Constant matters: Fine-grained Complexity of Differentially Private Continual Observation Using Completely Bounded Norms

Monika Henzinger [*] Jalaj Upadhyay [†]

Abstract

We study fine-grained error bounds for differentially private algorithms for averaging and counting in the continual observation model. For this, we use the completely bounded spectral norm (cb norm) from operator algebra. For a matrix W, its cb norm is defined as

$$\|W\|_{\mathsf{cb}} = \max_{Q} \left\{ \frac{\|Q \bullet W\|}{\|Q\|} \right\},$$

where $Q \bullet W$ denotes the Schur product and $\|\cdot\|$ denotes the spectral norm. We bound the cb norm of two fundamental matrices studied in differential privacy under the continual observation model: the counting matrix M_{count} and the averaging matrix M_{average}. For M_{count}, we give lower and upper bound whose additive gap is $1 + \frac{1}{\pi}$. Our factorization also has two desirable properties sufficient for streaming setting: the factorization contains of lower-triangular matrices and the number of distinct entries in the factorization is exactly T. This allows us to compute the factorization on the fly while requiring the curator to store a T-dimensional vector. For M_{average}, we show an additive gap between the lower and upper bound of ≈ 0.64.

1 Introduction

In the recent past, many large scale applications of data analysis involved repeated computations because of the incidence of an infectious disease [App21, CDC20], typically with a goal of preparing some appropriate response. However, privacy of the result (positive or negative infection) of a user is equally important. In such an application, the system is required to continually produce outputs while preserving a robust privacy guarantee such as differential privacy. Such an application is also not just a made-up example. In the first work on differential privacy under continual observation (which coincided with an H1N1 outbreak in 2009), Dwork et al. [DNPR10] used the following motivating example:

> "Consider a website for H1N1 self-assessment. Individuals can interact with the site to learn whether symptoms they are experiencing may be indicative of the H1N1 flu. The user fills in some demographic data (age, zipcode, sex), and responds to queries about his symptoms (fever over 100.4° F?, sore throat?, duration of symptoms?). We would like to continually analyze aggregate information of consenting users in order to monitor regional health conditions, with the goal, for example, of organizing improved flu response. Can we do this in a differentially private fashion with reasonable accuracy (despite the fact that the system is continually producing outputs)?"

[*] University of Vienna. A part of the work was done as the Stanford University Distinguished Visiting Austrian Chair. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 101019564 "The Design of Modern Fully Dynamic Data Structures (MoDynStruct)" and from the Austrian Science Fund (FWF) project "Fast Algorithms for a Reactive Network Layer (ReactNet)", P 33775-N, with additional funding from the *netidee SCIENCE Stiftung*, 2020–2024. Email: monika.henzinger@univie.ac.at

[†] Rutgers University. This work was supported by the Decanal Research grant from Rutgers University. Email: jalaj.upadhyay@rutgers.edu

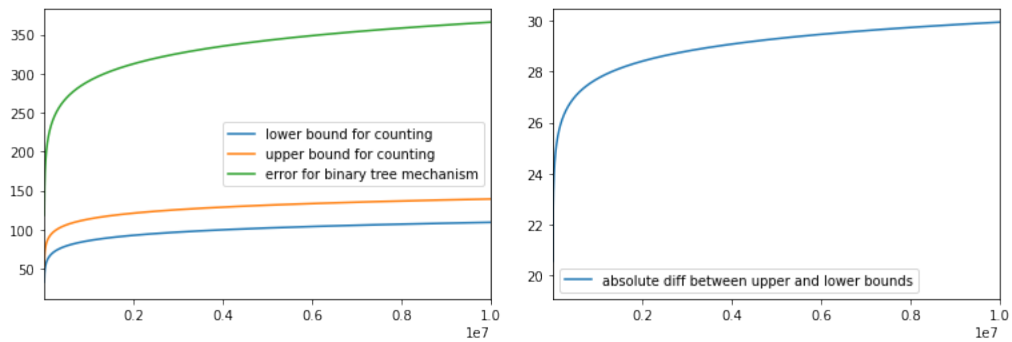

Figure 1: The lower and upper bound graphs for counting. The left graph is for additive error (with $\epsilon = 0.5, \delta = 10^{-8}$) and comparison with binary tree mechanism and the right graph gives the difference between the upper and lower bound.

For a stream of zeros and ones of length T Dwork et al. [DNPR10] and Chan et al. [CSS11] showed that there exists a differentially private mechanism, called the *binary mechanism*, for counting under continual observation with an additive error of $O(\log^{3/2} T)$. However, the constant has never been explicitly mentioned. Given the wide applications of binary counting in many downstream tasks [BFM$^+$13, CR21, CLSX12, FHO21, HQYC21, JRSS21, Upa19, UUA21, UU21], constants can define whether the output is useful or not in practice. In fact, from the practitioner point of view, the biggest problem is an interpretation of the asymptotic nature of error given in $O(\cdot)$ notation. As Figure 1 illustrates, with the binary tree mechanism, the error is $\geqslant 250$ as soon as the number of reporting events[1] goes beyond $T = 10^6$. Other approaches that use dynamic versions of off-the-shelf theoretically accurate privacy preserving algorithms for linear queries have constants that are of order of 512, making them almost useless in practice [Cum21]. This is because, for many organisations that deploy life-saving and sparse resources, a large error is a big deterrence and the blatantly non-private approach of rounding to nearest multiple of 100 becomes more appealing. With this in mind, we ask the following central question:

> Can we get a fine-grained bounds on constants in the error of differentially private algorithms under continual observation?

The problem of reducing the additive error for counting under continual observation has been pursued by many works (see [WCZ$^+$21] and references therein). Most of these works use some heuristics or use some smoothening techniques to improve their bound on the additive error[2]. These approaches make the algorithm not scalable, especially when dealing with high-scale deployments, like that of exposure-notification system (ENS) that has to operate when the stream length is in order of 1 million. Therefore, our focus is to design a scalable mechanism that gives a fine-grained error bound.

One of the scalable techniques that provably reduces the error on linear queries is query matrix optimization known as *workload optimizer* (see [MMHM21] and references therein). There are many techniques presented for this, one of them being the *factorization mechanism* [ENU20, MNT20]. It can be shown that the additive error (denoted as ℓ_∞ error) of the factorization mechanism for linear queries represented by a *workload matrix* M using the Gaussian mechanism scales as follows:

$$\frac{C\,\|L\|_{2\to\infty}\,\|R\|_{1\to 2}\,\sqrt{\log(1/\delta)\log|\mathcal{Q}|}}{\epsilon}, \quad \text{where } \|A\|_{a\to b} := \max_{x\in\mathbb{R}\setminus\{0\}} \frac{\|Ax\|_b}{\|x\|_a}, \tag{1}$$

(ϵ, δ) are the privacy budget, $M = LR$ is a factorization of the workload matrix, $|\mathcal{Q}|$ is the number of queries made, and $C > 0$ is a constant defined by the concentration bound on the Gaussian distribution. The quantity $\|L\|_{2\to\infty}\,\|R\|_{1\to 2}$ is known as the *completely bounded norm* (abbreviated as cb-norm and denoted by $\|W\|_{\mathsf{cb}}$)

[1] The United States of America alone has done more than 750 million COVID tests to date.

[2] The approach of Honaker requires to solve a linear program to estimate the weights to be placed on every nodes in the binary mechanism.

in operator algebra [Pau82] and *factorization norm* (denoted by $\gamma_2(W)$) in functional analysis and computer science [LMSS07]. The idea of factorization through Hilbert space to call it the factorization norm can be traced back to the book of Pisier [Pis86] that cites the Steinspring representation of cb maps by Paulsen as a factorization theorem [Pau21, PS85]. In this paper, we use the notation $\|M\|_{\mathsf{cb}}$ to be consistent with the fact that the cb norm is a norm [Pau82] and with the more accepted choice of notation in operator algebra from where the concepts is borrowed.

As for the error, $C\frac{\sqrt{\log(1/\delta)\log|\mathcal{Q}|}}{\epsilon}$ in equation (1) is due to the concentration bound on the Gaussian distribution followed by the union bound and can be concretely computed. In other words, to get a concrete additive error, we need to find a factorization $M = LR$ such that the quantity $\|L\|_{2\to\infty}\|R\|_{1\to 2}$ is not just minimum, but can be computed concretely. Therefore, for the rest of the paper, we only focus on $\|M\|_{\mathsf{cb}}$. In particular, we focus on the cb norm of two workload matrices corresponding to two fundamental queries in the continual observation model: counting and average.

We borrow the notations from Watrous [Wat18]. We use the notation $v[i]$ to denote the i-th coordinate of a vector v. For a matrix A, we use $A[i,j]$ to denote its (i,j)-th coordinate, $A[:,i]$ to denote its i-th column, $A[i,:]$ to denote its i-th row, $\|A\|_{\mathsf{tr}}$ to denote its trace norm of square matrix, and $\|A\|$ to denote its operator norm.

Equipped with this notation, we can formally define our problem. The stream can be seen as a binary vector, $v \in \{0,1\}^T$, and at any time $t \leqslant T$, we are concerned with the first t coordinates of v. This can be equivalently written as the i-th entry of $M_{\mathsf{count}}v$ and $M_{\mathsf{average}}v$, where the matrix M_{count} and M_{average} are $T \times T$ lower triangular matrices as follows:

$$M_{\mathsf{count}}[i,j] = \begin{cases} 1 & i \geqslant j \\ 0 & i < j \end{cases}, \quad M_{\mathsf{average}}[i,j] = \begin{cases} \frac{1}{i} & i \geqslant j \\ 0 & i < j \end{cases} \tag{2}$$

Our goal is to give upper and lower bounds for $\|M_{\mathsf{count}}\|_{\mathsf{cb}}$ and $\|M_{\mathsf{average}}\|_{\mathsf{cb}}$. This in turn implies bounds on the constants for the additive error of counting or computing the average of a stream of numbers under continual observation using the factorization mechanism.

Our goal is to compute explicit factorization of both M_{count} and M_{average} such that the factorization are lower triangular matrices. The requirement of lower triangular matrices is a sufficient condition for continual observation. However, it is also beneficial in the sense that it allows us to perform exact steps required in the mechanism while getting an explicit bound on the additive error of the mechanism.

1.1 Our results.

Bounding $\|M_{\mathsf{count}}\|_{\mathsf{cb}}$. The question of finding the optimal value of $\|M_{\mathsf{count}}\|_{\mathsf{cb}}$ was also raised in the conference version of Matousek, Nikolov and Talwar [MNT20]. In their IMRN version, they cite a result by Mathias [Mat93a, Corollary 3.5], which shows the following:

$$\frac{T+1}{2T}\widehat{\gamma}(T) \leqslant \|M_{\mathsf{count}}\|_{\mathsf{cb}} \leqslant \frac{\widehat{\gamma}(T)+1}{2}, \quad \text{where } \widehat{\gamma}(T) = \frac{1}{T}\sum_{j=1}^{T}\frac{1}{\sin(2j-1)\pi/T}.$$

It is known that

$$\lim_{T\to\infty}\widehat{\gamma}(T) \to \frac{2\log(T)}{\pi}$$

from the above. That is, the value of $\|M_{\mathsf{average}}\|_{\mathsf{cb}}$ achieves its optimum in limit; however, since $\widehat{\gamma}(T)$ is always greater than or equal to $\frac{2\log(T)}{\pi}$, this implies that there is a slack between the two bounds for the range of T one would be interested in practice. Finally, the proof of Mathias [Mat93a] relies on the dual characterization of cb-norm, and, therefore, does not give an explicit factorization.

Using a completely different proof, we give a lower and upper bound on $\|M_{\mathsf{count}}\|_{\mathsf{cb}}$ that has a direct dependence on T and the additive gap between the upper and lower bound is $\approx 1 + \frac{1}{\pi}$:

Theorem 1. *Let $M_{\mathsf{count}} \in \{0,1\}^{T\times T}$ be the matrix defined in eq. (2). Then, we have*

$$\frac{1}{\pi}\sqrt{(\log T)^2 - 2\log(T) + 2 - \frac{1}{T}} \leqslant \|M_{\mathsf{count}}\|_{\mathsf{cb}} \leqslant 1 + \frac{\log(T-1)}{\pi}. \tag{3}$$

We note that an asymptotic lower bound on the $\|M_{\mathsf{count}}\|_{\mathsf{cb}}$ is often attributed to Forster et al. [FSSS03] in computer science. Forster et al.'s lower bound is more general, but for the special case of M_{count}, it is a weak version of the lower bound mentioned above[3]

Our upper bound constructs an explicit factorization of M_{count}; therefore, the optimal way (up to an additive constant of $1+\frac{1}{\pi}\approx 1.30$) to do the counting is to use our factorization. We note that it is possible to get better factorization if entries of L and R are the quadratic extension of real field, $\mathbb{R}(\sqrt{-1})$; however, we prefer to work on the real field. To show the tightness of our concrete bounds in Theorem 1, we plot the graphs corresponding to the lower and upper bounds in Figure 1.

In a concurrent, independent work, McMahan et al. [MRT22] used similar techniques of matrix factorization to show a bound in the expected ℓ_2^2 norm (equation 3 in their paper). In contrast, we give a bound in ℓ_∞ norm. We also give an explicit factorization while they state their result in terms of solving an expensive convex program (see Section 1.2 for more detailed comparison).

Our explicit factorization of $M_{\mathsf{count}} = L^\top R$ is a square factorization and is such that both $L^\top$ and R are lower-triangular matrices. This allows us to perform binary counting in the continual observation model efficiently.

The upper bound can be derived in many ways. One direct approach would be to find appropriate Kraus operators of a linear map and then use the characterization by Haagerup and Pisier [HP93] of completely bounded norm in terms of Haagerup norm. This approach yields an upper bound of $1+\frac{\log(T)}{\pi}$; however, the proof of the result gets very technical and uses techniques from operator algebra. Moreover, it does not directly give lower triangular factorization matrices $L^\top$ and R.

Instead we use the characterization given by Paulsen [Pau82], which gives us a factorization in terms of lower triangular matrices. More precisely, using three basic trigonometric identities, we show that the (i,j)-th entry of R and $L^\top$ is an integral of even power of the cosine function, $\frac{2}{\pi}\int\limits_0^{\pi/2}\cos^{i-j}(\theta)\mathrm{d}\theta$ for $i \geqslant j$. This choice of matrices leads to the upper bound in eq. (3). Furthermore, it makes the analysis very simple with the most technical part requiring bounding a function related to the derivative of the truncated Reimann zeta function at $s=0$. Bounding this function reduces to understanding a recurrence relation that yields a monotonically decreasing sequence.

Our lower bound uses a characterization of completely bounded norm by Haagerup [Haa80].

Bounding $\|M_{\mathsf{average}}\|_{\mathsf{cb}}$. Our second result pertains to an almost tight bound on the matrix M_{average}:

Theorem 2. *Let $M_{\mathsf{average}} \in \mathbb{R}^{T\times T}$ be the matrix defined in eq. (2). Then $\|M_{\mathsf{average}}\| \geqslant 1$. Further, there is an explicit factorization $M_{\mathsf{average}} = L^\top R$ such that, for $T \geqslant 2$, we have*

$$\|L\|_{2\to\infty}\|R\|_{1\to 2} \leqslant \frac{\pi^2 T(T+1)}{3(2T+1)^2}. \tag{4}$$

To get some context for the above bound, the upper bound is $\frac{\pi^2}{6}$ in the limiting case $T\to\infty$. In other words, we have an additive gap of $\frac{\pi^2}{\sqrt{6}}-1\approx 0.64$ between the lower and upper bound. We leave it as an open problem to close this additive gap on $\|M_{\mathsf{average}}\|_{\mathsf{cb}}$.

The proof of the theorem requires us to bound the sum of the first T terms in the Reimann zeta function, $\zeta(s)$, at $s=2$. Euler showed that $\zeta(2)=\frac{\pi}{\sqrt{6}}$ as $T\to\infty$ [Eul06]. However, to get a bound that is a function of T, we need to compute the partial sum. There are many proofs for $\zeta(2)=\frac{\pi^2}{6}$ and the reader might wonder if it is possible to modify one of those proofs. However, most commonly known proofs do not give any estimate for partial sum. For example, the first proof by Euler's looks at the MacLaurin expansion of sin and his second proof looked at Reimann zeta function for even s and its characterization in terms of Bernoulli's number [Eul06]. Similarly, proofs using Fourier expansion or Parseval identity also directly deals with the infinite sum. To get the finite sum, we revisit the proof by Cauchy [Cau21]. Cauchy's original proof uses

[3] A weaker lower bound of $\frac{\log(T)}{2\pi}$ can be computed using Spreyer's answer on math stack-exchange https://math.stackexchange.com/q/2057845. This bound is weaker than the bound studied in nest algebra [Dav84]: $\|M_{\mathsf{count}}\|_{\mathsf{cb}} \geqslant \frac{4}{5\pi}\log(T)$..

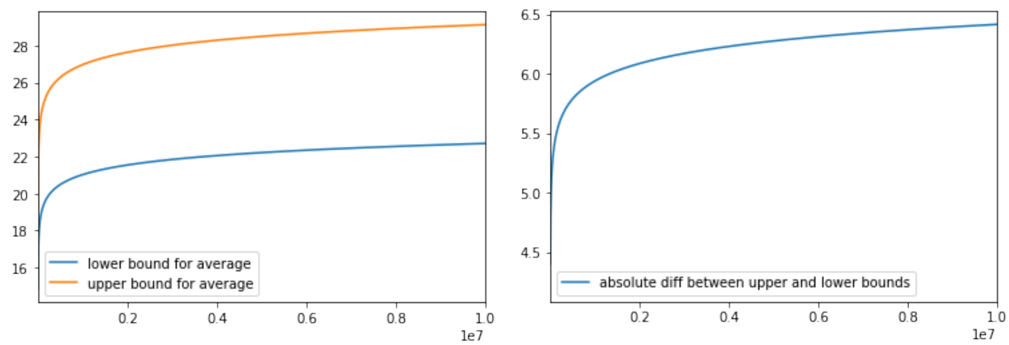

Figure 2: The lower and upper bound graphs for average. The left graph is for additive error (with $\epsilon = 0.5, \delta = 10^{-8}$) and the right graph gives the difference between the upper and lower bound on the absolute additive theoretical error.

Cauchy residue theorem; however, using de Moivre's and Vieta's formulas in Note VIII of Cauchy's "Cours d'Analyse," the proof (see the proof of Theorem 3) implies that

$$\frac{\pi}{2T+1}\sqrt{\frac{T(2T-1)}{3}} \leq \sqrt{1+\left(\frac{1}{2}\right)^2+\left(\frac{1}{3}\right)^2+\cdots+\left(\frac{1}{T}\right)^2} \leqslant \frac{2\pi}{2T+1}\sqrt{\frac{T(T+1)}{6}}.$$

Taking the limit $T \to \infty$, the sandwich theorem gives us the value of $\zeta(2)$.

As in the case of M_{count}, our lower bound follows using the characterization of Haagerup [Haa80].

1.2 Detailed Comparison with Previous Works

The binary tree mechanism of Chan et al. [CSS11] and Dwork et al. [DNPR10] and its improvement by Honaker [Hon15] can be seen as a factorization. This has been independently noticed by McMahan et al. [MRT22]. While Chan et al. [CSS11] and Dwork et al. [DNPR10] do allow computation on streaming data, Honaker's optimization [Hon15] does not because for a partial sum $[1, t]$ it also uses the information stored at the nodes formed after time t. Therefore, for this comparison with related work, we do not discuss the Honaker's optimization [Hon15]. Moreover, Honaker's optimization is for minimizing the expected ℓ_2^2 error. The other approaches used for binary counting under continual observation (see [WCZ$^+$21] and references therein) use some form of smoothening of the output. Considering this as a post-processing, we do not discuss it any further.

The most relevant work with ours is the concurrent work by McMahan et al [MRT22] that also looks at concrete bounds on performing counting under continual observation. The work of McMahan et al. [MRT22] is motivated by performing optimization privately on streamed data. Therefore, they bound the expected mean squared error (i.e., in ℓ_2^2 norm) on privately computing a running sum. On the other hand, we bound the absolute additive error (i.e., in ℓ_∞ norm). Further, they characterize optimal factorization for counting while we give explicit factorization for both counting and computing average under continual observation. As a result, we do not have to solve a convex program, but compute the entries of the factorization using a recurrence relation (for M_{count}) and solving $T(T+1)/2$ linear equations (for M_{average}). Finally, our explicit factorization for M_{count} has a nice property that there are exactly T distinct entries (instead of possibly T^2 entries in McMahan et al. [MRT22]) in the factorization. This has large impact on computation in practice.

Acknowledgements. We would like to thank Vern Paulsen for many insightful discussions regarding completely bounded norms and Haagerup's results along with his copy of Haagerup's manuscript [Haa80], Sarvagya Upadhyay for discussions on operator theory in quantum information and pointing to the semidefinite programs of Watrous [Wat09, Wat12], Sasho Nikolov for referring us to Mathias [Mat93a] and sharing a copy of IMRN version of [MNT20], and Abhradeep Thakurta for helpful discussion regarding [MRT22].

2 Concepts from operator algebra

We give a brief overview of concepts from operator algebra, namely, completely bounded norm and Toeplitz matrices, to the level required for this paper. Historically, completely bounded norm has been extensively studied in operator algebra [Haa80, HP93, Pau82, Pau86, Pau02]. Completely bounded trace norm (also known as *diamond norm* and equivalent, up to taking adjoint of the mapping to the completely bounded spectral norm [HP93, Pau21]) are used naturally in quantum information theory [AKN98, ABP19, CPR00, PW09, Wat07] since Kitaev [Kit97] noted that can be used to quantify distance between quantum channel, mathematical physics [BD15, DJKR06, HLP+18, She91], and fundamental physics [CR94, JPPG+10, TMB03, Wal94]. Recently, these norms have been recently studied in computer science for proving communication complexity lower bounds [LMSS07, LS09] and analyzing differentially private algorithms [ENU20, MNT20, Nik14].

The most popular definitions (and characterization) of cb norm are as follows:

$$\|W\|_{\mathsf{cb}} = \min_{W=LR}\left\{\|L\|_{2\to\infty}\|R\|_{1\to2}\right\} = \min_{W=L^\top R}\left\{c(L)c(R)\right\} = \max_{Q}\left\{\frac{\|Q\bullet W\|}{\|Q\|}\right\} = \max_{\substack{u,v\\ \|u\|=\|v\|=1}}\left\|W\bullet uv^\top\right\|_{\mathsf{tr}}.$$

Here, $c(X)$ denotes the maximum ℓ_2 column norm of the matrix X and $Q\bullet W$ is the Schur product [Sch11]. The second and third equality was shown by Haagerup [Haa80] and Mattias [Mat93b][4]. In fact, the characterization with respect to the trace norm can be derived from the duality of completely bounded spectral norm and completely bounded trace norm (or diamond norm). Some other characterization of completely bounded norm have been also studied, a partial list includes that in terms of Stinespring representations [PS85], Choi-Jamiołkowski representation [Wat12], Haagerup norm [Haa80, HP93] and fidelity by combining Alberti theorem and Uhlmann theorem.

Some of these characterizations have been instrumental in giving efficient algorithms for computing cb norm. For example, one can use the technique developed by Cowen et al. [CFJ+96]. Their algorithm is based on the primal-dual based algorithm by Watson [Wat96] that computes a lower bound on cb-norm. In particular, Cowen et al. [CFJ+96] uses the factorization theorem of Haagerup [Haa80] to show that the convergent of Watson's algorithm actually gives a very tight upper bound, too. Alternatively, one can use the methods developed by Johnston, Kribs, and Paulsen [JKP09] and Zarikian [Zar06] using the characterization of cb norm in terms of Haagerup norm [HP93]. For the special case of Hermittian matrices, one can also use the Watson's algorithm [Wat96] with Wittstock's decomposition theorem [Wit81]. These are practical iterative methods, but their rates of convergence is unknown. The only known algorithms with provable rate of convergence we are aware of is by Watrous [Wat09, Wat12] using various semi-definite formulations and Ben-Aroya and Ta-Shma [BATS09] using convex optimization.

We use *Toeplitz operator* defined over Hardy space. In real analysis, *Hardy spaces* are spaces of distributions on the real line, which are the boundary values of the holomorphic functions of the complex Hardy spaces. In complex analysis, they are spaces of holomorphic functions on the unit disk. In short, they consists of functions whose mean squared value on unit circle remains bounded as we reach the boundary. They are natural to deal where Lebesgue spaces are not well behaved [CW09]). A bounded operator on the Hardy space is Toeplitz if and only if its matrix representation, in the standard basis has constant diagonals. In other words, Toeplitz operators are just multiplication followed by projection onto the Hardy space. We refer the interested reader to the excellent book by Conway [Con00] for more in depth overview of operator theory and monograph by Paulsen [Pau86] for completely bounded norm.

3 Proofs of the Main Results

We first collect two useful lemma used in our proofs.

Lemma 1. *Let m be an even integer. Define*

$$S_m = \left(\frac{1}{2}\right)\left(\frac{3}{4}\right)\cdots\left(\frac{m-1}{m}\right)\frac{\pi}{2}.$$

[4]Paulsen attributed the first equality to Haagerup in his monograph [Pau86, Section 7.7].

Then $S_m \leqslant \sqrt{\frac{\pi}{2m}}$.

Proof of Lemma 1. One can use the fact that S_m is closely related to the derivative of the truncated Reimann zeta function, $\zeta(s)$, at $s = 0$. Another approach would be to use a trigonometric equality involving the integration of higher powers of cosine function. We take the latter approach for cleaner calculation. In particular, we know that for even values of m,

$$S_m := \int_0^{\pi/2} \cos^m(x)\mathrm{d}x = \left(\frac{1}{2}\right)\left(\frac{3}{4}\right)\cdots\left(\frac{m-1}{m}\right)\frac{\pi}{2} = S_{m-2} - \frac{1}{m-1}S_m,$$

Expanding on the recurrence relation, we can evaluate

$$S_m = \frac{(m)!}{2^m((m/2)!)^2}\cdot\frac{\pi}{2}$$

for even m. From here on, we can use many techniques. One technique would be to use Stirling approximation and set $m = 2i$ to arrive the result. We instead take a real analytic approach to simply the computation and keep the proof clean.

From the above, we can deduce that the sequences are equivalent in the terms of real analysis, i.e., $S_{n+1} \sim S_n$. Further, for all m, $S_{m+2} \leqslant S_{m+1} \leqslant S_m$ since the sequence is decreasing. This implies that

$$\frac{S_{m+2}}{S_m} \leqslant \frac{S_{m+1}}{S_m} \leqslant 1.$$

Now by the recurrence relation, we have $(m+1)S_m \leqslant (m+2)S_{m+1}$. By the sandwich theorem, we conclude that $\frac{S_{m+1}}{S_m} \to 1$ as $n \to \infty$, and hence $S_{m+1} \sim S_m$ in real-analytic terms. By examining $S_m S_{m+1}$, we thus obtain that

$$S_m \leqslant \sqrt{\frac{\pi}{2m}}.$$

This completes the proof. ☐

Theorem 3 (Cauchy [Cau21])**.** *Let $T \in \mathbb{N}$ be a natural number. Then*

$$\frac{\pi}{2T+1}\sqrt{\frac{T(2T-1)}{3}} \leqslant \sqrt{1+\left(\frac{1}{2}\right)^2+\left(\frac{1}{3}\right)^2+\cdots+\left(\frac{1}{T}\right)^2} \leqslant \frac{\pi}{2T+1}\sqrt{\frac{2T(T+1)}{3}}.$$

Proof. The theorem can be derived from Cauchy's proof for the value of Reimann zeta function of order 2, $\zeta(2)$. The original proof by Cauchy uses Cauchy residue theorem. Here, we give a self contained proof using basic complex analysis and trigonometric identities along the line of Cauchy [Cau21].

Let $0 < \theta < \pi/2$ and $n = 2T+1$ be a positive odd integer. Let $\imath = \sqrt{-1}$. Since n is an integer, using de Moivre's theorem, we have $(\cos\theta + \imath\sin\theta)^n = \cos(n\theta) + \imath\sin(n\theta)$. Dividing by $\sin^n\theta$, we have

$$\begin{aligned}\frac{\cos(n\theta)+\imath\sin(n\theta)}{\sin^n\theta} &= \frac{(\cos\theta+\imath\sin\theta)^n}{\sin^n\theta} = \left(\frac{\cos\theta+\imath\sin\theta}{\sin\theta}\right)^n = (\cot\theta+\imath)^n \\ &= \binom{n}{0}\cot^n\theta + \binom{n}{1}(\cot^{n-1}\theta)i + \cdots + \binom{n}{n-1}(\cot\theta)i^{n-1} + \binom{n}{n}i^n \\ &= \left(\binom{n}{0}\cot^n\theta - \binom{n}{2}\cot^{n-2}\theta \pm \cdots\right) + \imath\left(\binom{n}{1}\cot^{n-1}\theta - \binom{n}{3}\cot^{n-3}\theta \pm \cdots\right).\end{aligned}$$

Equating the imaginary parts gives the identity

$$\frac{\sin(n\theta)}{\sin^n\theta} = \binom{n}{1}\cot^{n-1}\theta - \binom{n}{3}\cot^{n-3}\theta \pm \cdots.$$

Consider $\theta_k = \frac{k\pi}{2T+1}$ for integer $1 \leqslant k \leqslant T$. Then $(2T+1)\theta_k = n\theta_k = k\pi$ is a multiple of π. Thus $\sin(n\theta_k) = 0$ and $\sin\theta_k \neq 0$ since $k \leqslant T$ is not a multiple of $2T+1$. This implies

$$\binom{2T+1}{1}\cot^{2T}\theta_k - \binom{2T+1}{3}\cot^{2T-2}\theta_k \pm \cdots + (-1)^T\binom{2T+1}{2T+1} = 0$$

Using the fact that $\cot^2(\cdot)$ is an injective function on $(0, \pi/2)$, we can say that the $\cot^2\theta_k$ values are the roots of the polynomial

$$p(x) = \binom{2T+1}{1}x^T - \binom{2T+1}{3}x^{T-1} \pm \cdots + (-1)^T\binom{2T+1}{2T+1}.$$

Since we are working in the integral domain, the generalization of Vieta's formula to the ring implies that the sum of the roots is just the ratio of the first two coefficients of the polynomial. Therefore,

$$\cot^2\theta_1 + \cot^2\theta_2 + \cdots + \cot^2\theta_T = \frac{\binom{2T+1}{3}}{\binom{2T+1}{1}} = \frac{2T(2T-1)}{6} = \frac{T(2T-1)}{3}.$$

Using $\cot^2\theta \leqslant \theta^{-2} \leqslant 1 + \cot^2\theta$ for $0 \leqslant \theta < 2$ and the fact that $0 < \theta_k < \pi/2$ for $1 \leqslant k \leqslant T$, this implies that

$$\frac{T(2T-1)}{3} < \left(\frac{2T+1}{\pi}\right)^2 + \left(\frac{2T+1}{2\pi}\right)^2 + \cdots + \left(\frac{2T+1}{T\pi}\right)^2 < \frac{T(2T+2)}{3}$$

Rearranging the terms completes the proof. □

3.1 Proof of Upper bounds

In this section, we will prove all the upper bounds in our main results.

3.1.1 Proof of Upper bounds in Theorem 1

Let us first consider the case of counting. It is one of the most basic statistics one would like to compute and has been studied in the continual release model by Dwork et al. [DNPR10] and Chan et al. [CSS11]. Both these works used the binary tree mechanism to perform binary counting. Since this work, many other applications of this basic framework of counting in the continual release has been used, including but not limited to, batch to online conversion in private learning, heavy hitter estimation.

To prove an upper bound, we use the following trigonometric identities:

1. For any $\theta \in [-\pi, \pi]$, $\sin^2(\theta) + \cos^2(\theta) = 1$.
2. For even m, $\frac{2}{\pi}\int_0^{\pi/2}\cos^m(\theta)\mathrm{d}\theta = \left(\frac{1}{2}\right)\left(\frac{3}{4}\right)\cdots\left(\frac{m-1}{m}\right)$.
3. For all $\theta \in [-\pi, \pi]$, $\cos(2\theta) = \cos^2(\theta) - \sin^2(\theta) = 2\cos^2(\theta) - 1$.

We now define an n-dimensional vector a such that its k-th entry is

$$a[k] = \frac{2}{\pi}\int_0^{\pi/2}\cos^{2k}(\theta)\mathrm{d}\theta = \left(\frac{1}{2}\right)\left(\frac{3}{4}\right)\cdots\left(\frac{2k-1}{2k}\right).$$

Note that the entries of the vector satisfy a nice recurrence relation and hence very easy to compute on the fly:

$$a[k] = \left(\frac{2k-1}{2k}\right)a[k-1].$$

Using these three identities, it is simple calculus and trigonometry to verify that $M_{\mathsf{count}} = L^\top R$, where L and R are defined as follows:

$$R[i,j] = L[j,i] = \begin{cases} 1 & i = j \\ 0 & i < j \\ a[k] & i - j = k \end{cases}. \tag{5}$$

It is straightforward to see that the number of distinct entries in R and L is n. Now the maximum column norm of $L^\top$ (and that of R) can be bounded as follows:

$$c(L) = \sqrt{1 + \sum_{i=1}^{T-1} \left(\left(\frac{1}{2} \left(\frac{3}{4}\right) \cdots \left(\frac{2i-1}{2i}\right) \right)^2 \right)} = \sqrt{1 + \sum_{i=1}^{T-1} \bar{S}_{2i}^2}, \quad \text{where } \bar{S}_{2i} := \frac{1}{2}\left(\frac{3}{4}\right) \cdots \left(\frac{2i-1}{2i}\right).$$

Setting $\bar{S}_m = \frac{2}{\pi} S_m$ and $m = 2i$ in Lemma 1, we have

$$c(L) = \sqrt{1 + \sum_{i=1}^{T-1} \left(\left(\frac{1}{2}\right) \left(\frac{3}{4}\right) \cdots \left(\frac{2i-1}{2i}\right) \right)^2} \leqslant \sqrt{1 + \sum_{i=1}^{T-1} \frac{1}{\pi i}} \leqslant \sqrt{1 + \frac{1}{\pi} \log(T-1)}.$$

Since $c(L) = c(R)$, we have the upper bound in eq. (3).

3.1.2 Proof of Upper Bounds in Theorem 2

In practice, if we want to continually output the average, we would first perform the counting and then divide by the current time-stamps. This would invariable lead to an error that depends on $\log^{3/2} T$.

We now move to proving eq. (4). One can prove the upper bound by using dual characterization of cb norm. However, such a bound only gives an upper bound on the best additive error possible for computing T successive average, but does not help in designing mechanism for computing average under continual observation. This is because for continual observation, we want an explicit factorization into matrices. Further, we would like the factorization to be lower triangular for efficiency reasons. Restricting ourself to lower triangular may not result in optimal factorization; however, we argue that the square root factorization would result in an additive error that is very close to the optimal. It is not clear how we can give an explicit clean and closed form expression for the entries of the square root factorization, but we give a analytic way to compute this factorization and argue that such a factorization would results in smaller error.

In particular, we wish to compute a factorization $M_{\mathsf{average}} = L^\top R$ such that $L^\top = R$ and R is a lower triangular matrix with non-negative entries. The entries of the coordinates can be computed by solving $\frac{T(T+1)}{2}$ system of equations in $\frac{T(T+1)}{2}$ unknowns. The system of equations consists of degree-2 polynomials, but it can be reduced to a linear system by solving the unknowns in the row-i before solving the row-$(i+1)$. Since the entries of L and R are non-negative, $c(R) = c(L^\top)$ equals the first column of the square root factorization of M_{average} and $0 < L[j,1] = R[j,1] < \frac{1}{j}$ for all $1 \leqslant j \leqslant T$. Therefore, $c(R)^2 \leqslant \sum_{j=1}^{T} \frac{1}{j^2}$ and we have the upper bound in equation (4):

$$c(R) = c(L^\top) \leqslant \sqrt{1 + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{3}\right)^2 + \cdots + \left(\frac{1}{T}\right)^2}$$

using Theorem 3.

3.2 Proof of Lower Bounds

3.2.1 Proof of Lower Bounds in Theorem 1

For the lower bound on $\|M_{\mathsf{count}}\|_{\mathsf{cb}}$, we appeal to operator algebra. For this, we use the characterization given by Haagerup [Haa80]:

$$\|M_{\mathsf{count}}\|_{\mathsf{cb}} = \max_{Q} \left\{ \frac{\|Q \bullet M_{\mathsf{count}}\|}{\|Q\|} \right\}. \tag{6}$$

We have the flexibility of choosing a candidate matrix Q and we have two main goals:

1. A upper bound on its operator norm and lower bound on the operator norm of its Schur product with $M_{\textsf{count}}$ can be computed easily.
2. We wish to get a lower bound that depends on $\log(T)$.

The end goal of $\log(T)$ dictates and limits our choice of the candidate matrix Q. Since computing the operator norm implies adding the square of coordinates of the vector formed by applying the operator on a vector and our end goal is to get $\log(T)$ term, we aim to get i-th coordinate to be $\log(i)$. A lower bound on the summation of $\log^2(i)$ over $1 \leqslant i \leqslant T$ can be easily found using the trapezoid rule of integration.

Now operator norm is the maximum over all choices of unit vectors, we can choose a vector v that allows us to compute $\|(Q \bullet M_{\textsf{count}})v\|_2$ easily. With $\log(T)$ end goal, the choices in v are either pick the i-th coordinate as $\sqrt{\frac{\log(T)}{i}}$ or v as normalized all one vector. However, the choice of v such that $v[i] = \sqrt{\frac{\log(T)}{i}}$ is not appropriate because we cannot exactly compute $\|v\|_2$, but rather we only know its limiting value (due to the upper and lower bound on the sum on Harmonic series).

Combining the thought process described above, our only choice is to find a matrix Q that introduces a Harmonic series for each coordinate. Since $Q \bullet M_{\textsf{count}}$ is a co-ordinate wise product, we use the following choice for candidate matrix,

$$Q := \begin{pmatrix} 0 & 1 & \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{T-1} \\ -1 & 0 & 1 & \frac{1}{2} & \cdots & \frac{1}{T-2} \\ \vdots & \vdots & \vdots & \vdots & \ldots & \vdots \\ -\frac{1}{T-1} & -\frac{1}{T-2} & -\frac{1}{T-3} & -\frac{1}{T-4} & \cdots & 0 \end{pmatrix} \tag{7}$$

It is possible to explicitly compute an upper bound on the singular value of our candidate Q from the first principle, but there is a simpler argument from Toeplitz operator that does not require tedious computation. Our choice of Q is a principal submatrix of a Toeplitz operator with a symbol f such that $f(e^{\imath\theta}) = \imath(\pi - \theta)$. With this observation, we can now use the results known in operator theory. In particular, let T_f be the Toeplitz operator on the underlying Hardy's space with the symbol f defined above. Since for any Toeplitz's operator, its singular value coincides with the ℓ_∞ of its symbol [BG00], we have

$$\left\| \mathcal{T}_f \right\| = \pi \tag{8}$$

Combining eq. (6) and eq. (8) gives us

$$\|M_{\textsf{count}}\|_{\textsf{cb}} \geqslant \frac{\|Q \bullet M_{\textsf{count}}\|}{\pi} = \frac{1}{\pi} \max_{v \in \mathbb{R}^T} \frac{\|(Q \bullet M_{\textsf{count}})\, v\|_2}{\|v\|_2}.$$

That is, to get a lower bound on $\|M_{\textsf{count}}\|_{\textsf{cb}}$, we need to lower bound $\|Q \bullet M_{\textsf{count}}\|$. For this, we pick a vector v such that we can easily manage $\|(Q \bullet M_{\textsf{count}})v\|_2$. Our choice of such v is the following vector

$$v = \begin{pmatrix} \frac{1}{\sqrt{T}} & \cdots \frac{1}{\sqrt{T}} \end{pmatrix}^\top. \tag{9}$$

With this choice and the definition of Q and v in eqs. (7) and (9), respectively, we have

$$\|Q \bullet M_{\textsf{average}} v\|_2 = \frac{1}{\sqrt{T}} \left(\sum_{j=1}^{T-1} \left(\sum_{k=1}^{j} \frac{1}{k} \right)^2 \right)^{\frac{1}{2}} \geqslant \frac{1}{\sqrt{T}} \sqrt{\underbrace{\sum_{j=1}^{T-1} (\log(j+1))^2}_{S(T)}} \tag{10}$$

To bound the term inside the square root, we look at the following related definite integral:

$$\begin{aligned} S(T) \geqslant I(T) = \int_0^{T-1} (\log(x+1))^2 \, \mathsf{d}x &= \Big[(x+1)\log^2(x+1) - 2(x+1)\log(x+1) + 2x - 2 \Big]_0^{T-1} \\ &= T(\log T)^2 - 2T\log(T) + 2T - 2. \end{aligned}$$

Plugging it in eq. (10) gives us the the left hand side of eq. (3):

$$\|M_{\mathsf{count}}\|_{\mathsf{cb}} \geqslant \frac{1}{\pi}\sqrt{\frac{T(\log T)^2 - 2T\log(T) + 2T - 2}{T}}.$$

3.2.2 Proof for Lower bound in Theorem 2

The lower bound on $\|M_{\mathsf{average}}\|_{\mathsf{cb}}$ is very simple. In particular, by picking $Q = \mathbb{I}$, we get

$$\|M_{\mathsf{average}}\|_{\mathsf{cb}} = \max_{Q} \frac{\|Q \bullet M_{\mathsf{average}}\|}{\|Q\|} \geqslant \frac{\|D\|}{\|Q\|} = 1,$$

where $D = \mathsf{diag}(e)$ is the diagonal matrix formed by the eigenvalues of M_{average}, $e = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{T} \end{pmatrix}$.

References

[ABP19] Srinivasan Arunachalam, Jop Briët, and Carlos Palazuelos. Quantum query algorithms are completely bounded forms. *SIAM Journal on Computing*, 48(3):903–925, 2019.

[AKN98] Dorit Aharonov, Alexei Kitaev, and Noam Nisan. Quantum circuits with mixed states. In *Proceedings of the thirtieth annual ACM symposium on Theory of computing*, pages 20–30, 1998.

[App21] Apple. `https://covid19.apple.com/contacttracing`, 2021.

[BATS09] Avraham Ben-Aroya and Amnon Ta-Shma. On the complexity of approximating the diamond norm. *arXiv preprint arXiv:0902.3397*, 2009.

[BD15] Tirthankar Bhattacharyya and Michael A Dritschel. *Operator Algebras and Mathematical Physics: 24th International Workshop in Operator Theory and Its Applications, Bangalore, December 2013*, volume 247. Birkhäuser, 2015.

[BFM+13] Jean Bolot, Nadia Fawaz, Shan Muthukrishnan, Aleksandar Nikolov, and Nina Taft. Private decayed predicate sums on streams. In *Proceedings of the 16th International Conference on Database Theory*, pages 284–295. ACM, 2013.

[BG00] Albrecht Böttcher and Sergei M Grudsky. *Toeplitz matrices, asymptotic linear algebra and functional analysis*, volume 67. Springer, 2000.

[Cau21] Augustin Louis Baron Cauchy. *Cours d'analyse de l'École Royale Polytechnique*, volume 1. Imprimerie royale, 1821.

[CDC20] CDC. `https://www.cdc.gov/coronavirus/2019-ncov/daily-life-coping/contact-tracing.html`, 2020.

[CFJ+96] C. C. Cowen, P. A. Ferguson, D. K. Jackman, E. A. Sexauer, C. Vogt, and H. J. Woolf. Finding norms of hadamard multipliers. *Linear algebra and its applications*, 247:217–235, 1996.

[CLSX12] T-H Hubert Chan, Mingfei Li, Elaine Shi, and Wenchang Xu. Differentially private continual monitoring of heavy hitters from distributed streams. In *International Symposium on Privacy Enhancing Technologies Symposium*, pages 140–159. Springer, 2012.

[Con00] John B Conway. *A course in operator theory*. American Mathematical Soc., 2000.

[CPR00] Andrew M Childs, John Preskill, and Joseph Renes. Quantum information and precision measurement. *Journal of modern optics*, 47(2-3):155–176, 2000.

[CR94] Alain Connes and Carlo Rovelli. Von neumann algebra automorphisms and time-thermodynamics relation in generally covariant quantum theories. *Classical and Quantum Gravity*, 11(12):2899, 1994.

[CR21] Adrian Rivera Cardoso and Ryan Rogers. Differentially private histograms under continual observation: Streaming selection into the unknown. *arXiv preprint arXiv:2103.16787*, 2021.

[CSS11] T.-H. Hubert Chan, Elaine Shi, and Dawn Song. Private and continual release of statistics. *ACM Trans. Inf. Syst. Secur.*, 14(3):26:1–26:24, 2011.

[Cum21] Rachel Cummings. Personal communication, 2021.

[CW09] Ronald R Coifman and Guido Weiss. Extensions of hardy spaces and their use in analysis. In *Fundamental Papers in Wavelet Theory*, pages 295–371. Princeton University Press, 2009.

[Dav84] Kenneth R Davidson. Similarity and compact perturbations of nest algebras. 1984.

[DJKR06] Igor Devetak, Marius Junge, Christoper King, and Mary Beth Ruskai. Multiplicativity of completely bounded p-norms implies a new additivity result. *Communications in mathematical physics*, 266(1):37–63, 2006.

[DNPR10] Cynthia Dwork, Moni Naor, Toniann Pitassi, and Guy N. Rothblum. Differential privacy under continual observation. In *Proceedings of the 42nd ACM Symposium on Theory of Computing, STOC 2010, Cambridge, Massachusetts, USA, 5-8 June 2010*, pages 715–724, 2010.

[ENU20] Alexander Edmonds, Aleksandar Nikolov, and Jonathan Ullman. The power of factorization mechanisms in local and central differential privacy. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, pages 425–438, 2020.

[Eul06] Leonhard Euler. Translation with notes of euler's paper remarques sur un beau rapport entre les series des puissances tant directes que reciproques. *Memoires de l'academie des sciences de Berlin*, 2006.

[FHO21] Hendrik Fichtenberger, Monika Henzinger, and Wolfgang Ost. Differentially private algorithms for graphs under continual observation. *arXiv preprint arXiv:2106.14756*, 2021.

[FSSS03] Jürgen Forster, Niels Schmitt, Hans Ulrich Simon, and Thorsten Suttorp. Estimating the optimal margins of embeddings in euclidean half spaces. *Machine Learning*, 51(3):263–281, 2003.

[Haa80] Uffe Haagerup. Decomposition of completely bounded maps on operator algebras, 1980.

[HLP+18] Samuel J Harris, Rupert H Levene, Vern I Paulsen, Sarah Plosker, and Mizanur Rahaman. Schur multipliers and mixed unitary maps. *Journal of Mathematical Physics*, 59(11):112201, 2018.

[Hon15] James Honaker. Efficient use of differentially private binary trees. *Theory and Practice of Differential Privacy (TPDP 2015), London, UK*, 2015.

[HP93] Uffe Haagerup and Gilles Pisier. Bounded linear operators between c^*-algebras. *Duke Mathematical Journal*, 71(3):889–925, 1993.

[HQYC21] Ziyue Huang, Yuan Qiu, Ke Yi, and Graham Cormode. Frequency estimation under multiparty differential privacy: One-shot and streaming. *arXiv preprint arXiv:2104.01808*, 2021.

[JKP09] Nathaniel Johnston, David W Kribs, and Vern I Paulsen. Computing stabilized norms for quantum operations via the theory of completely bounded maps. *Quantum Information & Computation*, 9(1):16–35, 2009.

[JPPG+10] Marius Junge, Carlos Palazuelos, David Pérez-García, Ignacio Villanueva, and Michael M Wolf. Operator space theory: a natural framework for bell inequalities. *Physical review letters*, 104(17):170405, 2010.

[JRSS21] Palak Jain, Sofya Raskhodnikova, Satchit Sivakumar, and Adam Smith. The price of differential privacy under continual observation. *arXiv preprint arXiv:2112.00828*, 2021.

[Kit97] Aleksei Yur'evich Kitaev. Quantum computations: algorithms and error correction. *Uspekhi Matematicheskikh Nauk*, 52(6):53–112, 1997.

[LMSS07] Nati Linial, Shahar Mendelson, Gideon Schechtman, and Adi Shraibman. Complexity measures of sign matrices. *Combinatorica*, 27(4):439–463, 2007.

[LS09] Nati Linial and Adi Shraibman. Lower bounds in communication complexity based on factorization norms. *Random Structures & Algorithms*, 34(3):368–394, 2009.

[Mat93a] Roy Mathias. The hadamard operator norm of a circulant and applications. *SIAM journal on matrix analysis and applications*, 14(4):1152–1167, 1993.

[Mat93b] Roy Mathias. Matrix completions, norms and hadamard products. *Proceedings of the American Mathematical Society*, 117(4):905–918, 1993.

[MMHM21] Ryan McKenna, Gerome Miklau, Michael Hay, and Ashwin Machanavajjhala. Hdmm: Optimizing error of high-dimensional statistical queries under differential privacy. *arXiv preprint arXiv:2106.12118*, 2021.

[MNT20] Jiří Matoušek, Aleksandar Nikolov, and Kunal Talwar. Factorization norms and hereditary discrepancy. *International Mathematics Research Notices*, 2020(3):751–780, 2020.

[MRT22] Brendan McMahan, Keith Rush, and Abhradeep Thakurta. Private online prefix sums via optimal matrix factorizations. *arXiv preprint arXiv:2202.08312*, 2022.

[Nik14] Aleksandar Nikolov. *New computational aspects of discrepancy theory*. PhD thesis, Rutgers University-Graduate School-New Brunswick, 2014.

[Pau82] Vern I Paulsen. Completely bounded maps on c^*-algebras and invariant operator ranges. *Proceedings of the American Mathematical Society*, 86(1):91–96, 1982.

[Pau86] Vern I Paulsen. *Completely bounded maps and dilations*. New York, 1986.

[Pau02] Vern Paulsen. *Completely bounded maps and operator algebras*. Number 78. Cambridge University Press, 2002.

[Pau21] Vern Paulsen. Personal communication, 2021.

[Pis86] Gilles Pisier. *Factorization of linear operators and geometry of Banach spaces*. Number 60. American Mathematical Soc., 1986.

[PS85] Vern I Paulsen and Ching Yun Suen. Commutant representations of completely bounded maps. *Journal of Operator Theory*, pages 87–101, 1985.

[PW09] Marco Piani and John Watrous. All entangled states are useful for channel discrimination. *Physical Review Letters*, 102(25):250501, 2009.

[Sch11] Jssai Schur. Remarks on the theory of restricted "a ned bilinear forms with infinitely many ä mutable ones. 1911.

[She91] Albert Jeu-Liang Sheu. Quantization of the poisson su (2) and its poisson homogeneous space—the 2-sphere. *Communications in mathematical physics*, 135(2):217–232, 1991.

[TMB03] Shinji Tsujikawa, Roy Maartens, and Robert Brandenberger. Non-commutative inflation and the cmb. *Physics Letters B*, 574(3-4):141–148, 2003.

[Upa19] Jalaj Upadhyay. Sublinear space private algorithms under the sliding window model. In *International Conference on Machine Learning*, pages 6363–6372, 2019.

[UU21] Jalaj Upadhyay and Sarvagya Upadhyay. A framework for private matrix analysis in sliding window model. In *International Conference on Machine Learning*, pages 10465–10475. PMLR, 2021.

[UUA21] Jalaj Upadhyay, Sarvagya Upadhyay, and Raman Arora. Differentially private analysis on graph streams. In *International Conference on Artificial Intelligence and Statistics*, pages 1171–1179. PMLR, 2021.

[Wal94] Robert M Wald. *Quantum field theory in curved spacetime and black hole thermodynamics*. University of Chicago press, 1994.

[Wat96] GA Watson. Estimating hadamard multiplier norms: with application t, o t, riangular truncation. *Linear Algebra. Appl*, 234:163–172, 1996.

[Wat07] John Watrous. Distinguishing quantum operations having few kraus operators. *arXiv preprint arXiv:0710.0902*, 2007.

[Wat09] John Watrous. Semidefinite programs for completely bounded norms. *arXiv preprint arXiv:0901.4709*, 2009.

[Wat12] John Watrous. Simpler semidefinite programs for completely bounded norms. *arXiv preprint arXiv:1207.5726*, 2012.

[Wat18] John Watrous. *The theory of quantum information*. Cambridge university press, 2018.

[WCZ+21] Tianhao Wang, Joann Qiongna Chen, Zhikun Zhang, Dong Su, Yueqiang Cheng, Zhou Li, Ninghui Li, and Somesh Jha. Continuous release of data streams under both centralized and local differential privacy. In *Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security*, pages 1237–1253, 2021.

[Wit81] Gerd Wittstock. Ein operatorwertiger hahn-banach satz. *Journal of Functional Analysis*, 40(2):127–150, 1981.

[Zar06] Vrej Zarikian. Alternating-projection algorithms for operator-theoretic calculations. *Linear algebra and its applications*, 419(2-3):710–734, 2006.