Nearly Optimal Private Convolution

Nadia Fawaz[1], S. Muthukrishnan[2], and Aleksandar Nikolov[2]

[1] Technicolor, Palo Alto, CA
[2] Rutgers University

Abstract. We study algorithms for computing the convolution of a private input x with a public input h, while satisfying the guarantees of (ε, δ)-differential privacy. Convolution is a fundamental operation, intimately related to Fourier Transforms. In our setting, the private input may represent a time series of sensitive events or a histogram of a database of confidential personal information. Convolution then captures important primitives including linear filtering, which is an essential tool in time series analysis, and aggregation queries on projections of the data. We give an algorithm for computing convolutions which satisfies (ε, δ)-differentially privacy and is nearly optimal *for every public* h, i.e. is instance optimal with respect to the public input. We prove optimality via spectral lower bounds on the *hereditary discrepancy* of convolution matrices. Our algorithm is very efficient – it is essentially no more computationally expensive than a Fast Fourier Transform. [3]

1 Introduction

Much useful data contains sensitive information about individuals (or the actions they take): typical examples are census data, data from medical studies, and financial data. While analyzing such sensitive datasets is valuable for scientific studies, policy and decision making, care must be taken to protect the privacy of the individuals represented in the data. Simple measures such as removing personally identifying attributes, replacing names with pseudonyms and publishing only aggregate statistics have proved inadequate protection from sophisticated linkage attacks [27,25,26]. An extreme solution would be to remove all sensitive information from the datasets, but this approach can destroy the utility of the data: a medical study without disease incidence rates would be useless, for example. In recent years *differential privacy* [10] has become a standard framework in which to reason about trade offs between privacy and utility, and this is the framework we adopt in this paper.

We study the *noise complexity* of a special class of queries. Consider a database representing users of N different types, or a time series of events that occurred over N time steps. We may encode the database as a vector $\mathbf{x}$ indexed by $\{1, \ldots, N\}$, where x_i gives the number of users of type i, in the database example, or x_i is the count of events that occurred at time step i. We say that

[3] The full version of this paper can be found at `http://arxiv.org/abs/1301.6447`

two such vectors $\mathbf{x}$ and $\mathbf{x}'$ are *neighbors* when $\|\mathbf{x} - \mathbf{x}'\|_1 \leq 1$. Neighboring input vectors correspond to databases that differ in at most a single user/event. Informally, an algorithm is differentially private if its output distribution is almost identical for neighboring inputs. More precisely, a randomized algorithm $\mathcal{A}$ satisfies (ε, δ)-differential privacy if for all neighbors $\mathbf{x}, \mathbf{x}' \in [0,1]^n$, and all measurable subsets T of the range of $\mathcal{A}$, we have

$$\Pr[\mathcal{A}(\mathbf{x}) \in T] \leq e^{\varepsilon} \Pr[\mathcal{A}(\mathbf{x}') \in T] + \delta,$$

where probabilities are taken over the randomness of $\mathcal{A}$.

In this work we are interested in *workloads* of M *linear queries*, given as a matrix $\mathbf{A}$; the intended output for the workload is $\mathbf{Ax}$. Differential privacy necessitates randomization and approximation for all non-trivial workloads; we discuss accuracy in terms of *mean squared error* (MSE) as a measure of approximation: the expected average of squared error over all M queries. The MSE achieved by an algorithm is the worst MSE the algorithm achieves on any input database.

The queries in a workload $\mathbf{A}$ can have different degrees of correlation, and this poses different challenges for the private approximation algorithm. In one extreme, when $\mathbf{A}$ is a set of $\Omega(N)$ independently sampled random $\{0,1\}$ (i.e. counting) queries differentially private algorithm needs to incur at least $\Omega(N)$ squared error per query on average [9]. On the other hand, if $\mathbf{A}$ consists of the same counting query repeated M times, we only need to add $O(1)$ noise per query [10]. While these two extremes are well understood, relatively less is known about workloads of queries with some, but not perfect, correlation.

The *convolution*[4] $y = x * y$ of the private input sequence x with a public sequence h is defined as

$$y_i = \sum_{j=0}^{N-1} h_j x_{i-j \bmod N}.$$

If we view the input sequence as a vector $\mathbf{x}$, and define the circulant *convolution matrix* $\mathbf{H} = (h_{N+j-i \bmod N})_{i,j \in \{0,\ldots,N-1\}}$, we see the convolution map is equivalent to computing the N linear queries $\mathbf{Hx}$. Each query is a circular shift of the previous one, and, therefore, the queries are far from independent but not identical either. Convolution is a fundamental operation that arises in algebraic computations such as polynomial multiplication, in signal analysis, and has well known connection to Fourier transforms. Of primary interest to us, it is a natural primitive in various applications:

– linear filters in the analysis of time series data can be cast as convolutions; as example applications, linear filtering can be used to isolate cycle components in time series data from spurious variations, and to compute time-decayed statistics of the data;

[4] Here we define circular convolution, but, as discussed in the paper, our results generalize to other types of convolution, which are defined similarly.

– when user type in the database is specified by d binary attributes, aggregate queries such as k-wise marginals and generalizations to other predicate queries can be represented as convolutions.

Privacy concerns arise naturally in these applications: the time series data can contain records of sensitive events, such as financial transactions, records of user activity, etc.; some of the attributes in a database can be sensitive, for example when dealing with databases of medical data.

We give the first (ε, δ)-differentially private algorithm which is nearly *query*-optimal: it achieves MSE which is not much smaller than the smallest MSE that any (ε, δ)-differentially private algorithm can achieve on the given convolution query.[5]

To prove the optimality of our algorithm, we need to prove *optimal lower bounds* on the noise complexity of private algorithms for computing convolutions. We use the recent discrepancy-based noise lower bounds of Muthukrishnan and Nikolov [24]. We use a characterization of combinatorial discrepancy in terms of determinants of submatrices discovered by Lovász, Spencer, and Vesztergombi [23], together with ideas by Hardt and Talwar [18]. A main technical ingredient in the proof of our lower bound is a connection between the discrepancy of a matrix $\mathbf{A}$ and the discrepancy of $\mathbf{PA}$ where $\mathbf{P}$ is an orthogonal projection operator.

Related work. The problem of computing private convolutions has not been considered in the literature before. However, there is a fair amount of work on the more general problem of computing arbitrary linear queries, as well as some work on special cases of convolution maps.

Bolot et al. [4] give algorithms for various decayed sum queries: window sums, exponentially and polynomially decayed sums. Any decayed sum function is a type of linear filter, and, therefore, a special case of convolution. Thus, our current work gives a nearly optimal (ε, δ)-differentially private approximation for *any decayed sum function*. Moreover, as far as mean squared error is concerned, our algorithms give improved error bounds for the window sums problem: constant squared error per query. However, unlike [4], we only consider the offline batch-processing setting, as opposed to the online continual observation setting.

The work of Barak et al. [1] on computing k-wise marginals concerns a restricted class of convolutions (see Section 5). Moreover, Kasiviswanathan [19] show a noise lower bound for k-wise marginals which is tight in the worst case. Our work is a generalization: we are able to give nearly optimal approximations to a wider class of queries, and our lower and upper bounds nearly match for any convolution.

Li and Miklau [21,22] proposed the class of extended matrix mechanisms, building on prior work on the matrix mechanism [20], and showed how to efficiently compute the optimal mechanism from the class. Since our mechanism is a special instance of the extended matrix mechanism, the algorithms of Li and

[5] We note that while our algorithm is instance optimal with respect to queries, the measure of error we use is still worst-case over databases.

Miklau have at most as much error as our algorithm. They also derived a spectral lower bound [22] on the extended matrix mechanism; their results further imply that the spectral lower bound is tight *for the extended mechanism* for workloads corresponding to convolutions. However, unlike our lower bounds, this has no direct implication for private algorithms which are not an instantiation of the matrix mechanism.

Independently and concurrently with our work, Cormode et al. [8] considered adding optimal non-uniform noise to a fixed transform of the private database. Similarly to [8], we gain significantly in efficiency over the general extended matrix mechanism by fixing a specific transform (in our case the Fourier transform) of the data and computing a closed form expression for the optimal noise magnitudes. Our lower bounds show that, somewhat surprisingly, this simplification of Cormode et al. in fact comes without loss of generality for *any* set of convolution queries.

In the setting of $(\epsilon, 0)$-differential privacy, [18,2] prove nearly optimal upper and lower bounds on approximating $\mathbf{Ax}$ for any matrix $\mathbf{A}$. Prior to our work a similar result was not known for the weaker notion of approximate privacy, i.e. (ε, δ)-differential privacy. After a preliminary version of this paper was made available, our results were generalized by Nikolov, Talwar, and Zhang [28] to give nearly optimal algorithms for computing any linear map A under (ε, δ)-differential privacy. However, this comes at the cost of higher computational complexity: even the algorithm from [28], which is more efficient than the algorithms from [18,2], has running time $\Omega(N^3)$, as it needs to approximate the minimum enclosing ellipsoid of an N-dimensional convex body. By contrast our algorithm's running time is dominated by the running time of the Fast Fourier Transform, i.e. $O(N \log N)$, making it suitable for practical applications.

A related line of research exploits sparsity assumptions on the private database in order to reduce error [3,11,16,28]. Using techniques from learning theory, more efficient algorithms for sparse databases have been designed for the set of marginal queries [15,17,7,29,6]. As we do not limit the database size, our results are not directly comparable. Also, our lower bounds already hold when the database size (which in our notation corresponds to $\|\mathbf{x}\|_1$) is at least the number of linear queries, and in that regime our algorithm is nearly optimal, and cannot be significantly improved in terms of noise complexity. Finally, note that the optimal error for *a subset of all marginal queries* may be less than linear in database size, and our algorithms will give near optimal error for the *specific subset* of interest.

Recent work [17,7,29] on privately answering marginal queries has taken the approach of treating the database as a function from queries to the reals, and approximating this function by a small degree polynomial. This technique bears some resemblance to our approach for generalized marginals: we compute the Fourier transform of the database privately and spend most of the privacy budget on lower order Fourier coefficients, since they carry the most information.

Organization. We begin with preliminaries on differential privacy and convolution operators. In section 3 we derive our main lower bound result, and in

section 4 we describe and analyze our nearly optimal algorithm. In section 5 we describe applications of our main results.

2 Preliminaries

Notation: $\mathbb{N}$, $\mathbb{R}$, and $\mathbb{C}$ are the sets of non-negative integers, real, and complex numbers respectively. By log we denote the logarithm in base 2 while by ln we denote the logarithm in base e. Matrices and vectors are represented by boldface upper and lower cases, respectively. $\mathbf{A}^T$, $\mathbf{A}^*$, $\mathbf{A}^H$ stand for the transpose, the conjugate and the transpose conjugate of $\mathbf{A}$, respectively. The trace and the determinant of $\mathbf{A}$ are respectively denoted by $\mathrm{tr}(\mathbf{A})$ and $\det(\mathbf{A})$. $\mathbf{A}_{m:}$ denotes the m-th row of matrix $\mathbf{A}$, and $\mathbf{A}_{:n}$ its n-th column. $\mathbf{A}|_S$, where $\mathbf{A}$ is a matrix with N columns and $S \subseteq [N]$, denotes the submatrix of $\mathbf{A}$ consisting of those columns corresponding to elements of S. $\lambda_{\mathbf{A}}(1), \ldots, \lambda_{\mathbf{A}}(n)$ represent the eigenvalues of an $n \times n$ matrix $\mathbf{A}$. $\mathbf{I}_N$ is the identity matrix of size N. $\mathrm{E}[\cdot]$ is the statistical expectation operator. $\mathrm{Lap}(x, s)$ denotes the Laplace distribution centered at x with scale s, i.e. the distribution of the random variable $x + \eta$ where η has probability density function $p(y) \propto \exp(-|y|/s)$.

2.1 Fourier Eigen-decomposition of Convolution

In this section, we recall the definition of the Fourier basis, and the eigen-decomposition of circular convolution in this basis.

Definition 1. *The normalized Discrete Fourier Transform (DFT) matrix of size N is defined as*

$$\mathbf{F}_N = \left(\frac{1}{\sqrt{N}} \exp\left(-\frac{j2\pi\, m\, n}{N} \right) \right)_{m,n \in \{0,\ldots,N-1\}}. \tag{1}$$

Note that $\mathbf{F}_N$ is symmetric ($\mathbf{F}_N = \mathbf{F}_N^T$) and unitary ($\mathbf{F}_N \mathbf{F}_N^H = \mathbf{F}_N^H \mathbf{F}_N = \mathbf{I}_N$).

We denote by $\mathbf{f}_m = N^{-1/2}(1, e^{\frac{j2\pi\, m}{N}}, \ldots, e^{\frac{j2\pi\, m\, (N-1)}{N}})^T \in \mathbb{C}^N$ the m-th column of the inverse DFT matrix $\mathbf{F}_N^H$. Or alternatively, $\mathbf{f}_m^H$ is the m-th row of $\mathbf{F}_N$. The normalized DFT of a vector $\mathbf{h}$ is simply given by $\hat{\mathbf{h}} = \mathbf{F}_N \mathbf{h}$.

Theorem 1 ([14]). *Any circulant matrix $\mathbf{H}$ can be diagonalized in the Fourier basis $\mathbf{F}_N$: the eigenvectors of $\mathbf{H}$ are given by the columns $(\mathbf{f}_m)_{m \in \{0,\ldots,N-1\}}$ of the inverse DFT matrix $\mathbf{F}_N^H$, and the associated eigenvalues $\{\lambda_m\}_{m \in \{0,\ldots,N-1\}}$ are given by $\sqrt{N}\hat{\mathbf{h}}$, i.e. by the DFT of the first column $\mathbf{h}$ of $\mathbf{H}$:*

$$\forall m \in \{0, \ldots, N-1\}, \quad \mathbf{H}\mathbf{f}_m = \lambda_m \mathbf{f}_m$$
$$\text{where} \quad \lambda_m = \sqrt{N}\hat{h}_m = \sum_{n=0}^{N-1} h_n e^{-\frac{j2\pi\, m\, n}{N}}.$$

Equivalently, in the Fourier domain, the circular convolution matrix $\mathbf{H}$ becomes a diagonal matrix $\hat{\mathbf{H}} = \mathrm{diag}\{\sqrt{N}\hat{\mathbf{h}}\}$.

Corollary 1 *Consider the circular convolution* $\mathbf{y} = \mathbf{H}\mathbf{x}$ *of* $\mathbf{x}$ *and* $\mathbf{y}$*. Let* $\hat{\mathbf{x}} = \mathbf{F}_N\mathbf{x}$ *and* $\hat{\mathbf{h}} = \mathbf{F}_N\mathbf{h}$ *denote the normalized DFT of* $\mathbf{x}$ *and* $\mathbf{h}$*. In the Fourier domain, the circular convolution becomes a simple entry-wise multiplication of the components of* $\sqrt{N}\hat{\mathbf{h}}$ *with the components of* $\hat{\mathbf{x}}$*:* $\hat{\mathbf{y}} = \mathbf{F}_N\,\mathbf{y} = \hat{\mathbf{H}}\,\hat{\mathbf{x}}$*.*

2.2 Accuracy

Definition 2. *Given a vector* $\mathbf{h} \in \mathbb{R}^N$ *which defines a convolution matrix* $\mathbf{H}$*, the mean (expected) squared error (*MSE*) of an algorithm* $\mathcal{A}$ *is defined as*

$$\text{MSE} = \sup_{\mathbf{x}\in\mathbb{R}^N} \frac{1}{N}\text{E}[\|\mathcal{A}(\mathbf{x}) - \mathbf{H}\mathbf{x}\|_2^2].$$

Note that MSE measures the mean expected squared error *per output component*. Note further that MSE is a function of both the algorithm and the public convolution matrix, but is defined to be worst-case over private inputs.

3 Lower Bounds

In this section we derive a spectral lower bound on mean squared error of differentially private approximation algorithms for circular convolution. We prove that this bound is nearly tight for every fixed $\mathbf{h}$ in the Section 4. The lower bound is stated as Theorem 2.

Theorem 2. *Let* $\mathbf{h} \in \mathbb{R}^N$ *be an arbitrary real vector and let us relabel the Fourier coefficients of* $\mathbf{h}$ *so that* $|\hat{h}_0| \geq \ldots \geq |\hat{h}_{N-1}|$*. For all sufficiently small* ε *and* δ*, the expected mean squared error* MSE *of any* (ε, δ)*-differentially private algorithm* $\mathcal{A}$ *that approximates* $\mathbf{h} * \mathbf{x}$ *is at least*

$$\text{MSE} = \Omega\left(\max_{K=1}^{N} \frac{K^2\hat{h}_{K-1}^2}{N\log^2 N}\right). \tag{2}$$

For the remainder of the paper, we define the notation specLB($\mathbf{h}$) for the right hand side of (2), i.e. $\text{specLB}(\mathbf{h}) = \max_{K=1}^{N} \frac{K^2\hat{h}_{K-1}^2}{N\log^2 N}$.

3.1 Discrepancy Preliminaries

We define (ℓ_2) hereditary discrepancy as

$$\text{herdisc}(\mathbf{A}) = \max_{W\subseteq[N]} \min_{\mathbf{v}\in\{-1,+1\}^W} \|\mathbf{A}\mathbf{v}\|_2.$$

The following result connects discrepancy and differential privacy:

Theorem 3 ([24]). *Let* $\mathbf{A}$ *be an* $M \times N$ *complex matrix and let* $\mathcal{A}$ *be an* (ε, δ)*-differentially private algorithm for sufficiently small constant* ε *and* δ*. There exists a constant* C *and a vector* $\mathbf{x} \in \{0,1\}^N$ *such that* $\text{E}[\|\mathcal{A}(\mathbf{x}) - \mathbf{A}\mathbf{x}\|_2^2] \geq C\frac{\text{herdisc}(\mathbf{A})^2}{\log^2 N}$*.*

The determinant lower bound for hereditary discrepancy due to Lovász, Spencer, and Vesztergombi gives us a spectral lower bound on the noise required for privacy.

Theorem 4 ([23]). *There exists a constant C' such that for any complex $M \times N$ matrix $\mathbf{A}$, $\operatorname{herdisc}(\mathbf{A}) \geq C' \max_{K,\mathbf{B}} \sqrt{K} |\det(\mathbf{B})|^{1/K}$, where K ranges over $[\min\{M, N\}]$ and $\mathbf{B}$ ranges over $K \times K$ submatrices of $\mathbf{A}$.*

Corollary 1. *Let $\mathbf{A}$ be an $M \times N$ complex matrix and let $\mathcal{A}$ be an (ε, δ)-differentially private algorithm for sufficiently small constant ε and δ. There exists a constant C and a vector $\mathbf{x} \in \{0,1\}^N$ such that, for any $K \times K$ submatrix $\mathbf{B}$ of $\mathbf{A}$, $\mathrm{E}[\|\mathcal{A}(\mathbf{x}) - \mathbf{A}\mathbf{x}\|_2^2] \geq C \frac{K |\det(\mathbf{B})|^{2/K}}{\log^2 N}$.*

3.2 Proof of Theorem 2

We exploit the power of the determinant lower bound of Corollary 1 by combining the simple but very useful observation that projections do not increase mean squared error with a lower bound on the maximum determinant of a submatrices of a rectangular matrix. We present these two ingredients in sequence and finish the section with a proof of Theorem 2.

Lemma 1. *Let $\mathbf{A}$ be an $M \times N$ complex matrix and let $\mathcal{A}$ be an (ε, δ)-differentially private algorithm for sufficiently small constant ε and δ. There exists a constant C and a vector $\mathbf{x} \in \{0,1\}^N$ such that for any $L \times M$ projection matrix $\mathbf{P}$ and for any $K \times K$ submatrix $\mathbf{B}$ of $\mathbf{PA}$, $\mathrm{E}[\|\mathcal{A}(\mathbf{x}) - \mathbf{A}\mathbf{x}\|_2^2] \geq C \frac{K |\det(\mathbf{B})|^{2/K}}{\log^2 N}$.*

The proof of the lemma is based on the observation that $\mathcal{A}$ can be used to answer linear queries $\mathbf{B}\mathbf{x}$ by computing $\mathbf{y} = \mathcal{A}(x)$ and outputting (a subset of the coordinates of) $\mathbf{P}\mathbf{x}$. The MSE of this new mechanism is no larger than the error of $\mathcal{A}$. Details can be found in the full version of the paper.

Our main technical tool is a linear algebraic fact connecting the determinant lower bound for $\mathbf{A}$ and the determinant lower bound for any projection of $\mathbf{A}$.

Lemma 2. *Let $\mathbf{A}$ be an $M \times N$ complex matrix with singular values $\lambda_1 \geq \ldots \geq \lambda_N$ and let $\mathbf{P}$ be a projection matrix onto the span of the left singular vectors corresponding to $\lambda_1, \ldots, \lambda_K$. There exists a constant C and $K \times K$ submatrix $\mathbf{B}$ of $\mathbf{PA}$ such that*

$$|\det(\mathbf{B})|^{1/K} \geq C \sqrt{\frac{K}{N}} \left(\prod_{i=1}^{K} \lambda_i \right)^{1/K}$$

Proof. Let $\mathbf{C} = \mathbf{PA}$ and consider the matrix $\mathbf{D} = \mathbf{C}\mathbf{C}^H$. It has eigenvalues $\lambda_1^2, \ldots, \lambda_K^2$, and therefore $\det(\mathbf{D}) = \prod_{i=1}^{K} \lambda_i^2$. On the other hand, by the Binet-Cauchy formula for the determinant, we have

$$\det(\mathbf{D}) = \det(\mathbf{C}\mathbf{C}^H) = \sum_{S \in \binom{[N]}{K}} \det(\mathbf{C}|_S)^2 \leq \binom{N}{K} \max_{S \in \binom{[N]}{K}} \det(\mathbf{C}|_S)^2.$$

Rearranging and raising to the power $1/2K$, we get that there exists a $K \times K$ submatrix of $\mathbf{C}$ such that $|\det(\mathbf{B})|^{1/K} \geq \binom{N}{K}^{-1/2K} \left(\prod_{i=1}^{K} \lambda_i\right)^{1/K}$. Using the bound $\binom{N}{K} \leq \left(\frac{Ne}{K}\right)^K$ completes the proof.

We can now prove our main lower bound theorem by combining Lemma 1 and Lemma 2.

Proof (of Theorem 2). As usual, we will express $\mathbf{h} * \mathbf{x}$ as the linear map $\mathbf{Hx}$, where $\mathbf{H}$ is the convolution matrix for $\mathbf{h}$. By Lemma 1, it suffices to show that for each K, there exists a projection matrix $\mathbf{P}$ and a $K \times K$ submatrix $\mathbf{B}$ of $\mathbf{PH}$ such that $|\det(B)|^{1/K} \geq \Omega(\sqrt{K}|\hat{h}_K|)$. Recall that the eigenvalues of $\mathbf{H}$ are $\sqrt{N}\hat{h}_0, \ldots, \sqrt{N}\hat{h}_{N-1}$, and, therefore, the i-th singular value of $\mathbf{H}$ is $\sqrt{N}|\hat{h}_{i-1}|$. By Lemma 2, there exists a constant C, a projection matrix P, and a submatrix $\mathbf{B}$ of $\mathbf{PH}$ such that

$$|\det(\mathbf{B})|^{1/K} \geq C\sqrt{\frac{K}{N}} \left(\prod_{i=0}^{K-1} \sqrt{N}|\hat{h}_i|\right)^{1/K} \geq C\sqrt{K}|\hat{h}_K|.$$

This completes the proof.

4 Upper Bounds

Next we describe an algorithm which is nearly optimal for (ε, δ)-differential privacy. This algorithm is derived by formulating the error of a natural class of private algorithms as a convex program and finding a closed form solution. The class of private algorithms we consider is those which add independent Gaussian noise to the Fourier coefficients of the private input $\mathbf{x}$. This is a special case of the extended matrix mechanism [21]; working with a less general algorithm is what allows us to derive a closed form for the optimal algorithm. At the same time, the error of our algorithm matches the lower bound on extended matrix mechanisms from [22].

Consider the class of algorithms, which first add independent Laplacian noise variables $z_i = \mathrm{Lap}(0, b_i)$ to the Fourier coefficients $\hat{x}_i$ to compute $\tilde{x}_i = \hat{x}_i + z_i$, and then output $\tilde{\mathbf{y}} = \mathbf{F}_N^H \hat{\mathbf{H}} \tilde{\mathbf{x}}$. This class of algorithms is parameterized by the vector $\mathbf{b} = (b_0, \ldots, b_{N-1})$; a member of the class will be denoted $\mathcal{A}(\mathbf{b})$ in the sequel. The question we address is: For given $\varepsilon, \delta > 0$, how should the noise parameters $\mathbf{b}$ be chosen such that the algorithm $\mathcal{A}(\mathbf{b})$ achieves (ε, δ)-differential privacy in $\mathbf{x}$ for ℓ_1 neighbors, while minimizing the mean squared error MSE? It turns out that by convex programming duality we can derive a closed form expression for the optimal $\mathbf{b}$, and moreover, the optimal $\mathcal{A}(\mathbf{b})$ is nearly optimal *among all* (ε, δ)*-differentially private algorithms*. The optimal parameters are used in Algorithm 1.

Algorithm 1 FOURIER MECHANISM

Set $\gamma = \frac{2\ln(1/\delta)\|\hat{\mathbf{h}}\|_1}{\varepsilon^2 N}$
Compute $\hat{\mathbf{x}} = \mathbf{F}_N\mathbf{x}$ and $\hat{\mathbf{h}} = \mathbf{F}_N\mathbf{x}$.
for all $i \in \{0, \ldots, N-1\}$ **do**
  if $|\hat{h}_i| > 0$ **then**
    Set $z_i = \mathrm{Lap}\left(\sqrt{\frac{\gamma}{|\hat{h}_i|}}\right)$
  else if $|\hat{h}_i| = 0$ **then**
    Set $z_i = 0$
  end if
  Set $\tilde{x}_i = \hat{x}_i + z_i$.
  Set $\bar{y}_i = \sqrt{N}\hat{h}_i\tilde{x}_i$.
end for
Output $\tilde{\mathbf{y}} = \mathbf{F}_N^H\bar{\mathbf{y}}$

Theorem 5. *Algorithm 1 satisfies (ε, δ)-differential privacy, and achieves expected mean squared error*

$$\mathrm{MSE} = 4\frac{\ln(1/\delta)}{\varepsilon^2 N}\|\hat{\mathbf{h}}\|_1^2. \tag{3}$$

Moreover, Algorithm 1 runs in time $O(N \log N)$.

The proof of Theorem 5 is omitted from the current version of the paper. Next, we show that it implies that Algorithm 1 is almost optimal *for any given* $\mathbf{h}$.

Theorem 6. *For any $\mathbf{h}$, Algorithm 1 satisfies (ε, δ)-differential privacy and achieves expected mean squared error $O\left(\mathrm{specLB}(\mathbf{h})\frac{\log^2 N \log^2 |I| \ln(1/\delta)}{\varepsilon^2}\right)$.*

Proof. Assume that $|\hat{h}_0| > |\hat{h}_1| > \ldots > |\hat{h}_{N-1}|$. Then, by definition of $I = \{0 \le i \le N-1 : |\hat{h}_i| > 0\}$, we have $|\hat{h}_j| = 0$, for all $j > |I| - 1$. Thus,

$$\begin{aligned}\|\hat{\mathbf{h}}\|_1 = \sum_{i=0}^{|I|-1}|\hat{h}_i| = \sum_{i=1}^{|I|}\frac{1}{i}i|\hat{h}_{i-1}| &\le \left(\sum_{i=1}^{|I|}\frac{1}{i}\right)\sqrt{N}\log N\sqrt{\mathrm{specLB}(\mathbf{h})} \\ &= H_{|I|}\sqrt{N}\log N\sqrt{\mathrm{specLB}(\mathbf{h})}, \end{aligned} \tag{4}$$

where $H_m = \sum_{i=1}^m \frac{1}{i}$ denotes the m-th harmonic number. Recalling that $H_m = O(\log m)$, and combining the bound (4) with the expression of the MSE (3) yields the desired bound.

5 Generalizations and Applications

In this section we describe some generalizations and applications of our lower bounds and algorithms for private convolution. Next we sketch applications to

computing running sums, and linear filters motivated by analysis of time series data. Applications to computing compressible convolution maps, and computing generalized marginals on data cubes (which are an example of compressible convolutions) are described in the full version of the paper.

5.1 Running Sum

Running sums can be defined as the circular convolution $x' * h$ of the sequences $h = (1, \ldots, 1, 0, \ldots, 0)$, where there are N ones and N zeros, and $x' = (x, 0, \ldots, 0)$, where the private input x is padded with N zeros. An elementary computation reveals that $\hat{h}_1 = \sqrt{N}$ and $\hat{h}_i = O(N^{-1/2})$ for all $i > 1$. By Theorem 5, Algorithm 1 computes running sums with mean squared error $O(1)$ (ignoring dependence on ϵ and δ), improving on the bounds of [5,12,30] in the mean squared error regime.

5.2 Linear Filters in Time Series Analysis

Linear filtering is a fundamental tool in analysis of time-series data. A time series is modeled as a sequence $x = (x_t)_{t=-\infty}^{\infty}$, supported on a finite set of time steps. A filter converts the time series into another time series. A linear filter does so by computing the convolution of x with a series of *filter coefficients* w, i.e. computing $y_t = \sum_{i=-\infty}^{\infty} w_i x_{t-i}$. For a finitely supported x, y can be computed using circular convolution by restricting x to its support set and padding with zeros on both sides.

We consider the case where x is a time series of sensitive *events*. Each element x_i is a count of events or sum of values of individual transactions that have occurred at time step i. When we deal with values of transactions, we assume that individual transactions have much smaller value than the total. We emphasize that the definition of differential privacy with respect to x defined this way corresponds to *event-level privacy*.

We consider applications to financial analysis, but our methods are applicable to other instances of time series data, e.g. we may also consider network traffic logs or a time series of movie ratings on an online movie streaming service. We can perform almost optimal differentially private linear filtering by casting the filter as a circular convolution. For more references and detailed description, we refer the reader to the full version of our paper and the book of Gençan, Selçuk, and Whitcher [13].

6 Conclusion

We derive nearly tight upper and lower bounds on the error of (ε, δ)-differentially private algorithms for computing convolutions. Our lower bounds rely on recent general lower bounds based on discrepancy theory and elementary linear algebra; our upper bound is a simple computationally efficient algorithm. We also

sketch several applications of private convolutions, in time series analysis and in computing generalizes marginal queries on a d-attribute database.

Since our algorithm for computing convolutions has running time $O(N \log N)$, we conjecture that there exists an $\tilde{O}(Nn)$ time algorithm for computing convolutions with optimal error when the database size is at most n. This would improve on the more general algorithm from [28], which has running time $O(M^2Nn)$.

References

1. Barak, B., Chaudhuri, K., Dwork, C., Kale, S., McSherry, F., and Talwar, K. Privacy, accuracy, and consistency too: a holistic solution to contingency table release. In *Proceedings of the twenty-sixth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems* (2007), ACM, pp. 273–282.
2. Bhaskara, A., Dadush, D., Krishnaswamy, R., and Talwar, K. Unconditional differentially private mechanisms for linear queries. In *Proceedings of the 44th symposium on Theory of Computing* (New York, NY, USA, 2012), STOC '12, ACM, pp. 1269–1284.
3. Blum, A., Ligett, K., and Roth, A. A learning theory approach to non-interactive database privacy. In *STOC '08: Proceedings of the 40th annual ACM symposium on Theory of computing* (New York, NY, USA, 2008), ACM, pp. 609–618.
4. Bolot, J., Fawaz, N., Muthukrishnan, S., Nikolov, A., and Taft, N. Private decayed sum estimation under continual observation. *Arxiv preprint arXiv:1108.6123* (2011).
5. Chan, T., Shi, E., and Song, D. Private and continual release of statistics. In *ICALP* (2010).
6. Chandrasekaran, K., Thaler, J., Ullman, J., and Wan, A. Faster private release of marginals on small databases. *arXiv preprint arXiv:1304.3754* (2013).
7. Cheraghchi, M., Klivans, A., Kothari, P., and Lee, H. Submodular functions are noise stable. In *Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms* (2012), SIAM, pp. 1586–1592.
8. Cormode, G., Procopiuc, C. M., Srivastava, D., and Yaroslavtsev, G. Accurate and efficient private release of datacubes and contingency tables.
9. Dinur, I., and Nissim, K. Revealing information while preserving privacy. In *Proceedings of the twenty-second ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems* (2003), ACM, pp. 202–210.
10. Dwork, C., Mcsherry, F., Nissim, K., and Smith, A. Calibrating noise to sensitivity in private data analysis. In *TCC* (2006).
11. Dwork, C., Naor, M., Reingold, O., Rothblum, G. N., and Vadhan, S. On the complexity of differentially private data release: efficient algorithms and hardness results. In *Proceedings of the 41st annual ACM symposium on Theory of computing* (2009), ACM, pp. 381–390.
12. Dwork, C., Pitassi, T., Naor, M., and Rothblum, G. Differential privacy under continual observation. In *STOC* (2010).
13. Gençay, R., Selçuk, F., and Whitcher, B. *An Introduction to Wavelets and Other Filtering Methods in Finance and Economics*. Elsevier Academic Press, 2002.
14. Gray, R. M. Toeplitz and circulant matrices: a review. *Foundations and Trends in Communications and Information Theory 2*, 3 (2006), 155–239.

15. Gupta, A., Hardt, M., Roth, A., and Ullman, J. Privately releasing conjunctions and the statistical query barrier. In *Proceedings of the 43rd annual ACM symposium on Theory of computing* (2011), ACM, pp. 803–812.
16. Hardt, M., and Rothblum, G. A multiplicative weights mechanism for privacy-preserving data analysis. *Proc. 51st Foundations of Computer Science (FOCS). IEEE* (2010).
17. Hardt, M., Rothblum, G., and Servedio, R. Private data release via learning thresholds. In *Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms* (2012), SIAM, pp. 168–187.
18. Hardt, M., and Talwar, K. On the geometry of differential privacy. In *Proceedings of the 42nd ACM symposium on Theory of computing* (2010).
19. Kasiviswanathan, S., Rudelson, M., Smith, A., and Ullman, J. The price of privately releasing contingency tables and the spectra of random matrices with correlated rows. In *Proceedings of the 42nd ACM symposium on Theory of computing* (2010), ACM, pp. 775–784.
20. Li, C., Hay, M., Rastogi, V., Miklau, G., and McGregor, A. Optimizing linear counting queries under differential privacy. In *Proceedings of the twenty-ninth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems* (New York, NY, USA, 2010), PODS '10, ACM, pp. 123–134.
21. Li, C., and Miklau, G. An adaptive mechanism for accurate query answering under differential privacy. *PVLDB 5*, 6 (2012), 514–525.
22. Li, C., and Miklau, G. Measuring the achievable error of query sets under differential privacy. *CoRR abs/1202.3399* (2012).
23. Lovász, L., Spencer, J., and Vesztergombi, K. Discrepancy of set-systems and matrices. *European Journal of Combinatorics 7*, 2 (1986), 151–160.
24. Muthukrishnan, S., and Nikolov, A. Optimal private halfspace counting via discrepancy. *Proceedings of the 44th ACM symposium on Theory of computing* (2012).
25. Narayanan, A., Shi, E., and Rubinstein, B. Link prediction by de-anonymization: How we won the kaggle social network challenge. In *Neural Networks (IJCNN), The 2011 International Joint Conference on* (2011), IEEE, pp. 1825–1834.
26. Narayanan, A., and Shmatikov, V. Robust de-anonymization of large sparse datasets. In *Security and Privacy, 2008. SP 2008. IEEE Symposium on* (2008), IEEE, pp. 111–125.
27. Narayanan, A., and Shmatikov, V. De-anonymizing social networks. In *Security and Privacy, 2009 30th IEEE Symposium on* (2009), Ieee, pp. 173–187.
28. Nikolov, A., Talwar, K., and Zhang, L. The geometry of differential privacy: the sparse and approximate cases.
29. Thaler, J., Ullman, J., and Vadhan, S. Faster algorithms for privately releasing marginals. *Automata, Languages, and Programming* (2012), 810–821.
30. Xiao, X., Wang, G., and Gehrke, J. Differential privacy via wavelet transforms.