7

Machine Learning with Differential Privacy

Anand D. Sarwate

7.1 Introduction: What Is Private Machine Learning?

In this chapter we take up the problem of machine learning for private or sensitive data. The phrase "privacy-preserving machine learning" can refer to myriad models for privacy and learning. Machine learning is a term that metastasized to encompass a large variety of approaches to the problem of inferring structure in data. While many "classical" methods in statistics have been rebranded as "machine learning", a useful distinction is that the latter places slightly more emphasis on the computational or algorithmic aspects of the inference problem. Many machine learning methods also attempt to be "distribution-free" in the sense that they try to make very few assumptions on the model generating the data.

In *privacy-preserving machine learning* the goal is to use and learn from sensitive data gathered from individuals in a way that respects the privacy of those individuals. The approach taken in different applications will depend on the way in which privacy is conceptualized. For example, legal or statutory definitions of privacy may differ by country or jurisdiction and certain types of data, such as financial or medical data, may be subject to different protections. An important aspect of defining privacy within a particular domain or application is to describe the *threat model*: what information can be "leaked" and what kinds of harms are caused by that "leakage". In some cases there may be a prescribed procedure for "de-identifying" data without explicitly defining the threat model. Some of these approaches may allow for privacy harms if they are not robust to the presence of side information or more sophisticated attacks [50].

We will therefore focus on *differentially private machine learning*, in which we want to use sensitive data to learn some property of the population or to build a predictive model without making assumptions on how the data were generated. The "distribution-free" framing using in machine learning is most compatible with the framework of differential privacy (DP) [49], defined in Chapter 1: DP also does not make probabilistic assumptions on the private data. In parametric statistics, we assume the data follows a distribution from a known parametric family and ask for inferences to be reliable for "typical" realizations of the data. Nonparametric models still assume that the data are drawn from a given distribution. The robust statistics [73] model is closer to the approach taken in differential privacy [11, 48, 123], where the privacy guarantees have to hold for every possible data set.

In the differential privacy threat model, we assume an adversary can observe the output of this machine learning (ML) algorithm. Privacy is a property of the algorithm: a privacy harm occurs if an adversary can infer whether a particular individual's data was used as an input to the algorithm. A differentially private (DP) algorithm uses *randomization* to hinder this adversarial inference by introducing uncertainty into the output of the algorithm: if the output has similar likelihood regardless of whether that individual's data was used,

DOI: 10.1201/9781003185284-9

then the adversary will be unable to accurately decide if they were indeed in the input data. The privacy risk in differential privacy quantifies this uncertainty for the adversary by characterizing the error-tradeoff in the hypothesis test between "the individual's data was used" and "the individual's data was not used" [82, 148]. However, the uncertainty in the output may also affect the *utility* of the ML algorithm. A randomized output from a DP algorithm may be less accurate, have additional bias, or exhibit higher variance compared to a non-private learning algorithm. This *privacy-utility tradeoff* is a central object of study for differentially private machine learning. From the statistics or ML perspective, a DP algorithm will require more data (samples) to achieve a target utility than the non-private algorithm: this is the cost of privacy. The relevant background for the discussion in this chapter can be found in Chapters 1, 4 and 5.

As a concrete example, suppose the data holder has data consisting of feature-label pairs and wants to learn a linear classifier/predictor using regularized logistic regression. The non-private ML algorithm using a textbook approach to produce a vector of coefficients would not guarantee differential privacy: if a single data point (individual) were removed, the algorithm would produce a different output, meaning that an adversary can infer whether or not that individual was present. Instead, a DP logistic regression method would introduce some noise (randomization) during the computation. The simplest approach could be to run the non-private ML method and add noise to the coefficient vector. The noisy coefficient vector will in general have worse classification performance than the non-private ML algorithm: this is the cost in terms of utility.

Machine learning methods are the driving force behind many recent advances in data analysis and artificial intelligence and are likely to play a major role in future applications in medical/scientific research, social science, and policy. Many of these applications involve learning from private or sensitive data collected about individuals. The institutions or entities holding this information have obligations (legal and ethical) to protect the confidentiality of this data. Differential privacy may be an appropriate approach for trading off privacy and utility when the desired analysis is to infer some feature of the population of individuals. For such applications we would hope that the presence or absence of an individual would not affect the value of that feature very much, making differential privacy compatible with the task. In our logistic regression example above, if we have a large number of data points, the resulting coefficients should not be too sensitive to the presence or absence of a single individual.

In this chapter we will discuss common approaches to standard problems in supervised and unsupervised learning. Readers interested in classical parametric inference methods (frequentist and Bayesian) can consult Chapters 8. In order to keep the discussion contained, we will mostly focus on methods for "centralized" differential privacy, rather than "local" (see Chapter 5) or "federated" data models (see Chapters 15 and 19). We will also generally restrict discussion to "pure" ϵ-DP and "approximate" (ϵ, δ)-DP rather than variations and extensions of the DP framework (see Chapter 4). Finally, while there is a large body of work on learning theory with privacy and its connections to other areas of mathematics, we will primarily focus more on algorithms for which there has been empirical validation. Many of the methods we discuss do have theoretical performance guarantees in terms of utility and the interested reader can find the details in those references. Even though an approach may have strong theoretical guarantees, data sets with small sample size may be challenging because generic methods make few assumptions and the utility metric may not match all applications. However, by incorporating domain expertise into differentially private methods or using two-step procedures which first test if the data have a "nice" property and then apply a method tuned to that property can often work quite well. This idea, which can be thought of as a form of he propose-test-release framework of Dwork and Lei [48], is a promising avenue for adapting generic approaches to specific applications.

7.1.1 Preliminaries

There have been several early implementations of DP algorithms by technology companies [4, 8, 25, 26, 34, 42, 47, 54, 135], although these mainly focus on counts and histograms. A major development was the use of differential privacy to publish the 2020 US Decennial Census [2, 3]. The type of machine learning algorithms we discuss here focus on different statistical tasks such as clustering, dimension reduction, predictive modeling, and deep learning.

In the context of machine learning algorithms, we will assume that the algorithm is given n data samples coming from n individuals and that two data sets are considered neighboring if they differ in a single individual. For unsupervised learning methods we will consider unlabeled data $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ and for supervised learning we will look at labeled data $((\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \ldots, (\mathbf{x}_n, y_n))$. We will generally assume that the features $\mathbf{x}_i \in \mathbb{R}^d$, unless otherwise specified.

The basic definitions of differential privacy and its properties have been described in Chapters 1 and 5. When looking at differentially private machine learning, it's important to specify what the output of the algorithm is. For example, when using principle component analysis (PCA) on a positive semidefinite (covariance or second moment) matrix $A \in \mathbb{R}^{d \times d}$, we compute the singular value decomposition $A = U \Sigma U^{\top}$. We may be either wish to output a rank-k approximation $A^{(k)}$ of A (where $k < d$) or to simply produce the k singular vectors corresponding to the k largest singular values: these are sufficient to compute a projection of the original data in $\mathbb{R}^d$ to a lower dimensional representation in $\mathbb{R}^k$. Differentially private PCA algorithms may produce one or the other. As a second example, in regression analyses we may be interested in feature/variable selection as well as producing the regression coefficients. Using cross-validation requires computing estimates using a validation set. If this set also contains private data, the cross-validation must also be computed using differential privacy.

7.1.2 Some Important Considerations

Privacy-preserving machine learning is a rapidly developing field and this chapter can at best provide a sampling of the problems and solutions that have been proposed. The focus here is on "generic" problems because most of the research has been on differential privacy for general machine learning and statistics methods. As the field has developed, algorithms adapted for specific application domains have been proposed. However, these methods typically use the core ideas presented here with some application-specific modifications.

There are several key challenges in the practical application of differentially private machine learning methods which we do not address in this chapter.

- **Data preprocessing:** Most works on differential privacy assume that the input data are already "clean" in the sense that standard preprocessing (imputation of missing values, standardizing or normalizing features/covariates, screening for outliers) is not necessary. However, many preprocessing methods involve computing statistics on the data and hence may incur privacy costs. For example, imputing missing values may involve taking the mean of the non-missing values. Very recently, some more sophisticated approaches for DP imputation have been proposed [29, 41, 44].

- **(Hyper) parameter tuning:** Many modern machine learning methods have several hyperparameters whose optimal values may depend strongly on the type of data or the specific data set being used. Typically these are chosen using cross-validation. However, cross-validation involves running the algorithm multiple times on the data, leading to additional privacy loss. Choosing privacy risk parameters to achieve acceptable utility may also require tuning. There is a body of work on private tuning methods [35, 39, 78, 96] which we do not address here.
- **Computational issues:** It is well-known that standard implementations of floating point arithmetic do not satisfy differential privacy [16, 33, 63, 75–77, 102]. This has led to the recent developments of fixed point implementations. For example, the discrete Gaussian [33] was used for the differentially private data release in the 2020 Decennial Census [3] in the United States.

The ultimate goal in designing differentially private machine learning algorithms is to guarantee "end-to-end" privacy in which all steps in the analysis that operate on private data are made differentially private. We can then use composition rules (see Chapters 1, 12 and 14) to compute the final privacy loss. While some works have taken on this bigger challenge, most of the work to date has focused on individual components of a complete pipeline. In this chapter we will focus on describing the concepts behind differentially private solutions to the core problem at the heart of different machine learning problems. This necessarily leaves a gap between what is described here and practice, as the preceding discussion notes. We therefore encourage readers interested in applying differential privacy to statistical products to consult with privacy experts in the design of their systems.

7.2 Unsupervised Learning

In *unsupervised learning* problems the goal is to find patterns or structure in data that is unlabeled. This could take the form of estimating the underlying distribution (or its properties), finding correlations between features, or clustering the data points. In this section, we will assume we have a data set $\mathcal{D} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ of unlabeled points from n individuals with $\mathbf{x}_i \in \mathbb{R}^d$ and we want to learn some structure in the data.

7.2.1 Clustering

In clustering, the goal is to partition the data into groups (clusters) such that similar points are grouped together. Typically, each cluster is associated with a cluster center and we are given a function $c(\mathbf{x}, \mathbf{y})$ that measures the cost of assigning a point $\mathbf{x}$ to a cluster with center $\mathbf{y}$. Given k cluster centers $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_k\}$, for each point $\mathbf{x}_i$ we find the lowest cost $\min_j c(\mathbf{x}_i, \mathbf{y}_j)$. The goal is to find the cluster centers that minimize the total cost. Many cost functions are possible but a common choice is the squared Euclidean error $\|\mathbf{x} - \mathbf{y}\|^2$: this is sometimes called the k-means problem. There is a long line of work on developing differentially private clustering methods (see Stemmer [128] for a detailed discussion). Non-private k-means algorithms alternate between estimating candidate centers and assigning data points to the centers over many iterations. This repeated use of the private data poses challenges for differentially private processing.

Stemmer and Kaplan [129] reduce the number of iterations by partitioning the data into disjoint sets and using a private local search algorithm to produce a large set of candidate clusters centers which will contain k good cluster centers. While having nice theoretical guarantees and spurring more recent work [81, 105, 119] there is still very little empirical exploration of these approaches. Feldman et al. [56] use the idea of private coresets [55] to design a clustering algorithm with good theoretical properties and apply it to mobility data. Alternative approaches with less theory but more empirical support include sampling using the exponential mechanism [143], grid-based methods [130, 131], and sketching with Laplace noise [118]. These latter methods address the challenge presented by k-means by computing a differentially private summary of the data and relying on postprocessing invariance.

7.2.2 Bayesian Machine Learning

A body of work has been developed on Bayesian machine learning with differential privacy, including graphical model estimation for directed models [150, 159, 160] and undirected models [22]. For some inference problems, sampling from the posterior distribution may be differentially private [45, 46, 58, 59, 100, 146]. While the single sample guarantees differential privacy, it is often computationally challenging to sample from these posterior distributions and single samples may not be what we want. One sampling approach is Markov Chain Monte Carlo (MCMC), but MCMC methods can only give samples from a distribution approximating the desired one, which may not have the same privacy guarantee. Modifying the MCMC method can yield a differential privacy guarantee either through stochastic gradient methods [95, 146, 154] or privatizing the acceptance test [68].

In the special case of exponential family models, we can guarantee privacy by adding noise to the sufficient statistics [22, 58, 163]. Because the noise model is known, it is then possible to do approximate inference with knowledge of the privacy mechanism [22, 23, 62, 86, 150] (see Chapter 8 for more discussion). Using noisy sufficient statistics is also a component of differentially private expectation maximization (EM) algorithms [112] and more generally in variational inference methods for conjugate exponential families in which noise can be added to the expected sufficient statistics [113] or by using private optimization of the evidence-based lower bound (ELBO) [80]. EM is often used for mixture models, but DP approaches designed for specific problems such as Gaussian mixtures [84, 106] may not extend to general mixtures. For inference in general models, private EM algorithms focus on privatizing the M-step using private optimization by either perturbing the gradient [80, 139] or iterative hard thresholding [162] based on DP hard thresholding [52].

7.2.3 Dimension Reduction

Consider an analyst who wishes to find latent factors in a sensitive data set $\mathcal{D} = \{\mathbf{x}_i\}$. A standard approach to latent factor analysis is principle component analysis (PCA), one of the most widely applied methods in data analysis. At its heart, PCA is a dimensionality reduction technique that maps high-dimensional vector-valued observations $\{\mathbf{x}_i\}$ into lower-dimensional vectors $\{\hat{\mathbf{x}}_i\}$ through a linear map. Dimensionality reduction is important because the utility of some differentially private algorithms can depend strongly on the dimension [30, 51, 127]. In PCA, we take the singular value decomposition of the matrix of observations or the eigenvalue decomposition of the second-moment or covariance matrix. If we start with the second-moment matrix of the data:

$$A = \sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^\top = \mathbf{X}\mathbf{X}^\top. \tag{7.1}$$

We decompose the matrix A using the SVD:

$$A = U \Sigma U^\top, \tag{7.2}$$

where $\Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_d)$ and $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_d$. The best rank-$k$ approximation of A (minimizing squared error or Frobenius norm) is

$$A^{(k)} = U^{(k)} \Sigma^{(k)} U^{(k)\top}, \tag{7.3}$$

where $U^{(k)}$ has the first k columns of U and $\Sigma^{(k)} = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_k)$.

Under differential privacy, it is important to specify what the output of our algorithm is. For PCA we can either output the low-rank approximation $A^{(k)}$, an approximation to $\mathbf{X}$, or simply the latent factors in $U^{(k)}$. One approach is to sample a random $U^{(k)}$ or $A^{(k)}$. The exponential mechanism [99] can be used to sample an ϵ-DP approximation to $U^{(k)}$ but must be implemented using Markov Chain Monte Carlo (MCMC) [36, 149]. This approach can also be extended to functional data [13]. In general, because sampling is only approximately correct, the privacy guarantee can only be approximately guaranteed: this is a challenge for many sampling-based methods for differential privacy based on the exponential mechanism because sampling from Gibbs distributions is often challenging [72]. An alternative approach is to sample eigenvalues and eigenvectors [6, 85]. Taking a cue from Bayesian statistics, Sheffet [121] proposed sampling differentially private approximation to A from the inverse-Wishart distribution, which is a conjugate prior for estimating the covariance of a multivariate Gaussian.

If we relax the privacy requirement to approximate (ϵ, δ)-privacy, adding Gaussian noise at various points in the PCA calculation can provide good results: approaches include adding noise to A itself [53], introducing noisy data [121], or adding noise in the computation of the eigenvectors [64–66]. From an empirical and theoretical standpoint, the first method of simply adding noise to A [53] works well and comes with confidence intervals [120]. It is also a standard approach for algorithms which use PCA as a substep even though the variance of the noise must depend on the ambient dimension of the data $\{\mathbf{x}_i\}$ and not k.

A completely different approach to dimension reduction is to choose a random subspace and project the very high dimensional observations into this subspace. These approaches come with favorable theoretical guarantees and are the backbone of "sketching" approaches to machine learning for streaming data, where samples come sequentially and the estimator is memory-limited [151]. Perhaps surprisingly, because this approach is itself randomized, certain variants already guarantee differential privacy [27] and more sophisticated versions can yield more robust estimators [9].

Since the privacy guarantee has to hold for any database, the utility bounds are often dictated by pathological examples that do not occur in real data. The challenge is then to make "reasonable assumptions" on the data being "nice" in order to get better results. Exploiting domain-specific knowledge about structure in the data is often critical in applications but challenging to incorporate in differentially private algorithms. Such assumptions can allow the use of smooth sensitivity [61]. Recent work [122] uses the "subsample and aggregate" model [106] to provide an algorithm for estimating a low-dimensional structure to the data set (if it exists).

7.3 Supervised Learning with Private Optimization

We next turn to supervised learning and prediction problems. In supervised learning we are given a set of labeled data $\mathcal{D} = \{(\mathbf{x}_i, y_i) : i = 1, 2, \ldots, n\}$ where $\mathbf{x}_i \in \mathcal{X}$ are features/covariates for the ith individual and $y_i \in \mathcal{Y}$ is the corresponding label/response. In classification problems $\mathcal{Y}$ is a set of discrete labels, whereas scalar regression problems may take $\mathcal{Y} = \mathbb{R}$. The goal of the ML algorithm is to learn a predictor, which is a map $f : \mathcal{X} \to \mathcal{Y}$, where f is assumed to be from some class of models $\mathcal{F}$. For example, in linear regression $\mathcal{F}$ may contain all linear functions of $\mathbf{x}$ or $\mathcal{F}$ may be all functions computable by a neural network with a given architecture. We will focus on parametric problems in which $\mathcal{F}$ can be parameterized by a vector $\mathbf{w} \subset \mathbb{R}^p$. In linear regression we have $p = d$ and $\mathbf{w}$ is vector of regression coefficients, whereas in neural networks we take $\mathbf{w}$ as the weights in the network, where often $p \gg d$.

7.3.1 Differentially Private Empirical Risk Minimization

A good predictor is one which can accurately estimate the label y for vectors $\mathbf{x}$ that will be seen in the future. Assuming that future data are sampled from the same population as $\mathcal{D}$, we want to minimize an expected risk $\ell(\mathbf{w}, \mathbf{x}, y)$:

$$\mathbf{w}^* = \underset{\mathbf{w}}{\operatorname{argmin}}\, \mathbb{E}\left[\ell(\mathbf{w}, \mathbf{x}, Y)\right]. \tag{7.4}$$

This optimization problem is called *empirical risk minimization (ERM)*. The data distribution of $(\mathbf{X}, Y)$ is unknown, so as a proxy we usually minimize the empirical risk

$$\mathbf{w}_{\text{ERM}} = \underset{\mathbf{w}}{\operatorname{argmin}}\, \frac{1}{n} \sum_{i=1}^{n} \ell(\mathbf{w}, \mathbf{x}_i, y_i). \tag{7.5}$$

Given a particular statistical model (for example, a generalized linear model (GLM)) we can use choose loss function to be the negative log likelihood. Structural assumptions on $\mathbf{w}$ can be enforced by using *regularized* empirical risk minimization:

$$L(\mathbf{w}, \mathcal{D}) = \frac{1}{n} \sum_{i=1}^{n} \ell(\mathbf{w}, \mathbf{x}_i, y_i) + \lambda R(\mathbf{w}), \tag{7.6}$$

where $R(\cdot)$ is chosen to penalize the "complexity". If we focus on linear models $\ell(\mathbf{w}, \mathbf{x}, y) = \ell(\mathbf{w}^\top \mathbf{x}, y)$, many standard statistical/machine learning methods fall into this category:

- **Ridge regression:** quadratic loss $\ell(\mathbf{w}^\top \mathbf{x}, y) = (y - \mathbf{w}^\top \mathbf{x})^2$, ℓ_2 regularizer $R(\mathbf{w}) = \|\mathbf{w}\|_2^2$.
- **Lasso:** quadratic loss $\ell(\mathbf{w}^\top \mathbf{x}, y) = (y - \mathbf{w}^\top \mathbf{x})^2$, regularizer $R(\mathbf{w}) = \|\mathbf{w}\|_1$.
- **Support vector machines (SVMs):** hinge loss $\ell(\mathbf{w}^\top \mathbf{x}, y) = \max(0, 1 - y\mathbf{w}^\top \mathbf{x})$, ℓ_2 regularizer $R(\mathbf{w}) = \|\mathbf{w}\|_2^2$.
- **Logistic regression:** loss $\ell(\mathbf{w}^\top \mathbf{x}, y) = \ln(1 + e^{-y\mathbf{w}^\top \mathbf{x}})$, ℓ_2 regularizer $R(\mathbf{w}) = \|\mathbf{w}\|_2^2$.

Finding a predictor using ERM entails solving a numerical optimization problem. If the objective function is convex and the space $\mathcal{F}$ is convex then standard convex optimization methods can be used to estimate the predictor. However, as mentioned in the introduction to the chapter, this approach will in general not guarantee differential privacy: the output of the method under two neighboring databases will be different so an adversary can detect whether the output was generated from one database or the other. Given the data set $\mathcal{D}$ we want to release the prediction model $\mathbf{w}$ in a differentially private manner.

To introduce randomness (and hence privacy) into the ERM problem, several approaches have been proposed. In output perturbation [35, 38, 116], we compute the non-private $\mathbf{w}$ and add noise to it:

$$\mathbf{w}_{\text{out}} = \operatorname*{argmin}_{\mathbf{w}} L(\mathbf{w}, \mathcal{D}) + \mathbf{Z}, \tag{7.7}$$

where $\mathbf{Z}$ is noise whose distribution is chosen to guarantee ϵ or (ϵ, δ) differential privacy. To choose the noise level we have to compute the sensitivity of the ERM problem. In objective perturbation, the function to be minimized is randomized – randomly changing the optimization problem induces randomness in the minimizer [20, 35, 38, 88]. The most common example is adding a random linear term to the objective function:

$$\mathbf{w}_{\text{obj}} = \operatorname*{argmin}_{\mathbf{w}} \left(L(\mathbf{w}, \mathcal{D}) + \mathbf{Z}(\mathbf{w}) \right). \tag{7.8}$$

For example, for ϵ-DP logistic regression with $R(\mathbf{w}) = \frac{1}{2} \|\mathbf{w}\|_2^2$, if we draw $\mathbf{Z}$ from the distribution

$$p(\mathbf{z}) \propto \exp\left(-\frac{n\lambda\epsilon}{2} \|\mathbf{z}\|_2 \right), \tag{7.9}$$

then

$$\hat{\mathbf{w}} = \left(\operatorname*{argmin}_{\mathbf{w}} \frac{1}{n} \sum_{i=1}^{n} \ln(1 + e^{-y_i \mathbf{w}^\top \mathbf{x}_i}) + \frac{\lambda}{2} \|\mathbf{w}\|_2^2 \right) + \mathbf{Z} \tag{7.10}$$

guarantees ϵ-differential privacy. By drawing $\mathbf{Z}$ from a Gaussian distribution, the output can be made (ϵ, δ) differentially private [88]. The multivariate distribution in (7.9) is isotropic with a magnitude that has a Gamma distribution with shape d and scale $\frac{2}{n\lambda\epsilon}$ and is sometimes referred to as "gamma noise". More generally, optimizing the choice of sensitivity measure and choosing an appropriate noise distribution can yield differentially private ERM algorithms that can be tuned to prior information or assumptions on the data set [14].

Other approaches to ERM have been proposed. The exponential mechanism comes with better theoretical guarantees, although it is prohibitive computationally [20]. An approach using genetic algorithms uses very different perturbation methods [161] but has not been further developed. Another approach is to use randomized approximations of the loss functions, for example through polynomial approximation. Methods such as the functional mechanism [157] or Bernstein mechanism [5] are also varieties of objective perturbation because they randomize the objective function to guarantee privacy. Another form of output perturbation tries to discretize the parameter space to optimize classification error [104]. The discretization is computationally expensive but for smaller data sets is quite promising because it seems to yield better empirical performance.

Theoretical results for DP ERM methods show how privacy affects the utility as measured by excess empirical risk, population risk, or parameter error. These results hold under various analytic assumptions (differentiability, Lipschitz continuity, (strong) convexity) about the loss function, its gradients, and regularizer, assumptions on the data (bounds on the norm), and the constraint set for $\mathbf{w}$. For example, a modified version of objective perturbation is differentially private when the loss function is twice continuously differentiable and the regularizer is convex (but need not be differentiable) and the nonprivate objective is strongly convex [79, 88]. Output perturbation is differentially private under a similar set of assumptions [79, 152, 158]. Motivated by the loss functions typically used in deep learning (see Section 7.4), some works explicitly consider nonconvex optimization problems [138, 141, 142, 158].

Empirically, several works have shown that approaches using objective perturbation generally produce better approximations than output perturbation [35, 38, 79]. Relaxing to (ϵ, δ)-DP rather than ϵ-DP also results in higher utility estimators in practice. Practical considerations often preclude the use of algorithms with better theoretical guarantees but high computational complexity: an algorithm may be polynomial-time ("efficient" in the parlance of theoretical computer science) but the actual running time may be prohibitive when the dimension of the data and/or number of samples is large. For example, a version of the exponential mechanism has better theoretical guarantees [20] under ϵ-DP but a running time proportional to n^3, where n is the number of data points. Ultimately, however, the choice of a "best" algorithm may depend strongly on the actual data set and desired privacy or utility levels.

7.3.2 Differentially Private Optimization Algorithms

Both output and objective perturbation assume that the optimization can be performed exactly. While mathematically convenient, the actual implementation of these algorithms with finite-precision computing raises a number of mathematical and engineering challenges. Firstly, using standard floating point libraries and numerical computing packages can result in privacy leakage (see Chapter 9 for a discussion). Secondly, numerical methods for optimization problems only yield approximate solutions. Thus the actual distribution of the output of the mechanism, such as the random minimizer $\mathbf{w}_{\text{obj}}$ in (7.8), will not be the same as that analyzed in the privacy result. This latter issue was addressed by Iyengar et al. [78] who propose a method called Approximate Minima Perturbation which uses objective and output perturbation, adding a noisy term to the objective function and then to the output and tuning the noise parameters to balance the privacy and utility guarantees.

A different approach is to introduce noise into the optimization procedure itself. Numerical methods for optimization are often iterative and by introducing noise during the iterations we can maintain a privacy guarantee without requiring exact minimization. Many numerical methods for problems such as (7.6) update using the gradient of the objective function [107]. A very simple gradient descent (GD) method would compute the gradient of $L(\mathbf{w}, \mathcal{D})$ with respect to $\mathbf{w}$ and update $\mathbf{w}_t = \mathbf{w}_{t-1} - \eta_t \nabla_{\mathbf{w}} L(\mathbf{w}, \mathcal{D})$, where η_t is called the step size or learning rate. This approach can be used to minimize the ERM problem (regularized or unregularized).

Differentially private gradient descent (DP-GD) perturbs the gradient calculation at each update [20, 125, 134, 150], using bounds on the norm of the gradient vector. More precisely, for unconstrained ERM, given a data set $\mathcal{D}$ and a constraint set $\mathcal{C} \subseteq \mathbb{R}^p$, assuming that the ℓ_2 norm of the gradient $\left\| \nabla_{\mathbf{w}} \ell(\mathbf{w}, \mathbf{x}_i, y_i) \right\| \leq L$, the algorithm performs the following iterations (the initial point can be random or set to $\mathbf{0}$):

$$\mathbf{g}_t = \frac{1}{n}\sum_{i=1}^{n} \nabla_{\mathbf{w}} \ell(\mathbf{w}_t, \mathbf{x}_i, y_i) + \mathbf{Z}_t \tag{7.11}$$

$$\mathbf{w}_{t+1} = \Pi_{\mathcal{C}}\left(\mathbf{w}_t - \eta_t \mathbf{g}_t\right) \tag{7.12}$$

where $\mathbf{Z}_t \sim \mathcal{N}(0, \sigma^2 I_p)$ and is independent at each iterations. When running T iteration, by choosing $\sigma^2 = \frac{2L^2 T \log(1/\delta)}{(n\epsilon)^2}$ we can guarantee (ϵ, δ)-differential privacy [1, 103]. It is worth noting an important change in perspective here: we specify the variance of the noise first and then compute the privacy guarantee: since there are many (ϵ, δ) pairs which use the same σ^2 we have more flexibility. Using gamma noise in (7.9) we can get an ϵ-DP guarantee.

Noisy (projected) gradient descent and variants have been analyzed to characterize properties such as convergence rates, consistency, minimax optimality, and asymptotic normality for a variety of loss functions [12, 31, 32, 126]. Differentially private algorithms for more complex numerical optimization methods have also been proposed and studied, including Frank-Wolfe [133], Mirror Descent [134], adaptive gradient descent [10, 92], and Newton's method [12]. Finally, stochastic gradient descent [20, 125, 134], which we discuss in Section 7.4, has become the de-facto approach to private optimization for large-scale problems.

7.3.3 Examples, Alternatives, and Extensions

Differentially private ERM gives algorithms for classification and regression problems in supervised learning. While many of the algorithms were validated on benchmark data sets and are hence "generic", specific applications or problem settings may require more tailored solutions or extending existing theory. An example of the application-specific algorithms is a differentially private logistic regression with elastic net regularization that was developed for use on genomic data [156], which generalizes the scope of the objective perturbation approach. Examples of the theoretical extensions are differentially private confidence intervals or significance testing for regression [12, 18, 114, 120, 147].

For the "simplest" problem of linear regression, Cai and Zhang [32] provide an improved projected gradient descent method which runs in fewer iterations than the general ERM approach. However, for this problem there are many different approaches [144], including adding noise to sufficient statistics, using subsampling and aggregation [106, 124], or sampling from Bayesian posteriors [45, 59, 100, 101, 146], in addition to the optimization-based methods we discuss here. Wang [144] proposes new Bayesian and frequentist approaches that generally improve on existing methods across a wide range of data sets [144]. These new methods are "adaptive" in the sense that the first find a DP estimate of a property of the distribution (the smallest eigenvalue of the Gram matrix) and then proceed to use this to tune the DP estimator of the regression coefficients, similar to the propose-test-release model [48].

Sparsity constraints are often a first step in model or feature selection problems [39, 94, 133, 137]. Differential privacy for sparse linear regression has been studied quite extensively and several different approaches have been proposed and analyzed, including generalizations of objective perturbation [88], "boostrapping"-style estimation [137], a differentially private version of iterative hard thresholding [32], and DP optimization [133]. More structured sparsity constraints like group sparsity [15] have also been considered [91, 156].

Many of the general approaches to Bayesian inference from Section 7.2.2 can be applied to regression problems [58, 80, 95, 154, 162]. Algorithms specifically for Bayesian regression generally work by perturbing sufficient statistics [24, 69, 90]. One challenge in linear regression is that both the features/covariates and labels/responses have to be protected, so there are multiple sufficient statistics to protect corresponding to the covariances and variances. One important question for practical settings is how to split the privacy budget between these private estimates [69]. Another challenge is sampling from the posterior, which can be done by approximating the posterior [69] or approximating the likelihood of the statistics and using a Gibbs sampler [24]. For GLMs which do not admit sufficient statistics, polynomial approximations can be used [74] which gives a differentially private method by adding noise to statistics for the approximation (usually a Gaussian) [90].

Robust statistics [73] also has connections to differential privacy [11, 37, 48, 93, 123]: finding a robust M-estimator also involves minimizing an objective function where the log likelihood is replaced with losses that have smaller influence functions. Recent work has developed output perturbation [11] and objective perturbation [123] approaches to for differentially private M-estimation with applications to robust versions of linear and logistic regression.

7.4 Large-Scale Machine Learning

The last decade has seen a massive expansion of machine learning research, primarily driven by the performance of deep neural networks (NNs) and the development of software tools and libraries which make it easier for practitioners to use these methods in a wide range of applications. A major challenge in these applications is the size and scale involved, with hundreds of millions of training points used to train models with hundreds of millions of parameters.[1] Many studies have proposed privacy attacks on ML models (see recent surveys [50, 70]). The published function of the data here is generally the weights for the neural network: the architecture itself is generally assumed to be public. Differentially private deep learning is an active area of research as different domain practitioners try to evaluate if reasonable privacy-utility tradeoffs are possible for their applications. We will discuss the dominant approach to privately training deep learning problems [1] and the associated issues which arise when applying privacy-preserving optimization methods to large-scale problems.

7.4.1 Private Stochastic Gradient Descent

For applications where the number of data points n is very large, optimizing (7.6) using GD can be computationally prohibitive since computing $\nabla_{\mathbf{w}} L(\mathbf{w}, \mathcal{D})$ involves taking gradients for each term in the empirical risk. The most common approach in ML applications and software systems therefore is to use stochastic gradient descent (SGD), in which at each iteration the optimizer chooses a (small) set $\mathcal{B}_t$ (a "minibatch") of samples and calculates the gradient on the batch:

$$\mathbf{g}_t = \frac{1}{|\mathcal{B}_t|} \sum_{i \in \mathcal{B}_t} \nabla_{\mathbf{w}} \ell(\mathbf{w}, \mathbf{x}_i, y_i). \tag{7.13}$$

[1] These numbers are ever increasing.

The algorithm can then update the weights $\mathbf{w}_t = \mathbf{w}_{t-1} - \eta_t \mathbf{g}_t$. Differentially private SGD (DP-SGD) adds noise to $\mathbf{g}_t$, which involves computing the sensitivity of the minibatched gradient [1, 20, 125]. The central component of DP-SGD is to privately estimate the mean of the gradient of the objective function: sampling a minibatch uniformly with replacement results and adding zero-mean noise for privacy results in an unbiased estimate of the gradient.

The theoretical properties of DP-SGD for convex optimization problems are fairly well-understood [19]. One subtlety in DP-SGD algorithms is the way in which the batch is selected: sampling data points can enhance privacy [17, 87, 145] but some care must be taken in the design of the algorithm since the privacy implications of sampling with replacement can be different from sampling without replacement [17]. As with non-private SGD, many variants of the simple DP-SGD algorithm above have been proposed to get better privacy-utility tradeoffs. Many of these works try to make the gradient updates adaptive by changing the batch size [57] or using methods to adapt to the local geometry or the problem [10, 89, 132, 164]: these improved estimators come with better theoretical guarantees and/or empirical performance.

7.4.2 Deep Learning with Differential Privacy

Neural networks are almost always trained using SGD, making DP-SGD a natural candidate for private deep learning [1]. Many applications use fairly high-dimensional data, which often means that DP methods have to introduce larger perturbations/more noise, decreasing utility. This can be helped by introducing more constraints on the data or loss functions; for DP-SGD we need to control the gradient norms. Unfortunately, practical NN training can often suffer from the "exploding gradient" problem. This is often ascribed to heavy-tail phenomena in many data sets. While some recent work has focused on changing the private estimator of the gradient under heavy-tail assumptions [71, 83, 136, 140], the most common way that practitioners enforce bounds on the gradient is by "clipping" (renormalizing) large gradients to control the sensitivity [1, 7, 40, 111, 115, 126], even though this can introduce bias in the gradient estimates and prevent the solution from converging to the same point as the original problem [12].

Most training algorithms perform multiple passes over the data ("training epochs") in which each point is used once. Although privacy can be amplified by random subsampling (see Chapter 5) each clipped minibatch gradient is a differentially private query, which means the total privacy risk needs to be controlled using more advanced composition techniques, namely Rényi differential privacy [103] and the moments accountant [1, 98, 145]. When using (ϵ, δ)-DP, a typical approach to algorithm design is to set the variance of the noise first and then computes the privacy guarantee in terms of the number of iterations, allowing for a more flexible tradeoff between the two types of privacy risk ϵ and δ. Even with all of this, in many empirical studies getting the training and test errors close to the nonprivate baseline requires relatively large values of ϵ. This is to be expected because training a neural network often involves several epochs, and even with the benefit of the measure concentration results using Rényi differential privacy [103], the privacy risk still increases with each epoch.

The high cost of training can be mitigated in scenarios where public data are available. An alternative approach to private training uses public data to train a differentially private model: the private dataset is split into disjoint sets which are used for training classifiers. These are used with public unlabeled data to privately train another model. A simple form of this was used for linear classifiers in a neuroimaging application [117] but in deep learning is called the Private Aggregation of Teacher Ensembles (PATE) framework [109, 110]. In that framework the "teacher" classifiers are trained on the private data in a non-private manner. However, they are only used to produce labels for the public data – a differentially private aggregation method then provides a consensus label for each sample. This approach is less general than DP-SGD because it requires public data and is restricted to supervised learning problems.

From a practical standpoint, applying differentially private deep learning almost always has to be done through existing software systems and frameworks for machine learning [153]. Both TensorFlow[2] [97] and PyTorch[3] [155] support differentially private training of ML models. These systems, both based on Python, support training with DP-SGD and provide built-in privacy accounting using the moments accountant. The newest software suite for differentially private algorithms, OpenDP[4] does not yet support many machine learning methods, but may do so in future releases.

A practitioner interested in applying differentially private machine learning (and especially deep learning) should be careful about the implicit assumptions made by using differential privacy as a framework. Evaluating methods on common benchmark data sets is important to compare approaches, but many benchmark data sets do not come from applications where privacy is a primary concern. Thus a DP method for computer vision may be tested on its ability to distinguish different birds [60] but the real application where privacy is a concern could be face recognition. A more fundamental challenge is that in DP we typically assume data records are associated to individuals and a DP algorithm tries to make it difficult to infer if a particular record is associated to an individual. In many data sets and applications this association between individuals is not so straightforward. For example, Brown et al. [28] surveyed the growing literature on privacy-preserving language modeling and find a mismatch between the model of differential privacy and its application to natural language data.

7.5 Conclusion

In this chapter we surveyed some of the work on differentially private machine learning and different approaches to privacy-preserving algorithms for unsupervised and supervised learning. The astute reader will note that the underlying mechanisms used for many of these methods are the same basic techniques discussed in Chapters 4, 5 and 6. Typically, the main challenges in making a machine learning algorithm differentially private are to determine the best point in the computation apply a standard mechanism and to analyze the impact of that mechanism on the utility. In many cases, the algorithms take the privacy parameters (ϵ, δ) as inputs and the effective privacy-utility must be determined empirically for each data set. For end-to-end privacy, even this exploration of the privacy-utility tradeoff must be done in a differentially private manner, which is a form of private model selection that we did not address here.

Since being introduced in 2006 [49], differential privacy has been widely adopted in machine learning research and new differentially private versions of machine learning algorithms are published every year [43]. In addition, new software frameworks are

[2] https://github.com/tensorflow/privacy
[3] https://github.com/pytorch/opacus
[4] https://opendp.org/

emerging to help analysts design and deploy differentially private algorithms [21, 67, 97, 108, 155]. A team designing an analytics pipeline including DP machine learning should include an expert in differentially private engineering to help address the issues raised in Section 7.1.2 regarding preprocessing, parameter tuning, and numerical implementation. By understanding the privacy-utility tradeoff in a wider variety of applications, such implementations can help spur future algorithm development to further narrow the gap between theory and practice.

Acknowledgments

The work of A. D. Sarwate was supported by the US National Institutes of Health under award 2R01DA040487 and by the National Science Foundation under awards SaTC-1617849, CCF-1453432, and CNS-2148104.

References

1. Martín Abadi, Andy Chu, Ian Goodfello, H. Brendan McMahan, Ilya Mironov, Kunal Talwar, and Li Zhang. Deep learning with differential privacy. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communication Security (CCS '16)*, pp. 303–318, Vienna, Austria, October 24–28 2016. ACM.
2. John Abowd, Robert Ashmead, Simson Garfinkel, Daniel Kifer, Philip Leclerc, Ashwin Machanavajjhala, Brett Moran, William Sexton, and Pavel Zhuravlev. Census topdown algorithm: Differentially private data, incremental schemas, and consistency with public knowledge. Technical report, US Census Bureau, 2019.
3. John M. Abowd. The u.s. census bureau adopts differential privacy. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, London, UK, August 19–23 2018.
4. Ahmet Aktay, Shailesh Bavadekar, Gwen Cossoul, John Davis, Damien Desfontaines, Alex Fabrikant, Evgeniy Gabrilovich, Krishna Gadepalli, Bryant Gipson, Miguel Guevara, Chaitanya Kamath, Mansi Kansal, Ali Lange, Chinmoy Mandayam, Andrew Oplinger, Christopher Pluntke, Thomas Roessler, Arran Schlosberg, Tomer Shekel, Swapnil Vispute, Mia Vu, Gregory Wellenius, Brian Williams, and Royce J Wilson. Google COVID-19 community mobility reports: Anonymization process description (version 1.1). Technical Report arXiv:2004.04145 [cs.CR], ArXiV, 11 2020.
5. Francesco Aldà and Benjamin I. P. Rubinstein. The bernstein mechanism: Function release under differential privacy. In Satinder Singh and Shaul Markovitch, editors, *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, February 4–9, 2017, San Francisco, California, USA*, pp. 1705–1711. AAAI Press, 2017.
6. Kareem Amin, Travis Dick, Alex Kulesza, Andres Munoz Medina, and Sergei Vassilvitskii. Private covariance estimation via iterative eigenvector sampling. In *2018 NIPS Workshop in Privacy-Preserving Machine Learning*, vol. 250, 2018.
7. Galen Andrew, Om Thakkar, H. Brendan McMahan, and Swaroop Ramaswamy. Differentially private learning with adaptive clipping. Technical Report arXiv:1905.03871 [cs.LG], ArXiV, 3 2021.
8. Apple Differential Privacy Team. Learning with privacy at scale. available on Apple's website.
9. Raman Arora, Vladimir Braverman, and Jalaj Upadhyay. Differentially private robust low-rank approximation. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and

R. Garnett, editors, *Advances in Neural Information Processing Systems*, vol. 31. Curran Associates, Inc., 2018.

10. Hilal Asi, John C. Duchi, Alireza Fallah, Omid Javidbakht, and Kunal Talwar. Private adaptive gradient methods for convex optimization. In Marina Meila and Tong Zhang, editors, *Proceedings of the 38th International Conference on Machine Learning, ICML 2021, 18–24 July 2021, Virtual Event*, volume 139 of *Proceedings of Machine Learning Research*, pp. 383–392. PMLR, 2021.
11. Marco Avella-Medina. Privacy-preserving parametric inference: A case for robust statistics. *Journal of the American Statistical Association*, 116(534):969–983, 2021.
12. Marco Avella-Medina, Casey Bradshaw, and Po-Ling Loh. Differentially private inference via noisy optimization. Technical Report arXiv:2103.11003 [math.ST], ArXiV, 11 2021.
13. Jordan Awan, Ana Kenney, Matthew Reimherr, and Aleksandra Slavković. Benefits and pitfalls of the exponential mechanism with applications to Hilbert spaces and functional PCA. In Kamalika Chaudhuri and Ruslan Salakhutdinov, editors, *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pp. 374–384. PMLR, 09–15 Jun 2019.
14. Jordan Awan and Aleksandra Slavković. Structure and sensitivity in differential privacy: Comparing k-norm mechanisms. *Journal of the American Statistical Association*, 116(534):935–954, 2021.
15. Francis R. Bach. Learning with submodular functions: A convex optimization perspective. *Foundations and Trends in Machine Learning*, 6(2-3):145–373, 2013.
16. Victor Balcer and Salil Vadhan. Differential privacy on finite computers. *Journal of Privacy and Confidentiality*, 9(2), 2019.
17. Borja Balle, Gilles Barthe, and Marco Gaboardi. Privacy amplification by subsampling: Tight analyses via couplings and divergences. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, vol. 31. Curran Associates, Inc., 2018.
18. Andrés F. Barrientos, Jerome P. Reiter, Ashwin Machanavajjhala, and Yan Chen. Differentially private significance tests for regression coefficients. *Journal of Computational and Graphical Statistics*, 28(2):440–453, 2019.
19. Raef Bassily, Vitaly Feldman, Kunal Talwar, and Abhradeep Guha Thakurta. Private stochastic convex optimization with optimal rates. In Hanna M. Wallach, Hugo Larochelle, Alina Beygelzimer, Florence d'Alché-Buc, Emily B. Fox, and Roman Garnett, editors, *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8–14, 2019, Vancouver, BC, Canada*, pp. 11279–11288, 2019.
20. Raef Bassily, Adam Smith, and Abhradeep Thakurta. Private empirical risk minimization: Efficient algorithms and tight error bounds. In *2014 IEEE 55th Annual Symposium on Foundations of Computer Science*, pp. 464–473, 2014.
21. Skye Berghel, Philip Bohannon, Damien Desfontaines, Charles Estes, Sam Haney, Luke Hartman, Michael Hay, Ashwin Machanavajjhala, Tom Magerlein, Gerome Miklau, Amritha Pai, William Sexton, and Ruchit Shrestha. Tumult analytics: a robust, easy-to-use, scalable, and expressive framework for differential privacy. Technical Report 2212.04133 [cs.CR], ArXiV, 2022.
22. Garrett Bernstein, Ryan McKenna, Tao Sun, Daniel Sheldon, Michael Hay, and Gerome Miklau. Differentially private learning of undirected graphical models using collective graphical models. In Doina Precup and Yee Whye Teh, editors, *Proceedings of the 34th International Conference on Machine Learning, ICML 2017, Sydney, NSW, Australia, 6–11 August 2017*, volume 70 of *Proceedings of Machine Learning Research*, pp. 478–487. PMLR, 2017.
23. Garrett Bernstein and Daniel Sheldon. Differentially private bayesian inference for exponential families. In Samy Bengio, Hanna M. Wallach, Hugo Larochelle, Kristen Grauman, Nicolò Cesa-Bianchi, and Roman Garnett, editors, *Advances in Neural Information Processing Systems 31: Annual Conference on Neural Information Processing Systems 2018, NeurIPS 2018, December 3–8, 2018, Montréal, Canada*, pp. 2924–2934, 2018.
24. Garrett Bernstein and Daniel 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*, vol. 32. Curran Associates, Inc., 2019.

25. A. Bhowmick, J. Duchi, J. Freudiger, G. Kapoor, and R. Rogers. Protection against reconstruction and its applications in private federated learning. Technical Report arXiv:1812.00984 [stat.ML], ArXiV, 6 2018.
26. Andrea Bittau, Úlfar Erlingsson, Petros Maniatis, Ilya Mironov, Ananth Raghunathan, David Lie, Mitch Rudominer, Ushasree Kode, Julien Tinnes, and Bernhard Seefeld. Prochlo: Strong privacy for analytics in the crowd. In *Proceedings of the 26th Symposium on Operating Systems Principles, SOSP '17*, pp. 441–459, New York, NY, USA, 2017. Association for Computing Machinery.
27. Jeremiah Blocki, Avrim Blum, Anupam Datta, and Or Sheffet. The Johnson-Lindenstrauss Transform itself preserves differential privacy. In *IEEE 53rd Annual Symposium on Foundations of Computer Science (FOCS)*, pp. 410–419, October 2012.
28. Hannah Brown, Katherine Lee, Fatemehsadat Mireshghallah, Reza Shokri, and Florian Tramèr. What does it mean for a language model to preserve privacy? In *FAccT '22: 2022 ACM Conference on Fairness, Accountability, and Transparency, Seoul, Republic of Korea, June 21–24, 2022*, pp. 2280–2292. ACM, 2022.
29. Zhiqi Bu, Zongyu Dai, Yiliang Zhang, and Qi Long. Misnn: Multiple imputation via semi-parametric neural networks. In *Advances in Knowledge Discovery and Data Mining: 27th Pacific-Asia Conference on Knowledge Discovery and Data Mining, PAKDD 2023, Osaka, Japan, May 25–28, 2023, Proceedings, Part I*, pp. 430–442, Berlin, Heidelberg, 2023. Springer-Verlag.
30. Mark Bun, Jonathan Ullman, and Salil Vadhan. Fingerprinting codes and the price of approximate differential privacy. *SIAM Journal on Computing*, 47(5):1888–1938, 2018.
31. T. Tony Cai, Yichen Wang, and Linjun Zhang. The cost of privacy in generalized linear models: Algorithms and minimax lower bounds. Technical Report arXiv:2011.03900 [stat.ML], ArXiV, 12 2020.
32. 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.
33. Clément L Canonne, Gautam Kamath, and Thomas Steinke. The discrete gaussian for differential privacy. In H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, vol. 33, pp. 15676–15688. Curran Associates, Inc., 2020.
34. Moses Charikar, Kevin Chen, and Martin Farach-Colton. Finding frequent items in data streams. In Peter Widmayer, Stephan Eidenbenz, Francisco Triguero, Rafael Morales, Ricardo Conejo, and Matthew Hennessy, editors, *Automata, Languages and Programming*, volume 2380 of *Lecture Notes in Computer Science*, pp. 693–703, Berlin, Heidelberg, 2002. Springer, Berlin Heidelberg.
35. K. Chaudhuri, C. Monteleoni, and A. D. Sarwate. Differentially private empirical risk minimization. *Journal of Machine Learning Research*, 12:1069–1109, 2011.
36. K. Chaudhuri, A. D. Sarwate, and K. Sinha. A near-optimal algorithm for differentially-private principal components. *Journal of Machine Learning Research*, 14:2905–2943, 2013.
37. Kamalika Chaudhuri and Daniel Hsu. Sample complexity bounds for differentially private learning. In Sham Kakade and Ulrike von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory (COLT '11)*, volume 19 of *JMLR Workshop and Conference Proceedings*, pp. 155–186, Budapest, Hungary, 6 2011.
38. Kamalika Chaudhuri and Claire Monteleoni. Privacy-preserving logistic regression. In D. Koller, D. Schuurmans, Y. Bengio, and L. Bottou, editors, *Advances in Neural Information Processing Systems*, vol. 21. Curran Associates, Inc., 2008.
39. Kamalika Chaudhuri and Staal A. Vinterbo. A stability-based validation procedure for differentially private machine learning. In C.J.C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems 26*. Curran Associates, Inc., December 2013.
40. Xiangyi Chen, Steven Z. Wu, and Mingyi Hong. Understanding gradient clipping in private sgd: A geometric perspective. In H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, vol. 33, pp. 13773–13782. Curran Associates, Inc., 2020.

41. Chris Clifton, Eric J. Hanson, Keith Merrill, and Shawn Merrill. Differentially private k-nearest neighbor missing data imputation. *ACM Transactions on Privacy and Security*, 25(3), 2022.
42. Graham Cormode and S. Muthukrishnan. An improved data stream summary: The count-min sketch and its applications. In Martín Farach-Colton, editor, *LATIN 2004: Theoretical Informatics*, volume 2976 of *Lecture Notes in Computer Science*, pp. 29–38, Berlin, Heidelberg, 2004. Springer, Berlin Heidelberg.
43. Rachel Cummings, Damien Desfontaines, David Evans, Roxana Geambasu, Matthew Jagielski, Yangsibo Huang, Peter Kairouz, Gautam Kamath, Sewoong Oh, Olga Ohrimenko, Nicolas Papernot, Ryan Rogers, Milan Shen, Shuang Song, Weijie Su, Andreas Terzis, Abhradeep Thakurta, Sergei Vassilvitskii, Yu-Xiang Wang, Li Xiong, Sergey Yekhanin, Da Yu, Huanyu Zhang, and Wanrong Zhang. Challenges towards the next frontier in privacy. Technical Report 2304.06929 [cs.CR], ArXiV, 2023.
44. Soumojit Das, Jorg Drechsler, Keith Merrill, and Shawn Merrill. Imputation under differential privacy. Technical Report 2206.15063 [cs.DB], ArXiV, 7 2022.
45. Christos Dimitrakakis, Blaine Nelson, Aikaterini Mitrokotsa, and Benjamin I. P. Rubinstein. Robust and private bayesian inference. In Peter Auer, Alexander Clark, Thomas Zeugmann, and Sandra Zilles, editors, *Algorithmic Learning Theory - 25th International Conference, ALT 2014, Bled, Slovenia, October 8–10, 2014. Proceedings*, volume 8776 of *Lecture Notes in Computer Science*, pp. 291–305. Springer, 2014.
46. Christos Dimitrakakis, Blaine Nelson, Zuhe Zhang, Aikaterini Mitrokotsa, and Benjamin I. P. Rubinstein. Differential privacy for bayesian inference through posterior sampling. *J. Mach. Learn. Res.*, 18:11:1–11:39, 2017.
47. Bolin Ding, Janardhan Kulkarni, and Sergey Yekhanin. Collecting telemetry data privately. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems 30*, pp. 3571–3580. Curran Associates, Inc., 2017.
48. Cynthia Dwork and Jing Lei. Differential privacy and robust statistics. In *Proceedings of the 41st Annual ACM Symposium on Theory of Computing (STOC '09)*, pp. 371–380, New York, 2009. ACM.
49. Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. In Shai Halevi and Tal Rabin, editors, *Theory of Cryptography*, volume 3876 of *Lecture Notes in Computer Science*, pp. 265–284, Berlin, Heidelberg, March 4–7, 2006. Springer.
50. Cynthia Dwork, Adam Smith, Thomas Steinke, and Jonathan Ullman. Exposed! a survey of attacks on private data. *Annual Review of Statistics and Its Application*, 4:61–84, 2017.
51. Cynthia Dwork, Adam Smith, Thomas Steinke, Jonathan Ullman, and Salil Vadhan. Robust traceability from trace amounts. In *2015 IEEE 56th Annual Symposium on Foundations of Computer Science*, pp. 650–669, 2015.
52. Cynthia Dwork, Weijie Su, and Li Zhang. Differentially private false discovery rate control. *Journal of Privacy and Confidentiality*, 11(2), 2021.
53. Cynthia Dwork, Kunal Talwar, Abhradeep Thakurta, and Li Zhang. Analyze Gauss: optimal bounds for privacy-preserving PCA. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing (STOC)*, pp. 11–20, 2014.
54. U. Erlingsson, V. Pihur, and A. Korolova. RAPPOR: Randomized aggregatable privacy-preserving ordinal response. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*, Scottsdale, Arizona, USA, 11 2014.
55. Dan Feldman, Amos Fiat, Haim Kaplan, and Kobbi Nissim. Private coresets. In Michael Mitzenmacher, editor, *Proceedings of the 41st Annual ACM Symposium on Theory of Computing, STOC 2009, Bethesda, MD, USA, May 31–June 2, 2009*, pp. 361–370. ACM, 2009.
56. Dan Feldman, Chongyuan Xiang, Ruihao Zhu, and Daniela Rus. Coresets for differentially private k-means clustering and applications to privacy in mobile sensor networks. In Pei Zhang, Prabal Dutta, and Guoliang Xing, editors, *Proceedings of the 16th ACM/IEEE International Conference on Information Processing in Sensor Networks, IPSN 2017, Pittsburgh, PA, USA, April 18–21, 2017*, pp. 3–15. ACM, 2017.

57. Vitaly Feldman, Tomer Koren, and Kunal Talwar. Private stochastic convex optimization: optimal rates in linear time. In Konstantin Makarychev, Yury Makarychev, Madhur Tulsiani, Gautam Kamath, and Julia Chuzhoy, editors, *Proccedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing, STOC 2020, Chicago, IL, USA, June 22–26, 2020*, pp. 439–449. ACM, 2020.
58. James Foulds, Joseph Geumlek, Max Welling, and Kamalika Chaudhuri. On the theory and practice of privacy-preserving bayesian data analysis. In *Proceedings of the 2016 Conference on Uncertainty in Artificial Intelligence*, 2016.
59. Joseph Geumlek, Shuang Song, and Kamalika Chaudhuri. Rényi differential privacy mechanisms for posterior sampling. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, vol. 30. Curran Associates, Inc., 2017.
60. Aditya Golatkar, Alessandro Achille, Yu-Xiang Wang, Aaron Roth, Michael Kearns, and Stefano Soatto. Mixed differential privacy in computer vision. In *IEEE / CVF Computer Vision and Pattern Recognition Conference (CVPR 2022)*, 2022.
61. Alon Gonem and Ram Gilad-Bachrach. Smooth sensitivity based approach for differentially private pca. In Firdaus Janoos, Mehryar Mohri, and Karthik Sridharan, editors, *Proceedings of Algorithmic Learning Theory*, volume 83 of *Proceedings of Machine Learning Research*, pp. 438–450. PMLR, 07–09 Apr 2018.
62. Ruobin Gong. Exact inference with approximate computation for differentially private data via perturbations. Technical Report arXiv:1909.12237 [stat.CO], ArXiV, 10 2019.
63. Samuel Haney, Damien Desfontaines, Luke Hartman, Ruchit Shrestha, and Michael Hay. Precision-based attacks and interval refining: how to break, then fix, differential privacy on finite computers. Technical Report 2207.13793 [cs.CR], ArXiV, 2022.
64. Moritz Hardt. Robust subspace iteration and privacy-preserving spectral analysis. Technical Report arXiv:1311.2495 [cs.DS], ArXiV, November 2013.
65. Moritz Hardt and Eric Price. The noisy power method: A meta algorithm with applications. In Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, vol. 27. Curran Associates, Inc., 2014.
66. Moritz Hardt and Aaron Roth. Beyond worst-case analysis in private singular vector computation. In *Proceedings of the Forty-Fifth Annual ACM Symposium on Theory of Computing (STOC)*, pp. 331–340, New York, June 2013. ACM.
67. Michael Hay, Marco Gaboardi, and Salil Vadhan. A programming framework for OpenDP. In *6th Workshop on the Theory and Practice of Differential Privacy (TPDP 2020)*, 2020.
68. Mikko A. Heikkilä, Joonas Jälkö, Onur Dikmen, and Antti Honkela. Differentially private markov chain monte carlo. In Hanna M. Wallach, Hugo Larochelle, Alina Beygelzimer, Florence d'Alché-Buc, Emily B. Fox, and Roman Garnett, editors, *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8–14, 2019, Vancouver, BC, Canada*, pp. 4115–4125, 2019.
69. Antti Honkela, Mrinal Das, Arttu Nieminen, Onur Dikmen, and Samuel Kaski. Efficient differentially private learning improves drug sensitivity prediction. *Biology Direct*, 13(1):1, 2018.
70. Hongsheng Hu, Zoran Salcic, Lichao Sun, Gillian Dobbie, Philip S. Yu, and Xuyun Zhang. Membership inference attacks on machine learning: A survey. *ACM Comput. Surv.*, Jan 2022.
71. Lijie Hu, Shuo Ni, Hanshen Xiao, and Di Wang. High dimensional differentially private stochastic optimization with heavy-tailed data. Technical Report, arXiv:2107.11136 [cs.LG], ArXiV, 9 2021.
72. Zhiyi Huang and Sampath Kannan. The exponential mechanism for social welfare: Private, truthful, and nearly optimal. In *53rd Annual IEEE Symposium on Foundations of Computer Science, FOCS 2012, New Brunswick, NJ, USA, October 20–23, 2012*, pp. 140–149. IEEE Computer Society, 2012.
73. Peter J. Huber and Elvezio M. Ronchetti. *Robust Statistics, Second Edition*. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc, New York, USA, 2009.
74. Jonathan H. Huggins, Ryan P. Adams, and Tamara Broderick. PASS-GLM: polynomial approximate sufficient statistics for scalable bayesian GLM inference. In Isabelle Guyon, Ulrike von

Luxburg, Samy Bengio, Hanna M. Wallach, Rob Fergus, S. V. N. Vishwanathan, and Roman Garnett, editors, *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, December 4–9, 2017, Long Beach, CA, USA*, pp. 3611–3621, 2017.

75. Christina Ilvento. Implementing differentially private integer partitions. Presented at the 2020 Workshop on the Theory and Practice of Differential Privacy, 11 2020.
76. Christina Ilvento. Implementing sparse vector. Presented at the 2020 Workshop on the Theory and Practice of Differential Privacy, 11 2020.
77. Christina Ilvento. Implementing the exponential mechanism with base-2 differential privacy. In *Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security, CCS '20*, pp. 717–742, New York, 2020. Association for Computing Machinery.
78. Roger Iyengar, Joseph P. Near, Dawn Song, Om Thakkar, Abhradeep Thakurta, and Lun Wang. Towards practical differentially private convex optimization. In *2019 IEEE Symposium on Security and Privacy (SP)*, pp. 299–316, 2019.
79. Prateek Jain and Abhradeep Guha Thakurta. (Near) dimension independent risk bounds for differentially private learning. In Eric P. Xing and Tony Jebara, editors, *Proceedings of the 31st International Conference on Machine Learning*, volume 32 of *Proceedings of Machine Learning Research*, pp. 476–484, Bejing, China, 22–24 Jun 2014. PMLR.
80. Joonas Jälkö, Onur Dikmen, and Antti Honkela. Differentially private variational inference for non-conjugate models. Technical Report arXiv:1610.08749 [stat.ML], ArXiV, 4 2017.
81. Matthew Jones, Huy L. Nguyen, and Thy D Nguyen. Differentially private clustering via maximum coverage. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(13):11555–11563, May 2021.
82. Peter Kairouz, Sewoong Oh, and Pramod Viswanath. The composition theorem for differential privacy. *IEEE Transactions on Information Theory*, 63(6):4037–4049, 6 2017.
83. Gautam Kamath, Xingtu Liu, and Huanyu Zhang. Improved rates for differentially private stochastic convex optimization with heavy-tailed data. In Kamalika Chaudhuri, Stefanie Jegelka, Le Song, Csaba Szepesvári, Gang Niu, and Sivan Sabato, editors, *International Conference on Machine Learning, ICML 2022, 17-23 July 2022, Baltimore, Maryland, USA*, volume 162 of *Proceedings of Machine Learning Research*, pp. 10633–10660. PMLR, 2022.
84. Gautam Kamath, Or Sheffet, Vikrant Singhal, and Jonathan R. Ullman. Differentially private algorithms for learning mixtures of separated gaussians. In Hanna M. Wallach, Hugo Larochelle, Alina Beygelzimer, Florence d'Alché-Buc, Emily B. Fox, and Roman Garnett, editors, *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8–14, 2019, Vancouver, BC, Canada*, pp. 168–180, 2019.
85. Mikhail Kapralov and Kunal Talwar. On differentially private low rank approximation. In *Proceedings of the Twenty-Fourth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA '13)*, pp. 1395–1414, New Orleans, LA, January 2013.
86. Vishesh Karwa and Aleksandra Slavković. Inference using noisy degrees: Differentially private β-model and synthetic graphs. *The Annals of Statistics*, 44(1):87–112, 2016.
87. Shiva Prasad Kasiviswanathan, Homin K. Lee, Kobbi Nissim, Sofya Raskhodnikova, and Adam Smith. What can we learn privately? *SIAM Journal on Computing*, 40(3):793–826, 6 2011.
88. Daniel Kifer, Adam Smith, and Abhradeep Thakurta. Private convex empirical risk minimization and high-dimensional regression. In Shie Mannor, Nathan Srebro, and Robert C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory (COLT '12)*, volume 23 of *JMLR Workshop and Conference Proceedings*, pp. 25.1–25.40, Edinburgh, Scotland, 2012.
89. Janardhan Kulkarni, Yin Tat Lee, and Daogao Liu. Private non-smooth ERM and SCO in subquadratic steps. In Marc'Aurelio Ranzato, Alina Beygelzimer, Yann N. Dauphin, Percy Liang, and Jennifer Wortman Vaughan, editors, *Advances in Neural Information Processing Systems 34: Annual Conference on Neural Information Processing Systems 2021, NeurIPS 2021, December 6–14, 2021, virtual*, pp. 4053–4064, 2021.

90. Tejas Kulkarni, Joonas Jälkö, Antti Koskela, Samuel Kaski, and Antti Honkela. Differentially private bayesian inference for generalized linear models. In Marina Meila and Tong Zhang, editors, *Proceedings of the 38th International Conference on Machine Learning, ICML 2021, 18-24 July 2021, Virtual Event*, volume 139 of *Proceedings of Machine Learning Research*, pp. 5838–5849. PMLR, 2021.
91. K S Sesh Kumar and Marc Peter Deisenroth. Differentially private empirical risk minimization with sparsity-inducing norms. Technical Report arXiv:1905.04873 [cs.LG], ArXiV, 2019.
92. Jaewoo Lee and Daniel Kifer. Concentrated differentially private gradient descent with adaptive per-iteration privacy budget. In Yike Guo and Faisal Farooq, editors, *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD 2018, London, UK, August 19–23, 2018*, pp. 1656–1665. ACM, 2018.
93. Jing Lei. Differentially private m-estimators. In John Shawe-Taylor, Richard S. Zemel, Peter L. Bartlett, Fernando C. N. Pereira, and Kilian Q. Weinberger, editors, *Advances in Neural Information Processing Systems 24: 25th Annual Conference on Neural Information Processing Systems 2011. Proceedings of a meeting held 12–14 December 2011, Granada, Spain*, pp. 361–369, 2011.
94. Jing Lei, Anne-Sophie Charest, Aleksandra Slavkovic, Adam Smith, and Stephen Fienberg. Differentially private model selection with penalized and constrained likelihood. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 181(3):609–633, 2018.
95. Bai Li, Changyou Chen, Hao Liu, and Lawrence Carin. On connecting stochastic gradient MCMC and differential privacy. In Kamalika Chaudhuri and Masashi Sugiyama, editors, *The 22nd International Conference on Artificial Intelligence and Statistics, AISTATS 2019, 16–18 April 2019, Naha, Okinawa, Japan*, volume 89 of *Proceedings of Machine Learning Research*, pp. 557–566. PMLR, 2019.
96. Jingcheng Liu and Kunal Talwar. Private selection from private candidates. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing, STOC 2019*, pp. 298–309, New York, 2019. Association for Computing Machinery.
97. H. Brendan McMahan, Galen Andrew, Ulfar Erlingsson, Steve Chien, Ilya Mironov, Nicolas Papernot, and Peter Kairouz. A general approach to adding differential privacy to iterative training procedures. In *NeurIPS 2018 Workshop on Privacy Preserving Machine Learning*, 2018.
98. H. Brendan McMahan, Daniel Ramage, Kunal Talwar, and Li Zhang. Learning differentially private recurrent language models. In *6th International Conference on Learning Representations, ICLR 2018, Vancouver, BC, Canada, April 30–May 3, 2018, Conference Track Proceedings*. OpenReview.net, 2018.
99. Frank McSherry and Kunal Talwar. Mechanism design via differential privacy. In *48th Annual IEEE Symposium on Foundations of Computer Science (FOCS '07)*, pp. 94–103, 10 2007.
100. Kentaro Minami, Hiromi Arai, Issei Sato, and Hiroshi Nakagawa. Differential privacy without sensitivity. In Daniel D. Lee, Masashi Sugiyama, Ulrike von Luxburg, Isabelle Guyon, and Roman Garnett, editors, *Advances in Neural Information Processing Systems 29: Annual Conference on Neural Information Processing Systems 2016, December 5-10, 2016, Barcelona, Spain*, pp. 956–964, 2016.
101. Darakhshan J Mir. *Differential privacy: an exploration of the privacy-utility landscape*. PhD Thesis, Rutgers, The State University of New Jersey, 2013.
102. Ilya Mironov. On significance of the least significant bits for differential privacy. In *Proceedings of the 2012 ACM Conference on Computer and Communications Security (CCS)*, pp. 650–661, 2012.
103. Ilya Mironov. Rényi differential privacy. In *2017 IEEE 30th Computer Security Foundations Symposium (CSF)*, pp. 263–275, 8 2017.
104. Seth Neel, Aaron Roth, Giuseppe Vietri, and Steven Wu. Oracle efficient private non-convex optimization. In Hal Daumé III and Aarti Singh, editors, *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 7243–7252. PMLR, 13–18 Jul 2020.
105. Huy L. Nguyen, Anamay Chaturvedi, and Eric Z Xu. Differentially private k-means via exponential mechanism and max cover. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(10):9101–9108, May 2021.

106. Kobbi Nissim, Sofya Raskhodnikova, and Adam Smith. Smooth sensitivity and sampling in private data analysis. In *Proceedings of the Thirty-Ninth Annual ACM Symposium on Theory of Computing (STOC '07)*, pp. 75–84, New York, 2007. ACM.
107. Jorge Nocedal and Stephen J. Wright. *Numerical Optimization*. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2006.
108. OpenDP. Smartnoise samples: Differential privacy examples, notebooks and documentation, 2021.
109. Nicolas Papernot, Martín Abadi, Úlfar Erlingsson, Ian J. Goodfellow, and Kunal Talwar. Semi-supervised knowledge transfer for deep learning from private training data. In *5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24–26, 2017, Conference Track Proceedings*. OpenReview.net, 2017.
110. Nicolas Papernot, Shuang Song, Ilya Mironov, Ananth Raghunathan, Kunal Talwar, and Úlfar Erlingsson. Scalable private learning with PATE. In *6th International Conference on Learning Representations, ICLR 2018, Vancouver, BC, Canada, April 30–May 3, 2018, Conference Track Proceedings*. OpenReview.net, 2018.
111. Nicolas Papernot, Abhradeep Thakurta, Shuang Song, Steve Chien, and Úlfar Erlingsson. Tempered sigmoid activations for deep learning with differential privacy. In *Thirty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2021, Thirty-Third Conference on Innovative Applications of Artificial Intelligence, IAAI 2021, The Eleventh Symposium on Educational Advances in Artificial Intelligence, EAAI 2021, Virtual Event, February 2-9, 2021*, pp. 9312–9321. AAAI Press, 2021.
112. Mijung Park, James Foulds, Kamalika Chaudhuri, and Max Welling. DP-EM: Differentially Private Expectation Maximization. In Aarti Singh and Jerry Zhu, editors, *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, volume 54 of *Proceedings of Machine Learning Research*, pp. 896–904. PMLR, 4 2017.
113. Mijung Park, James R. Foulds, Kamalika Chaudhuri, and Max Welling. Variational bayes in private settings (VIPS). *J. Artif. Intell. Res.*, 68:109–157, 2020.
114. Víctor Peña and Andrés F. Barrientos. Differentially private methods for managing model uncertainty in linear regression models. Technical Report arXiv:2109.03949 [stat.ME], ArXiV, 2021.
115. Venkatadheeraj Pichapati, Ananda Theertha Suresh, Felix X. Yu, Sashank J. Reddi, and Sanjiv Kumar. Adaclip: Adaptive clipping for private sgd. Technical Report arXiv:1908.07643 [cs.LG], ArXiV, 10 2019.
116. Benjamin I. P. Rubinstein, Peter L. Bartlett, Ling Huang, and Nina Taft. Learning in a large function space: Privacy-preserving mechanisms for SVM learning. *Journal of Privacy and Confidentiality*, 4(1):65–100, 2012.
117. Anand D. Sarwate, Sergey M. Plis, Jessica A. Turner, Mohammad R. Arbabshirani, and Vince D. Calhoun. Sharing privacy-sensitive access to neuroimaging and genetics data: a review and preliminary validation. *Frontiers in Neuroinformatics*, 8(35), 2014.
118. V. Schellekens, A. Chatalic, F. Houssiau, Y.-A. de Montjoye, L. Jacques, and R. Gribonval. Differentially private compressive k-means. In *ICASSP 2019–2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 7933–7937, 2019.
119. Moshe Shechner, Or Sheffet, and Uri Stemmer. Private k-means clustering with stability assumptions. In Silvia Chiappa and Roberto Calandra, editors, *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*, volume 108 of *Proceedings of Machine Learning Research*, pp. 2518–2528. PMLR, 26–28 Aug 2020.
120. Or Sheffet. Differentially private ordinary least squares. In Doina Precup and Yee Whye Teh, editors, *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pp. 3105–3114. PMLR, 6–11 Aug 2017.
121. Or Sheffet. Old techniques in differentially private linear regression. In Aurélien Garivier and Satyen Kale, editors, *Proceedings of the 30th International Conference on Algorithmic Learning Theory*, volume 98 of *Proceedings of Machine Learning Research*, pp. 789–827. PMLR, 22–24 Mar 2019.
122. Vikrant Singhal and Thomas Steinke. Privately learning subspaces. Technical Report 2106.00001 [cs.CR], ArXiV, 8 2021.

123. Aleksandra Slavković and Roberto Molinari. Perturbed m-estimation: A further investigation of robust statistics for differential privacy. Technical Report arXiv:2108.08266 [cs.CR], ArXiV, 8 2021.
124. 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*, pp. 813–822, New York, 2011. Association for Computing Machinery.
125. Shuang Song, Kamalika Chaudhuri, and Anand D. Sarwate. Stochastic gradient descent with differentially private updates. In *Proceedings of the 2013 Global Conference on Signal and Information Processing (GlobalSIP 2013)*, pp. 245–248, December 2013.
126. Shuang Song, Thomas Steinke, Om Thakkar, and Abhradeep Thakurta. Evading the curse of dimensionality in unconstrained private glms. In Arindam Banerjee and Kenji Fukumizu, editors, *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics*, volume 130 of *Proceedings of Machine Learning Research*, pp. 2638–2646. PMLR, 13–15 Apr 2021.
127. Thomas Steinke and Jonathan Ullman. Between pure and approximate differential privacy. *Journal of Privacy and Confidentiality*, 7(2), 2016.
128. Uri Stemmer. Locally private *k*-means clustering. In *Proceedings of the 2020 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pp. 548–559, 2020.
129. Uri Stemmer and Haim Kaplan. Differentially private k-means with constant multiplicative error. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, vol. 31. Curran Associates, Inc., 2018.
130. Dong Su, Jianneng Cao, Ninghui Li, Elisa Bertino, and Hongxia Jin. Differentially private k-means clustering. In *Proceedings of the Sixth ACM Conference on Data and Application Security and Privacy, CODASPY '16*, pp. 26–37, New York, NY, USA, 2016. Association for Computing Machinery.
131. Dong Su, Jianneng Cao, Ninghui Li, Elisa Bertino, Min Lyu, and Hongxia Jin. Differentially private k-means clustering and a hybrid approach to private optimization. *ACM Transactions on Privacy and Security*, 20(4), 2017.
132. Jinyan Su, Lijie Hu, and Di Wang. Faster rates of private stochastic convex optimization. In Sanjoy Dasgupta and Nika Haghtalab, editors, *Proceedings of The 33rd International Conference on Algorithmic Learning Theory*, volume 167 of *Proceedings of Machine Learning Research*, pp. 995–1002. PMLR, 29 Mar–01 Apr 2022.
133. Kunal Talwar, Abhradeep Thakurta, and Li Zhang. Nearly optimal private LASSO. In Corinna Cortes, Neil D. Lawrence, Daniel D. Lee, Masashi Sugiyama, and Roman Garnett, editors, *Advances in Neural Information Processing Systems 28: Annual Conference on Neural Information Processing Systems 2015, December 7–12, 2015, Montreal, Quebec, Canada*, pp. 3025–3033, 2015.
134. Kunal Talwar, Abhradeep Thakurta, and Li Zhang. Private empirical risk minimization beyond the worst case: The effect of the constraint set geometry. Technical Report arXiv:1411.5417v3 [cs.LG], ArXiV, 11 2016.
135. Jun Tang, Aleksandra Korolova, Xiaolong Bai, Xueqiang Wang, and Xiaofeng Wang. Privacy loss in Apple's implementation of differential privacy on MacOS 10.12. Technical Report arXiv:1709.02753 [cs.CR], ArXiV, 9 2017.
136. Youming Tao, Yulian Wu, Xiuzhen Cheng, and Di Wang. Private stochastic convex optimization and sparse learning with heavy-tailed data revisited. In Luc De Raedt, editor, *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI 2022, Vienna, Austria, 23–29 July 2022*, pp. 3947–3953. ijcai.org, 2022.
137. Abhradeep Thakurta and Adam D. Smith. Differentially private feature selection via stability arguments, and the robustness of the lasso. In Shai Shalev-Shwartz and Ingo Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12-14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Workshop and Conference Proceedings*, pp. 819–850. JMLR.org, 2013.
138. Di Wang, Changyou Chen, and Jinhui Xu. Differentially private empirical risk minimization with non-convex loss functions. In Kamalika Chaudhuri and Ruslan Salakhutdinov, editors, *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pp. 6526–6535. PMLR, 09–15 Jun 2019.

139. Di Wang, Jiahao Ding, Zejun Xie, Miao Pan, and Jinhui Xu. Differentially private (gradient) expectation maximization algorithm with statistical guarantees. Technical Report arXiv:2010.13520 [cs.LG], ArXiV, 2020.
140. Di Wang, Hanshen Xiao, Srinivas Devadas, and Jinhui Xu. On differentially private stochastic convex optimization with heavy-tailed data. In *Proceedings of the 37th International Conference on Machine Learning, ICML 2020, 13–18 July 2020, Virtual Event*, volume 119 of *Proceedings of Machine Learning Research*, pp. 10081–10091. PMLR, 2020.
141. Di Wang and Jinhui Xu. Differentially private empirical risk minimization with smooth non-convex loss functions: A non-stationary view. In *The Thirty-Third AAAI Conference on Artificial Intelligence, AAAI 2019, The Thirty-First Innovative Applications of Artificial Intelligence Conference, IAAI 2019, The Ninth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2019, Honolulu, Hawaii, USA, January 27 - February 1, 2019*, pp. 1182–1189. AAAI Press, 2019.
142. Di Wang, Minwei Ye, and Jinhui Xu. Differentially private empirical risk minimization revisited: Faster and more general. In Isabelle Guyon, Ulrike von Luxburg, Samy Bengio, Hanna M. Wallach, Rob Fergus, S. V. N. Vishwanathan, and Roman Garnett, editors, *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, December 4-9, 2017, Long Beach, CA, USA*, pp. 2722–2731, 2017.
143. Yining Wang, Yu-Xiang Wang, and Aarti Singh. Differentially private subspace clustering. In C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, vol. 28. Curran Associates, Inc., 2015.
144. Yu-Xiang Wang. Revisiting differentially private linear regression: optimal and adaptive prediction & estimation in unbounded domain. In Amir Globerson and Ricardo Silva, editors, *Proceedings of the Thirty-Fourth Conference on Uncertainty in Artificial Intelligence, UAI 2018, Monterey, California, USA, August 6–10, 2018*, pp. 93–103. AUAI Press, 2018.
145. Yu-Xiang Wang, Borja Balle, and Shiva Prasad Kasiviswanathan. Subsampled renyi differential privacy and analytical moments accountant. In Kamalika Chaudhuri and Masashi Sugiyama, editors, *Proceedings of the Twenty-Second International Conference on Artificial Intelligence and Statistics*, volume 89 of *Proceedings of Machine Learning Research*, pp. 1226–1235. PMLR, 16–18 Apr 2019.
146. Yu-Xiang Wang, Stephen Fienberg, and Alex Smola. Privacy for free: Posterior sampling and stochastic gradient monte carlo. In Francis Bach and David Blei, editors, *Proceedings of the 32nd International Conference on Machine Learning*, volume 37 of *Proceedings of Machine Learning Research*, pp. 2493–2502, Lille, France, 7 2015. PMLR.
147. Yue Wang, Daniel Kifer, and Jaewoo Lee. Differentially private confidence intervals for empirical risk minimization. *Journal of Privacy and Confidentiality*, 9(1), 2019.
148. L. Wasserman and S. Zhou. A statistical framework for differential privacy. *Journal of the American Statistical Association*, 105(489):375–389, 2010.
149. Lu Wei, Anand D. Sarwate, Jukka Corander, Alfred Hero, and Vahid Tarokh. Analysis of a privacy-preserving PCA algorithm using random matrix theory. In *Proceedings of the 2016 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pp. 1335–1339, Washington, DC, USA, 7–9 December 2016.
150. Oliver Williams and Frank McSherry. Probabilistic inference and differential privacy. In J.D. Lafferty, C.K.I. Williams, J. Shawe-Taylor, R.S. Zemel, and A. Culotta, editors, *Advances in Neural Information Processing Systems 23*, pp. 2451–2459. Curran Associates, Inc., 2010.
151. David P. Woodruff. Sketching as a tool for numerical linear algebra. *Foundations and Trends in Theoretical Computer Science*, 10(1–2):1–157, 2014.
152. Xi Wu, Matthew Fredrikson, Wentao Wu, Somesh Jha, and Jeffrey F. Naughton. Revisiting differentially private regression: Lessons from learning theory and their consequences. *CoRR*, abs/1512.06388, 2015.
153. Xi Wu, Fengan Li, Arun Kumar, Kamalika Chaudhuri, Somesh Jha, and Jeffrey Naughton. Bolt-on differential privacy for scalable stochastic gradient descent-based analytics. In *Proceedings of the 2017 ACM International Conference on Management of Data, SIGMOD '17*, pp. 1307–1322, New York, 2017. Association for Computing Machinery.

154. Sinan Yıildıirıim and Beyza Ermiş. Exact mcmc with differentially private moves. *Statistics and Computing*, 29(5):947–963, 2019.
155. Ashkan Yousefpour, Igor Shilov, Alexandre Sablayrolles, Davide Testuggine, Karthik Prasad, Mani Malek, John Nguyen, Sayan Ghosh, Akash Bharadwaj, Jessica Zhao, Graham Cormode, and Ilya Mironov. Opacus: User-friendly differential privacy library in PyTorch. *arXiv preprint arXiv:2109.12298*, 2021.
156. Fei Yu, Michal Rybar, Caroline Uhler, and Stephen E Fienberg. Differentially-private logistic regression for detecting multiple-snp association in gwas databases. In *International Conference on Privacy in Statistical Databases*, pp. 170–184. Springer, 2014.
157. J. Zhang, Z. Zhang, X. Xiao, Y. Yang, and M. Winslett. Functional mechanism: Regression analysis under differential privacy. *Proceedings of the VLDB Endowment*, 5(11):1364–1375, 2012.
158. Jiaqi Zhang, Kai Zheng, Wenlong Mou, and Liwei Wang. Efficient private ERM for smooth objectives. In Carles Sierra, editor, *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI 2017, Melbourne, Australia, August 19–25, 2017*, pp. 3922–3928. ijcai.org, 2017.
159. Jun Zhang, Graham Cormode, Cecilia M. Procopiuc, Divesh Srivastava, and Xiaokui Xiao. Privbayes: private data release via Bayesian networks. In Curtis E. Dyreson, Feifei Li, and M. Tamer Özsu, editors, *International Conference on Management of Data, SIGMOD 2014, Snowbird, UT, USA, June 22–27, 2014*, pp. 1423–1434. ACM, 2014.
160. Jun Zhang, Graham Cormode, Cecilia M. Procopiuc, Divesh Srivastava, and Xiaokui Xiao. PrivBayes: Private data release via bayesian networks. *ACM Transactions on Database Systems*, 42(4):1–41, 2017.
161. Jun Zhang, Xiaokui Xiao, Yin Yang, Zhenjie Zhang, and Marianne Winslett. Privgene: differentially private model fitting using genetic algorithms. In Kenneth A. Ross, Divesh Srivastava, and Dimitris Papadias, editors, *Proceedings of the ACM SIGMOD International Conference on Management of Data, SIGMOD 2013, New York, NY, USA, June 22–27, 2013*, pp. 665–676. ACM, 2013.
162. Zhe Zhang and Linjun Zhang. High-dimensional differentially-private EM algorithm: Methods and near-optimal statistical guarantees. Technical Report arXiv:2104.00245 [stat.ML], ArXiV, 9 2021.
163. Zuhe Zhang, Benjamin I. P. Rubinstein, and Christos Dimitrakakis. On the differential privacy of bayesian inference. In Dale Schuurmans and Michael P. Wellman, editors, *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, February 12-17, 2016, Phoenix, Arizona, USA*, pp. 2365–2371. AAAI Press, 2016.
164. Yingxue Zhou, Steven Wu, and Arindam Banerjee. Bypassing the ambient dimension: Private SGD with gradient subspace identification. In *9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3–7, 2021*. OpenReview.net, 2021.