Journal of Privacy and Confidentiality
Vol. 12 (2) 2022

Submitted Dec 2021
Published Nov 2022

EXACT INFERENCE WITH APPROXIMATE COMPUTATION FOR DIFFERENTIALLY PRIVATE DATA VIA PERTURBATIONS

RUOBIN GONG

Department of Statistics, Rutgers University, New Brunswick, NJ 08854
e-mail address: rg915@stat.rutgers.edu

Abstract. This paper discusses how two classes of approximate computation algorithms can be adapted, in a modular fashion, to achieve exact statistical inference from differentially private data products. Considered are approximate Bayesian computation for Bayesian inference, and Monte Carlo Expectation-Maximization for likelihood inference. Up to Monte Carlo error, inference from these algorithms is exact with respect to the joint specification of both the analyst's original data model, and the curator's differential privacy mechanism. Highlighted is a duality between approximate computation on exact data, and exact computation on approximate data, which can be leveraged by a well-designed computational procedure for statistical inference.

1. Introduction

Differential privacy (Dwork et al., 2006) advances statistical disclosure limitation by putting forth a formal and practical framework. In addition to grounding the concept of privacy on a mathematical footing, differential privacy distinguishes itself from traditional approaches by offering transparent probabilistic mechanisms, whose specifications can be made public without sabotaging the privacy guarantee. Differential privacy has been adapted by major data curators in industry, research organizations and government. As a prime example, the U.S. Census Bureau deploys differential privacy to protect the 2020 Decennial Census data products (Abowd et al., 2022). The P.L. 94-171 redistricting data files were released on August 12, 2021 (U.S. Census Bureau, 2021).

In this work, we adopt the perspective of a data analyst operating under the *dissemination* mode of data access (Hotz et al., 2022). A data curator such as the Census Bureau collects potentially sensitive data and releases differentially private data products to the analyst. The analyst in turn conducts statistical inference for their quantities of interest based on the privatized data. The analyst's goal is to draw trustworthy inference from the statistical model they wish to fit, knowing that the data have undergone privacy protection. This may not be a trivial task. The curator instills differential privacy in the data product via a data processing mechanism. Naïvely treating processed data as if they are unprocessed may result in erroneous and misleading statistical inference. With the wide adoption of differential

Key words and phrases: approximate Bayesian computation (ABC); Expectation Maximization (EM); ignorability; Monte Carlo; privacy-efficiency tradeoff; statistical inference.

www.journalprivacyconfidentiality.org
DOI:10.29012/jpc.797

privacy for disclosure limitation, social scientists and policy researchers are faced with the challenge to revise their preferred statistical analyses to account for the privacy mechanism, however complex they may be. To keep up with advances in privacy protection, we need theoretically sound and computationally efficient statistical methodologies to supplant their predecessors (Hansen, 2018).

This paper discusses the adaptation of two classes of approximate computation algorithms, *approximate Bayesian computation* (ABC) and *Monte Carlo Expectation-Maximization* (MCEM), to obtain exact Bayesian and likelihood statistical inferences based on differentially private data products. The word *exact* means that, up to Monte Carlo error, the resulting inference corresponds precisely to the joint statistical model that accounts for both the analyst's specifications and the differential privacy mechanism. This paper draws a concrete connection between the novel disclosure limitation mechanisms that obey differential privacy, and the vast reserve of computational strategies available for likelihood and Bayesian statistical inference. The hope is that users of traditional, non-differentially private data can smoothly transition their existing methodologies to suit novel, differentially private data products while maintaining statistical validity. The two methods discussed in this work are applicable to a wide range of existing models, dispensing with the need to analytically recompute the new joint model to account for the privacy mechanism. Both classes of algorithms discussed in this paper do not assume specific structures of the likelihood, prior, and privacy mechanism. Indeed, the likelihood approach only requires that the analyst's original model is suitable for EM, and the Bayesian approach only requires that the original likelihood can be simulated and that the prior is proper. Should specific and convenient model structures be available, the proposed mechanisms would be amenable to adaptations that enhance computational efficiency.

The remainder of this paper is organized as follows. Section 2 lays out the mathematical formalism and notation for differential privacy and perturbation mechanisms. Section 3 proposes a rejection ABC algorithm, and shows that with kernel and bandwidth chosen to correspond to the perturbation mechanism underlying the privatized data, it produces exact posterior inference in the form of independent and identically distributed samples from the true posterior distribution. Section 4 discusses an importance sampling implementation of Monte Carlo EM for likelihood inference. The validity of both approximate computation methods derives from the fact that their tuning elements can be chosen in accordance with the differentially private perturbation mechanism that is used to generate the privatized data product. Section 5 provides two numerical demonstrations of Bayesian and likelihood inference for privatized count data, and a differentially private adaptation of the Lalonde dataset for inference on job training program efficacy. Section 6 concludes with a discussion on the duality between *approximate computation on exact data* and *exact computation on approximate data*, and the various challenges to the efficiency of these proposals.

2. Differential privacy and perturbation mechanism

Differential privacy aims to protect the confidential information of individual respondents in a dataset $\boldsymbol{x} \in \mathcal{X}$, without undue sacrifice of accuracy in learning about aggregate features of the underlying population as represented by $\boldsymbol{x}$. Here, an aggregate feature is a query $\boldsymbol{s} : \mathcal{X} \to \mathbb{R}^p$, a deterministic function of $\boldsymbol{x}$, such as the sample average, variance, quantiles and so on. Queries are the means through which analysts learn from the dataset. Counting queries, including histograms and contingency tables which are ordered multivariate counts

over a partition of $\boldsymbol{x}$, constitute a most useful class of queries. It is the main query type for the 2020 U.S. Census data products, tabulated across various geographic levels such as states, counties, and Census blocks.

Differential privacy is realized via a probabilistic mechanism based on the intended query. A differentially private query reflects as truthfully as possible the status of $\boldsymbol{x}$, while behaving similarly should it be calculated based on any neighboring dataset of $\boldsymbol{x}$. The notion of differential privacy is defined in probabilistic terms.

Definition 1 (differential privacy; Dwork et al., 2006)**.** A random function $\boldsymbol{s}_{\text{dp}} : \mathcal{X} \to \mathbb{R}^p$ is (ϵ, δ)-differentially private if for all neighboring datasets $(\boldsymbol{x}, \boldsymbol{x}') \in \mathcal{X}^2$ and all $A \in \mathscr{B}(\mathbb{R}^p)$,

$$Pr\left(\boldsymbol{s}_{\text{dp}}\left(\boldsymbol{x}'\right) \in A\right) \leq e^{\epsilon} \cdot Pr\left(\boldsymbol{s}_{\text{dp}}\left(\boldsymbol{x}\right) \in A\right) + \delta. \tag{2.1}$$

$\boldsymbol{s}_{\text{dp}}$ is ϵ-differentially private if it is $(\epsilon, 0)$-differentially private.

The pair $(\boldsymbol{x}, \boldsymbol{x}') \in \mathcal{X}^2$ are neighboring datasets if they differ by precisely one entry, either by adding or dropping one respondent, or by taking a different value (as used in the definition of *bounded* differential privacy; Dwork et al., 2006). When operating on neighboring datasets, the random function $\boldsymbol{s}_{\text{dp}}$ induces pairs of probability measures, associated respectively with $\boldsymbol{s}_{\text{dp}}(\boldsymbol{x})$ and $\boldsymbol{s}_{\text{dp}}(\boldsymbol{x}')$, that are close to each other. The degree of closeness is controlled by the *privacy loss budget* ϵ and δ. In the extreme case that both are zero, the two measures must be equal on every Borel set A, which for general $\boldsymbol{x}$ can only happen if $\boldsymbol{s}_{\text{dp}}$ does not depend on the data at all. In other words, differential privacy requires that the distribution of $\boldsymbol{s}_{\text{dp}}$ to be stable within the small neighborhood around the observable dataset.

Differential privacy is a property of the random function $\boldsymbol{s}_{\text{dp}}$. Many widely employed differentially private mechanisms take the form of perturbation mechanisms.

Definition 2. For a dataset $\boldsymbol{x} \in \mathcal{X}$ and a deterministic function $\boldsymbol{s} : \mathcal{X} \to \mathbb{R}^p$, the random function $\boldsymbol{s}_{\text{dp}}$ is a *perturbation mechanism* based on $\boldsymbol{s}$ if

$$\boldsymbol{s}_{\text{dp}}\left(\boldsymbol{x}\right) \mid \boldsymbol{s}\left(\boldsymbol{x}\right) \sim \eta_{\text{dp}}\left(\,\cdot \mid \boldsymbol{s}\left(\boldsymbol{x}\right)\right), \tag{2.2}$$

for η_{dp} a known conditional probability distribution. In particular, $\boldsymbol{s}_{\text{dp}}$ is an *additive perturbation mechanism* based on $\boldsymbol{s}$ if

$$\boldsymbol{s}_{\text{dp}}\left(\boldsymbol{x}\right) = \boldsymbol{s}\left(\boldsymbol{x}\right) + h\boldsymbol{u}, \tag{2.3}$$

where the noise component $\boldsymbol{u}$ is a p-dimensional random variable with known distribution η, $\mathbb{E}(\boldsymbol{u}) = \boldsymbol{0}$, and $h > 0$ is a scale (or bandwidth) parameter.

The differentially private query $\boldsymbol{s}_{\text{dp}}$ is a noisy version of its deterministic counterpart $\boldsymbol{s}$. The protection of privacy is achieved through randomly perturbing what would otherwise be a deterministic query calculated based on $\boldsymbol{x}$. The subscript "dp" in $\boldsymbol{s}_{\text{dp}}$ emphasizes that it instantiates the privacy mechanism η_{dp}, rather than the data generation mechanism of $\boldsymbol{x}$, as the analyst might posit. The perturbation mechanism embodied by $\boldsymbol{s}_{\text{dp}}$ is said to be *unbiased* if it satisfies $\mathbb{E}\left(\boldsymbol{s}_{\text{dp}}\left(\boldsymbol{x}\right) \mid \boldsymbol{s}\left(\boldsymbol{x}\right)\right) = \boldsymbol{s}\left(\boldsymbol{x}\right)$. Additive perturbation mechanisms, by construction of (2.3), are unbiased. Furthermore, if the scale parameter h does not depend on the confidential dataset $\boldsymbol{x}$, the mechanism may be called a *data-independent* mechanism (Li et al., 2015). Note that the additive perturbation mechanism resembles the classical measurement error model (Carroll et al., 2006), where the noise $h\boldsymbol{u}$ has a known distribution, and the noisy measurement $\boldsymbol{s}_{\text{dp}}$ is observed precisely once. Appendix A gives three examples of widely used differentially private mechanisms, with additive perturbation using Gaussian

and Laplace noises. Their definitions invoke three notions of *functional sensitivity*, (A.1)-(A.3), which we generally denote as $\Delta(\boldsymbol{s})$, to capture the idea that certain $\boldsymbol{s}$ is more revealing of individual information in $\boldsymbol{x}$ than others. It is crucial that the scale parameter of the additive perturbation mechanism is chosen as a function of both the sensitivity of $\boldsymbol{s}$ and the privacy budget, i.e. $h = h(\epsilon, \delta, \Delta(\boldsymbol{s}))$. Additional examples of additive differentially private mechanisms include the generalized Cauchy (Nissim et al., 2007), double Geometric (Schein et al., 2019), correlated multivariate Gaussian (Nikolov et al., 2013) and the k-norm (Hardt and Talwar, 2010; Bhaskara et al., 2012) mechanisms. Examples of non-additive perturbation mechanisms include the randomized response mechanism (Warner, 1965), exponential mechanism (McSherry and Talwar, 2007), objective perturbation (Chaudhuri et al., 2011; Kifer et al., 2012), among others.

A primary strength of differential privacy over traditional disclosure limitation frameworks is its *transparency*, which means that the specification of the perturbation mechanism η_{dp} may be fully revealed to the data analyst (and indeed the public) while keeping the privacy guarantee intact. For additive mechanisms, this specification consists of $\boldsymbol{u}$'s distribution η, scale parameter h, and the privacy loss budget ϵ and δ. Perturbation mechanisms can be correctly accounted for in the probabilistic modeling of privatized data. Despite the necessary sacrifice of statistical efficiency, likelihood and Bayesian models utilizing privatized data can still retain validity, in the sense that any inference drawn based on $\boldsymbol{s}$ can still be drawn based on $\boldsymbol{s}_{\text{dp}}$ correctly while accounting for its generative process. As Section 3 will discuss, for Bayesian analysis, an ABC rejection algorithm guarantees the exactness of draws from the true posterior distribution, when properly tuned according to the parameters of the perturbation mechanism. The nature of the privatized query makes ABC an appealing choice for posterior computation, even when the model is not as complex as to necessitate its use.

3. Exact Bayesian inference with differentially private data

In the absence of privacy protection, suppose a Bayesian model was posited based on the confidential query $\boldsymbol{s}$ as a function of $\boldsymbol{x}$. Let $\boldsymbol{s}(\boldsymbol{x}) \mid \theta \sim \pi(\boldsymbol{s} \mid \theta)$ be the confidential data likelihood, and $\theta \sim \pi_0(\theta)$ the prior distribution for θ. The posterior distribution of θ given $\boldsymbol{s}$ is

$$\pi(\theta \mid \boldsymbol{s}) \propto \pi_0(\theta)\, \pi(\boldsymbol{s} \mid \theta). \tag{3.1}$$

If the query $\boldsymbol{s}$ isn't privacy-protected, quantities calculated based on (3.1), either analytically or via simulation, would conclude the Bayesian analysis. With the privacy protection mechanism in place, however, we no longer observe the confidential query $\boldsymbol{s}$, but rather the privatized (perturbed) query $\boldsymbol{s}_{\text{dp}}$ as a single realization of the privacy mechanism (2.2). The joint distribution of θ and $\boldsymbol{s}_{\text{dp}}$ is

$$\pi(\theta, \boldsymbol{s}_{\text{dp}}) = \int \pi(\theta, \boldsymbol{s})\, \eta_{\text{dp}}(\boldsymbol{s}_{\text{dp}} \mid \boldsymbol{s})\, d\boldsymbol{s},$$

marginalized over the latent $\boldsymbol{s}$. This identity holds because the conditional distribution of $\boldsymbol{s}_{\text{dp}}$ given $\boldsymbol{s}$ and θ is free of θ, as it is precisely the known perturbation mechanism: $\pi(\boldsymbol{s}_{\text{dp}} \mid \boldsymbol{s}, \theta) = \eta_{\text{dp}}(\boldsymbol{s}_{\text{dp}} \mid \boldsymbol{s})$. The posterior distribution of θ given $\boldsymbol{s}_{\text{dp}}$ is

$$\pi(\theta \mid \boldsymbol{s}_{\text{dp}}) = \int \frac{\pi(\boldsymbol{s}, \boldsymbol{s}_{\text{dp}}, \theta)}{\pi(\boldsymbol{s}_{\text{dp}})} d\boldsymbol{s} = \frac{\pi_0(\theta) \int \eta_{\text{dp}}(\boldsymbol{s}_{\text{dp}} \mid \boldsymbol{s})\, \pi(\boldsymbol{s} \mid \theta)\, d\boldsymbol{s}}{\int \pi_0(\theta) \int \eta_{\text{dp}}(\boldsymbol{s}_{\text{dp}} \mid \boldsymbol{s})\, \pi(\boldsymbol{s} \mid \theta)\, d\boldsymbol{s} d\theta}. \tag{3.2}$$

Algorithm 1: Rejection ABC algorithm with differentially private queries

Data: Privatized query $\boldsymbol{s}_{\text{dp}}$, perturbation mechanism η_{dp};
Result: A set of parameter values $\{\theta_i\}_{i=1}^N$;
for *each* $i = 1, \ldots, N$ **do**
1. Simulate $\theta_i \sim \pi_0(\theta)$;
2. Simulate $\boldsymbol{s}_i \sim \pi(\boldsymbol{s} \mid \theta_i)$;
3. Accept θ_i with probability $c\eta_{\text{dp}}(\boldsymbol{s}_{\text{dp}} \mid \boldsymbol{s}_i)$ where $c^{-1} = \max \eta_{\text{dp}}$, otherwise go to step 1;

end

As (3.2) is the true posterior distribution for θ given the observable information, analytical or simulated computation based on (3.2) would conclude the exact Bayesian analysis. However, computation of (3.2) may not be trivial, as part of it involves the observed likelihood $\int \eta_{\text{dp}}(\boldsymbol{s}_{\text{dp}} \mid \boldsymbol{s})\, \pi(\boldsymbol{s} \mid \theta)\, d\boldsymbol{s}$, which is an integral of the product between the confidential data likelihood and the privacy mechanism. The challenge is exacerbated by the fact that the confidential likelihood is specified by the data analyst, whereas the privacy mechanism is specified by the data curator. These choices are typically independent of one another, and either of them may already be complex and computationally demanding on its own.

Algorithm 1 presents a recipe to generate independent and identically distributed samples from the exact posterior distribution (3.2). It demands little in terms of the tractability of the confidential likelihood. The only requirement is that for given values of θ, one can simulate data from $\pi(\boldsymbol{s} \mid \theta)$, but otherwise it need not be available in closed form. Algorithm 1 is a type of ABC algorithm, which was designed to supply practical solutions to large-scale models for which the likelihood may be implicit or intractable and have posteriors that lack closed-form expressions. ABC brought computational feasibility to stochastic differential equation models for complex dynamic systems in population genetics (Beaumont et al., 2002), systems biology (Toni et al., 2008) and ecology (Wood, 2010), albeit ABC posteriors are typically only approximate relative to the true target posterior. However, as will be shown in Theorem 3.1 and discussed in Section 6, the employment of ABC for differentially private data serendipitously eradicates the "approximate" nature of the resulting posterior samples, which otherwise would be the case if the data were noise-free.

Theorem 3.1. *Let $\pi(\boldsymbol{s} \mid \theta)$ be the likelihood for the unobserved confidential query $\boldsymbol{s}$, $\pi_0(\theta)$ a proper prior distribution, and $\eta_{dp}(\boldsymbol{s}_{dp} \mid \boldsymbol{s})$ a perturbation mechanism. Then, Algorithm 1 samples independently and identically from the exact posterior distribution $\pi(\theta \mid \boldsymbol{s}_{dp})$ defined in (3.2).*

Proof of Theorem 3.1 can be found in Appendix B. Key to the validity of Theorem 3.1 is that the differentially private perturbation mechanism is ignorable for θ (Little and Rubin, 2014), or in other words, the unobserved confidential query $\boldsymbol{s}$ is sufficient with respect to the complete likelihood $\pi(\boldsymbol{s}, \boldsymbol{s}_{\text{dp}} \mid \theta)$. Traditional statistical disclosure limitation mechanisms may or may not enjoy ignorability, a matter further complicated by their non-transparency to impact the quality of downstream statistical analysis (Abowd and Schmutte, 2016). By contrast, the ignorability property of differential privacy enables exact statistical inference and may substantially simplify the computational task.

An intuitive connection with traditional ABC can be drawn if we restrict attention to the case of additive perturbation. As defined in (2.3), assume $\eta_{\text{dp}}(\boldsymbol{s}_{\text{dp}} \mid \boldsymbol{s}) = \eta((\boldsymbol{s}_{\text{dp}} - \boldsymbol{s})/h)$

where $\eta(\cdot)$ is the density of the additive noise $\boldsymbol{u}$ and h a scale parameter, both known precisely to the analyst. Algorithm 1 adopts the kernel density η, properly scaled by a factor of c, with bandwidth $h = h(\epsilon, \delta, \Delta(\boldsymbol{s}))$ and center $\boldsymbol{s}_{\text{dp}}$ to be its acceptance probability at step 3, thus reduces to a classic rejection ABC algorithm with tuning parameters (i.e. kernel and bandwidth) set to match precisely the additive perturbation mechanism employed to generate $\boldsymbol{s}_{\text{dp}}$.

One way to understand Theorem 3.1 is that the privacy mechanism plays the role of the "random summary statistic" in the noisy ABC algorithm of Fearnhead and Prangle (2012). Noisy ABC is calibrated with respect to the joint Bayesian model, whereas ABC typically isn't. However, the kernel and bandwidth in noisy ABC are merely parameters to fine-tune the tradeoff between approximation error and the Monte Carlo error in the posterior, which in turn controls the efficiency of the sampler. In contrast, both the kernel and the bandwidth of Algorithm 1 are dictated externally by the perturbation mechanism and the privacy loss budget. The computational tradeoff and the privacy tradeoff are "bundled" together: specifying the parameters of ABC also specifies those of the privacy mechanism, and vice versa.

The overall acceptance probability of Algorithm 1 is $\pi\left(\boldsymbol{s}_{\text{dp}}\right) / \max \eta_{\text{dp}}$, or the model evidence evaluated at $\boldsymbol{s}_{\text{dp}}$ divided by the modal density of η_{dp} (see Appendix B). This means that rejection can be frequent if model evidence is low, such as when the prior and the observed likelihood are in disagreement (termed *prior-data conflict*; Evans and Moshonov, 2006), or if the privacy bandwidth h is too small.

To address the concern, Algorithm 1 can be adapted to work with a variety of alternative ABC sampling techniques to produce consistent posterior estimates for functions of interest. As an example, we discuss an importance sampling variation to Algorithm 1 as follows. At step 1 of each iteration, sample $\theta_i \sim g(\theta)$, a proposal distribution that is positive wherever the prior $\pi_0(\theta)$ is positive. At step 3, no rejection is performed, but instead θ_i is assigned a weight

$$\omega_i = \omega(\boldsymbol{s}_i, \theta_i) = \eta_{\text{dp}}\left(\boldsymbol{s}_{\text{dp}} \mid \boldsymbol{s}_i\right) \pi_0(\theta_i) / g(\theta_i).$$

The algorithm returns weighted draws $\{\theta_i, \omega_i\}_{i=1}^N$. For a square-integrable function of interest $a(\theta)$, the weighted average estimator converges in probability to its posterior expectation given $\boldsymbol{s}_{\text{dp}}$ as $N \to \infty$ (Liu, 2008):

$$\frac{\sum_{i=1}^N \omega_i a\left(\theta_i\right)}{\sum_{i=1}^N \omega_i} \xrightarrow{p} \frac{\mathbb{E}_g\left(\omega\left(\theta, \boldsymbol{s}\right) a\left(\theta\right)\right)}{\mathbb{E}_g\left(\omega\left(\theta, \boldsymbol{s}\right)\right)} = \mathbb{E}\left(a\left(\theta\right) \mid \boldsymbol{s}_{\text{dp}}\right), \tag{3.3}$$

where $\mathbb{E}_g\left(\cdot\right)$ is with respect to the joint distribution $g(\theta)\pi(\boldsymbol{s} \mid \theta)$, and $\mathbb{E}\left(\cdot \mid \boldsymbol{s}_{\text{dp}}\right)$ is with respect to the true posterior in (3.2). The proposal distribution $g(\cdot)$ can be chosen to minimize the variance of the estimator in (3.3), such as a density that is close in shape to $a(\theta)\pi_0(\theta)$ (Liu, 2008). Further adaptations of and beyond ABC, such as hybrid importance-rejection sampling (Fearnhead and Prangle, 2012), rejection control (Sisson et al., 2018, ch.4), Markov chain Monte Carlo (Marjoram et al., 2003) and sequential Monte Carlo (Sisson et al., 2007) can be developed likewise, while the consistency result of (3.3) remains standing.

4. Exact likelihood inference with differentially private data

This section discusses a Monte Carlo Expectation-Maximization (EM; Dempster et al., 1977; Wei and Tanner, 1990) implementation for likelihood inference with differentially private data. Under the classic setting, when a likelihood involves both observed and latent data, EM

seeks the maximum likelihood estimate of the parameter by iteratively integrating the log likelihood over the conditional predictive distribution of the latent data given the observed data and a current parameter value (the E-step), and maximizing the parameter value over this integral (the M-step).

In the context of differential privacy, the complete data is $(\boldsymbol{s}, \boldsymbol{s}_{\text{dp}})$, in which the latent data is the confidential query $\boldsymbol{s}$, and the observed data is the privatized query $\boldsymbol{s}_{\text{dp}}$. In the special case of additive perturbation, $\boldsymbol{s}_{\text{dp}} = \boldsymbol{s} + h\boldsymbol{u}$ is a convolution of $\boldsymbol{s}$ and the noise component $\boldsymbol{u}$. The complete likelihood is written as $L(\theta; \boldsymbol{s}, \boldsymbol{s}_{\text{dp}}) \propto \pi(\boldsymbol{s}, \boldsymbol{s}_{\text{dp}} \mid \theta)$, as defined in Section 3. The EM algorithm for maximum likelihood inference for θ given the differentially private $\boldsymbol{s}_{\text{dp}}$ is schematically described in Algorithm 2.

Algorithm 2: EM algorithm for differentially private queries

Data: Privatized query $\boldsymbol{s}_{\text{dp}}$, initial $\theta^{(0)}$;
Result: A local maximizer $\theta^{(t^*)}$;
while $\Delta\left(\theta^{(t)}, \theta^{(t-1)}\right) > tol.$ **do**
E-step: Evaluate the expectation of the complete log likelihood with respect to the conditional predictive distribution of $\boldsymbol{s}$ given $\boldsymbol{s}_{\text{dp}}$ and the current maximizer $\theta^{(t)}$:

$$\begin{aligned} Q(\theta; \theta^{(t)}) &= \mathbb{E}\left(\log L(\theta; \boldsymbol{s}, \boldsymbol{s}_{\text{dp}}) \mid \boldsymbol{s}_{\text{dp}}, \theta^{(t)}\right) \\ &= \mathbb{E}\left(\log \pi(\boldsymbol{s} \mid \theta) \mid \boldsymbol{s}_{\text{dp}}, \theta^{(t)}\right) + \text{const.}; \end{aligned} \tag{4.1}$$

M-step: Calculate $\theta^{(t+1)} := \text{argmax}_\theta Q(\theta; \theta^{(t)})$, and set $t := t + 1$;
end

The E- and M-steps are iterated until convergence, that is when $\theta^{(t)}$ stabilizes so that its distance (somehow measured) from the previous iteration, $\text{dist}\left(\theta^{(t)}, \theta^{(t-1)}\right)$, is sufficiently small. It is worth noting that the constant term in (4.1) is equal to

$$\mathbb{E}\left(\log \eta_{\text{dp}}(\boldsymbol{s}_{\text{dp}} \mid \boldsymbol{s}) \mid \boldsymbol{s}_{\text{dp}}, \theta^{(t)}\right),$$

which can be ignored within the EM algorithm. This is because, as discussed in Section 3, the privacy mechanism η_{dp} is known and independent of θ, and so is the conditional predictive expectation of its log density.

As alluded to in Section 1, for likelihood modeling of differentially private data, the confidential data likelihood and the privacy mechanism are typically specified by separate parties without coordination with one another. Thus in general, one cannot expect the observed data likelihood (which is an integral of their product) to come from an exponential family (cf. Park et al., 2017), nor be able to perform both the E- and the M-steps analytically. Monte Carlo implementation of one or both steps may be needed, which amounts to implementing the E-step of Algorithm 2 via an importance sampling scheme. We describe this scheme in Algorithm 3. The set of weighted samples $\{\boldsymbol{s}_i, \omega_i\}_{i=1}^N$ produced by Algorithm 3 may be used in two ways, depending on whether the confidential data likelihood is or is not from an exponential family. We discuss both cases below.

Algorithm 3: E-step via importance sampling for differentially private queries

Data: Privatized query $\boldsymbol{s}_{\text{dp}}$, perturbation mechanism η_{dp};
Result: A set of weighted samples $\{\boldsymbol{s}_i, \omega_i\}_{i=1}^N$, to be used for (4.2)-(4.7);
for *the* t^{th} *E-step of Algorithm 2,* **do**
1. Simulate $\boldsymbol{s}_i \sim \pi(\boldsymbol{s} \mid \theta^{(t)})$;
2. Calculate $\omega_i = \eta_{\text{dp}}(\boldsymbol{s}_{\text{dp}} \mid \boldsymbol{s}_i)$;

end

4.1. **Confidential data with exponential family likelihood.** In the simpler scenario that the confidential data likelihood $\pi(\boldsymbol{s} \mid \theta)$ as specified by the analyst belongs to the exponential family, it admits a sufficient statistic to the parameter θ which we denote as $b(\boldsymbol{s})$. The function $Q(\theta; \theta^{(t)})$ in (4.1) can be written as an explicit function of θ and

$$\mathbb{E}\left(b(\boldsymbol{s}) \mid \boldsymbol{s}_{\text{dp}}, \theta^{(t)}\right), \tag{4.2}$$

the conditional expectation of $b(\boldsymbol{s})$ given $\boldsymbol{s}_{\text{dp}}$ and the current maximizer $\theta^{(t)}$. With this simplification, however, (4.2) may still not be estimable in closed form, in which case we utilize the set of weighted samples $\{\boldsymbol{s}_i, \omega_i\}_{i=1}^N$ produced by Algorithm 3 to consistently estimate it at every iteration t. Indeed, as $N \to \infty$, the weighted estimator

$$\frac{\sum_{i=1}^N \omega_i b(\boldsymbol{s}_i)}{\sum_{i=1}^N \omega_i} \tag{4.3}$$

converges in probability to (4.2). For the E-step of the $(t+1)$st iteration, $\theta^{(t+1)}$ can be found by maximizing $Q(\theta; \theta^{(t)})$, replacing (4.2) therein with (4.3). The effective sample size at the tth iteration is

$$\text{ESS}^{(t)}(N) = N\pi^2\left(\boldsymbol{s}_{\text{dp}} \mid \theta^{(t)}\right)\mathbb{E}^{-1}_{\boldsymbol{s}\mid\theta^{(t)}}\left(\eta^2_{\text{dp}}(\boldsymbol{s}_{\text{dp}} \mid \boldsymbol{s})\right), \tag{4.4}$$

where the subscript "$\boldsymbol{s} \mid \theta^{(t)}$" signifies that the expectation is taken with respect to the current approximation to the confidential data likelihood, or equivalently, the proposal distribution of the E-step importance sampler. Derivation of (4.4) may be found in Appendix C.

In Algorithm 3, the $\boldsymbol{s}_i$'s are simulated from the current approximation to the analyst's confidential data likelihood, and the weights ω_i's are separately determined by the curator's privacy mechanism. Similar in spirit to Algorithm 1, this separation allows the computation to easily accommodate independently derived choices of data likelihood and privacy mechanisms, and does not require the evaluation or integration of quantities that are nontrivial functions of both. Whenever appropriate, however, one may modify Algorithm 3 to sample from the conditional predictive distribution in more efficient ways. For example, with rejection or Markov chain-based samplers, $\boldsymbol{s}_i$ follows a proposal distribution and $\omega_i = 1$ if $\boldsymbol{s}_i$ is accepted and 0 otherwise (McCulloch, 1997; Booth and Hobert, 1999). One may also perform importance sampling where $\boldsymbol{s}_i \sim \pi\left(\boldsymbol{s} \mid \boldsymbol{s}_{\text{dp}}, \theta^{(t-1)}\right)$, the approximation to the conditional predictive distribution at the previous iteration, and $\omega_i = \pi\left(\boldsymbol{s} \mid \boldsymbol{s}_{\text{dp}}, \theta^{(t)}\right)/\pi\left(\boldsymbol{s} \mid \boldsymbol{s}_{\text{dp}}, \theta^{(t-1)}\right)$ the ratio between the current and previous approximations, thereby reweighting and recycling the multiply-imputed $\boldsymbol{s}_i$'s to save computational effort (Quintana et al., 1999). One may also resample the simulated $\boldsymbol{s}_i$'s according to their associated weights to obtain an unweighted rejection sample, as long as the goal is to construct as accurate as possible an estimate for (4.2) as part of the E-step.

4.2. **Confidential data with general likelihood.** If the confidential data likelihood does not come from an exponential family, $Q(\theta;\theta^{(t)})$ of (4.1) may not reduce to a straightforward expression involving θ and (4.2). In this case, the E-step requires a full approximation to $Q(\theta;\theta^{(t)})$ as a mixture of augmented log likelihoods, constructed as follows.

Let $\{\boldsymbol{s}_i,\omega_i\}_{i=1}^N$ be a weighted sample from the target distribution $\pi\left(\boldsymbol{s}\mid\boldsymbol{s}_{\text{dp}},\theta^{(t)}\right)$, the tth approximation to the conditional predictive distribution. Specifically $\{\boldsymbol{s}_i,\omega_i\}_{i=1}^N$ can be the importance sample generated by Algorithm 3, or by one of its variations described above. Then,

$$\hat{Q}(\theta;\theta^{(t)}) = m\sum_{i=1}^{N}\omega_i\log\pi(\boldsymbol{s}_i\mid\theta) \tag{4.5}$$

serves as a consistent approximation to $Q(\theta;\theta^{(t)})$. The constant $m^{-1}=\sum_{i=1}^N\omega_i$ in (4.5) is inconsequential to the maximizer in the ensuing M-step, as long as the ω_i's do not involve the unknown parameter θ. That is indeed the case since, again, the perturbation mechanism is ignorable for θ. Writing $\lambda_\theta(\boldsymbol{s})=\nabla_\theta\log\pi\left(\boldsymbol{s}\mid\theta\right)$, the observed score function $\nabla_\theta\log\pi\left(\boldsymbol{s}_{\text{dp}}\mid\theta^{(t)}\right)$ can be approximated at the tth iteration according to

$$\mathbb{E}\left(\lambda_\theta(\boldsymbol{s})\mid\boldsymbol{s}_{\text{dp}},\theta^{(t)}\right)\dot{=}\, m\sum_{i=1}^{N}\omega_i\lambda_\theta(\boldsymbol{s}_i). \tag{4.6}$$

The observed Fisher information can also be approximated according to

$$-\nabla_\theta^2\log\pi\left(\boldsymbol{s}_{\text{dp}}\mid\theta^{(t)}\right)\dot{=}$$
$$m\sum_{i=1}^{N}\omega_i\left\{-\nabla_\theta\lambda_\theta\left(\boldsymbol{s}_i\right)-\lambda_\theta\left(\boldsymbol{s}_i\right)\lambda_\theta\left(\boldsymbol{s}_i\right)^\top\right\}+m^2\sum_{i=1}^{N}\sum_{j=1}^{N}\omega_i\omega_j\lambda_\theta\left(\boldsymbol{s}_i\right)\lambda_\theta\left(\boldsymbol{s}_j\right)^\top. \tag{4.7}$$

Derivations of the observed score function and observed Fisher information can be found in Appendix D. Both (4.6) and (4.7) may be used for quantifying the inferential uncertainty under the normal approximation to the likelihood (Meilijson, 1989), as well as accelerating and assessing convergence for Newton-type implementations of the M-step. The approximations given above rely only the first and second derivatives of the confidential likelihood be evaluable at the simulated s_i's.

For any EM algorithm (and not just Monte Carlo EM) to be applicable to likelihood inference from differentially private data, one must be able to evaluate the confidential data likelihood $\pi(\boldsymbol{s}\mid\theta)$, to the extent that maximization of the Q function can be done at least numerically. The vast literature on Monte Carlo EM has much to offer in terms of options for implementing both the E- and the M-steps with better convergence rates, sampling efficiency, or under computational capacity constraints, and for adapting modeling scenarios to differentially private data. The additive perturbation mechanism of (2.3) is a special instance of a linear mixed effects model, which is particularly suitable for Monte Carlo EM and has been studied extensively in the literature, e.g. Wolfinger and O'Connell (1993); McCulloch (1997).

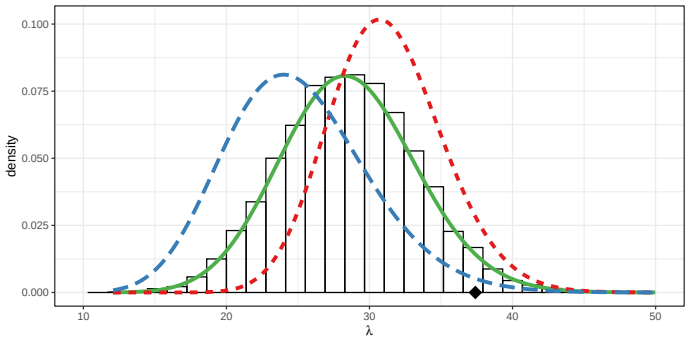

Figure 1: Algorithm 1 produces exact draws (black histogram, $N = 10^4$) from the true posterior (green density), which is different from the naïve posterior (red dotted density) which treats the observed $\boldsymbol{s}_{\text{dp}} = 37.4$ (black diamond) as if without privatization. The blue dashed density is the prior.

5. Numerical Demonstrations

5.1. **Bayesian and likelihood inference from privatized count data.** In this simple example, we consider modeling the number of respondents from a sample $\boldsymbol{x}$ in possession of a certain feature. $\boldsymbol{s}(\cdot)$ is the univariate counting query, for which we posit the sampling model $\boldsymbol{s}(\boldsymbol{x}) \mid \theta \sim Pois(\theta)$. θ is the population expectation parameter for which we wish to draw Bayesian and likelihood inference.

First consider a Bayesian model for θ. We implement rejection ABC as described in Algorithm 1 to draw from the exact Bayesian posterior based on the privatized count $\boldsymbol{s}_{\text{dp}}$. Suppose $\boldsymbol{s}_{\text{dp}}$ is produced by the ϵ-differentially private Laplace mechanism (Example 1 in Appendix A), where the additive noise follows $\boldsymbol{u} \sim Lap_p(1)$ with bandwidth $h = \epsilon^{-1}$. As with general ABC samplers, Algorithm 1 can work with arbitrary choices of prior and likelihood that need not be conjugate, so long as the prior is proper. For the purpose of illustration, we consider the prior $\theta \sim Gamma(\alpha, \beta)$, where α and β are fixed hyperparameters, so that an analytically tractable posterior can be obtained for visual comparison.

Figure 1 depicts both the correct and the naïve analyses, with hyperparameters $\alpha = 25, \beta = 1$, privacy loss budget $\epsilon = 0.2$, and $\boldsymbol{s}_{\text{dp}} = 37.4$. The true analytical posterior (green solid density), normalized via numerical integration, coincides with the differentially private ABC posterior histogram tabulated from 10^4 draws from Algorithm 1. The correct analysis differs substantially from the incorrect naïve posterior (red dotted density), which treats $\boldsymbol{s}_{\text{dp}}$ as if it were an observed confidential query. (The latter posterior amounts to the posterior from the standard Gamma-Poisson conjugate model.) Compared to the correct posterior, the naïve posterior succumbs less to the shrinkage effect imposed by the prior. It assigns a heavier weight of evidence to the observed value of $\boldsymbol{s}_{\text{dp}}$, more than it deserves. It is furthermore overly concentrated at the mode, underestimating the posterior uncertainty associated with θ.

Appendix E reports additional experiments that employ Gamma prior distributions with hyperparameters $\alpha = 2, 5, 50, 75$ and $\beta = 1$. It is worth noting that when the privatized observation $\boldsymbol{s}_{\text{dp}}$ appears highly unlikely under the chosen prior (or is *in conflict* with it, in other words), the correct posterior heavily discounts the contribution by $\boldsymbol{s}_{\text{dp}}$. Such is the case when

Table 1: Acceptance rate of Algorithm 1 under various priors ; $\boldsymbol{s}_{\text{dp}} = 37.4$

prior: $\theta \sim Gamma(\alpha, 1)$	prior predictive: $\mathbb{E}_\theta(\boldsymbol{s}_{\text{dp}})$	acceptance rate (%)	s.e. (%)
$\alpha = 2$	2	0.09	0.02
$\alpha = 5$	5	0.21	0.06
(Figure 1) $\alpha = 25$	25	16.24	0.35
$\alpha = 50$	50	19.83	0.31
$\alpha = 75$	75	0.64	0.07

$\alpha = 2$ or 5, as seen in Figure 3 (a) and (b): the correct posterior is in close alignment with the prior and differs significantly from the naïve posterior. As alluded to in Section 3, prior-data conflict presents a challenge for ABC algorithms in general, because forward sampling tends to explore the area with higher prior predictive concentration. A realized observation far from that area would result in a low acceptance rate. To see this, Table 1 reports the average acceptance rates and their standard errors over 20 direct repetitions of Algorithm 1 under various choices of Gamma priors. In comparison with the observed query, these priors range from congruent to conflicting, as can be seen from the varied differences between $\boldsymbol{s}_{\text{dp}}$ and its prior predictive expectation: $\mathbb{E}_\theta(\boldsymbol{s}_{\text{dp}}) = \int \boldsymbol{s}_{\text{dp}} \int \eta_{\text{dp}}\left(\boldsymbol{s}_{\text{dp}} \mid \boldsymbol{s}\right) \int \pi\left(\boldsymbol{s} \mid \theta\right) \pi_0\left(\theta\right) d\theta d\boldsymbol{s} d\boldsymbol{s}_{\text{dp}}$.

Maximum likelihood estimation for θ is carried out as follows. The confidential data likelihood is the Poisson density. Importance sampling as described in Algorithm 3 is used to construct estimates for (4.2) at every iteration of the E-step, followed by an analytical M-step. Appendix E describes details of the implementation using three stages of successively more stringent tolerance levels. With $\theta^{(1)} = 1$, the algorithm converges to the maximizer $\hat{\theta} = 37.237$, with observed Fisher information estimated to be 1.582×10^{-2}. If $\boldsymbol{s}_{\text{dp}}$ were erroneously treated as the confidential data, the MLE for θ would've been 37.4, and the observed Fisher information would've been 2.674×10^{-2}, or 69% larger than the correct estimate, again displaying an underestimation of inferential uncertainty. The reduction of Fisher information content reflects a loss of statistical efficiency induced by the privatization mechanism, and is expected in typical inference problems whenever confidential data are replaced with their privatized counterparts. Details of the above calculations can be found in Appendix E.

5.2. **Lalonde dataset.** The Lalonde dataset (LaLonde, 1986) was built from the randomized trial of the National Supported Work (NSW) Demonstration and nonexperimental comparison data, for the purpose of studying the efficacy of the job training program on recipients' future earnings. The dataset, with a total of 185 treated and 260 control units, is well-studied in the causal inference and econometrics literatures using regression and propensity matching methods, see e.g. Heckman and Hotz (1989); Dehejia and Wahba (1999, 2002). We employ the example here to illustrate a Bayesian analysis that compares the 1978 earnings of the treatment and control groups, if ϵ-differentially private versions of the key descriptive statistics were released instead.

Let z_i be the observed indicator for whether subject i received treatment ($z_i = 1$) or control ($z_i = 0$), and y_i their earning in 1978 (in $\$1k$). The full parameter of the model is $\theta = \left(\tau, \mu, \sigma_t^2, \sigma_c^2\right)$, in which τ is the difference in average earnings between the treatment and control groups, and is the primary parameter of interest. We posit independent priors for

elements of θ, as well as the sampling model

$$y_i \mid z_i, \theta \sim N\left(\tau z_i + \mu, \sigma_t^2 z_i + \sigma_c^2 (1 - z_i)\right).$$

For the sake of simplicity, we do not consider additional covariates that distinguish the treatment and control subjects.

Among the descriptive statistics that the publisher plans to release, relevant to the inferential task at hand are the within-group sample means and sample variances: $\boldsymbol{s} = \left(\bar{y}_t, \bar{y}_c, s_t^2, s_c^2\right)$. Together they make up the sufficient statistic for the full parameter θ. The top row of Figure 2 displays the posterior inference for θ by repeatedly fitting this model in `RStan` using the actual value of $\boldsymbol{s}$. Discrepancies among the ten boxplots within each figure, all of them minor, are due to Monte Carlo errors. According to the model, there is a discernible positive treatment effect since the posterior mass of τ is overwhelmingly positive.

Suppose that the data publisher releases ϵ-differentially private version of sample means and variances. Since the mean and the variance are real-valued functions, they do not have a finite global sensitivity Δ_{GS} as defined in (A.1), hence the Laplace mechanism cannot apply directly to them. To circumvent this issue, the publisher may *clamp* the underlying query, that is to enforce its value to stay within a bounded range. For simplicity's sake, suppose that the clamping range on individual income is conservatively set, say to between zero and $\$100k$, and the treatment and control groups are guaranteed to exceed 100 people. This effectively restricts the global sensitivity of $\bar{y}_t$ and $\bar{y}_c$ to 1 and that of s_t^2 and s_c^2 to 100. For reference, the maximum observed individual income in the dataset is $\$60.3k$, and the treatment and control group are respectively of sizes $n_t = 185$ and $n_c = 260$, ensuring that all confidential query values fall well within the clamping range. The benefit of conservative clamping is that the privatized statistics would not require truncation correction, even though it amounts to an inefficient privacy budget allocation strategy. Further suppose two separate privacy loss budgets of $\epsilon = 1/3$ and $100/6$ are respectively expended on the sample means and variances, through Laplace mechanisms employing independent zero-mean noise components with bandwidths $h^{-1} = 1/3$ for each of the sample means $\bar{y}_t, \bar{y}_c$, and $h^{-1} = 1/6$ for each of the sample variances s_t^2, s_c^2.

The middle and bottom rows of Figure 2 respectively display posterior inferences from naïvely fitting the original model (i.e. disregarding the privacy mechanism) in `RStan`, and correctly fitting the exact posterior (i.e. accounting for the privacy mechanism) using rejection ABC of Algorithm 1. Both methods were fitted to the same ten independent realizations of $\boldsymbol{s}_{\text{dp}}$ from the Laplace mechanism. Discrepancies among the ten boxplots within each figure in these two rows are due to the random privacy mechanism and to Monte Carlo errors – the latter to a much lesser extent. We see that with the correct analysis, posterior uncertainty for all parameters are substantially inflated, in part due to the highly inefficient allocation of the privacy loss budget. As a result, we can no longer conclude that the treatment effect is significant in either direction. However, the posterior quantiles overlap substantially with their counterparts from the original posterior on the top row, indicating that the cost of privacy manifests more as an estimation precision loss rather than bias. This stand in contrast against the naïve analysis which delivers tight, yet idiosyncratically displaced, posterior masses. Details of this analysis can be found in Appendix F.

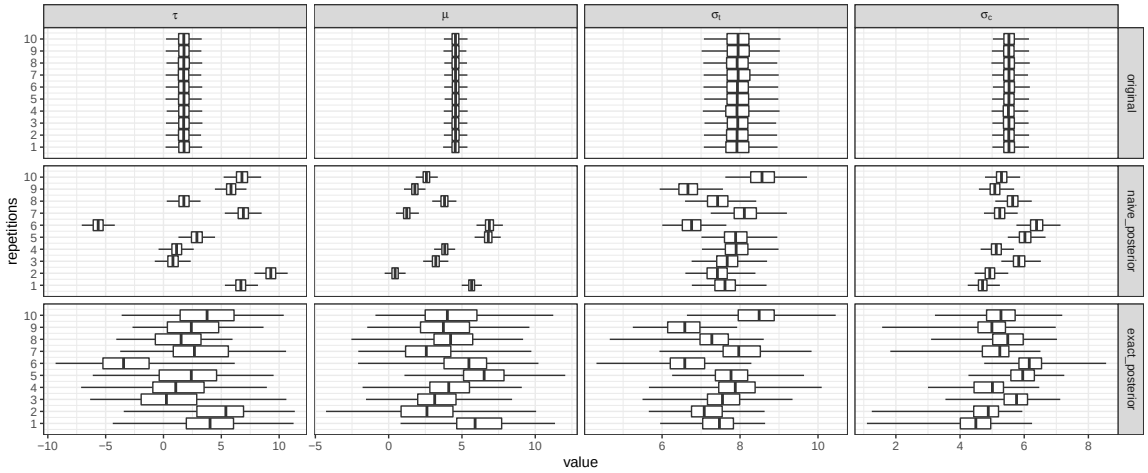

Figure 2: Boxplots of $(1\%, 25\%, 50\%, 75\%, 99\%)$ posterior quantiles of $\left(\tau, \mu, \sigma_t^2, \sigma_c^2\right)$. Top row: ten repeated `RStan` fittings of the original model to the original data $\boldsymbol{s}$; Mid row: naïve `RStan` fittings of the original model to ten realizations of $\boldsymbol{s}_{\text{dp}}$ via the Laplace mechanism; Bottom row: exact posterior fittings using rejection ABC (Algorithm 1) on the same ten $\boldsymbol{s}_{\text{dp}}$ realizations as above.

6. Conclusion and Discussion

Modern likelihood and Bayesian inference face the challenge of model complexity. They appeal to Monte Carlo and approximate methods to carry out needed computation, even if the resulting inferences are only approximate with respect to the full statistical model. This paper discussed how approximate computation algorithms, specifically ABC and Monte Carlo EM, can be adapted to obtain exact Bayesian and likelihood statistical inferences based on differentially private data products. In both cases, the tuning elements of the approximate computation algorithms are chosen to accord with the specifications of the differentially private perturbation mechanism, which can be made transparent to the data analyst. Both methods are applicable to a wide range of modeling scenarios, and may help data users transition existing methodologies to apply to differentially private data products while maintaining the statistical validity of their analysis.

When no privacy mechanism is involved, ABC algorithms exhibits a bias whenever they cannot enforce an exact match between the observed and the simulated data (Nunes and Balding, 2010; Drovandi et al., 2011; Gleim and Pigorsch, 2013; Barnes et al., 2012; Bernton et al., 2019), which is typically the case in practice. The justification of ABC relies on that in the limit as the bandwidth $h \to 0$, the ABC posterior $\pi_{ABC}(\theta \mid \boldsymbol{s})$ approaches the true posterior $\pi(\theta \mid \boldsymbol{s})$ (Blum et al., 2013; Sisson et al., 2018). However in practice, h cannot be too small in order for the algorithm to generate an adequate number of samples, trading off a larger approximation error with a smaller Monte Carlo error.

The statistical insight underscored by this paper is the duality

approximate computation on exact data $\quad\leftrightarrow\quad$ *exact computation on approximate data.*

Differentially private data is approximate data. The perturbation mechanism with which the data were treated serves coincidentally as the attributable cause of the approximation error. When differentially private data are employed, the Monte Carlo error becomes the sole kind of error attributable to the ABC algorithm, and vanishes as $N \to +\infty$ as any other consistent method of simulation.

The pursuit of differential privacy pits a direct tradeoff against statistical efficiency (Duchi et al., 2018). But the efficiency-privacy tradeoff as a statistical consideration is interweaved with the approximation-exactness tradeoff as a computational consideration, a sentiment that is shared by explorations of other simulation-based Bayesian computational algorithms with differentially private data, including stochastic gradient Monte Carlo (Wang et al., 2015) and Gibbs sampling (Foulds et al., 2016). For ABC algorithms, to insist on maximal statistical efficiency necessitates computational approximation. Whereas the act of data perturbation not only gains differential privacy, but also computational exactness for free. Both the ABC and Monte Carlo EM approaches adapt to differentially private data using the same logic, by setting the tuning parameters governing their numerical performance based on the privacy parameters. Tailoring an algorithm according to the data generative specification exploits the alignment between the statistical and computational tradeoffs, hitting two birds with one stone, so to speak.

There are several computational challenges to the practical implementation of the proposed frameworks. These challenges are of two types: those intrinsic to ABC and other forward sampling techniques, and those induced by the privacy mechanism. A weakness in either of these aspects may impact the computational efficiency of these proposals, or in the worst case, render them infeasible. We discuss the two types of challenges below.

A data analyst operating under the dissemination mode of data access is on the receiving end of data products which are designed and privatized by the data curator. As this paper discusses the migration of existing statistical methodology to accommodate privacy-protected data products, we assume that the analyst knows how to perform their preferred analysis on the data product *were it not* privatized, i.e. if the curator releases $\boldsymbol{s}$ rather than $\boldsymbol{s}_{\text{dp}}$. That is, $\pi(\theta \mid \boldsymbol{s})$ in (3.1) is taken to be the ultimate posterior the analyst targets. Depending on the model, however, the analyst may or may not prefer to use ABC or other forward sampling techniques to draw inference from $\pi(\theta \mid \boldsymbol{s})$. The strength of ABC lies in its ability to handle intractable likelihoods, but it presents several limitations. In the construction of the current paper, the intended query function $\boldsymbol{s}$ (and hence the private query $\boldsymbol{s}_{\text{dp}}$) may be multi-dimensional, where each dimension is generated in isolation, in conjunction, or sequentially. In particular, we do not preclude the identity function, $\boldsymbol{s}(\boldsymbol{x}) = \boldsymbol{x}$, in which case the privacy perturbation is performed element-wise on the full dataset for publication, such as may be encountered in the *local* differential privacy setting.

Whenever the full data likelihood does not admit a low-dimensional sufficient summary to $\boldsymbol{s}$, the computational efficiency of both proposed algorithms will likely suffer. For classic ABC, the synthetic data matching step (step 3 in Algorithm 1) will be computationally wasteful. The ABC literature explores the use of approximate summary statistics (Beaumont et al., 2002; Joyce and Marjoram, 2008; Wegmann et al., 2009) to achieve dimension reduction and efficient matching. Unless carefully designed, however, general approximate summary reduction to $\boldsymbol{s}$ will complicate the expression of the privacy kernel η_{dp}, and will destroy the "exact" nature of the proposed algorithm. The lack of sufficient reduction challenges the feasibility of other modes of computational for privacy-aware Bayesian inference as well; see e.g. Bernstein and Sheldon (2018, 2019). The question remains with the data curator: in

anticipation of a broad range of data analysis needs, how to choose the query $\boldsymbol{s}$ that provides better statistical utility *and* computational efficiency?

Another limitation of ABC methods is that their performance depends on the prior and the nature of the state space. As mentioned previously, ABC must work with proper prior distributions. This minimal requirement speaks nothing about the algorithm's efficiency. As the numerical experiment in Section 5.1 demonstrates, the acceptance probability of Algorithm 1 is low when the observed data is in conflict with the prior. The remedy is to devote more sampling resource to areas of the parameter space for which the data exhibit more support. This is a tautology of sorts, since the area we seek is precisely the area with high posterior density, which may be particularly difficult to locate when the parameter space is high dimensional, and when the prior distribution is diffuse (despite being proper).

There are also computational challenges brought forth by the privacy mechanism. Since both proposed algorithms require the transparency of $\eta_{\text{dp}}(\boldsymbol{s}_{\text{dp}} \mid \cdot)$, any act that deprives the analyst's ability to evaluate this quantity also hinders the proposed computational schemes. Two notable causes to diminished transparency of the privacy mechanism are clamping and *post-processing*. As discussed in Section 5.2, the curator performs clamping when the query has unbounded global sensitivity. While naïve and conservative clamping (such as presented in Section 5.2) requires little additional work from the analyst, carefully designed clamping procedures typically involve the underlying confidential dataset in a nontrivial fashion (see e.g. Biswas et al., 2020). The resulting privacy mechanism may not be simply captured by an analytically tractable η_{dp}. In addition, the post-processing of differentially private data products may also complicate an otherwise simple expression of η_{dp}. Such is the case if the post-processing operation depends nontrivially on aspects of the observed data. For example, the TopDown algorithm imposes *invariants* on the differentially private noisy measurements via optimization-based post-processing (Abowd et al., 2022). As a result, the output of the algorithm does not permit a straightforward probabilistic description, which threatens its *congeniality* as a building block in the data processing pipeline (Gong and Meng, 2020). From the statistical point of view, a transparent privacy mechanism is instrumental to the feasibility of conducting exact statistical inference from privacy-protected data (Gong, 2022). To ensure transparency of the privacy mechanism is yet another challenging task that lies with the data curator.

Acknowledgment

The author wishes to thank Xiao-Li Meng for inspiring discussions, as well as John Abowd, Gary King, Zhiqiang Tan, and three anonymous reviewers for helpful comments. The author gratefully acknowledges research support by the National Science Foundation (DMS-1916002).

References

J. M. Abowd and I. M. Schmutte. Economic analysis and statistical disclosure limitation. *Brookings Papers on Economic Activity*, 2015(1):221–293, 2016. https://doi.org/10.1353/eca.2016.0004.

J. M. Abowd, R. Ashmead, R. Cumings-Menon, S. Garfinkel, M. Heineck, C. Heiss, R. Johns, D. Kifer, P. Leclerc, A. Machanavajjhala, B. Moran, W. Sexton, M. Spence, and P. Zhuravlev. The 2020 Census Disclosure Avoidance System TopDown Algorithm. *Harvard Data Science Review*, (Special Issue 2), 2022. https://doi.org/10.1162/99608f92.529e3cb9.

C. P. Barnes, S. Filippi, M. P. Stumpf, and T. Thorne. Considerate approaches to constructing summary statistics for abc model selection. *Statistics and Computing*, 22(6):1181–1197, 2012. https://doi.org/10.1007/s11222-012-9335-7.

M. A. Beaumont, W. Zhang, and D. J. Balding. Approximate Bayesian computation in population genetics. *Genetics*, 162(4):2025–2035, 2002. https://doi.org/10.1093/genetics/162.4.2025.

G. Bernstein and D. 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. https://proceedings.neurips.cc/paper/2018/file/08040837089cdf46631a10aca5258e16-Paper.pdf.

G. Bernstein and D. R. Sheldon. Differentially private bayesian linear regression. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019. https://proceedings.neurips.cc/paper/2019/file/f90f2aca5c640289d0a29417bcb63a37-Paper.pdf.

E. Bernton, P. E. Jacob, M. Gerber, and C. P. Robert. Approximate Bayesian computation with the Wasserstein distance. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 81(2):235–269, 2019. https://doi.org/10.1111/rssb.12312.

A. Bhaskara, D. Dadush, R. Krishnaswamy, and K. Talwar. Unconditional differentially private mechanisms for linear queries. In *Proceedings of the forty-fourth annual ACM symposium on Theory of computing*, pages 1269–1284, 2012. https://doi.org/10.1145/2213977.2214089.

S. Biswas, Y. Dong, G. Kamath, and J. Ullman. Coinpress: Practical private mean and covariance estimation. In H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 14475–14485. Curran Associates, Inc., 2020. https://proceedings.neurips.cc/paper/2020/file/a684eceee76fc522773286a895bc8436-Paper.pdf.

A. Blum, C. Dwork, F. McSherry, and K. Nissim. Practical privacy: the SuLQ framework. In *Proceedings of the twenty-fourth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*, pages 128–138, 2005. https://doi.org/10.1145/1065167.1065184.

M. G. Blum, M. A. Nunes, D. Prangle, S. A. Sisson, et al. A comparative review of dimension reduction methods in approximate Bayesian computation. *Statistical Science*, 28(2):189–208, 2013. https://doi.org/10.1214/12-STS406.

J. G. Booth and J. P. Hobert. Maximizing generalized linear mixed model likelihoods with an automated Monte Carlo EM algorithm. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 61(1):265–285, 1999. https://doi.org/10.1111/1467-9868.

00176.

B. S. Caffo, W. Jank, and G. L. Jones. Ascent-based Monte Carlo expectation–maximization. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2):235–251, 2005. https://doi.org/10.1111/j.1467-9868.2005.00499.x.

R. J. Carroll, D. Ruppert, L. A. Stefanski, and C. M. Crainiceanu. *Measurement error in nonlinear models: a modern perspective*. Chapman and Hall/CRC, 2006. https://doi.org/10.1201/9781420010138.

K. Chaudhuri, C. Monteleoni, and D. Sarwate. Differentially private empirical risk minimization. In *Journal of Machine Learning Research*, volume 12, pages 1069–1109, 2011.

R. H. Dehejia and S. Wahba. Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs. *Journal of the American statistical Association*, 94(448): 1053–1062, 1999.

R. H. Dehejia and S. Wahba. Propensity score-matching methods for nonexperimental causal studies. *Review of Economics and statistics*, 84(1):151–161, 2002. https://doi.org/10.1162/003465302317331982.

A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, 39(1):1–22, 1977. https://doi.org/10.1111/j.2517-6161.1977.tb01600.x.

C. C. Drovandi, A. N. Pettitt, and M. J. Faddy. Approximate Bayesian computation using indirect inference. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 60(3):317–337, 2011. http://www.jstor.org/stable/41262278.

J. C. Duchi, M. I. Jordan, and M. J. Wainwright. Minimax optimal procedures for locally private estimation. *Journal of the American Statistical Association*, 113(521):182–201, 2018. https://doi.org/10.1080/01621459.2017.1389735.

C. Dwork, F. McSherry, K. Nissim, and A. Smith. Calibrating noise to sensitivity in private data analysis. In *Theory of cryptography conference*, pages 265–284. Springer, 2006. https://doi.org/10.1007/978-3-540-32732-5_32.

M. Evans and H. Moshonov. Checking for prior-data conflict. *Bayesian analysis*, 1(4): 893–914, 2006. https://doi.org/10.1214/06-BA129.

P. Fearnhead and D. Prangle. Constructing summary statistics for approximate Bayesian computation: semi-automatic approximate Bayesian computation. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 74(3):419–474, 2012. https://doi.org/10.1111/j.1467-9868.2011.01010.x.

J. Foulds, J. Geumlek, M. Welling, and K. Chaudhuri. On the theory and practice of privacy-preserving Bayesian data analysis. In *Proceedings of the Thirty-Second Conference on Uncertainty in Artificial Intelligence*, UAI'16, pages 192–201, Arlington, Virginia, United States, 2016. AUAI Press. ISBN 978-0-9966431-1-5. http://dl.acm.org/citation.cfm?id=3020948.3020969.

A. Gleim and C. Pigorsch. Approximate Bayesian computation with indirect summary statistics. Technical report, University of Bonn, Bonn, Germany, 2013. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.665.5503&rep=rep1&type=pdf.

R. Gong. Transparent Privacy is Principled Privacy. *Harvard Data Science Review*, (Special Issue 2), jun 24 2022. https://doi.org/10.1162/99608f92.b5d3faaa.

R. Gong and X.-L. Meng. Congenial differential privacy under mandated disclosure. In *Proceedings of the 2020 ACM-IMS on Foundations of Data Science Conference*, FODS'20, pages 59–70, New York, NY, USA, 2020. Association for Computing Machinery. https://doi.org/10.1145/3412815.3416892.

M. Hansen. To reduce privacy risks, the census plans to report less accurate data. *New York Times*, Dec 2018. www.nytimes.com/2018/12/05/upshot/to-reduce-privacy-risks-the-census-plans-to-report-less-accurate-data.html.

M. Hardt and K. Talwar. On the geometry of differential privacy. In *Proceedings of the Forty-Second ACM Symposium on Theory of Computing*, STOC'10, pages 705–714, New York, NY, USA, 2010. Association for Computing Machinery. https://doi.org/10.1145/1806689.1806786.

J. J. Heckman and V. J. Hotz. Choosing among alternative nonexperimental methods for estimating the impact of social programs: The case of manpower training. *Journal of the American statistical Association*, 84(408):862–874, 1989. https://doi.org/10.1080/01621459.1989.10478848.

V. J. Hotz, C. R. Bollinger, T. Komarova, C. F. Manski, R. A. Moffitt, D. Nekipelov, A. Sojourner, and B. D. Spencer. Balancing data privacy and usability in the federal statistical system. *Proceedings of the National Academy of Sciences*, 119(31):e2104906119, 2022. https://doi.org/10.1073/pnas.2104906119.

P. Joyce and P. Marjoram. Approximately sufficient statistics and Bayesian computation. *Statistical applications in genetics and molecular biology*, 7(1), 2008. https://doi.org/10.2202/1544-6115.1389.

D. Kifer, A. Smith, and A. Thakurta. Private convex empirical risk minimization and high-dimensional regression. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 25.1–25.40, Edinburgh, Scotland, 25–27 Jun 2012. PMLR. https://proceedings.mlr.press/v23/kifer12.html.

R. J. LaLonde. Evaluating the econometric evaluations of training programs with experimental data. *The American economic review*, pages 604–620, 1986. https://www.jstor.org/stable/1806062.

C. Li, G. Miklau, M. Hay, A. McGregor, and V. Rastogi. The matrix mechanism: optimizing linear counting queries under differential privacy. *The VLDB journal*, 24(6):757–781, 2015. https://doi.org/10.1007/s00778-015-0398-x.

R. Little and D. Rubin. *Statistical Analysis with Missing Data*. Wiley Series in Probability and Statistics. Wiley, 2014. https://doi.org/10.1002/9781119013563.

J. S. Liu. *Monte Carlo strategies in scientific computing*. Springer Science & Business Media, 2008. https://doi.org/10.1007/978-0-387-76371-2.

T. A. Louis. Finding the observed information matrix when using the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, 44(2):226–233, 1982. https://doi.org/10.1111/j.2517-6161.1982.tb01203.x.

P. Marjoram, J. Molitor, V. Plagnol, and S. Tavaré. Markov chain Monte Carlo without likelihoods. *Proceedings of the National Academy of Sciences*, 100(26):15324–15328, 2003. https://doi.org/10.1073/pnas.0306899100.

C. E. McCulloch. Maximum likelihood algorithms for generalized linear mixed models. *Journal of the American statistical Association*, 92(437):162–170, 1997. https://doi.org/10.1080/01621459.1997.10473613.

F. McSherry and K. Talwar. Mechanism design via differential privacy. In *48th Annual IEEE Symposium on Foundations of Computer Science (FOCS'07)*, pages 94–103, 2007. https://doi.org/10.1109/FOCS.2007.66.

I. Meilijson. A fast improvement to the EM algorithm on its own terms. *Journal of the Royal Statistical Society: Series B (Methodological)*, 51(1):127–138, 1989. https://doi.org/10.1111/j.2517-6161.1989.tb01754.x.

A. Nikolov, K. Talwar, and L. Zhang. The geometry of differential privacy: the sparse and approximate cases. In *Proceedings of the forty-fifth annual ACM symposium on Theory of computing*, pages 351–360. ACM, 2013. https://doi.org/10.1145/2488608.2488652.

K. Nissim, S. Raskhodnikova, and A. Smith. Smooth sensitivity and sampling in private data analysis. In *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*, pages 75–84. ACM, 2007. https://doi.org/10.1145/1250790.1250803.

M. A. Nunes and D. J. Balding. On optimal selection of summary statistics for approximate Bayesian computation. *Statistical applications in genetics and molecular biology*, 9(1), 2010. https://doi.org/10.2202/1544-6115.1576.

M. Park, J. Foulds, K. Choudhary, and M. Welling. DP-EM: Differentially Private Expectation Maximization. In A. Singh and J. Zhu, editors, *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, volume 54 of *Proceedings of Machine Learning Research*, pages 896–904. PMLR, 20–22 Apr 2017. https://proceedings.mlr.press/v54/park17c.html.

F. A. Quintana, J. S. Liu, and G. E. del Pino. Monte Carlo EM with importance reweighting and its applications in random effects models. *Computational statistics & data analysis*, 29(4):429–444, 1999. https://doi.org/10.1016/S0167-9473(98)00075-9.

A. Schein, Z. S. Wu, A. Schofield, M. Zhou, and H. Wallach. Locally private Bayesian inference for count models. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97, pages 5638–5648. PMLR, 2019. https://proceedings.mlr.press/v97/schein19a.html.

S. A. Sisson, Y. Fan, and M. M. Tanaka. Sequential Monte Carlo without likelihoods. *Proceedings of the National Academy of Sciences*, 104(6):1760–1765, 2007. https://doi.org/10.1073/pnas.0607208104.

S. A. Sisson, Y. Fan, and M. Beaumont. *Handbook of approximate Bayesian computation*. Chapman and Hall/CRC, 2018.

T. Toni, D. Welch, N. Strelkowa, A. Ipsen, and M. P. Stumpf. Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of the Royal Society Interface*, 6(31):187–202, 2008. https://doi.org/10.1098/rsif.2008.0172.

U.S. Census Bureau. 2020 Census: Redistricting file (Public Law 94-171) dataset (Aug 12, 2021), 2021. https://www.census.gov/data/datasets/2020/dec/2020-census-redistricting-summary-file-dataset.html.

Y.-X. Wang, S. Fienberg, and A. Smola. Privacy for free: Posterior sampling and stochastic gradient Monte Carlo. In F. Bach and D. Blei, editors, *Proceedings of the 32nd International Conference on Machine Learning*, volume 37 of *Proceedings of Machine Learning Research*, pages 2493–2502, Lille, France, 07–09 Jul 2015. PMLR. https://proceedings.mlr.press/v37/wangg15.html.

S. L. Warner. Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, 60(309):63–69, 1965. https://doi.org/10.1080/01621459.1965.10480775.

D. Wegmann, C. Leuenberger, and L. Excoffier. Efficient approximate Bayesian computation coupled with Markov chain Monte Carlo without likelihood. *Genetics*, 182(4):1207–1218, 2009. https://doi.org/10.1534/genetics.109.102509.

G. C. Wei and M. A. Tanner. A Monte Carlo implementation of the EM algorithm and the poor man's data augmentation algorithms. *Journal of the American statistical Association*, 85(411):699–704, 1990. https://doi.org/10.1080/01621459.1990.10474930.

R. D. Wilkinson. Approximate Bayesian computation (ABC) gives exact results under the assumption of model error. *Statistical applications in genetics and molecular biology*, 12 (2):129–141, 2013. https://doi.org/10.1515/sagmb-2013-0010.

R. Wolfinger and M. O'Connell. Generalized linear mixed models a pseudo-likelihood approach. *Journal of statistical Computation and Simulation*, 48(3-4):233–243, 1993. https://doi.org/10.1080/00949659308811554.

S. N. Wood. Statistical inference for noisy nonlinear ecological dynamic systems. *Nature*, 466(7310):1102, 2010. https://doi.org/10.1038/nature09319.

Appendix A. Examples of additive perturbation DP mechanisms

Example 1 (ϵ-Laplace mechanism; Dwork et al. (2006))**.** In (2.3), let $\boldsymbol{u} \sim \text{Lap}_p(1)$, the p-dimensional product of independent and identically distributed standard Laplace variables, and $h = \epsilon^{-1}\Delta_{GS}(\boldsymbol{s})$, where

$$\Delta_{GS}(\boldsymbol{s}) = \sup_{\boldsymbol{x},\boldsymbol{x}'} \left\{ \left\| \boldsymbol{s}\left(\boldsymbol{x}\right) - \boldsymbol{s}\left(\boldsymbol{x}'\right) \right\| : d\left(\boldsymbol{x},\boldsymbol{x}'\right) = 1 \right\}, \tag{A.1}$$

is the global sensitivity of $\boldsymbol{s}$, with $\|\cdot\|$ denoting the ℓ_1 norm. Then, $\boldsymbol{s}_{\text{dp}}$ is ϵ-differentially private.

Example 2 ((ϵ,δ)-Laplace mechanism; Nissim et al. (2007))**.** In (2.3), let $\boldsymbol{u} \sim \text{Lap}_p(1)$, $h = \epsilon^{-1}\Delta_\xi(\boldsymbol{s},\boldsymbol{x})$, and $\xi = \epsilon\left\{4\left(p + \log\left(2/\delta\right)\right)\right\}^{-1}$, where

$$\Delta_\xi(\boldsymbol{s},\boldsymbol{x}) = \sup_{\boldsymbol{x}'} \left\{ e^{-\xi d(\boldsymbol{x},\boldsymbol{x}')} \Delta_{LS}\left(\boldsymbol{s},\boldsymbol{x}'\right) : \boldsymbol{x}' \in \mathcal{X} \right\} \tag{A.2}$$

is the ξ-smooth sensitivity ($\xi > 0$) of $\boldsymbol{s}$ at $\boldsymbol{x}$, and

$$\Delta_{LS}(\boldsymbol{s},\boldsymbol{x}) = \sup_{\boldsymbol{x}'} \left\{ \left\| \boldsymbol{s}\left(\boldsymbol{x}\right) - \boldsymbol{s}\left(\boldsymbol{x}'\right) \right\| : d\left(\boldsymbol{x},\boldsymbol{x}'\right) = 1 \right\} \tag{A.3}$$

is the local sensitivity of $\boldsymbol{s}$ at $\boldsymbol{x}$. Then, $\boldsymbol{s}_{\text{dp}}$ is (ϵ,δ)-differentially private.

Example 3 (Gaussian mechanism; Blum et al. (2005); Nissim et al. (2007))**.** In (2.3), let $\boldsymbol{u} \sim N(\boldsymbol{0},\mathbf{I}_p)$ the p-dimensional standard multivariate Normal variable, $h = \epsilon^{-1}5\sqrt{2\log(2/\delta)}\Delta_\xi(\boldsymbol{s},\boldsymbol{x})$, and $\xi = \epsilon\left\{4\left(p + \log\left(2/\delta\right)\right)\right\}^{-1}$. Then, $\boldsymbol{s}_{\text{dp}}$ is (ϵ,δ)-differentially private.

The above examples invoke three notions of functional sensitivity (A.1)-(A.3), generally denoted as $\Delta(\boldsymbol{s})$, to capture the idea that certain choices of $\boldsymbol{s}$ may be more revealing of individual information in $\boldsymbol{x}$ than others. The global sensitivity measures the extent to which $\boldsymbol{s}$ varies between all conceivable pairs of neighboring datasets, whether or not realized in the observed sample. For example, the global sensitivity of the counting query is 1. On the other hand, the local sensitivity of $\boldsymbol{s}$ measures its maximum variability among neighboring datasets to a given observed dataset $\boldsymbol{x}$. The smooth sensitivity strikes a balance between the two, by providing an upper bound on the local sensitivity at $\boldsymbol{x}$ in such a way that the bound does not vary too quickly as a function of $\boldsymbol{x}$. It is crucial that the scale parameter of the additive perturbation mechanism is chosen as a function of both the sensitivity of $\boldsymbol{s}$ as well as the privacy budget, that is, $h = h(\epsilon,\delta,\Delta(\boldsymbol{s}))$.

Appendix B. Proof of Theorem 3.1

Proof. Let I be the indicator of the event that a draw of θ is accepted. The joint distribution of all quantities produced by the ith iteration is $\tilde{\pi}(\theta,\boldsymbol{s},I) = \pi_0(\theta)\pi(\boldsymbol{s} \mid \theta)\tilde{\pi}(I \mid \boldsymbol{s})$, where $\tilde{\pi}(I \mid \boldsymbol{s})$ is the Bernoulli mass function with proportion parameter $c\eta_{\text{dp}}\left(\boldsymbol{s}_{\text{dp}} \mid \boldsymbol{s}\right)$. The marginal distribution of an accepted θ sample is

$$\tilde{\pi}\left(\theta \mid I = 1\right) = \int \frac{\tilde{\pi}\left(\theta,\boldsymbol{s},I=1\right)}{\tilde{\pi}\left(I=1\right)} d\boldsymbol{s} = \frac{\int \pi_0\left(\theta\right)\pi\left(\boldsymbol{s} \mid \theta\right) c\eta_{\text{dp}}\left(\boldsymbol{s}_{\text{dp}} \mid \boldsymbol{s}\right) d\boldsymbol{s}}{\int\int \pi_0\left(\theta\right)\pi\left(\boldsymbol{s} \mid \theta\right) c\eta_{\text{dp}}\left(\boldsymbol{s}_{\text{dp}} \mid \boldsymbol{s}\right) d\boldsymbol{s} d\theta}, \tag{B.1}$$

which is equal to $\pi\left(\theta \mid \boldsymbol{s}_{\text{dp}}\right)$ as defined in (3.2). From here, one can see that the overall acceptance probability of Algorithm 1 is

$$\tilde{\pi}(I=1) = \pi\left(\boldsymbol{s}_{\text{dp}}\right) / \max \eta_{\text{dp}}(\cdot).$$

□

Note that under the special case of additive perturbation, the proof of Theorem 3.1 parallels Theorem 1 of Wilkinson (2013). However, there is an important conceptual difference. In Wilkinson (2013), the conditioning query is a query that was observed noiselessly, but construed as if subject to additive error. The ABC-induced posterior of θ therein, while essentially identical to (3.2), is not the true posterior of θ but that of a "best model input $\hat{\theta}$" given $\boldsymbol{s}_{\mathrm{dp}}$. With $\boldsymbol{s}_{\mathrm{dp}}$ being a privatized query, no pretense is necessary in treating it as observed with error, since it indeed was.

Appendix C. Effective sample size for Monte Carlo EM

In reference to Algorithm 3, at the tth iteration, the normalized version of the importance sampling weights is

$$\tilde{\omega}_i = c_{(t)}\eta_{\mathrm{dp}}\left(\boldsymbol{s}_{\mathrm{dp}} \mid \boldsymbol{s}_i\right) = c_{(t)}\omega_i$$

where $c_{(t)} = 1/\pi(\boldsymbol{s}_{\mathrm{dp}} \mid \theta^{(t)})$ is the reciprocal of the current approximation to the observed likelihood and is free of $\boldsymbol{s}_i$. The weighted estimator $\sum_{i=1}^{N} \tilde{\omega}_i b\left(\boldsymbol{s}_i\right)$ is a consistent estimator of (4.2) because

$$\begin{aligned}
\mathbb{E}\left(b\left(\boldsymbol{s}\right) \mid \boldsymbol{s}_{\mathrm{dp}}, \theta^{(t)}\right) &= \int b\left(\boldsymbol{s}\right)\pi\left(\boldsymbol{s} \mid \boldsymbol{s}_{\mathrm{dp}}, \theta^{(t)}\right) d\boldsymbol{s} \\
&= \int b\left(\boldsymbol{s}\right)\frac{\pi\left(\boldsymbol{s} \mid \theta^{(t)}\right)\eta_{\mathrm{dp}}\left(\boldsymbol{s}_{\mathrm{dp}} \mid \boldsymbol{s}\right)}{\int \pi\left(\boldsymbol{s} \mid \theta^{(t)}\right)\eta_{\mathrm{dp}}\left(\boldsymbol{s}_{\mathrm{dp}} \mid \boldsymbol{s}\right) d\boldsymbol{s}} d\boldsymbol{s} \\
&= \int \tilde{\omega}\left(\boldsymbol{s}\right) b\left(\boldsymbol{s}\right)\pi\left(\boldsymbol{s} \mid \theta^{(t)}\right) d\boldsymbol{s}.
\end{aligned}$$

We have that at the expectation of weights for the tth iteration is

$$\begin{aligned}
\mathbb{E}_{\boldsymbol{s}\mid\theta^{(t)}}\left(\tilde{\omega}\right) &= \int c_{(t)}\eta_{\mathrm{dp}}\left(\boldsymbol{s}_{\mathrm{dp}} \mid \boldsymbol{s}\right)\pi(\boldsymbol{s} \mid \theta^{(t)}) d\boldsymbol{s} \\
&= c_{(t)}\pi(\boldsymbol{s}_{\mathrm{dp}} \mid \theta^{(t)}) = 1,
\end{aligned}$$

where the subscript $\boldsymbol{s} \mid \theta^{(t)}$ signifies the expectation is evaluated with respect to the current approximation to the latent data likelihood, or equivalently, the proposal distribution of the importance sampler. Similarly,

$$\begin{aligned}
\mathrm{VAR}_{\boldsymbol{s}\mid\theta^{(t)}}\left(\tilde{\omega}\right) &= \mathbb{E}_{\boldsymbol{s}\mid\theta^{(t)}}\left(\tilde{\omega}^2\right) - \mathbb{E}^2_{\boldsymbol{s}\mid\theta^{(t)}}\left(\tilde{\omega}\right) \\
&= c^2_{(t)}\mathbb{E}_{\boldsymbol{s}\mid\theta^{(t)}}(\eta^2_{\mathrm{dp}}\left(\boldsymbol{s}_{\mathrm{dp}} \mid \boldsymbol{s}\right)) - 1.
\end{aligned}$$

This gives rise to the effective sample size

$$\begin{aligned}
\mathrm{ESS}^{(t)}\left(N\right) &= N/\left(1 + \mathrm{VAR}_{\boldsymbol{s}\mid\theta^{(t)}}\left(\tilde{\omega}\right)\right) \\
&= N\pi^2\left(\boldsymbol{s}_{\mathrm{dp}} \mid \theta^{(t)}\right)\mathbb{E}^{-1}_{\boldsymbol{s}\mid\theta^{(t)}}\left(\eta^2_{\mathrm{dp}}\left(\boldsymbol{s}_{\mathrm{dp}} \mid \boldsymbol{s}\right)\right).
\end{aligned}$$

See also section 2.5.3 of (Liu, 2008).

Appendix D. Observed score and Fisher information for Monte Carlo EM

We have that the observed data log likelihood

$$\log \pi\left(\boldsymbol{s}_{\mathrm{dp}} \mid \theta\right)=\log \int \pi\left(\boldsymbol{s}_{\mathrm{dp}} \mid \boldsymbol{s}\right) \pi\left(\boldsymbol{s} \mid \theta\right) d \boldsymbol{s},$$

thus the observed score

$$\begin{aligned}
\nabla_\theta \log \pi\left(\boldsymbol{s}_{\mathrm{dp}} \mid \theta\right) &= \frac{\int \pi\left(\boldsymbol{s}_{\mathrm{dp}} \mid \boldsymbol{s}\right) \nabla_\theta \pi\left(\boldsymbol{s} \mid \theta\right) d \boldsymbol{s}}{\int \pi\left(\boldsymbol{s}_{\mathrm{dp}} \mid \boldsymbol{s}\right) \pi\left(\boldsymbol{s} \mid \theta\right) d \boldsymbol{s}} \\
&= \frac{\int \frac{\pi\left(\boldsymbol{s}_{\mathrm{dp}} \mid \boldsymbol{s}\right) \nabla_\theta \pi(\boldsymbol{s} \mid \theta)}{\pi\left(\boldsymbol{s}_{\mathrm{dp}} \mid \boldsymbol{s}\right) \pi(\boldsymbol{s} \mid \theta)} \pi\left(\boldsymbol{s}_{\mathrm{dp}} \mid \boldsymbol{s}\right) \pi\left(\boldsymbol{s} \mid \theta\right) d \boldsymbol{s}}{\int \pi\left(\boldsymbol{s}_{\mathrm{dp}} \mid \boldsymbol{s}\right) \pi\left(\boldsymbol{s} \mid \theta\right) d \boldsymbol{s}} \\
&= \int \frac{\nabla_\theta \pi\left(\boldsymbol{s} \mid \theta\right)}{\pi\left(\boldsymbol{s} \mid \theta\right)} \frac{\pi\left(\boldsymbol{s}_{\mathrm{dp}}, \boldsymbol{s} \mid \theta\right)}{\int \pi\left(\boldsymbol{s}_{\mathrm{dp}}, \boldsymbol{s} \mid \theta\right) d \boldsymbol{s}} d \boldsymbol{s} \\
&= \int \nabla_\theta \log \pi\left(\boldsymbol{s} \mid \theta\right) \pi\left(\boldsymbol{s} \mid \boldsymbol{s}_{\mathrm{dp}}, \theta\right) d \boldsymbol{s} \\
&= \mathbb{E}\left(\nabla_\theta \log \pi\left(\boldsymbol{s} \mid \theta\right) \mid \boldsymbol{s}_{\mathrm{dp}}, \theta\right).
\end{aligned}$$

Writing $\lambda_\theta(\boldsymbol{s})=\nabla_\theta \log \pi\left(\boldsymbol{s} \mid \theta\right)$, we have that $\mathbb{E}\left(\lambda_\theta(\boldsymbol{s}) \mid \boldsymbol{s}_{\mathrm{dp}}, \theta^{(t)}\right)$ serves as the tth approximation to the observed score $\nabla_\theta \log \pi\left(\boldsymbol{s}_{\mathrm{dp}} \mid \theta^{(t)}\right)$, giving rise to the expression

$$\mathbb{E}\left(\lambda_\theta(\boldsymbol{s}) \mid \boldsymbol{s}_{\mathrm{dp}}, \theta^{(t)}\right) \approx m \sum_{i=1}^N \omega_i \lambda_\theta(\boldsymbol{s}_i).$$

Similarly the Hessian, or the negative of the observed Fisher information matrix, is

$$\begin{aligned}
\nabla_\theta^2 \log \pi\left(\boldsymbol{s}_{\mathrm{dp}} \mid \theta\right) &= \int \frac{\nabla_\theta^2 \pi\left(\boldsymbol{s} \mid \theta\right)}{\pi\left(\boldsymbol{s} \mid \theta\right)} \pi\left(\boldsymbol{s} \mid \boldsymbol{s}_{\mathrm{dp}}, \theta^{(t)}\right) d \boldsymbol{s} \\
&\quad - \left(\nabla_\theta \log \pi\left(\boldsymbol{s}_{\mathrm{dp}} \mid \theta\right)\right)\left(\nabla_\theta \log \pi\left(\boldsymbol{s}_{\mathrm{dp}} \mid \theta\right)\right)^\top \\
&= \mathbb{E}\left(\nabla_\theta^2 \log \pi\left(\boldsymbol{s} \mid \theta\right)+\nabla_\theta \log \pi\left(\boldsymbol{s} \mid \theta\right) \nabla_\theta \log \pi\left(\boldsymbol{s} \mid \theta\right)^\top \mid \boldsymbol{s}_{\mathrm{dp}}, \theta\right) \\
&\quad - \left(\nabla_\theta \log \pi\left(\boldsymbol{s}_{\mathrm{dp}} \mid \theta\right)\right)\left(\nabla_\theta \log \pi\left(\boldsymbol{s}_{\mathrm{dp}} \mid \theta\right)\right)^\top.
\end{aligned}$$

Substituting again $\lambda_\theta(\boldsymbol{s})$ and the expression for the observed score into the above equation, we have that the tth approximation to the observed Fisher information $-\nabla_\theta^2 \log \pi\left(\boldsymbol{s}_{\mathrm{dp}} \mid \theta^{(t)}\right)$ takes the form

$$\begin{aligned}
&\mathbb{E}\left(-\nabla_\theta \lambda_\theta\left(\boldsymbol{s}\right)-\lambda_\theta\left(\boldsymbol{s}\right) \lambda_\theta\left(\boldsymbol{s}\right)^\top \mid \boldsymbol{s}_{\mathrm{dp}}, \theta^{(t)}\right)+\mathbb{E}\left(\lambda_\theta\left(\boldsymbol{s}\right) \mid \boldsymbol{s}_{\mathrm{dp}}, \theta^{(t)}\right) \mathbb{E}\left(\lambda_\theta\left(\boldsymbol{s}\right) \mid \boldsymbol{s}_{\mathrm{dp}}, \theta^{(t)}\right)^\top \\
&\approx m \sum_{i=1}^N \omega_i\left\{-\nabla_\theta \lambda_\theta\left(\boldsymbol{s}_i\right)-\lambda_\theta\left(\boldsymbol{s}_i\right) \lambda_\theta\left(\boldsymbol{s}_i\right)^\top\right\}+m^2 \sum_{i=1}^N \sum_{j=1}^N \omega_i \omega_j \lambda_\theta\left(\boldsymbol{s}_i\right) \lambda_\theta\left(\boldsymbol{s}_j\right)^\top.
\end{aligned}$$

See also the appendix of Louis (1982).

Appendix E. Details of Section 5.1: privatized count inference

For the Bayesian analysis, by the ϵ-Laplace perturbation mechanism, the conditional distribution of $\boldsymbol{s}_{\text{dp}}$ given $\boldsymbol{s}$ is $Lap\left(\boldsymbol{s}, \epsilon^{-1}\right)$, which has density $\frac{\epsilon}{2}\exp\left(-\epsilon\left|\boldsymbol{s}_{\text{dp}}-\boldsymbol{s}\right|\right)$. By construction, $\boldsymbol{s}_{\text{dp}}$ is not an integer with probability one, hence

$$\begin{aligned}\pi\left(\boldsymbol{s}_{\text{dp}} \mid \theta\right) &= \int \pi\left(\boldsymbol{s} \mid \theta\right)\pi\left(\boldsymbol{s}_{\text{dp}} \mid \boldsymbol{s}\right)d\boldsymbol{s} \\ &\propto e^{-\theta}\left\{\sum_{s=0}^{\lfloor \boldsymbol{s}_{\text{dp}}\rfloor}\frac{\theta^{\boldsymbol{s}}}{\boldsymbol{s}!}e^{-\epsilon\boldsymbol{s}_{\text{dp}}+\epsilon\boldsymbol{s}} + \sum_{s=\lceil \boldsymbol{s}_{\text{dp}}\rceil}^{\infty}\frac{\theta^{\boldsymbol{s}}}{\boldsymbol{s}!}e^{\epsilon\boldsymbol{s}_{\text{dp}}-\epsilon\boldsymbol{s}}\right\}.\end{aligned}$$

Adopting the notations $\theta_\epsilon^+ = \theta e^{\epsilon}$ and $\theta_\epsilon^- = \theta e^{-\epsilon}$, the first sum within the brackets can be written as

$$e^{-\epsilon\boldsymbol{s}_{\text{dp}}}\sum_{s=0}^{\lfloor \boldsymbol{s}_{\text{dp}}\rfloor}\frac{\left(\theta_\epsilon^+\right)^{\boldsymbol{s}}}{\boldsymbol{s}!} = e^{\theta_\epsilon^+-\epsilon\boldsymbol{s}_{\text{dp}}}F_{\theta_\epsilon^+}\left(\lfloor \boldsymbol{s}_{\text{dp}}\rfloor\right)$$

where $F_\lambda\left(a\right)$ stands for the $Pois\left(\lambda\right)$ CDF evaluated at a. Similarly, the second sum can be written as

$$e^{\epsilon\boldsymbol{s}_{\text{dp}}}\sum_{s=\lceil \boldsymbol{s}_{\text{dp}}\rceil}^{\infty}\frac{\left(\theta_\epsilon^-\right)^{\boldsymbol{s}}}{\boldsymbol{s}!} = e^{\theta_\epsilon^-+\epsilon\boldsymbol{s}_{\text{dp}}}\left(1-F_{\theta_\epsilon^-}\left(\lfloor \boldsymbol{s}_{\text{dp}}\rfloor\right)\right).$$

Combining the above with the Gamma prior, $\pi_0\left(\theta\right)\propto\theta^{\alpha-1}e^{-\beta\theta}$, we have that the posterior $\pi\left(\theta \mid \boldsymbol{s}_{\text{dp}}\right)$ takes the form

$$\pi\left(\theta \mid \boldsymbol{s}_{\text{dp}}\right)\propto\theta^{\alpha-1}e^{-(\beta+1)\theta}\left[\frac{\Gamma\left(\lceil \boldsymbol{s}_{\text{dp}}\rceil, \theta_\epsilon^+\right)}{\Gamma\left(\lceil \boldsymbol{s}_{\text{dp}}\rceil\right)}e^{\theta_\epsilon^+-\epsilon\boldsymbol{s}_{\text{dp}}} + \frac{\gamma\left(\lceil \boldsymbol{s}_{\text{dp}}\rceil, \theta_\epsilon^-\right)}{\Gamma\left(\lceil \boldsymbol{s}_{\text{dp}}\rceil\right)}e^{\theta_\epsilon^-+\epsilon\boldsymbol{s}_{\text{dp}}}\right],$$

where $\theta_\epsilon^+ = \theta e^{\epsilon}$, $\theta_\epsilon^- = \theta e^{-\epsilon}$, $\lceil\cdot\rceil$ is the ceiling function, and $\Gamma\left(s,x\right) = \int_x^\infty r^{s-1}e^{-r}dr$ is the incomplete Gamma function with $\Gamma\left(s\right) = \Gamma\left(s,0\right)$ and $\gamma\left(s,x\right) = \Gamma\left(s\right) - \Gamma\left(s,x\right)$.

Figure 3 displays additional comparisons between the true posterior and the naïve posterior for the same privatized count under other choices of prior distributions in the Gamma family. Notice that when the observed count appears highly unlikely under a chosen prior (such as $Gamma(2,1)$ or $Gamma(5,1)$), a situation known as *prior-data conflict* (Evans and Moshonov, 2006), the correct posterior heavily discounts the contribution by the privatized observation. The discounting can be seen from the close alignment between the correct posterior (represented by either the solid green density or the black histogram) and the prior (blue dashed density), which in contrast differ drastically from the naïve posterior (red dotted density) in Figure 3 (a) and (b). The acceptance rate of Algorithm 1 in these situations are reported in Table 1.

For the implementation of the Monte Carlo EM, three stages of iterations were performed with successively more stringent tolerance levels ($\left|\theta^{(t)}-\theta^{(t-1)}\right| < 10^{-3}$, 10^{-4}, and 10^{-5}) and larger Monte Carlo sample size ($N = 10^3, 10^5$, and 10^7), using the stable maximizer from the last stage as the starting point. This is a crude rule to let N increase, hence the Monte Carlo error decrease, as $\theta^{(t)}$ approaches the true MLE. Advanced adaptive techniques, such as the ascent-based modification of Caffo et al. (2005), can be employed achieve better performance.

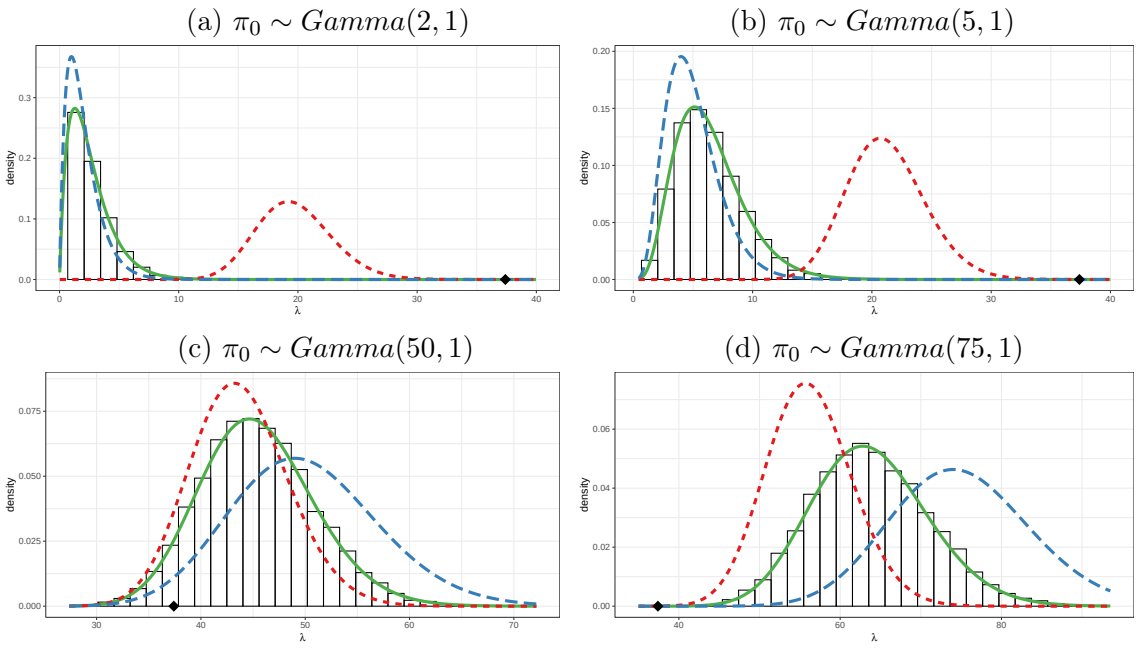

Figure 3: Comparisons between the true posterior (green density; approximated by exact draws as black histogram, $N = 10^4$) and the naïve posterior (red dotted density) treating observed $\boldsymbol{s}_{\text{dp}} = 37.4$ (black diamond) as if without privatization, under four different choices of prior distribution π_0 (blue dashed density).

Appendix F. Details of Section 5.2: Lalonde dataset

Let z_i be the indicator variable for whether subject i received treatment ($z_i = 1$) or control ($z_i = 0$), and y_i the earning in 1978 (in \$1k). The full parameter of the model is $\theta = \left(\tau, \mu, \sigma_t^2, \sigma_c^2\right)$, for which we posit independent priors

$$\theta \sim \pi_0(\tau) \times \pi_0(\mu) \times \pi_0\left(\sigma_t^2\right) \times \pi_0\left(\sigma_c^2\right),$$

where for concreteness, we use $\pi_0(\tau) \sim N(0,5)$, $\pi_0(\mu) \sim N(4,5)$, $\pi_0(\sigma_t) \sim Gamma(2, 0.2)$ and $\pi_0(\sigma_c) \sim Gamma(2, 0.2)$ for the analysis. The sampling model is

$$y_i \mid z_i, \theta \sim N\left(\tau z_i + \mu, \sigma_t^2 z_i + \sigma_c^2 (1 - z_i)\right),$$

where τ is the difference in average earnings between the treatment and control groups. Equivalently stated, treatment group earnings have the distribution $N\left(\mu + \tau, \sigma_t^2\right)$ and the control group earnings have distribution $N\left(\mu, \sigma_c^2\right)$.

The sufficient statistics for θ are the within-group mean and sample variances

$$\boldsymbol{s} = \left(\bar{y}_t, \bar{y}_c, s_t^2, s_c^2\right) = \left(\frac{1}{n_t}\sum_{i:z_i=1} y_i, \frac{1}{n_c}\sum_{i:z_i=0} y_i, \frac{1}{n_t - 1}\sum_{i:z_i=1}(y_i - \bar{y}_t)^2, \frac{1}{n_c - 1}\sum_{i:z_i=0}(y_i - \bar{y}_c)^2\right).$$

Due to statistical independence of the sample mean and variance of normal random variables, the likelihood can be equivalently represented by the generative model

$$\bar{y}_t, \bar{y}_c, s_t^2, s_c^2 \mid z, \theta \sim N\left(\mu + \tau, \frac{\sigma_t^2}{n_t}\right) \times N\left(\mu, \frac{\sigma_c^2}{n_c}\right) \times \frac{\sigma_t^2}{n_{t-1}}\chi^2_{n_t - 1} \times \frac{\sigma_c^2}{n_{c-1}}\chi^2_{n_c - 1}.$$

Through a conservative clamping treatment described in Section 5.2, the ϵ-differentially private statistic $\boldsymbol{s}_{\text{dp}}$ is obtained via a Laplace mechanism with independent Laplace noise components with bandwidth $h^{-1} = (1/3, 1/3, 1/6, 1/6)$ corresponding to $\boldsymbol{s}$. Since the clamping range well exceeds the anticipated range of observable data, we do not perform inferential correction for truncation. Both the original analysis using $\boldsymbol{s}$ (top row of Figure 2) and the naïve analysis using $\boldsymbol{s}_{\text{dp}}$ (bottom row of Figure 2) are carried out in `RStan`, whereas the correct analysis (middle row of Figure 2) is carried out using rejection ABC of Algorithm 1.