, 2004) *Let $\varphi \in {\mathcal{F}_t}^{2^\dagger}(d)$ and the function $\theta_t \mapsto 1$ lies in ${\mathcal{F}_t}^{2^\dagger}$, then*

$$\sqrt{N}\left[\frac{1}{N}\sum_{j=1}^{N}\varphi\left((\tilde{\theta}, x)^{(j,t)}\right) - \mathbb{E}_{\pi_{\epsilon_t}}(\varphi)\right] \xrightarrow{d} N(0, \widetilde{V}_t(\varphi)),$$

under the inductive hypothesis (6)*.*

Lemma 24 quantifies the MCSE from multinomial resampling. The asymptotic variance is the sum of (5) and the variance of φ with respect to the posterior π_{ϵ_t}. This also indicates that, at $t = T$, the algorithm should stop at the reweighting step to avoid additional randomness from resampling. Nevertheless, we need Lemma 23 to complete the mathematical induction for t from the inductive hypothesis at $t - 1$. ■

Lemma 13 (ESS estimate) *Let $w_i = \frac{\pi(X_i)}{h(X_i)}$ and $\widehat{ESS}_N = \frac{(\sum_{i=1}^N w_i)^2}{\sum_{i=1}^N w_i^2}$. Assume $(X_i)_{i\geqslant 1}$ has marginal distribution h and satisfies $\frac{1}{N}\sum_{i=1}^N w_i \xrightarrow{p} \mathbb{E}_h[w(X)]$ and $\frac{1}{N}\sum_{i=1}^N w_i^2 \xrightarrow{p} \mathbb{E}_h[w(X)^2]$, with $\mathbb{E}_h[w(X)^2] < \infty$. Then, as $N \to \infty$,*

$$\widehat{ESS}_N/N \xrightarrow{p} \frac{1}{1 + Var_h(w(X))}.$$

Proof By the moment assumptions, we have the following two weak convergence results: $\frac{1}{N}\sum_{i=1}^N w_i \xrightarrow{p} \mathbb{E}_h(w(X)) = 1$ and $\frac{1}{N}\sum_{i=1}^N w_i^2 \xrightarrow{p} \mathbb{E}_h(w^2(X)) = 1 + \mathrm{Var}_h(w(X))$. Then, by the continuous mapping theorem,

$$\widehat{ESS}_N/N = \frac{\left(\frac{1}{N}\sum_{i=1}^N w_i\right)^2}{\frac{1}{N}\sum_{i=1}^N w_i^2} \xrightarrow{p} \frac{1}{1 + \mathrm{Var}_h(w(X))}.$$

■

The ESS estimate is set up for the proof of Theorem 14.

Theorem 14 (Asymptotic Confidence Interval) *Let $\varphi : \Theta_t \times \mathcal{X}^n \to \mathbb{R}$ satisfy the recursive measurability and moment conditions stated in Appendix C, and suppose there exists $\zeta > 0$ such that $\mathbb{E}_{h_t}|w_t\varphi|^{2+\zeta} < \infty$. Then,*

1. *a consistent estimate of $V_t(\varphi)$ is*

$$\hat{V}_t(\varphi) := \frac{\sum_{i=1}^N \mathbf{w}^{(i,t)} \left[\varphi\left((\theta,x)^{(i,t)}\right) - \hat{\mathbb{E}}_t[\varphi]\right]^2}{\widehat{ESS_N}/N} + \sum_{i=1}^N \mathbf{w}^{(i,t)} \left[\mathbf{w}^{(i,t)} - \frac{1}{N}\right] \left[\varphi\left((\theta,x)^{(i,t)}\right) - \hat{\mathbb{E}}_t[\varphi]\right]^2;$$

2. *an asymptotic $(1-\alpha)$ confidence interval for $\mathbb{E}_t[\varphi]$ is*

$$\hat{\mathbb{E}}_t[\varphi] \pm z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{V}_t(\varphi)}{N}},$$

where $z_{1-\frac{\alpha}{2}}$ is the $\left(1-\frac{\alpha}{2}\right)$ quantile of a standard normal distribution.

Proof Let σ_t^2 be the variance of φ from the i.i.d. samples $(\theta,x)^{(i,t)}$ drawn from the intermediate private posterior π_{ϵ_t} and hence $\mathrm{Var}_{\pi_{\epsilon_t}}\left[\frac{1}{N}\sum_{i=1}^N \varphi(\theta,x)^{(i,t)}\right] = \frac{\sigma_t^2}{N}$. Then, consider

$$\sqrt{N}\left(\begin{bmatrix} \frac{1}{N}\sum_{i=1}^N \tilde{w}^{(i,t)}\varphi\left((\theta,x)^{(i,t)}\right) \\ \frac{1}{N}\sum_{i=1}^N \tilde{w}^{(i,t)} \end{bmatrix} - \begin{bmatrix} \gamma_t \mathbb{E}_t[\varphi] \\ \gamma_t \end{bmatrix}\right) \xrightarrow{d} \mathcal{N}\left(0_2, \Sigma_t = \begin{pmatrix} \mathrm{Var}_{h_t}(\tilde{w}_t\varphi) & \mathrm{Cov}_{h_t}(\tilde{w}_t\varphi, \tilde{w}_t) \\ \mathrm{Cov}_{h_t}(\tilde{w}_t\varphi, \tilde{w}_t) & \mathrm{Var}_{h_t}(\tilde{w}_t) \end{pmatrix}\right).$$

By first-order Delta method for $f_1(x,y) = \frac{x}{y}$ with gradient $\nabla f_1(x,y) = (\frac{1}{y}, \frac{-x}{y^2})$, we have

$$\sqrt{N}\left(\hat{\mathbb{E}}_t[\varphi] - \mathbb{E}_t[\varphi]\right) \xrightarrow{d} \mathcal{N}\left(0, \left(\frac{1}{\gamma_t}, \frac{-1}{\gamma_t}\mathbb{E}_t[\varphi]\right) \Sigma_t \left(\frac{1}{\gamma_t}, \frac{-1}{\gamma_t}\mathbb{E}_t[\varphi]\right)^\top\right),$$

and

$$
\begin{aligned}
V_t &= \left(\frac{1}{\gamma_t}, \frac{-1}{\gamma_t}\mathbb{E}_t[\varphi]\right) \Sigma_t \left(\frac{1}{\gamma_t}, \frac{-1}{\gamma_t}\mathbb{E}_t[\varphi]\right)^\top \\
&= \mathrm{Var}_{h_t}(w_t\varphi) - 2\mathbb{E}_t[\varphi]\mathrm{Cov}_{h_t}(w_t\varphi, w_t) + (\mathbb{E}_t[\varphi])^2 \mathrm{Var}_{h_t}(w_t) \\
&= \left[\mathbb{E}_{h_t}\left(w_t^2\varphi^2\right) - (\mathbb{E}_{h_t}(w_t\varphi))^2\right] - 2\mathbb{E}_t[\varphi]\left[\mathbb{E}_{h_t}\left(w_t^2\varphi\right) - \mathbb{E}_{h_t}\left(w_t\varphi\right)\mathbb{E}_{h_t}\left(w_t\right)\right] + (\mathbb{E}_t[\varphi])^2 \mathrm{Var}_{h_t}(w_t) \\
&= \left[\mathbb{E}_t\left(w_t\varphi^2\right) - (\mathbb{E}_t[\varphi])^2\right] - 2\mathbb{E}_t[\varphi]\left[\mathrm{Cov}_t(w_t, \varphi) + \mathbb{E}_t(w_t)\mathbb{E}_t[\varphi] - \mathbb{E}_t[\varphi]\right] + (\mathbb{E}_t[\varphi])^2 \mathrm{Var}_{h_t}(w_t) \qquad (7) \\
&= \left[\left\{\mathbb{E}_t(w_t)\left(\mathbb{E}_t[\varphi]\right)^2 + 2\mathrm{Cov}_t(w_t, \varphi)\mathbb{E}_t[\varphi] + \mathrm{Var}_t(\varphi)\mathbb{E}_t(w_t) + r\right\} - (\mathbb{E}_t[\varphi])^2\right] \\
&\quad - 2\mathbb{E}_t[\varphi]\left[\mathrm{Cov}_t(w_t, \varphi) + \mathbb{E}_t(w_t)\mathbb{E}_t[\varphi] - \mathbb{E}_t[\varphi]\right] + (\mathbb{E}_t[\varphi])^2 \mathrm{Var}_{h_t}(w_t) \qquad (8) \\
&= (\mathbb{E}_t[\varphi])^2\left[1 + \mathrm{Var}_{h_t}(w_t) - \mathbb{E}_t(w_t)\right] + \mathbb{E}_t(w_t)\mathrm{Var}_t(\varphi) + r \\
&= \left(1 + \mathrm{Var}_{h_t}(w_t)\right)\mathrm{Var}_t(\varphi) + r \qquad (9) \\
&= \frac{N}{ESS_N}\sigma_t^2 + r,
\end{aligned}
$$

where $\mathbb{E}_t(\cdot)$ and $\mathrm{Cov}_t(\cdot,\cdot)$ are integral with respect to $\pi_{\epsilon_t}(\theta, x \mid s_{dp})$.

For (7), we apply the fact $w_t = \frac{\pi_{\epsilon_t}}{h_t}$. For (8), we consider the second-order Taylor's approximation for $f_2(x,y) = xy^2$ with Hessian $H_{f_2}(x,y) = \begin{pmatrix} 0 & 2y \\ 2y & 2x \end{pmatrix}$ such that

$$\mathbb{E}(f_2(T)) = \mathbb{E}(f_2(\mathbb{E}T)) + \frac{1}{2}\mathrm{tr}(\mathrm{Cov}(T)H_{f_2}(\mathbb{E}T)) + r,$$

where $r = \mathbb{E}_t\left[\left(w_t - \mathbb{E}_t(w_t)\right)\left(\varphi - \mathbb{E}_t[\varphi]\right)^2\right]$ is the remainder for the non-zero third order term. For (9), we note that fact that $\mathbb{E}_t(w_t^2) = 1 + \mathrm{Var}_{h_t}(w_t)$.

Since $V_t = \frac{\sigma_t^2}{ESS_N/N} + r$ and $r = \mathbb{E}_t\left[\left(w_t - \mathbb{E}_t(w_t)\right)\left(\varphi - \mathbb{E}_t[\varphi]\right)^2\right]$, a consistent estimate $\hat{V}_t$ is

$$N\frac{\sum_{i=1}^N \mathbf{w}^{(i,t)}\left[\varphi\left((\theta,x)^{(i,t)}\right) - \hat{\mathbb{E}}_t[\varphi]\right]^2}{\widehat{ESS_N}} + \sum_{i=1}^N \mathbf{w}^{(i,t)}\left[\mathbf{w}^{(i,t)} - \frac{1}{N}\right]\left[\varphi\left((\theta,x)^{(i,t)}\right) - \hat{\mathbb{E}}_t[\varphi]\right]^2.$$

By Slutsky's theorem, we obtain the following CLT:

$$\sqrt{N}\left(\frac{\hat{\mathbb{E}}_t[\varphi] - \mathbb{E}_t[\varphi]}{\sqrt{\hat{V}_t}}\right) \xrightarrow{d} \mathcal{N}(0,1),$$

from which we can build an asymptotic $(1-\alpha)$ confidence interval for $\mathbb{E}_t[\varphi]$ as

$$\hat{\mathbb{E}}_t[\varphi] \pm z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{V}_t(\varphi)}{N}},$$

where $z_{1-\frac{\alpha}{2}}$ is the $\left(1-\frac{\alpha}{2}\right)$ quantile of a standard normal distribution. ■

References

Martin Abadi, Andy Chu, Ian Goodfellow, H Brendan McMahan, Ilya Mironov, Kunal Talwar, and Li Zhang. Deep learning with differential privacy. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, pages 308–318, 2016.

John M. Abowd, Robert Ashmead, Ryan Cumings-Menon, Simson Garfinkel, Micah Heineck, Christine Heiss, Robert Johns, Daniel Kifer, Philip Leclerc, Ashwin Machanavajjhala, Brett Moran, William Sexton, Matthew Spence, and Pavel Zhuravlev. The 2020 Census Disclosure Avoidance System TopDown Algorithm. *Harvard Data Science Review*, jun 24 2022.

Sergios Agapiou, Omiros Papaspiliopoulos, Daniel Sanz-Alonso, and Andrew M. Stuart. Importance sampling: Intrinsic dimension and computational cost. *Statistical Science*, 32(3):405–431, 2017.

Jordan Awan and Aleksandra Slavković. Differentially private uniformly most powerful tests for binomial data. *Advances in Neural Information Processing Systems*, 31, 2018.

Jordan Awan and Aleksandra Slavković. Structure and sensitivity in differential privacy: Comparing k-norm mechanisms. *Journal of the American Statistical Association*, 116 (534):935–954, April 2021. ISSN 0162-1459.

Jordan Awan and Zhanyu Wang. Simulation-based, finite-sample inference for privatized data. *Journal of the American Statistical Association*, 120(551):1669–1682, 2025.

Jordan Awan, Xi Chen, and Roberto Molinari. Large-sample Bayesian approximations for privatized data. *arXiv preprint arXiv:2604.24817*, 2026.

Jordan Alexander Awan and Aleksandra Slavkovic. Differentially private inference for binomial data. *Journal of Privacy and Confidentiality*, 10(1), 2020.

Mark A. Beaumont, Jean-Marie Cornuet, Jean-Michel Marin, and Christian P. Robert. Adaptive approximate Bayesian computation. *Biometrika*, 96(4):983 – 990, December 2009.

James O. Berger. *Statistical Decision Theory and Bayesian Analysis*. Springer series in statistics. Springer, New York, N.Y, 2nd ed. edition, 1985. ISBN 9781475742862.

James O Berger, Elías Moreno, Luis Raul Pericchi, M Jesús Bayarri, José M Bernardo, Juan A Cano, Julián De la Horra, Jacinto Martín, David Ríos-Insúa, Bruno Betrò, et al. An Overview of Robust Bayesian Analysis. *TEST*, 3(1):5–124, 1994.

Garrett Bernstein and Daniel R Sheldon. Differentially private Bayesian inference for exponential families. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018.

Garrett Bernstein and Daniel R Sheldon. Differentially private Bayesian linear regression. *Advances in Neural Information Processing Systems*, 32, 2019.

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

Tony Cai, Yichen Wang, and Linjun Zhang. The cost of privacy: Optimal rates of convergence for parameter estimation with differential privacy. *The Annals of Statistics*, 49(5): 2825–2850, 2021.

Statistics Canada. 2021 census of population, 2022.

James Carpenter, Peter Cliffordy, and Paul Fearnhead. An improved particle filter for non-linear problems. *IEE Proc. Radar, Sonar Navig.*, 146, 09 2000.

Kamalika Chaudhuri, Claire Monteleoni, and Anand D Sarwate. Differentially private empirical risk minimization. *Journal of Machine Learning Research*, 12(3), 2011.

Nicolas Chopin. Central limit theorem for sequential Monte Carlo methods and its application to Bayesian inference. *The Annals of Statistics*, 32(6):2385 – 2411, 2004.

Nicolas Chopin and Omiros Papaspiliopoulos. *An Introduction to Sequential Monte Carlo*. Springer Series in Statistics. Springer International Publishing, 2020. ISBN 978-3-030-47845-2.

Pierre Del Moral, Arnaud Doucet, and Ajay Jasra. Sequential Monte Carlo samplers. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 68(3):411–436, 2006.

Bolin Ding, Janardhan Kulkarni, and Sergey Yekhanin. Collecting telemetry data privately. *Advances in Neural Information Processing Systems*, 30, 2017.

Jinshuo Dong, Aaron Roth, and Weijie J. Su. Gaussian differential privacy. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 84(1):3–37, February 2022. ISSN 1369-7412.

Richard Durrett. *Probability: Theory and Examples*. Cambridge University Press, 2019.

Cynthia Dwork. Differential privacy. In *International Colloquium on Automata, Languages, and Programming*, pages 1–12. Springer, 2006.

Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. *Theory of Cryptography*, Vol. 3876:265–284, 01 2006.

Cynthia Dwork, Aaron Roth, et al. The algorithmic foundations of differential privacy. *Foundations and Trends in Theoretical Computer Science*, 9(3–4):211–407, 2014.

Úlfar Erlingsson, Vasyl Pihur, and Aleksandra Korolova. RAPPOR: Randomized aggregatable privacy-preserving ordinal response. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*, pages 1054–1067, 2014.

Cecilia Ferrando, Shufan Wang, and Daniel Sheldon. Parametric bootstrap for differentially private confidence intervals. In *International Conference on Artificial Intelligence and Statistics*, pages 1598–1618. PMLR, 2022.

Sarah Filippi, Chris P Barnes, Julien Cornebise, and Michael PH Stumpf. On optimality of kernels for approximate Bayesian computation using sequential Monte Carlo. *Statistical Applications in Genetics and Molecular Biology*, 12(1):87–107, 2013.

Marco Gaboardi and Ryan Rogers. Local private hypothesis testing: Chi-square tests. *arXiv preprint arXiv:1709.07155*, 2017.

Marco Gaboardi, Hyun Lim, Ryan Rogers, and Salil Vadhan. Differentially private chi-squared hypothesis testing: Goodness of fit and independence testing. In *International Conference on Machine Learning*, pages 2111–2120. PMLR, 2016.

John Geweke. Bayesian inference in econometric models using Monte Carlo integration. *Econometrica*, 57(6):1317–1339, 1989. ISSN 00129682, 14680262.

Ruobin Gong. Exact inference with approximate computation for differentially private data via perturbations. *Journal of Privacy and Confidentiality*, 12(2), 2022.

Moritz Hardt and Kunal Talwar. On the geometry of differential privacy. In *Proceedings of the Forty-Second ACM Symposium on Theory of Computing*, pages 705–714, 2010.

Nianqiao Ju, Jordan Awan, Ruobin Gong, and Vinayak Rao. Data augmentation MCMC for Bayesian inference from privatized data. *Advances in Neural Information Processing Systems*, 35:12732–12743, 2022.

Vishesh Karwa, Pavel N Krivitsky, and Aleksandra B Slavković. Sharing social network data: Differentially private estimation of exponential family random-graph models. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 66(3):481–500, 2017.

Augustine Kong. A note on importance sampling using standardized weights. *Technical Report No.348*, 1992.

Tejas Kulkarni, Joonas Jälkö, Antti Koskela, Samuel Kaski, and Antti Honkela. Differentially private Bayesian inference for generalized linear models. In *International Conference on Machine Learning*, pages 5838–5849. PMLR, 2021.

Jun S. Liu and Rong Chen. Sequential Monte Carlo methods for dynamic systems. *Journal of the American Statistical Association*, 93(443):1032–1044, 1998.

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

Iain Murray, Zoubin Ghahramani, and David JC MacKay. MCMC for doubly-intractable distributions. In *Proceedings of the 22nd Annual Conference on Uncertainty in Artificial Intelligence (UAI-06)*, pages 359–366. AUAI Press, 2006.

Aaron Schein, Zhiwei Steven Wu, Alexandra Schofield, Mingyuan Zhou, and Hanna Wallach. Locally private Bayesian inference for count models. In *International Conference on Machine Learning*, pages 5638–5648. PMLR, 2019.

Adam Smith. Privacy-preserving statistical estimation with optimal convergence rates. In *Proceedings of the 43rd Annual ACM Symposium on Theory of Computing (STOC '11)*, pages 813–822, 06 2011.

Thomas Steinke and Jonathan Ullman. Between pure and approximate differential privacy. *Journal of Privacy and Confidentiality*, 7(2), 2016. ISSN 2575-8527. Number: 2.

Kunal Talwar, Abhradeep Guha Thakurta, and Li Zhang. Nearly optimal private lasso. *Advances in Neural Information Processing Systems*, 28, 2015.

Tina Toni, David Welch, Natalja Strelkowa, Andreas Ipsen, and Michael P. H. Stumpf. Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of the Royal Society, Interface*, 6(31):187–202, February 2009. ISSN 1742-5689.

Salil Vadhan. The complexity of differential privacy. *Tutorials on the Foundations of Cryptography: Dedicated to Oded Goldreich*, pages 347–450, 2017.

Yue Wang, Daniel Kifer, Jaewoo Lee, and Vishesh Karwa. Statistical approximating distributions under differential privacy. *Journal of Privacy and Confidentiality*, 8(1), 2018.

Zhanyu Wang, Arin Chang, and Jordan Awan. Optimal debiased inference on privatized data via indirect estimation and parametric bootstrap. *arXiv preprint arXiv:2507.10746*, 2025a.

Zhanyu Wang, Guang Cheng, and Jordan Awan. Differentially private bootstrap: New privacy analysis and inference strategies. *Journal of Machine Learning Research*, 26 (257):1–57, 2025b.