# Differential Privacy: General Inferential Limits via Intervals of Measures

**James Bailie** JAMESBAILIE@G.HARVARD.EDU
*Department of Statistics, Harvard University, USA*

**Ruobin Gong** RUOBIN.GONG@RUTGERS.EDU
*Department of Statistics, Rutgers University, USA*

## Abstract

Differential privacy (DP) is a mathematical standard for assessing the privacy provided by a data-release mechanism. We provide formulations of pure $\epsilon$-differential privacy first as a Lipschitz continuity condition and then using an object from the imprecise probability literature: the interval of measures. We utilise this second formulation to establish bounds on the appropriate likelihood function for $\epsilon$-DP data – and in turn derive limits on key quantities in both frequentist hypothesis testing and Bayesian inference. Under very mild conditions, these results are valid for arbitrary parameters, priors and data generating models. These bounds are weaker than those attainable when analysing specific data generating models or data-release mechanisms. However, they provide generally applicable limits on the ability to learn from differentially private data – even when the analyst's knowledge of the model or mechanism is limited. They also shed light on the semantic interpretation of differential privacy, a subject of contention in the current literature.

**Keywords:** disclosure risk, prior-to-posterior semantics, Neyman-Pearson hypothesis testing, Lipschitz continuity, multiplicative distance, transparency, statistical disclosure control

## 1. Introduction

The world today is witnessing an explosive growth of large-scale datasets containing personal information. Demographic and economic surveys, biomedical studies and massive online service platforms facilitate understanding of human biological functions and socio-behavioral environments. At the same time, they pose the risk of exposing confidential information about data contributors. Breaches of privacy can happen counter-intuitively and without malice. For example, [37] demonstrated that even coarsely aggregated SNP (single-nucleotide polymorphisms [47]) data from genome-wide association studies (GWAS) can still reliably reveal individual participants. This unsettling revelation led to the decision by the U.S. National Institute of Health to remove aggregate SNP data from open-access databases [64]. This incident, and similar occurrences across science, government and industry [54, 28, 17, 57], have attracted public attention and sparked debate about privacy-preserving data curation and dissemination.

Commensurate with the increasing risk of privacy breaches, the recent decades have also seen rapid advances in formal approaches to statistical disclosure limitation (SDL). These methodologies supply a solid mathematical foundation for endeavors that enhance confidentiality protection without undue sacrifice to data quality. Notably, *differential privacy* (DP) [26] puts forth a rigorous and practical standard for assessing the level of privacy provided by a data release. Many large IT companies including Google [29], Apple [3], and Microsoft [19] have been early adopters of differential privacy. More recently, the U.S. Census Bureau deployed differential privacy to protect data publications of the 2020 Decennial Census [2]. The U.S. Internal Revenue Service is also exploring differentially private synthetic data methods for the publication of individual tax data [11]. These decisions by statistical agencies and corporations showcase the growing popularity of differential privacy among major data curators.

Innovations in privacy protection methods have prompted quantitative researchers to confront a new reality, as existing modes, practices and expectations of data access are subject to renewal. We highlight two points of tension in this development. First, DP promises *transparency*, in the sense that the design details about the protection method can be made public without compromising privacy. Transparency is one of the advantages of DP over tranditional SDC methods since it supports valid statistical inference by providing the analyst with the ability to model the privacy noise. However, this promise often falls short in practice, leaving the statistician with tied hands [33]. Second, following the high-profile adoption of differential privacy by the U.S. Census Bureau, a debate ensued concerning its interpretation, or *semantics*, as well as its reconciliation with other notions of statistical disclosure risk; see e.g. [44, 40, 46].

The current work takes multiple steps toward the resolution of these debates by examining DP via the lens of imprecise probabilities (IP). First we describe the classic notion of $\epsilon$-differential privacy as a Lipschitz continuity condition (Section 2). This enables a re-interpretation of

$\epsilon$-DP as an *interval of measures* [18] induced by the data-release mechanism (Section 3). From here, we derive some implications of this interpretation on the problem of statistical inference using privacy-protected data releases. These results concern the probability model of the observable privatised data (Section 4), as well as frequentist hypothesis testing (Section 5) and Bayesian posterior inference (Section 6) using these data. These results establish bounds on key inferential objects while having general validity under mild assumptions about the data model, privacy mechanism, and analyst's prior (when applicable).

Within the current literature on DP, studies that leverage tools from IP are sparse at best. We mention two branches of work that are known to us. In [50], the authors propose novel definitions of local differential privacy using the language of belief functions, and in subsequent work examine constraints on DP mechanisms as belief revision and update [51]. Also, [48] examine the issue of partial identification in inference from privacy-protected data, where in certain situations the identification set can be described with a belief function. The goals and approaches undertaken by the current paper are substantially different from these existing lines of work.

## 2. Differential Privacy

Define the data universe $\mathcal{X}$ as the set of all theoretically-possible observable datasets. Let $d$ be a metric on $\mathcal{X}$. Given data $x \in \mathcal{X}$, consider releasing some (potentially randomised) summary statistics $t \in \mathcal{T}$. To formalise this, equip $\mathcal{T}$ with a $\sigma$-algebra $\mathscr{F}$ and define a *data-release mechanism* as a function $M : \mathcal{X} \times [0,1] \to \mathcal{T}$ (whose output is the summary statistic $t$) such that $M(x,\cdot)$ is $(\mathcal{B}[0,1], \mathscr{F})$-measurable for each $x \in \mathcal{X}$. ($\mathcal{B}[0,1]$ is the Borel $\sigma$-algebra on $[0,1]$.) A distribution on the *seed* $U$ induces a probability on the summary statistic $t = M(x,U)$. Without loss of generality, we may take $U \sim \text{Unif}[0,1]$. Denote by $P_x$ the probability measure of $M(x,U)$ induced by $U$, taking $x$ as fixed.

The realised value of the seed $U$ and the observed data $x$ are assumed to remain secret, while all other details of $M$ (including the distribution of $U$) may – and should – be made public [33]. An attacker is tasked with inferring $x$ based on observing a draw $t = M(x,U) \sim P_x$. This set-up is analogous to fiducial inference, with $x$ taking the role of the parameters, $t$ the data, and $M$ the data-generating equation [35].

Differential privacy is a Lipschitz condition on $M$:

**Definition 1** *Given a data universe $\mathcal{X}$ equipped with a metric $d$, a data-release mechanism $M : \mathcal{X} \times [0,1] \to \mathcal{T}$ satisfies* $\epsilon$-differential privacy *if, for all $x, x' \in \mathcal{X}$,*

$$d_{MULT}(P_x, P_{x'}) \leq \epsilon d(x, x'), \tag{1}$$

*where*

$$d_{MULT}(P,Q) = \sup\left\{\left|\ln \frac{P(S)}{Q(S)}\right| : S \in \mathscr{F}\right\},$$

*is the* multiplicative distance[1] *between measures $P, Q$ on $(\mathcal{T}, \mathscr{F})$.*

The Lipschitz constant $\epsilon \geq 0$ is called the *privacy loss budget*. Larger $\epsilon$ intuitively corresponds to less privacy; smaller $\epsilon$ gives stronger privacy protection. A tenet of differential privacy (in contrast with other statistical disclosure risk frameworks) is that any dependence of $M(x,U)$ on $x$ implies non-negligible privacy loss $\epsilon > 0$. Since $d_{\text{MULT}}(P,Q) = 0$ if and only if $P = Q$, perfect privacy (i.e. $\epsilon = 0$) is only possible by releasing pure noise.

Two common choices of the metric $d$ on $\mathcal{X}$ are

A) the Hamming distance

$$d_{\text{HAM}}(x,x') = \begin{cases} \sum_{i=1}^{n} 1_{x_i \neq x'_i} & \text{if } |x| = |x'| = n, \\ \infty & \text{otherwise,} \end{cases}$$

where the data $x = (x_1, x_2, \ldots, x_n)$ are vectors and $|x|$ is the size of $x$; and

B) the symmetric difference metric

$$d_{\triangle}(x,x') = |x \setminus x'| + |x' \setminus x|,$$

where the data $x, x' \in \mathcal{X}$ are multisets and $x \setminus x'$ is the (multi-)set difference.[2]

Equation (1) with $d$ the Hamming distance is referred to as *bounded* DP and with the symmetric difference as *unbounded* DP.

The intuition behind differential privacy considers each record $x_i$ in the data $x$ as representing a single distinct individual. A distance $d(x,x') = 1$ then implies that $x$ and $x'$ differ according to the change in behaviour of a single individual – a change in the individual's response, for the Hamming distance; or a change in whether the individual responds or not, for the symmetric difference metric. $\epsilon$-DP implies that a single individual can change the summary statistic $M(x,U)$ by at most $\epsilon$, where "change" is interpreted probabilistically in terms of the multiplicative distance.

Under the mild assumption that $d$ is a graph distance with unit edges (Assumption 4, given in Section 3), the

[1] In defining $d_{\text{MULT}}$ we set $0/0 = 1$.
$d_{\text{MULT}}$ is strongly equivalent to the *density ratio metric* $\delta$ [61]:

$$d_{\text{MULT}}(P,Q) < \delta(P,Q) \leq 2 d_{\text{MULT}}(P,Q),$$

so that $\epsilon$-DP can be defined with $\delta$ in place of $d_{\text{MULT}}$, up to rescaling of $\epsilon$.

[2] We formally define a multiset $S$ as a function $\text{Dom}(S) \to \mathbb{N}^{\geq 0}$ with $S(a)$ denoting the number of times the element $a$ appears in $S$. The multiset difference is defined as $(S \setminus S')(a) = \max\{0, S(a) - S'(a)\}$ and the multiset cardinality is defined as $|S| = \sum_{a \in \text{Dom}(S)} S(a)$.

converse implication also holds. That is, $\epsilon$-DP is equivalent to the Lipschitz condition (1) holding when $d(x, x') = 1$. (This follows by the triangle inequality; for details see the proof of Theorem 5.) From herein, we restrict our attention to the set of such metrics, which includes $d_{\text{HAM}}$ and $d_{\triangle}$.

Since DP controls the change in $t$ due to perturbations in the data $x$, it can naturally be understood as a robustness property [23]. This insight hints at its connections to IP. Measuring the "change in $t$" with the multiplicative distance $d_{\text{MULT}}$ – as compared to more familiar metrics typically seen in the robustness literature such as Kolmogorov or total variation distance – is motivated by a strong notion of privacy as indistinguishability. The formulation of DP as an interval of measure makes this motivation clear. We therefore postpone the remainder of this discussion until the end of Section 3.

**Example 1 (Laplace mechanism [26])** *Consider the problem of releasing a sanitised version of a deterministic summary statistic $q : \mathcal{X} \to \mathbb{R}^k$. (The terms 'sanitised', 'privatised', 'privacy-protected' and 'privacy-preserving' are synonymous in the DP literature.) The Laplace mechanism adds noise with standard deviation proportional to the* global $\ell_1$-sensitivity *of $q$:*

$$\Delta(q) = \sup_{\substack{x,x' \in \mathcal{X} \\ d(x,x')=1}} \|q(x) - q(x')\|_1$$

*Specifically, define $M(x, L) = q(x) + bL$, where $b = \frac{\Delta(q)}{\epsilon}$ and $L$ is a $k$-vector of iid Laplace random variables with density $f(z) = 0.5\exp(-|z|)$. When $d(x, x') = 1$,*

$$\begin{aligned} P_x(S_1 \times \ldots \times S_k) &= \prod_{i=1}^{k} \left[ 0.5b^{-1} \int_{S_i} \exp\left( -\frac{|z - q_i(x)|}{b} dz \right) \right] \\ &\geq \exp\left( \frac{-\Delta(q)}{b} \right) \prod_{i=1}^{k} \left[ 0.5b^{-1} \int_{S_i} \exp\left( -\frac{|z - q_i(x')|}{b} dz \right) \right] \\ &= \exp(-\epsilon) P_{x'}(S_1 \times \ldots \times S_k). \end{aligned}$$

*We will see in Theorem 5 that this suffices to prove that $M$ is $\epsilon$-DP.*

**Example 2 (randomised response [60])** *Taking $\mathcal{X} = \bigcup_{n \in \mathbb{N}} \{0, 1\}^n$ as the data universe, the randomised response mechanism $M$ flips each bit $x_i$ with probability $p = (\exp \epsilon + 1)^{-1}$. That is, given a binary $n$-vector $x$ as input, $M$ outputs another binary $n$-vector with $i$-th component $x_i + B_i \mod 2$ where $B_1, B_2, \ldots \overset{iid}{\sim} \text{Bernoulli}(p)$. This mechanism is $\epsilon$-DP when $d = d_{\text{HAM}}$.*

*Moreover, $M$ is an example of* local *(non-interactive) DP [22], which requires $P_x$ to be a product measure $\prod_{i=1}^{n} P_{x_i}$ (where $n = |x|$ is the number of records in $x$). In the local interactive setting, the factors $P_{x_i}$ can depend not only on $x_i$ but also the previous outputs $t_j$ for $j < i$. The local model is typical of data collection by an untrusted entity (such as an IT company) where privacy protection must be applied to each record before it is seen.*

*DP without the constraint $P_x = \prod_{i=1}^{n} P_{x_i}$ is referred to as* central*, since the raw data can be aggregated by a central, trusted authority, such as a national statistical office.*

## 3. $\epsilon$-Differential Privacy as an Interval of Measures.

We introduce the definition of interval of measures, due to DeRobertis and Hartigan [18]:

**Definition 2** *Let $\Omega$ be the set of all $\sigma$-finite measures on $(\mathcal{T}, \mathscr{F})$. For $\mu, \nu \in \Omega$, write $\mu \leq \nu$ to denote that $\mu(S) \leq \nu(S)$ for all $S \in \mathscr{F}$.*

*Given $L, U \in \Omega$ with $L \leq U$, the convex set of measures*

$$\mathcal{I}(L, U) = \{\mu \in \Omega : L \leq \mu \leq U\}$$

*is called an* interval of measures. *$L$ and $U$ are called the* lower *and* upper measures*, respectively.*

As a direct consequence of the above definition, the odds ratio $P(A)/P(B)$ – for any $P \in \mathcal{I}(L, U)$ and any $A, B \in \mathscr{F}$ – is bounded between $L(A)/U(B)$ and $U(A)/L(B)$, whenever the ratios are well-defined.

An equivalent concept is the *density ratio class* [8], defined as follows: Fix some $\nu \in \Omega$ and pick $\nu$-densities $l \leq u$. The density ratio class $\mathcal{I}(l, u)$ consists of $\nu$-densities $f$ satisfying $l \leq f \leq u$. (This is equivalent to Definition 2 since every $\mu \in \mathcal{I}(L, U)$ is absolutely continuous with respect to $U$ and so will always have a $\nu$-density when $\nu = U$.) Density ratio class are often used as prior neighborhoods in robust Bayesian analysis due to their attractive properties [61]. Moreover, interval of measures can also represent neighborhoods of sampling distributions [49]. When used in conjunction with other prior neighborhoods it augments Bayesian robustness beyond prior robustness, without resorting to trivial posterior bounds.

Theorem 5 establishes an equivalence between the $\epsilon$-DP property of a data-release mechanism $M$ and the interval of measures $M$ induces.

**Definition 3** *Two datasets $x, x' \in \mathcal{X}$ are* connected *– or more precisely, d-connected – if $d(x, x') < \infty$. In this case, we say that $x$ is a* connection *of $x'$, and that the probability measures $P_x$ and $P_{x'}$ are* connected*. More generally, $S \subset \mathcal{X}$ is connected if all $x, x' \in S$ are.*

*The data universe $\mathcal{X}$ is partitioned into* connected components $[x] = \{x' \in \mathcal{X} \mid d(x, x') < \infty\}$.

Since the Lipschitz condition (1) is vacuous when $d(x, x') = \infty$, DP only constrains a mechanism $M$ to act

similarly on connected datasets $x, x'$; it makes no restrictions between $M(x, U)$ and $M(x', U)$ for unconnected $x, x'$. That is, there is no privacy guarantee of indistinguishability between unconnected datasets.

When $d = d_{\text{Ham}}$, any dataset $x, x'$ of different dimension (i.e. $|x| \neq |x'|$) are unconnected, so DP does not protect against, for example, an attacker determining $|x|$. Unconnected datasets also arise in the presence of *invariants* [6, 34].

**Assumption 4** *$d(x, x')$ is equal to the length of a shortest path between $x$ and $x'$ in a graph on $\mathcal{X}$ with unit-length edges.*

When $d(x, x') > 1$, the Lipschitz condition (1) is called *group privacy*. Since each $x_i$ represents an individual, condition (1) with $d(x, x') > 1$ is intuitively protecting multiple individuals simultaneously. We prove in Theorem 5 that Assumption 4 and individual-only privacy (i.e. condition (1) for $x, x'$ with $d(x, x') = 1$) together imply group privacy.

**Theorem 5** *Let $M : \mathcal{X} \times [0, 1] \to \mathcal{T}$ be a data-release mechanism with the seed $U \sim \text{Unif}[0, 1]$ inducing a probability $P_x$ on $M(x, U)$ (where $x$ is taken as fixed).*

*For $0 \leq \epsilon < \infty$, the following statements are equivalent given Assumption 4:*

*I $M$ is $\epsilon$-differentially private.*

*II $P_{x'}(S) \leq e^{\epsilon} P_x(S)$ for all $S \in \mathscr{F}$ and all $x, x' \in \mathcal{X}$ with $d(x, x') = 1$.*

*III For all $\delta \in \mathbb{N}$ and all $x, x' \in \mathcal{X}$ with $d(x, x') = \delta$,*

$$P_{x'} \in \mathcal{I}\big(L_{x,\delta\epsilon}, U_{x,\delta\epsilon}\big),$$

*where $L_{x,\delta\epsilon} = e^{-\delta\epsilon} P_x$ and $U_{x,\delta\epsilon} = e^{\delta\epsilon} P_x$.*

*IV For all $x \in \mathcal{X}$ and all measures $\nu \in \Omega$, if $P_x$ has a density $p_x$ with respect to $\nu$, then every $d$-connected $x' \in [x]$ also has a density $p_{x'}$ (with respect to $\nu$) satisfying*

$$p_{x'}(t) \in p_x(t) \exp(\pm \epsilon d(x, x')), \tag{2}$$

*for all $t \in \mathcal{T}$.*

In (2), the notation $a \in \exp(\pm b)$ is shorthand for

$$\exp(-b) \leq a \leq \exp(b).$$

II is the standard definition of $\epsilon$-differential privacy [26] and is listed here to justify our novel formulation of DP given in Definition 1. Without Assumption 4, group privacy is not implied by II. Hence Assumption 4 is needed only to extend II to provide group privacy; the equivalences between I, III and IV are automatic. Without Assumption 4 (which almost always holds in practice, such as for $d = d_{\text{Ham}}$ or $d_{\triangle}$), our definition of $\epsilon$-DP is more stringent than the standard formulation.

**Proof** "I $\Leftrightarrow$ II": Since $d$ is a graph distance, there is a path $x = x_0, \ldots, x_n = x'$ such that $d(x, x') = n$ and $d(x_i, x_{i+1}) = 1$. By the triangle inequality

$$d_{\text{Mult}}(P_x, P_{x'}) \leq \sum_{i=0}^{n-1} d_{\text{Mult}}(P_{x_i}, P_{x_{i+1}}).$$

Hence $\epsilon$-DP is equivalent to the Lipschitz condition (1) holding only when $d(x, x') = 1$. The equivalence between I and II then follows by the fact

$$e^{-\epsilon} P_x(S) \leq P_{x'}(S) \leq e^{\epsilon} P_x(S), \quad \forall S \in \mathscr{F},$$

if and only if $d_{\text{Mult}}(P_x, P_{x'}) \leq \epsilon$.

"I $\Leftrightarrow$ III" is immediate by noting $P \in \mathcal{I}(e^{-\epsilon} Q, e^{\epsilon} Q)$ if and only if $d_{\text{Mult}}(P, Q) \leq \epsilon$.

"III $\Leftrightarrow$ IV": The direction $\Rightarrow$ is straightforward since the densities in an interval of measure $\mathcal{I}(L, U)$ are bounded by the densities of $L$ and $U$. In the other direction, $P_x$ is always absolutely continuous with respect to itself, hence taking $P_x$ to be the dominating measure $\nu$, we have that (2) implies $P_{x'} \in \mathcal{I}(L_\epsilon, U_\epsilon)$. ■

IV is a strong property. It implies that, for an $\epsilon$-DP mechanism, all connected $P_x$ are mutually absolutely continuous. Further, for all connected $x, x' \in \mathcal{X}$, either $p_x(t)$ and $p_{x'}(t)$ are both zero or both non-zero (regardless of the dominating measure). Thus, if any $x$ is plausible (i.e. $p_x(t) > 0$) then all its connections $x' \in [x]$ are also plausible. This is a strong notion of privacy: regardless of the output $t = M(x, U)$, it's impossible for an attacker to distinguish between connected $x, x'$ with certainty. In other words, the fiducial distribution for $x$ is never degenerate (assuming that every $x$ has at least one connection).

This notion of privacy is the motivation for $d_{\text{Mult}}$ in place of more standard concepts in the robustness literature such as total variation distance or $\epsilon$-contamination classes. Indeed, indistinguishability requires that $p_x(t)/p_{x'}(t)$ is bounded away from zero and infinity, which is equivalent to $P_{x'} \in \mathcal{I}(aP_x, bP_x)$ for some $0 < a \leq 1 \leq b < \infty$. Yet Theorem 5 shows that $P_{x'} \in \mathcal{I}(aP_x, bP_x)$ if and only if

$$d_{\text{Mult}}(x, x') \leq \max(-\log a, \log b).$$

Therefore, $d_{\text{Mult}}$ is necessary to encode the idea of privacy as indistinguishability between connected $x, x'$.

This argument demonstrates that the Lipschitz condition (1) with another metric $\delta$ in place of $d_{\text{Mult}}$ will not ensure indistinguishability (except in the trivial case where $\alpha\delta \geq d_{\text{Mult}}$ for some constant $\alpha$). This is why all the common variants of DP – such as $(\epsilon, \delta)$-DP [25], zero-concentrated

DP [zCDP; 24, 13], and Rényi DP [53] – do not guarantee this strong notion of privacy, even though they may be preferred over pure $\epsilon$-DP for data utility reasons.

The observations of Theorem 5, specifically the equivalent characterization of $\epsilon$-DP via intervals of measures established by III and IV, bear important consequences for statistical inference from privacy-protected data. Notably, they impose meaningful bounds on both the probability of the privatised query and on relevant quantities in the frequentist and Bayesian inference from the privatised queries. These bounds are valid under arbitrary statistical models for the unknown confidential database, assuming only mild conditions on the models' support. The next three sections explore these consequences in detail.

## 4. Bounds on the Privatised Data Probability

Consider the situation of statistical inference, where a data analyst supplies a parametric model $\mathcal{P} = \{P_\theta \mid \theta \in \Theta\}$ of data generating distributions $P_\theta$. Nature generates data $X \sim P_\theta$ according to some unknown $\theta \in \Theta$. (We use capital $X$ to emphasise that the dataset is now random, whereas in the previous Sections, it was considered fixed.) In the typical (non-private) setting, the data analyst observes $X$ directly. In the private setting, the data analyst only sees the summary statistic $t = M(X, U) \sim P_X$ outputted from an $\epsilon$-DP data-release mechanism $M$. (We now require that the data universe $\mathcal{X}$ is equipped with a $\sigma$-algebra $\mathscr{G}$ and that every data-release mechanism $M$ is $(\mathscr{G} \otimes \mathcal{B}[0,1], \mathscr{F})$-measurable, where $\mathcal{B}[0,1]$ is the Borel $\sigma$-algebra on $[0,1]$.)

The relevant vehicle for inference in the private setting is the marginal probability of the observed data $t$:

$$P(t \in S \mid \theta) = \int_{\mathcal{X}} P_x(S) dP_\theta(x). \tag{3}$$

We call $P(t \in S \mid \theta)$ the *privatised data probability*. Viewed as a function of $\theta$, it is termed the *marginal likelihood* of $\theta$. All frequentist procedures compliant with likelihood theory and all Bayesian inference from privacy-protected data hinge on this function. The crucial role of (3) for inference from privacy-protected data was first recognized in the differential privacy literature by Williams and McSherry [63], and has since been utilized extensively to derive likelihood and Bayesian methodologies [e.g. 4, 5, 9, 10, 32, 42].

When $M$ is $\epsilon$-DP and $\mathcal{X}$ is $d$-connected, the existence of a density $p(t \mid \theta)$ is implied by Theorem 5. The following result proves this density exists in an interval of measures under the weaker assumption that (informally) the support of "$P(x \mid t, \theta)$" is $d$-connected. Other than this weak assumption, the following results hold for arbitrary data generating models $\{P_\theta\}$ and $\epsilon$-DP mechanisms $M$.

To state this assumption more precisely, define $\text{supp}(x \mid t, \theta)$ as the set of databases $x \in \mathcal{X}$ which could both generate $t$ and be generated by $P_\theta$. That is, $\text{supp}(x \mid t, \theta)$ is informally the intersection of $\text{supp}(P_\theta) \approx \{x \mid p_\theta(x) > 0\}$ and $\{x \mid p_x(t) > 0\}$. See Appendix A for an exact definition.

**Theorem 6** *Let $M$ be an $\epsilon$-DP mechanism. Suppose that $\text{supp}(x \mid t, \theta)$ is $d$-connected. Then, for any $x_* \in \text{supp}(x \mid t, \theta)$,*

$$p(t \mid \theta) \in p_{x_*}(t) \exp(\pm \epsilon d_*), \tag{4}$$

*where $d_* = \sup_{x \in \text{supp}(x \mid t, \theta)} d(x, x_*)$.*

*Now assume that $\text{supp}(x \mid t, \theta)$ is $d$-connected for $P(t \mid \theta)$-almost all $t \in \mathcal{T}$. Then*

$$P(t \mid \theta) \in \mathcal{I}(L_\epsilon, U_\epsilon) \tag{5}$$

*where $L_\epsilon$ and $U_\epsilon$ have densities*

$$\operatorname*{ess\,sup}_{x_* \in \text{supp}(x \mid t, \theta)} \exp(-\epsilon d_*) p_{x_*} \text{ and } \operatorname*{ess\,inf}_{x_* \in \text{supp}(x \mid t, \theta)} \exp(\epsilon d_*) p_{x_*}$$

*respectively.*

**Proof** Existence of $p(t \mid \theta)$ follows from the fact that all $P_x(t)$ with $x \in \text{supp}(x \mid t, \theta)$ are mutually absolutely continuous by Theorem 5. For the upper bound of (4),

$$\begin{aligned} p(t \mid \theta) &= \int_{\text{supp}(x \mid t, \theta)} p_x(t) dP_\theta(x) \\ &\leq \int_{\text{supp}(x \mid t, \theta)} e^{\epsilon d(x, x_*)} p_{x_*}(t) dP_\theta(x) \\ &\leq e^{\epsilon d_*} p_{x_*}(t). \end{aligned}$$

The lower bound follows similarly. The proof of (5) is left to Appendix B. ■

Surprisingly, the interval of measure $\mathcal{I}(L_\epsilon, U_\epsilon)$ in (5) depends on the data generating distribution $P_\theta$ only through $\text{supp}(x \mid t, \theta)$. When $\text{supp}(P_\theta)$ is constant, $\mathcal{I}(L_\epsilon, U_\epsilon)$ is completely free of $\theta$.

Theorem 6 is only meaningful when $d_* < \infty$. Typically $d(x, x')$ is unbounded for $x, x' \in \mathcal{X}$, which one might presume would imply that $d_* = \infty$. But $\text{supp}(P_\theta)$ can be much smaller than the data universe $\mathcal{X}$ when the analyst has prior knowledge of the data $X$. The analyst is free to restrict $\text{supp}(P_\theta)$ to the set of plausible datasets; the tighter this restriction, the stronger Theorem 6 is. For example, the analyst may have an upper bound $N$ on the number of records $|X|$. Moreover, $\text{supp}(x \mid t, \theta)$ can be much smaller than $\mathcal{X}$ when $t$ restricts the possible values of $X$, such as in the presence of invariants [34, 6]. For example in local DP, the number of records $|t|$ is invariant; this restricts $\text{supp}(x \mid t, \theta)$ to data $x$ satisfying $|x| = |t|$.

**Remark 7** *Theorem 6 only relies on the single assumption that $\text{supp}(x \mid t, \theta)$ is connected. This assumption is weak. In*

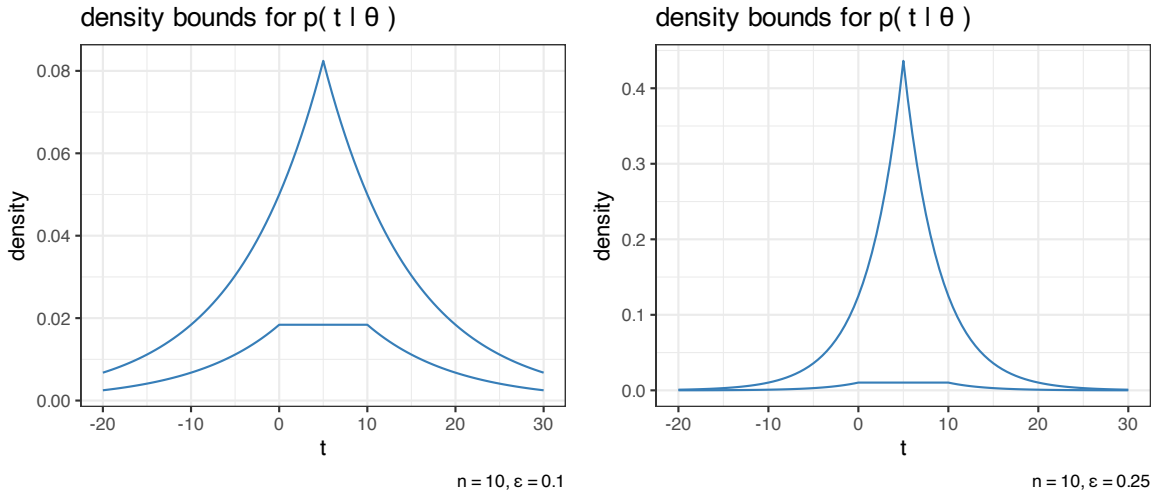


Figure 1: Upper and lower density bounds for the privatised data probability $p(t \mid \theta)$, as a function of $t$ (the privatised binary sum), for $\epsilon = 0.1$ (left) and $\epsilon = 0.25$ (right) from Example 3. These bounds do not depend on $\theta$ nor the assumed data model $P_\theta$. The bounds are tighter and more informative when privacy protection is more stringent (smaller $\epsilon$).

*fact, we can always augment the data-release mechanism $M$ so that this assumption is satisfied without increasing $M$'s privacy loss budget $\epsilon$. Specifically, the (deterministic) mechanism $x \mapsto [x]$ (which publishes the connected component $[x]$ of the observed data $x$) is trivially $\epsilon$-DP with $\epsilon = 0$. Publishing $[x]$ alongside $M(x, U)$ ensures that* $\text{supp}(x \mid t, \theta)$ *is always connected.*

We illustrate Theorem 6 with two examples.

**Example 3 (privatised binary sum)** *Suppose the database $x \in \mathcal{X} = \{0, 1\}^n$ consists of $n$ records of binary features, and its sum $q(x) = \sum_{i=1}^n x_i$ is to be queried. Consider sanitising $q(x)$ using the Laplace mechanism defined in Example 1. For every privacy loss budget $\epsilon > 0$ and every database $x$,*

$$p_x(t) = \frac{\epsilon}{2\Delta_q} \exp\left(\frac{\epsilon|t - q(x)|}{\Delta(q)}\right),$$

*where the global $\ell_1$-sensitivity $\Delta(q) = 1$ in this case. The data analyst posits an arbitrary statistical model $X \sim P_\theta$ for $\theta \in \Theta$, and considers the confidential and unknown database $x$ to be a realization from this model.*

*Figure 1 displays the upper and lower densities, corresponding respectively to $L_\epsilon$ and $U_\epsilon$ in Theorem 6, for the privatised data probability $p(t \mid \theta)$. The total record length $n$ is upper bounded by 10, so that $d_* = 10$. (In general, restrictions on* $\text{supp}(x \mid t, \theta)$ *such as this are imposed by the data analyst as they set the support for the data model $P_\theta$.) The left and right panels display bounds under two different settings of $\epsilon$. The bounds are tighter and more informative when privacy protection is more stringent ($\epsilon = 0.1$), and looser as the privacy loss budget increases ($\epsilon = 0.25$). Notice that these bounds for $p(t \mid \theta)$ are functions of the value of the privatised query $t$. In particular, they do not depend on $\theta$ nor the form of the posited data model $P_\theta$.*

**Example 4 (local DP)** *In the local (non-interactive) privacy model, the distribution $P_x$ of the published summary statistic $M(x, U)$ factors as $\prod_{i=1}^n P_{x_i}$, where $n = |x|$. This implies $|t| = |x|$. If $d = d_{H\text{AM}}$, as is typical for local DP, $d_* \leq |t|$ regardless of the choice of $x_*$. Hence, for any $x$,*

$$p(t \mid \theta) \in \prod_{i=1}^n p_{x_i}(t_i) \exp(\pm\epsilon n),$$

*under local $\epsilon$-DP. For randomised response,* $\min_{x_i} p_{x_i}(t_i) = (\exp\epsilon + 1)^{-1}$ *and* $\max_{x_i} p_{x_i}(t_i) = e^\epsilon(\exp\epsilon + 1)^{-1}$*, so that*

$$\frac{1}{(\exp\epsilon + 1)^{|t|}} \leq p(t \mid \theta) \leq \frac{\exp(|t|\epsilon)}{(\exp\epsilon + 1)^{|t|}}. \tag{6}$$

*The bounds in (6) depend on $t$ only through $|t|$ (the number of records), regardless of the records' values. Figure 2 displays these bounds as a function of $|t|$ for $\epsilon = 1$. As more records are released (larger $|t|$), both bounds tend to zero with a narrowing gap.*

## 5. Frequentist Privacy-Protected Inference

The interval of measures formulation of DP also shows that Neyman-Pearson hypothesis testing is restricted in the private setting, as demonstrated by the following Theorem.

**Theorem 8** *Consider testing $H_0 : \theta = \theta_0$ versus $H_1 : \theta = \theta_1$ for some $\theta_0 \neq \theta_1 \in \Theta$. Let $S_i = \text{supp}(x \mid t, \theta_i)$ and suppose that every $x \in S_0$ is $d$-connected to every $x' \in S_1$. In the private setting, the power of any level-$\alpha$ test is bounded above by $\alpha \exp(d_{**}\epsilon)$ where*

$$d_{**} = \sup_{x \in S_0, x' \in S_1} d(x, x').$$

**Proof** By IV of Theorem 5,

$$\begin{aligned}\frac{p(t \mid \theta_1)}{p(t \mid \theta_0)} &= \frac{\int_{S_1} p_x(t) dP_{\theta_1}(x)}{\int_{S_0} p_{x'}(t) dP_{\theta_0}(x')} \\ &= \int_{S_1} \left[ \int_{S_0} \frac{p_{x'}(t)}{p_x(t)} dP_{\theta_0}(x') \right]^{-1} dP_{\theta_1}(x) \\ &\in \exp(\pm \epsilon d_{**}).\end{aligned}$$

Let $R$ be the rejection region of a test with size $P(t \in R \mid \theta_0) \leq \alpha$ and let $\nu$ be the dominating measure of the densities $p(t \mid \theta_0)$ and $p(t \mid \theta_1)$. Then

$$\begin{aligned}P(t \in R \mid \theta_1) &= \int_R p(t \mid \theta_1) d\nu(t) \qquad (7) \\ &\leq \exp(d_{**}\epsilon) \int_R p(t \mid \theta_0) d\nu(t) \\ &\leq \alpha \exp(d_{**}\epsilon).\end{aligned}$$

■

Compare Theorem 8 to the hypothesis test $H_0 : x_{1:m} = y$ versus $H_1 : x_{1:m} = y'$ where $m \leq |x|$. This test models an attacker trying to distinguish the first $m$ records of the database. Wasserman and Zhou [62] showed that any level-$\alpha$ test of $x_{1:m}$ has power at most $\alpha \exp(\epsilon m)$ when $X_i$ are iid.

If the data analyst restricts $S_0$ and $S_1$ to datasets of length $m$, then typically $d_{**} = m$. Thus, any level-$\alpha$ test on the parameter $\theta$ has the same bound $\alpha \exp(\epsilon m)$ on its power (under an arbitrary data generating model, not just iid $X_i$). This highlights the fundamental tension between data privacy and data utility: bounding an attacker's power necessarily bounds the power of a legitimate analyst.

Theorem 8 strictly generalises the result of Wasserman and Zhou [62]. By taking $\Theta \subset \mathcal{X}$ and setting $P_\theta$ as degenerate point masses, we recover the set-up of an attacker's hypothesis test.[3] Thus, Theorem 8 is applicable to both the attacker testing $x$ (like [62]) and the analyst testing $\theta$ (with non-degenerate $P_\theta$).

[3]This ignores one minor technicality: the attacker may take some records as nuisance parameters, which they do not want to test. It is straightforward to generalise Theorem 8 to this situation. Without loss of generality, suppose $x_{m:n}$ are nuisance parameters when testing $x_{1:m-1}$ against $x'_{1:m-1}$. By assigning a conditional probability on $x_{m:n}$ satisfying $\pi(x_{m:n} \mid x_{1:m-1}) = \pi(x_{m:n} \mid x'_{1:m-1})$, the nuisance parameters can be integrated out in (7). This gives the same power bound $\alpha \exp(d_{**}\epsilon)$, except now with

$$d_{**} = \sup_{x_{m:n}} d\Big( [x_{1:m-1}, x_{m:n}], [x'_{1:m-1}, x_{m:n}] \Big).$$

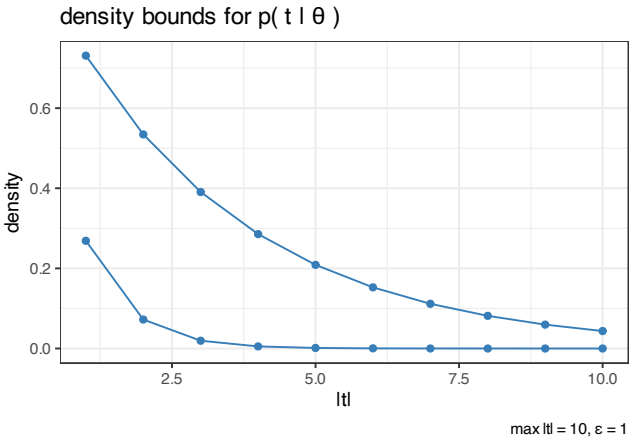


Figure 2: Upper and lower density bounds for $p(t \mid \theta)$ under randomised response (Example 4) with $\epsilon = 1$. These bounds are a function of $t$ only through $|t|$ (the number of observed records).

## 6. Bayesian Privacy-Protected Inference

Following the set-up from the previous Sections, we further assume that the analyst is Bayesian and places a (proper) prior $\pi$ on $\Theta$. This setting can be seen as a Bayesian hierarchical model where the raw data $X$ acts as latent parameter in the Markov chain $\theta \to X \to M(X, U)$.

The Theorem below establishes bounds on the analyst's prior predictive distribution $P(t \in S) = \iint P_x(t \in S) dP_\theta(x) d\pi(\theta)$ for the privatised data $t$.

We make the following Assumption throughout this Section.

**Assumption 9** *Suppose that*

$$\text{supp}(x \mid t) := \bigcup_{\theta \in \text{supp}(\pi)} \text{supp}(x \mid t, \theta),$$

*is $d$-connected for $P(t)$-almost all $t \in \mathcal{T}$. Further, assume the prior $\pi$ on $\theta$ is proper.*

By the same reasoning as in Remark 7, the first half of Assumption 9 is weak because it can always be satisfied by augmenting the data-release mechanism $M$ without additional privacy loss.

**Theorem 10** *The analyst's prior predictive probability for $t \sim M(X, U)$ (with $M$ an $\epsilon$-DP mechanism) satisfies*

$$\underline{p}_{\epsilon}(t) \leq p(t) \leq \overline{p}_{\epsilon}(t),$$

*for every $t \in \mathcal{T}$, where $\underline{p}_{\epsilon}$ and $\overline{p}_{\epsilon}$ are defined as*

$$\operatorname*{ess\,sup}_{x_* \in \text{supp}(x|t)} \exp(-\epsilon d_*) p_{x_*} \text{ and } \operatorname*{ess\,inf}_{x_* \in \text{supp}(x|t)} \exp(\epsilon d_*) p_{x_*}$$

*respectively, with $d_* = \sup_{x \in \text{supp}(x|t)} d(x, x_*)$.*

**Proof** Since $p(t) = \int_{\Theta} p(t \mid \theta) d\pi(\theta)$, Theorem 10 follows by showing $p(t \mid \theta)$ is bounded by $\underline{p}_{\epsilon}(t)$ and $\overline{p}_{\epsilon}(t)$. The proof of this is analogous to (5). ■

The prior predictive distribution $p(t)$ plays an important role in Bayesian inference and model checking. Before observing the data, $p(t)$ captures the analyst's implied specification on the data generation process. After observing the data, this quantity assessed at their value is called *model evidence* where low $p(t)$ reveals potential *conflict* between the data and the prior [30, 59]. In addition, it is also the normalizing constant for the posterior distribution $p(\theta \mid t)$ and is useful for computation.

As an illustration, we can see from Figure 1 of Example 3 that when $\epsilon = 0.1$, the prior predictive probability of the privatised query is lower-bounded at $\approx 0.02$ whenever $0 \leq t \leq 10$, and can never exceed $\approx 0.08$ even when $t = 5$. On the other hand, when privacy protection is less stringent ($\epsilon = 0.5$), the upper bound on the prior predictive probability increases to more than 0.4.

An important observation on Theorem 10 is the following: While $p(t)$ is a function of both the data model $P_\theta$ and the prior $\pi$, the density bounds $\underline{p}_{\epsilon}(t)$ and $\overline{p}_{\epsilon}(t)$ are free of both. In this sense, these bounds provide a non-trivial yet almost assumption-free prior predictive model sensitivity analysis. Non-trivial bounds on $p(t)$ are not possible in general; in this case they are a consequence of the data $t$ being $\epsilon$-DP.

The following Theorem provides general bounds limiting the learning of a Bayesian analyst.

**Theorem 11** *The analyst's posterior probability given (a realisation of an $\epsilon$-DP mechanism) $t$ satisfies*

$$\pi(\theta \mid t) \in \pi(\theta) \exp(\pm \epsilon d_{**}), \tag{8}$$

*where $d_{**} = \sup_{x, x' \in \text{supp}(x|t)} d(x, x')$.*

**Proof** As in the Proof of Theorem 8, we can show

$$\frac{p(t \mid \theta)}{p(t \mid \theta')} \in \exp(\pm \epsilon d_{**}),$$

for all $\theta, \theta' \in \Theta$. Plugging this into $\pi(\theta \mid t) = \frac{\pi(\theta) p(t|\theta)}{\int_{\Theta} p(t|\theta') d\pi(\theta')}$ gives the result. ■

Theorem 11 contributes to what is called the *prior-to-posterior semantics* of differential privacy, in the sense that (8) describes the extent to which a Bayesian agent's posterior about a parameter can depart from their prior when learning from an $\epsilon$-DP data product.[4] Analogous to the discussion in Section 5, Theorem 11 demonstrates the balance between restricting a Bayesian attacker while allowing for legitimate Bayesian learning: By setting $\Theta \subset \mathcal{X}$ and $P_\theta$ as degenerate point masses, we strictly generalise the result of Gong and Meng [34] which bounds an attacker's prior-to-posterior change in a single record $x_i$.[5] We therefore have illustrated the privacy-utility tradeoff which is fundamental to $\epsilon$-DP under both frequentist and Bayesian inference.

Theorem 11 is powerful because it holds for arbitrary specifications of the data model $P_\theta$ and is applicable to the agent's arbitrary (proper) prior $\pi(\theta)$. So long as $d_{**}$ is finite (see the discussion after Theorem 6 on why this is not unreasonable), the bounds in (8) are non-trivial.

With that said, whenever $d_{**}$ is large, the bounds provided by Theorem 11 are wide, rendering the results weakly informative at best. Indeed, rather than a pair of wide posterior bounds, the agent would be better off with a precise Bayesian posterior, which is theoretically derivable via the simple relation

$$\pi(\theta \mid t) \propto \pi(\theta) p(t \mid \theta), \tag{9}$$

where $p(t \mid \theta)$ can in turn be derived from the convolution of the data model $P_\theta$ and the privacy mechanism $P_x$ according to (3). In practice, however, direct computation or sampling from (9) is not always possible or feasible. Difficult situations include A) when the privacy mechanism $P_x$ is not fully transparent to the analyst due to its complex dependence on $x$, whether by design or by post-processing [33]; B) when the data model $P_\theta$ is intractable, such as if defined algorithmically or treated as a black-box; and C) when their convolution (3), an $n$-dimensional integral, is intractable. Under any of these situations, the analyst may still rely on Theorem 11 to obtain bounds on their posterior.

Despite their width, these bounds are optimal whenever $\epsilon$ is the smallest constant satisfying the Lipschitz condition (1). Without adding further assumptions on $M$, $P_\theta$, or $\pi$, these bounds cannot be shrunk. (This also applies to the

[4] An alternative type of semantics for differential privacy is the *posterior-to-posterior semantics* [20, 43], whose focus is on the extent to which a Bayesian agent's posterior may vary were it derived from privacy-protected queries based on different (counterfactual) confidential databases. Previous literature in differential privacy predominantly adopted posterior-to-posterior semantics [46]. However, prior-to-posterior semantics have recently attracted increasing attention as they circumvent counterfactuals and are closely connected with the literature on statistical disclosure risk [34, 40].

[5] By fixing some $x \in \mathcal{X}$ and setting $\pi(x_{-i}) = 1$, we get $d_{**} = 1$ and thereby rederive the result from [34].

bounds from Sections 4 and 5. We prove this in Section 7.) Yet they are not necessarily tight at a given $\theta$. This deficiency is an inevitable consequence of our analysis, which replaced the average case ($\int p_x(t)dP_\theta(x)$) with the extreme case ($p_{x_*}(t)\exp[\epsilon d_*]$). Such an analysis is necessarily loose whenever there is any variation away from the extreme. But the analysis cannot be tightened without making assumptions about the nature of this variation – i.e. by making further assumptions on $M$, $P_\theta$, or $\pi$.

We illustrate the posterior bounds of Theorem 11 with an example of Bayesian inference for a privatised count.

**Example 5 (privatised single count)** *Suppose the database consists of a single count record $x \in \mathbb{N}$. We wish to query the value of $x$ after it has been* clamped *to a pre-specified range $[a_0, a_1]$. That is, $q(x) = a_0$ if $x < a_0$, $q(x) = a_1$ if $x > a_1$, and $q(x) = x$ otherwise. In differentially private mechanism design, clamping is a necessary procedure when the intended query has otherwise unbounded global sensitivity. Under clamping, the sensitivity is reduced to $\Delta(q) = a_1 - a_0$.*

*The analyst's Bayesian model is*

$$\begin{aligned} \theta &\sim Gamma(\alpha, \beta), \\ x \mid \theta &\sim Pois(\theta), \\ t \mid x &\sim Lap\Big(q(x); \epsilon^{-1}\Delta(q)\Big). \end{aligned}$$

*For illustration, set $a_0 = 0$, $a_1 = 6$, $\alpha = 3$, $\beta = 1$. Figure 3 depicts in blue solid lines the upper and lower density bounds on the analyst's posterior distribution $p(\theta \mid t)$ as given by Theorem 11. With $\epsilon = 1$ and $d_{**} = 1$, they are equal to the $Gamma(3, 1)$ prior density (blue dashed line), scaled by $\exp(\pm 1)$. Overlaid in grey are Monte Carlo posterior densities $p(\theta \mid t^{(k)})$, $k = 1, \ldots, 10$, produced via the exact sampling algorithm proposed by [32]. Each $t^{(k)}$ is independently simulated from the prior predictive distribution of the above Bayesian model.*

Several aspects of Example 5 are worth noting. First, the posterior density bounds (solid blue) are functions of the analyst's chosen prior $\pi(\theta)$ and the privacy mechanism parameters $\epsilon$ and $d_{**}$ only. They are valid for any data model $P_\theta$ that the analyst wishes to employ, including (but not limited to) the Poisson data model that underlie the depicted precise posteriors densities $p(\theta \mid t^{(k)})$ in grey. On the other hand, while these precise posterior densities display moderate variations among each other, they do not depart much from the prior density (dashed blue). This is due to the heavy-handedness of the privacy mechanism employed for this analysis, resulting in poor statistical utility of the privatised count $t$. Indeed, the mechanism injects Laplace noise with standard deviation of $\sqrt{2}\epsilon^{-1}\Delta(q) = 8.48$ into a statistic clamped between $a_0 = 0$ and $a_1 = 6$. That

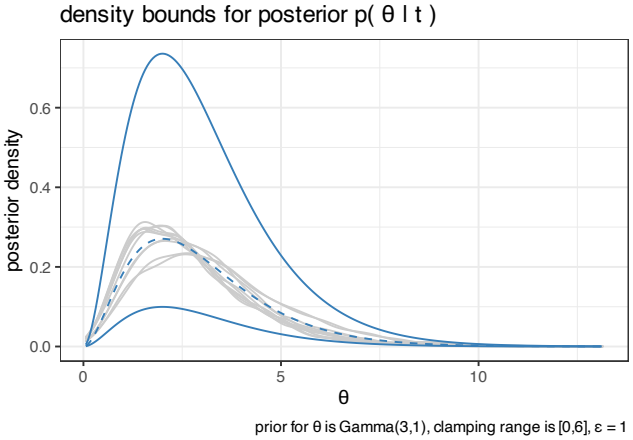


Figure 3: Upper and lower density bounds (solid blue) for the Bayesian posterior $p(\theta \mid t)$ from Example 5, based on a privatised single count. Dashed blue is density of the $Gamma(3, 1)$ distribution, the analyst's prior for $\theta$. In grey are simulation-based posterior densities based on 10 realizations of $t$ from its prior predictive distribution under the Poisson data model.

$t$ cannot be highly informative for the inferential problem at hand is correctly identified by the full Bayesian analysis which precisely accounts for the uncertainty induced by the privacy mechanism (grey lines). Furthermore, these precise posterior distributions are generally far from the bounds implied by Theorem 11; this re-enforces the shallowness of these bounds due to their validity for very general classes of the data model $P_\theta$ and priors $\pi$.

## 7. Optimality of this Paper's Results

The bounds presented in this paper cannot be improved without additional assumptions on the data-release mechanism $M$, the data generating model $P_\theta$ or the prior $\pi$. This is illustrated by the Laplace mechanism $M$ (Example 1) for the count query $q(x) = \sum_i x_i$. The density of $M$ is $p_x(t) = \frac{\epsilon}{2}\exp(-\epsilon|t - q(x)|)$ when $\mathcal{X} = \{0, 1\}^n$ and the metric $d$ on $\mathcal{X}$ is the Hamming distance $d_{\text{HAM}}$. The bound (4) in Theorem 6 is tight if $p_x(t) = \exp(\epsilon d_*)p_{x_*}(t)$ for $P_\theta$-almost all $x \in \text{supp}(x \mid t, \theta)$. Hence $p(t \mid \theta)$ can be arbitrarily close to $\exp(\epsilon d_*)p_{x_*}(t)$ as $P_\theta$ concentrates on $x = (0, \ldots, 0)$ with $x_* = (1, \ldots, 1)$ and $t \leq 0$. To see the tightness of Theorem 8, set $\theta_0 = (0, \ldots, 0)$ and $\theta_1 = (1, \ldots, 1)$ with $P_\theta(x)$ the point mass on $\theta = x$. The Neyman-Pearson (NP) test has rejection region $R = \{t > t_0\}$. For small enough $\epsilon$, we must have $t_0 \geq n = d_{**}$ (assuming $\alpha < 0.5$). Then,

$p(t \mid \theta_1) = \exp(\epsilon n) p(t \mid \theta_0)$ for all $t \in R$, which means the NP test has power exactly $\alpha \exp(\epsilon n)$. For Theorem 11, take $\Theta = [0, 1]$ with $\pi = \text{Unif}[0, 1]$ and $P_\theta(x)$ the point mass on $(1, \ldots, 1)$ if $\theta = 1$ and the point mass on $(0, \ldots, 0)$ otherwise. For $t > n$, we have $\pi(\theta = 1 \mid t) = \pi(\theta = 1) \exp(\epsilon n)$. Thus, the bound in Theorem 11 is achieved since $d_{**} = n$.

## 8. Discussion

The results we obtain in this paper make novel contributions to the differential privacy literature in the following ways. Firstly, the bounds we obtain in Theorems 6, 8, 10 and 11 are non-trivial, due to the validity of these results across a broad range of data models, privacy mechanisms and prior distributions. When the analyst has little knowledge or is only willing to make minimal assumptions about their model, these bounds are useful representations of the limits of statistical learning under privacy constraints. This draws a contrast with the existing DP literature, which has largely focused on asymptotic lower bounds or on constructing (asymptotically-)optimal data-release mechanisms for specific data use cases [56, 14, 16, 22, 7, 58, 27, 62, 5]. This literature aligns with *query-based access* [40] where the user can choose what statistics are released. Our results, on the other hand, are finite-sample and apply to the *dissemination mode* of data release where the mechanism is not tailored for the analyst's use case. This setting is typical of official statistics (e.g. censuses and surveys) and, more generally, data products with multiple users, and is more common in the research community than query-based access [40].

Secondly, the generality of our bounds implies that they are inherent consequences of $\epsilon$-DP itself. Specifically, these bounds stem only from the requirement that the mechanism $M$ is $\epsilon$-DP and not on any particularities of $P_\theta, M$ or $\pi$. That these bounds are typically wide in practice – as can be seen from Examples 3 and 5 – is in part due to the near-total lack of assumption under which they are derived. While these bounds can approach vacuity as the data size $n$ grows, in practical examples that need not be the case if, for example, the data analyst has probabilistic knowledge about the privacy mechanism (see e.g. Example 5) or the data space $\mathcal{X}$. For a given choice of $P_\theta, M$ and $\pi$, we may obtain tighter bounds than those in this paper. In addition to the asymptotic results in the aforementioned papers, sharp bounds for specific $P_\theta, M$ and $\pi$ may be derivable from the existing literature on measurement errors (*errors-in-variables*) in statistics and econometrics, particularly in the case of point identification problems (see e.g. [15, 38]).

Through the lens of Theorems 8 and 11, we obtain a valuable insight into the *privacy-utility trade-off* of $\epsilon$-DP [40]. Qualitatively speaking, there exists an inherent tension between protecting private information and deriving scientific knowledge. To date, quantitative approaches to this trade-off predominantly rely on the privacy loss budget as the sole metric to balance this trade-off [1, 41, 36]. However, from the suite of IP analyses presented here, we see that other building blocks – notably the metric structure $(\mathcal{X}, d)$ of the data universe, the associated database distances (such as $d_*$ and $d_{**}$), and the clamping parameters – are all relevant factors that, together with the privacy loss budget $\epsilon$, collectively determine the limits to statistical learning for attackers and scientists alike. Therefore, $\epsilon$ is not the only parameter of concern – and perhaps not even the central concern – when assessing and trading-off privacy and utility [6].

While this paper does reveal the tradeoff of $\epsilon$-DP, it narrowly conceives privacy and utility as the extent of statistical estimation attainable under either frequentist or Bayesian paradigms. There are, of course, many other aspects of utility that are worth examining – such as the ease of analysis, use of computational resources, facial validity and logical consistency [12, 39, 55] – and other paradigms (in particular decision theory) with which the concepts of privacy and utility can be quantified. In fact, both notions are multi-faceted and context-specific and, as one of our reviewers pointed out, a judicious conceptualisation of privacy and utility may improve their tradeoff's efficiency frontier. Acknowledging the complex makeup of this tradeoff, we advocate for future privacy mechanism design and analyses to treat the conceptualisation of privacy and utility – and the choice of $\mathcal{X}, d$ and $d_{\text{Mult}}$ – with scrutiny, given their scarcity in the current literature [6].

One direction for immediate future research is to explore IP characterisations of common variants of DP, in particular $(\epsilon, \delta)$-DP [25], zero-concentrated DP [24, 13], and Gaussian DP [21], which are popular in practice due to flexible privacy mechanism design, better privacy budget accounting and increased statistical efficiency. Since these variants, and others such as Pufferfish [45] and subspace DP [31], stem from reconceptualising $\mathcal{X}, d$ and the metric $d_{\text{Mult}}$ on $\Omega$ (the set of $\sigma$-finite measures on $(\mathcal{T}, \mathscr{F})$), a key question is how the $\Omega$-metric corresponding to a DP variant can be characterised as an IP object [6]. Our preliminary analysis shows that $(\epsilon, \delta)$-*probabilistic* DP [52] cannot be described by an interval of measures, at least not alone. However, it can be interpreted as restricting $P_{x'}$ to the union of $\delta$-total variation neighbourhoods of $P$, where the union is over all $P \in \mathcal{I}(e^{-\epsilon} P_x, e^{\epsilon} P_x)$ – i.e. the $\epsilon$-density ratio neighbourhood of $P_x$ (with $d(x, x') = 1$).

## Appendix A. Definition of $\text{supp}(x \mid t, \theta)$

Denote the support of $P_\theta$ by

$$\text{supp}(P_\theta) = \bigcap \{S \in \mathscr{G} \mid S \text{ closed}, P_\theta(S) = 1\} \subset \mathcal{X}. \quad (10)$$

Here we mean 'closed' with respect to an appropriate topology $\tau$ on $\mathcal{X}$, not necessarily the topology induced by the metric $d$. Typically $\tau$ would be such that $\mathscr{G}$ is the Borel $\sigma$-algebra of $\tau$. A standard example would be $\mathcal{X} = \bigcup_{n\in\mathbb{N}} \mathbb{R}^n$ (i.e. the universe of real-valued datasets with length $n \in \mathbb{N}$) with the topology induced by the map

$$\begin{aligned} \mathcal{X} &\to \mathbb{R}^{\mathbb{N}} \\ x &\mapsto (x, 0, 0, \ldots), \end{aligned}$$

where $\mathbb{R}^{\mathbb{N}}$ is equipped with the product topology.

(Generally we should not use the topology induced by $d$. Since $d$ is typically discrete, it would give $\operatorname{supp}(P_\theta) = \emptyset$ whenever $\mathcal{X}$ is uncountable and then Theorems 6 and 11 would be vacuous.)

Similar to (10), define $\operatorname{supp}(P_x) \subset \mathcal{T}$. Write $\operatorname{supp}_0(x \mid t) = \{x \mid t \in \operatorname{supp}(P_x)\}$ and finally define

$$\operatorname{supp}(x \mid t, \theta) = \operatorname{supp}(P_\theta) \cap \operatorname{supp}_0(x \mid t).$$

In general, the topologies on $\mathcal{X}$ and $\mathcal{T}$ should be chosen so that $\operatorname{supp}(x \mid t, \theta)$ is as small as possible without being empty, so that Theorems 6 and 11 are as strong as possible.

## Appendix B. Proof of (5)

By Theorem 5, we can fix a measure $\nu \in \Omega$ which dominates all $P_x$ for $x \in \operatorname{supp}(x \mid t, \theta)$. Take the essential supremum

$$\operatorname*{ess\,sup}_{x_* \in \operatorname{supp}(x \mid t, \theta)} \exp(-\epsilon d_*) p_{x_*} \tag{11}$$

with respect to $\nu$. (11) exists and is measurable as $\nu$ is $\sigma$-finite. Thus (11) is a $\nu$-density for some measure $L_\epsilon \in \Omega$. By (4),

$$p(t \mid \theta) \geq \operatorname*{ess\,sup}_{x_* \in \operatorname{supp}(x \mid t, \theta)} \exp(-\epsilon d_*) p_{x_*}$$

for $\nu$-almost all $t$. This proves one half of (5); the argument for the upper measure $U_\epsilon$ is analogous.

## Acknowledgments

We thank the four anonymous reviewers for their generous suggestions, which have greatly improved the paper. JB gratefully acknowledges partial financial support from the Australian-American Fulbright Commission and the Kinghorn Foundation.

## Author Contributions

Both authors conceived the project, and contributed to the research and the writing of the manuscript.

## References

[1] John M Abowd and Ian M Schmutte. An economic analysis of privacy protection and statistical accuracy as social choices. *American Economic Review*, 109 (1):171–202, 2019.

[2] 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*, (Special Issue 2), 2022.

[3] Apple Inc. Apple differential privacy technical overview, 2017. https://www.apple.com/privacy/docs/Differential_Privacy_Overview.pdf (accessed: Feb 18, 2023).

[4] Jordan A Awan and Aleksandra Slavković. Differentially private uniformly most powerful tests for binomial data. In *Advances in Neural Information Processing Systems 31*, pages 4208–4218. Curran Associates, Inc., 2018.

[5] Jordan A Awan and Aleksandra Slavković. Differentially private inference for binomial data. *Journal of Privacy and Confidentiality*, 10(1), 2020.

[6] James Bailie, Ruobin Gong, and Xiao-Li Meng. Can swapping be differentially private? A refreshment stirred, not shaken. *In preparation*, 2023+.

[7] Raef Bassily, Adam Smith, and Abhradeep Thakurta. Private empirical risk minimization: Efficient algorithms and tight error bounds. In *Proceedings of the 2014 IEEE 55th Annual Symposium on Foundations of Computer Science*, FOCS '14, pages 464–473, USA, October 2014. IEEE Computer Society. ISBN 978-1-4799-6517-5. doi: 10.1109/FOCS.2014.56.

[8] James O Berger. Robust Bayesian analysis: sensitivity to the prior. *Journal of Statistical Planning and Inference*, 25(3):303–328, 1990.

[9] Garrett Bernstein and Daniel R Sheldon. Differentially private Bayesian inference for exponential families. *Advances in Neural Information Processing Systems*, 31, 2018.

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

[11] Claire McKay Bowen, Victoria Bryant, Leonard Burman, John Czajka, Surachai Khitatrakun, Graham MacDonald, Robert McClelland, Livia Mucciolo, Madeline Pickens, Kyle Ueyama, et al. Synthetic individual income tax data: Methodology, utility, and privacy implications. In *Privacy in Statistical Databases: International Conference, PSD 2022, Paris, France, September 21–23, 2022, Proceedings*, pages 191–204. Springer, 2022.

[12] danah boyd and Jayshree Sarathy. Differential perspectives: Epistemic disconnects surrounding the U.S. Census Bureau's use of differential privacy. *Harvard Data Science Review*, (Special Issue 2), June 2022. doi: 10.1162/99608f92.66882f0e.

[13] Mark Bun and Thomas Steinke. Concentrated differential privacy: Simplifications, extensions, and lower bounds. In Martin Hirt and Adam Smith, editors, *Theory of Cryptography*, Lecture Notes in Computer Science, pages 635–658, Berlin, Heidelberg, 2016. Springer. doi: 10.1007/978-3-662-53641-4_24.

[14] T. 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.

[15] Raymond J Carroll and Peter Hall. Optimal rates of convergence for deconvolving a density. *Journal of the American Statistical Association*, 83(404):1184–1186, 1988.

[16] Julien Chhor and Flore Sentenac. Robust estimation of discrete distributions under local differential privacy. In *International Conference on Algorithmic Learning Theory*, pages 411–446. PMLR, 2023.

[17] Chris Culnane, Benjamin I. P. Rubinstein, and Vanessa Teague. Stop the Open Data Bus, We Want to Get Off. *arXiv:1908.05004 [cs]*, August 2019.

[18] Lorraine DeRobertis and John A Hartigan. Bayesian inference using intervals of measures. *The Annals of Statistics*, 9(2):235–244, 1981.

[19] Bolin Ding, Janardhan Kulkarni, and Sergey Yekhanin. Collecting telemetry data privately, 2017. https://www.microsoft.com/en-us/research/blog/collecting-telemetry-data-privately/ (accessed: Feb 18, 2023).

[20] Irit Dinur and Kobbi Nissim. Revealing information while preserving privacy. In *Proceedings of the twenty-second ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*, pages 202–210, 2003.

[21] Jinshuo Dong, Aaron Roth, and Weijie J Su. Gaussian differential privacy. *Journal of the Royal Statistical Society Series B*, 84(1):3–37, 2022.

[22] John C Duchi, Michael I Jordan, and Martin J Wainwright. Minimax optimal procedures for locally private estimation. *Journal of the American Statistical Association*, 113(521):182–201, 2018.

[23] Cynthia Dwork and Jing Lei. Differential privacy and robust statistics. In *Proceedings of the forty-first annual ACM symposium on Theory of computing*, pages 371–380. ACM, 2009.

[24] Cynthia Dwork and Guy N. Rothblum. Concentrated differential privacy. *arXiv:1603.01887*, March 2016.

[25] Cynthia Dwork, Krishnaram Kenthapadi, Frank McSherry, Ilya Mironov, and Moni Naor. Our data, ourselves: Privacy via distributed noise generation. In Serge Vaudenay, editor, *Advances in Cryptology - EUROCRYPT 2006*, Lecture Notes in Computer Science, pages 486–503, Berlin, Heidelberg, 2006. Springer. doi: 10.1007/11761679_29.

[26] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. In *Theory of cryptography conference*, pages 265–284. Springer, 2006.

[27] Cynthia Dwork, Kunal Talwar, Abhradeep Thakurta, and Li Zhang. Analyze Gauss: Optimal bounds for privacy-preserving principal component analysis. In *Proceedings of the Forty-Sixth Annual ACM Symposium on Theory of Computing*, STOC '14, pages 11–20, New York, NY, USA, May 2014. Association for Computing Machinery. doi: 10.1145/2591796.2591883.

[28] Cynthia Dwork, Adam Smith, Thomas Steinke, and Jonathan Ullman. Exposed! A survey of attacks on private data. *Annual Review of Statistics and Its Application*, 4(1):61–84, 2017. doi: 10.1146/annurev-statistics-060116-054123.

[29] 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.

[30] Michael Evans and Hadas Moshonov. Checking for prior-data conflict. *Bayesian Analysis*, 1(4):893–914, December 2006. doi: 10.1214/06-BA129.

[31] Jie Gao, Ruobin Gong, and Fang-Yi Yu. Subspace differential privacy. *Proceedings of the AAAI Conference on Artificial Intelligence*, 36(4):3986–3995, 2022.

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

[33] Ruobin Gong. Transparent privacy is principled privacy. *Harvard Data Science Review*, (Special Issue 2), June 2022. doi: 10.1162/99608f92.b5d3faaa.

[34] Ruobin Gong and Xiao-Li 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, October 2020. Association for Computing Machinery. ISBN 978-1-4503-8103-1. doi: 10.1145/3412815.3416892.

[35] Jan Hannig, Hari Iyer, Randy CS Lai, and Thomas CM Lee. Generalized fiducial inference: A review and new results. *Journal of the American Statistical Association*, 111(515):1346–1361, 2016.

[36] Ori Heffetz. What will it take to get to acceptable privacy-accuracy combinations? *Harvard Data Science Review*, (Special Issue 2), June 2022. doi: 10.1162/99608f92.5d9b1a8d.

[37] Nils Homer, Szabolcs Szelinger, Margot Redman, David Duggan, Waibhav Tembe, Jill Muehling, John V Pearson, Dietrich A Stephan, Stanley F Nelson, and David W Craig. Resolving individuals contributing trace amounts of DNA to highly complex mixtures using high-density SNP genotyping microarrays. *PLoS genetics*, 4(8):e1000167, 2008.

[38] Joel L. Horowitz and Charles F. Manski. Identification and robustness with contaminated and corrupted data. *Econometrica*, 63(2):281–302, 1995. ISSN 0012-9682. doi: 10.2307/2951627.

[39] V. Joseph Hotz and Joseph Salvo. A chronicle of the application of differential privacy to the 2020 Census. *Harvard Data Science Review*, (Special Issue 2), June 2022. ISSN 2644-2353, 2688-8513. doi: 10.1162/99608f92.ff891fe5.

[40] V Joseph Hotz, Christopher R Bollinger, Tatiana Komarova, Charles F Manski, Robert A Moffitt, Denis Nekipelov, Aaron Sojourner, and Bruce D Spencer. Balancing data privacy and usability in the federal statistical system. *Proceedings of the National Academy of Sciences*, 119(31):e2104906119, 2022.

[41] Justin Hsu, Marco Gaboardi, Andreas Haeberlen, Sanjeev Khanna, Arjun Narayan, Benjamin C. Pierce, and Aaron Roth. Differential Privacy: An economic method for choosing epsilon. In *Proceedings of the 2014 IEEE 27th Computer Security Foundations Symposium*, pages 398–410, Vienna, July 2014. IEEE. ISBN 978-1-4799-4290-9. doi: 10.1109/CSF.2014.35.

[42] Nianqiao Ju, Jordan A Awan, Ruobin Gong, and Vinayak A Rao. Data augmentation MCMC for Bayesian inference from privatized data. *Thirty-sixth Annual Conference on Neural Information Processing Systems*, 2022.

[43] Shiva P Kasiviswanathan and Adam Smith. On the 'semantics' of differential privacy: A Bayesian formulation. *Journal of Privacy and Confidentiality*, 6(1), 2014.

[44] Christopher T Kenny, Shiro Kuriwaki, Cory McCartan, Evan TR Rosenman, Tyler Simko, and Kosuke Imai. The use of differential privacy for Census data and its impact on redistricting: The case of the 2020 US Census. *Science Advances*, 7(41):eabk3283, 2021.

[45] Daniel Kifer and Ashwin Machanavajjhala. Pufferfish: A framework for mathematical privacy definitions. *ACM Transactions on Database Systems (TODS)*, 39 (1):1–36, 2014.

[46] Daniel Kifer, John M Abowd, Robert Ashmead, Ryan Cumings-Menon, Philip Leclerc, Ashwin Machanavajjhala, William Sexton, and Pavel Zhuravlev. Bayesian and frequentist semantics for common variations of differential privacy: Applications to the 2020 Census. *arXiv preprint arXiv:2209.03310*, 2022.

[47] Sobin Kim and Ashish Misra. SNP genotyping: Technologies and biomedical applications. *Annual Review of Biomedical Engineering*, 9(1):289–320, August 2007. ISSN 1523-9829, 1545-4274. doi: 10.1146/annurev.bioeng.9.060906.152037.

[48] Tatiana Komarova and Denis Nekipelov. Identification and formal privacy guarantees. *arXiv preprint arXiv:2006.14732*, 2020.

[49] Michael Lavine. Sensitivity in Bayesian statistics: The prior and the likelihood. *Journal of the American Statistical Association*, 86(414):396–399, 1991.

[50] Qiyu Li, Chunlai Zhou, Biao Qin, and Zhiqiang Xu. Local differential privacy for belief functions. *Proceedings of the AAAI Conference on Artificial Intelligence*, 36(9):10025–10033, June 2022. doi: 10.1609/aaai.v36i9.21241.

[51] Likang Liu, Keke Sun, Chunlai Zhou, and Yuan Feng. Two views of constrained differential privacy: Belief revision and update. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2023.

[52] Ashwin Machanavajjhala, Daniel Kifer, John M Abowd, Johannes Gehrke, and Lars Vilhuber. Privacy: Theory meets practice on the map. In *Proceedings of the 2008 IEEE 24th International Conference on Data Engineering*, pages 277–286. IEEE Computer Society, 2008.

[53] Ilya Mironov. Rényi differential privacy. *2017 IEEE 30th Computer Security Foundations Symposium (CSF)*, pages 263–275, August 2017. doi: 10.1109/CSF.2017.11.

[54] Arvind Narayanan and Vitaly Shmatikov. Robust deanonymization of large sparse datasets. In *2008 IEEE Symposium on Security and Privacy (Sp 2008)*, pages 111–125, May 2008. doi: 10.1109/SP.2008.33.

[55] Steven Ruggles, Catherine Fitch, Diana Magnuson, and Jonathan Schroeder. Differential privacy and Census data: Implications for social and economic research. *AEA Papers and Proceedings*, 109:403–408, May 2019. ISSN 2574-0768. doi: 10.1257/pandp.20191107.

[56] Adam Smith. Privacy-preserving statistical estimation with optimal convergence rates. In *Proceedings of the Forty-Third Annual ACM Symposium on Theory of Computing*, STOC '11, pages 813–822, New York, NY, USA, June 2011. Association for Computing Machinery. doi: 10.1145/1993636.1993743.

[57] Latanya Sweeney. Simple demographics often identify people uniquely. Data Privacy Working Paper 3, Carnegie Mellon University, Pittsburgh, 2000.

[58] Kunal Talwar, Abhradeep Guha Thakurta, and Li Zhang. Nearly optimal private LASSO. In *Advances in Neural Information Processing Systems*, volume 28. Curran Associates, Inc., 2015.

[59] Gero Walter and Thomas Augustin. Imprecision and prior-data conflict in generalized Bayesian inference. *Journal of Statistical Theory and Practice*, 3(1):255–271, 2009.

[60] Stanley L. Warner. Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, 60(309):63–69, March 1965. ISSN 0162-1459. doi: 10.1080/01621459.1965.10480775.

[61] Larry Wasserman. Invariance properties of density ratio priors. *The Annals of Statistics*, 20(4):2177–2182, 1992.

[62] Larry Wasserman and Shuheng Zhou. A statistical framework for differential privacy. *Journal of the American Statistical Association*, 105(489):375–389, 2010.

[63] Oliver Williams and Frank McSherry. Probabilistic inference and differential privacy. *Advances in Neural Information Processing Systems*, 23:2451–2459, 2010.

[64] Bin Yu. Stability. *Bernoulli*, 19(4):1484–1500, 2013.