EXPONENTIAL SEPARATION FOR ONE-WAY QUANTUM COMMUNICATION COMPLEXITY, WITH APPLICATIONS TO CRYPTOGRAPHY

DMITRY GAVINSKY*, JULIA KEMPE†, IORDANIS KERENIDIS‡, RAN RAZ§, AND RONALD DE WOLF¶

Abstract. We give an exponential separation between one-way quantum and classical communication protocols for a partial Boolean function (a variant of the Boolean Hidden Matching Problem of Bar-Yossef et al.) Earlier such an exponential separation was known only for a relational problem. The communication problem corresponds to a *strong extractor* that fails against a small amount of *quantum* information about its random source. Our proof uses the Fourier coefficients inequality of Kahn, Kalai, and Linial.

We also give a number of applications of this separation. In particular, we show that there are privacy amplification schemes that are secure against classical adversaries but not against quantum adversaries; and we give the first example of a key-expansion scheme in the model of bounded-storage cryptography that is secure against classical memory-bounded adversaries but not against quantum ones.

Key words. communication complexity, exponential separation, quantum communication, one-way communication, Hidden Matching problem, quantum cryptography, bounded-storage model, extractor, streaming model

AMS subject classifications. 81P68, 81P15, 94A05, 94A60, 68P30, 68Q01

1. Introduction.

One of the main goals of quantum computing is to exhibit problems where quantum computers are much faster (or otherwise better) than classical computers. Preferably exponentially better. The most famous example, Shor's efficient quantum factoring algorithm [32], constitutes a separation only if one is willing to believe that efficient factoring is impossible on a classical computer—proving this would, of course, imply P $\neq$ NP. One of the few areas where one can establish *unconditional* exponential separations is communication complexity.

Communication complexity is a central model of computation, first defined by Yao [36]. It has found applications in many areas [21]. In this model, two parties, Alice with input x and Bob with input y, collaborate to solve some computational problem that depends on both x and y. Their goal is to do this with minimal communication. The problem to be solved could be a function $f(x, y)$ or some relational problem where for each x and y, several outputs are valid. The protocols could be *interactive*

*NEC Laboratories America, Inc., Princeton, NJ, U.S.A. Supported in part by Canada's NSERC. Part of this work was done while this author was at the Institute for Quantum Computing, University of Waterloo.

†School of Computer Science, Tel Aviv University, Israel. Part of this work was done while this author was with LRI & CNRS, Univ. de Paris-Sud, Orsay, France. Supported in part by ACI Sécurité Informatique SI/03 511 and ANR AlgoQP grants of the French Research Ministry, by an Alon Fellowship of the Israeli Higher Council of Academic Research, by a grant of the Israeli Science Foundation and by the European Commission under the Integrated Project Qubit Applications (QAP) funded by the IST directorate as Contract Number 015848.

‡CNRS & LRI, Univ. de Paris-Sud, Orsay, France.

§Faculty of Mathematics, Weizmann Institute, Israel. Part of this work was done when this author visited Microsoft Research, Redmond, supported by ISF and BSF grants.

¶CWI, Amsterdam, The Netherlands. Supported by a Veni grant from the Netherlands Organization for Scientific Research (NWO) and also partially supported by the European Commission under the Integrated Project Qubit Applications (QAP) funded by the IST directorate as Contract Number 015848.

(two-way), in which case Alice and Bob take turns sending messages to each other; *one-way*, in which case Alice sends a single message to Bob who then determines the output; or *simultaneous*, where Alice and Bob each pass one message to a third party (the *referee*) who determines the output. The bounded-error *communication complexity* of the problem is the worst-case communication of the best protocol that gives (for every input x and y) a correct output with probability at least $1-\varepsilon$, for some fixed constant $\varepsilon \in [0, 1/2)$, usually $\varepsilon = 1/3$.

Allowing the players to use *quantum* resources can reduce the communication complexity significantly. Examples of problems where quantum communication gives exponential savings were given by Buhrman, Cleve, and Wigderson for one-way and interactive protocols with zero error probability [9]; by Raz for bounded-error interactive protocols [28]; and by Buhrman, Cleve, Watrous, and de Wolf for bounded-error simultaneous protocols [8]. The first two problems are partial Boolean functions, while the third one is a total Boolean function. However, the latter separation does not hold in the presence of public coins.[1] Bar-Yossef, Jayram, and Kerenidis [4] showed an exponential separation for one-way protocols and simultaneous protocols with public coins, but they only achieved this for a relational problem, called the *Hidden Matching Problem* (HMP). This problem can be solved efficiently by one quantum message of $\log n$ qubits, while classical one-way protocols need to send nearly $\sqrt{n}$ bits to solve it. However, Boolean functions are much more natural objects than relations both in the model of communication complexity and in the cryptographic settings that we consider later in this paper. Bar-Yossef et al. stated a Boolean version of their problem (a partial Boolean function) and conjectured that the same quantum-classical gap holds for this problem as well.

1.1. Exponential separation for a variant of Boolean Hidden Matching. In this paper we prove an exponential quantum-classical one-way communication gap for a variant of the Boolean Hidden Matching Problem of [4]. Let us first state a non-Boolean communication problem. Suppose Alice has an n-bit string x, and Bob has a sequence M of αn disjoint pairs $(i_1, j_1), (i_2, j_2), \ldots, (i_{\alpha n}, j_{\alpha n}) \in [n] \times [n]$, for some parameter $\alpha \in (0, 1/2]$. This M may be viewed as a partial matching on the graph whose vertices are the n bits $x_1, \ldots, x_n$. We call this an *α-matching*. Together, x and M induce an αn-bit string z defined by the parities of the αn edges:

$$z = z(x, M) = (x_{i_1} \oplus x_{j_1}), (x_{i_2} \oplus x_{j_2}), \ldots, (x_{i_{\alpha n}} \oplus x_{j_{\alpha n}}).$$

Suppose Bob wants to learn some information about z. Let $x \in \{0,1\}^n$ be uniformly distributed, and M be uniform over the set $\mathcal{M}_{\alpha n}$ of all α-matchings. Note that for any fixed M, a uniform distribution on x induces a uniform distribution on z. Hence Bob (knowing M but not x) knows nothing about z: from his perspective it is uniformly distributed. But now suppose Alice can send Bob a short message. How much can Bob learn about z, given that message and M?

The answer is very different depending on whether the message is quantum or classical. To state this difference, we need to introduce some terminology. For probability distributions p and q whose supports are subsets of a set S, define their *total variation distance* as

$$\| p - q \|_{tvd} = \sum_{i \in S} |p(i) - q(i)|. \tag{1.1}$$

[1] In fact, whether there exists a superpolynomial separation for a total Boolean function in the presence of public coins is one of the main open questions in the area of quantum communication complexity.

This distance is 0 if and only if $p = q$; it is 2 if and only if p and q have disjoint supports; and the value lies between 0 and 2 otherwise. Suppose we want to distinguish p from q, given a sample from one of the two. The best probability with which we can succeed is $\frac{1}{2} + \frac{\|p-q\|_{tvd}}{4}$. This well-known fact gives a clear intuitive meaning to the notion of total variation distance. Modifying the protocol of [4], it is easy to show that a short *quantum* message of about $\log(n)/2\alpha$ qubits allows Bob to learn a bit at a random position in the string z. This already puts a lower bound of 1 on the total variation distance between Bob's distribution on z and the uniform αn-bit distribution.

What about a short *classical* message? Using the Birthday Paradox, one can show that if Alice sends Bob about $\sqrt{n/\alpha}$ bits of x, then with constant probability there will be one edge (i_ℓ, j_ℓ) for which Bob receives both bits x_{i_ℓ} and x_{j_ℓ}. Since $z_\ell = x_{i_\ell} \oplus x_{j_\ell}$, this gives Bob a bit of information about z. Our key theorem says that this classical upper bound is essentially optimal: if Alice sends much fewer bits, then from Bob's perspective the string z will be close (in total variation distance) to uniformly distributed, so he does not even know one bit of z.

In order to be able to state this precisely, suppose Alice is deterministic and sends c bits of communication. Then her message partitions the set of 2^n x's into 2^c sets, one for each message. A typical message will correspond to a set A of about 2^{n-c} x's. Given this message, Bob knows the random variable X is drawn uniformly from this set A and he knows M, which is his input. Hence his knowledge of the random variable $Z = z(X, M)$ is fully described by the distribution

$$p_M(z) = \Pr[Z = z \mid \text{given } M \text{ and Alice's message}] = \frac{|\{x \in A \mid z(x, M) = z\}|}{|A|}.$$

Our main technical result says that if the communication c is much less than $\sqrt{n/\alpha}$ bits, then for a typical message and averaged over all matchings M, this distribution is very close to uniform in total variation distance. In other words, most of the time Bob knows essentially nothing about z.

THEOREM 1.1. *Let x be uniformly distributed over a set $A \subseteq \{0,1\}^n$ of size $|A| \geq 2^{n-c}$ for some $c \geq 1$, and let M be uniformly distributed over the set $\mathcal{M}_{\alpha n}$ of all α-matchings, for some $\alpha \in (0, 1/4]$. There exists a universal constant $\gamma > 0$ (independent of n, c, and α), such that for all $\varepsilon \in (0, 2]$: if $c \leq \gamma\varepsilon\sqrt{n/\alpha}$ then*

$$\mathbb{E}_M [\| p_M - U \|_{tvd}] \leq \varepsilon.$$

Note that the ε in this theorem is not the error probability of a protocol for a Boolean function, but an upper bound on the expected distance between Bob's distribution p_M and the uniform distribution. We prove Theorem 1.1 using the Fourier coefficients inequality of Kahn, Kalai, and Linial [17], which is a special case of the Bonami-Beckner inequality [7, 5]. We remark that Fourier analysis has been previously used in communication complexity by Raz [27] and Klauck [18].

This result allows us to turn the above communication problem into a partial Boolean function, as follows. Again we give Alice input $x \in \{0,1\}^n$, while Bob now receives two inputs: a partial matching M as before, and an αn-bit string w. The promise on the input is that w is either equal to $z = z(x, M)$, or to its complement $\overline{z}$ (i.e. z with all bits flipped). The goal is to find out which of these two possibilities is the case. We call this communication problem αPM, for "α-Partial Matching". As mentioned before, Alice can allow Bob to learn a random bit of z with high probability by sending him an $O(\log(n)/\alpha)$-qubit message. Knowing one bit z_ℓ of z suffices to compute the Boolean function: just compare z_ℓ with w_ℓ. In contrast, if Alice sends Bob much less than $\sqrt{n/\alpha}$ classical bits, then Bob still knows essentially nothing about z. In particular, he cannot decide whether $w = z$ or $w = \overline{z}$! This gives the following separation result for the classical and quantum one-way communication complexities (with error probability fixed to 1/3, say):

THEOREM 1.2. *Let $\alpha \in (0, 1/4]$. The classical bounded-error one-way communication complexity of the α-Partial Matching problem is $R^1(\alpha\mathrm{PM}) = \Theta(\sqrt{n/\alpha})$, while the quantum bounded-error one-way complexity is $Q^1(\alpha\mathrm{PM}) = O(\log(n)/\alpha)$*

Fixing α to 1/4, we obtain the promised exponential quantum-classical separation for one-way communication complexity of $O(\log n)$ qubits vs $\Omega(\sqrt{n})$ classical bits.

As noted by Aaronson [1, Section 5], Theorem 1.2 implies that his general simulation of bounded-error one-way quantum protocols by deterministic one-way protocols

$$D^1(f) = O(mQ^1(f) \log Q^1(f)),$$

is tight up to a polylogarithmic factor. Here m is the length of Bob's input. This simulation works for any partial Boolean function f. Taking f to be our αPM for $\alpha = 1/4$, one can show that $D^1(f) = \Theta(n)$, $m = \Theta(n \log n)$, and $Q^1(f) = O(\log n)$. It also implies that his simulation of quantum bounded-error one-way protocols by classical bounded-error one-way protocols

$$R^1(f) = O(mQ^1(f)),$$

cannot be considerably improved. In particular, the product on the right cannot be replaced by the sum: if we take $f = \alpha\mathrm{PM}$ with $\alpha = 1/\sqrt{n}$, then by Theorem 1.2 we have $R^1(f) \approx n^{3/4}$, $m \approx \sqrt{n} \log n$, and $Q^1(f) = O(\sqrt{n} \log n)$.

Remarks. The earlier conference version of this paper [13] had two different communication problems, establishing an exponential one-way separation for both of them in quite different ways. The present paper unifies these two approaches to something substantially simpler.

The original Boolean Hidden Matching Problem stated in [4] is our αPM with $\alpha = 1/2$ (i.e. M is a perfect matching). Theorem 1.2, on the other hand, assumes $\alpha \leq 1/4$ for technical reasons. By doing the analysis in Section 3 a bit more carefully, we can prove Theorem 1.2 for every α that is bounded away from 1/2. Note that if $\alpha = 1/2$, then the parity of $z = z(x, M)$ equals the parity of x, so by communicating the parity of x in one bit, Alice can give Bob one bit of information about z. The conference version of this paper showed that one can prove a separation for the case where M is a perfect matching if the promise is that w is "close" to z or its complement (instead of being *equal* to z or its complement). One can think of w in this case as a "noisy" version of $z = z(x, M)$ (or its complement), while the w of our current version can be thought of as starting from a perfect matching M', and then "erasing" some of the $n/2$ bits of the string $z(x, M')$ to get the αn-bit string z (or its complement).

The separation given here can be modified to a separation in the simultaneous message passing model, between the models of classical communication with shared entanglement and classical communication with shared randomness. Earlier, such a separation was known only for a relational problem [4, 14], not for a Boolean function.

1.2. Application: privacy amplification. Randomness *extractors* extract almost uniform randomness from an *imperfect* (i.e. non-uniform) source of randomness

X with the help of an independent uniform seed Y. With a bit of extra work (see Section 4), Theorem 1.1 actually implies that our function $z : \{0,1\}^n \times \mathcal{M}_{\alpha n} \to \{0,1\}^{\alpha n}$ is an extractor:

> If $X \in \{0,1\}^n$ is a random variable with min-entropy at least $n - \gamma\varepsilon\sqrt{n/\alpha}$ (i.e. $\max_x \Pr[X = x] \leq 2^{-(n-\gamma\varepsilon\sqrt{n/\alpha})}$) and Y is a random variable uniformly distributed over $\mathcal{M}_{\alpha n}$, then the random variable $Z := z(X,Y)$ is ε-close to the uniform distribution on $\{0,1\}^{\alpha n}$.

It is in fact a *strong* extractor: the pair (Y,Z) is ε-close to the uniform distribution on $\mathcal{M}_{\alpha n} \times \{0,1\}^{\alpha n}$.[2] Informally, this says that if there is a lot of uncertainty about X, then Z will be close to uniform even if Y is known.[3]

Extractors have found numerous applications in computer science, in particular in complexity theory (see e.g. [31] and the references therein) and cryptography. One important cryptographic application is that of *privacy amplification*, introduced in [6, 16]. In this setting two parties called *Alice* and *Bob* start with a shared random variable X, about which an *adversary* has partial knowledge. The parties' goal is to generate a secret key Z, about which the adversary would have very little information.

They can achieve this by communicating an independent uniform seed Y over a *public* channel, and using a strong extractor to generate the key $Z(X,Y)$. Our extractor guarantees that if the shared variable X, conditioned upon the adversary's knowledge, has min-entropy at least $n - \gamma\varepsilon\sqrt{n/\alpha}$, then the generated αn-bit key Z, conditioned upon adversary's knowledge, is ε-close to uniform. On the other hand, we show that this scheme is *insecure* against a *quantum* adversary who uses only $O(\log n)$ qubits of storage. This is the first example of a privacy amplification scheme that is safe against classical adversaries with up to $\Theta(\sqrt{n})$ bits of storage (with some small constant in the $\Theta(\cdot)$), but not against quantum adversaries with *exponentially less* quantum storage.

This dependence on whether the adversary has quantum or classical memory is quite surprising, particularly in light of the following two facts. First, privacy amplification based on two-universal hashing provides exactly the same security against classical and quantum adversaries. The length of the key that can be extracted is given by the min-entropy both in the classical ([6, 16]) and the quantum case ([19, 30], [29, Ch. 5]). Second, König and Terhal [20] have shown that for protocols that extract just *one* bit, the level of security against a classical and a quantum adversary (with the same information bound) is comparable.

1.3. Application: key-expansion in the bounded-storage model. In privacy amplification, we can ensure that the adversary has much uncertainty about the random variable X by assuming that he has only bounded storage. The idea of basing cryptography on storage-limitations of the adversary was introduced by Maurer [23] with the aim of implementing information-theoretically secure *key-expansion*. In this setting, a large random variable X is publicly but only temporarily available. Alice and Bob use a shared secret key Y to extract an additional key $Z = Z(X,Y)$ from X,

---

[2] Note that $\mathbb{E}_M\left[\| p_M - U \|_{tvd}\right] = \| (Y,Z) - U \|_{tvd}$, where '$U$' on left and right is uniform over different domains.

[3] It should be noted that the parameters of our extractor are quite bad, as far as these things go. First, the uniform input seed Y takes about $\alpha n \log n$ bits to describe, which is more than the αn bits that the extractor outputs; in a good extractor, we want the seed length to be much shorter than the output length. Second, our assumed lower bound on the initial min-entropy is quite stringent. Finally, the distance from uniform can be made polynomially small in n (by putting an $n - n^{1/2-\eta}$ lower bound on the min-entropy of X) but not exponentially small, which is definitely a drawback in cryptographic contexts. Still, this extractor suffices for our purposes here.

in such a way that the adversary has only limited information about the pair (Y, Z). "Limited information" means that the distribution on (Y, Z) is ε-close to uniform even when conditioned on the information about X that the adversary stored. Thus Alice and Bob have expanded their shared secret key from Y to (Y, Z). Aumann, Ding, and Rabin [3] were the first to prove a bounded-storage scheme secure, and essentially tight constructions have subsequently been found [11, 22, 33].

It is an important open question whether any of these constructions remain secure if the adversary is allowed to store *quantum* information. One may even conjecture that a bounded-storage protocol secure against classical adversaries with a certain amount of memory, should be roughly as secure against *quantum* adversaries with roughly the same memory bound. After all, Holevo's theorem [15] tells us that k qubits cannot contain more information than k classical bits. However, a key-expansion scheme based on our extractor refutes this conjecture. The scheme is essentially the same as the above privacy amplification scheme, but we describe it separately because the context is a bit different. Alice and Bob will compute $Z := z(X, Y)$ by applying our extractor to X and Y. If the adversary's memory is bounded by $\gamma\varepsilon\sqrt{n/\alpha}$ bits, then Z will be ε-close to uniform from the adversary's perspective. On the other hand, $O(\log n)$ qubits of storage suffice to learn one or more bits of information about Z, given Y, which shows that (Y, Z) is not good as a key against a quantum adversary. Thus we have an example of a key-expansion scheme that is secure against classical adversaries with nearly $\sqrt{n}$ bits of storage, but insecure against quantum adversaries even with exponentially less quantum storage.

1.4. Application: a separation in the streaming model. In the *streaming model* of computation, the input is given as a stream of bits and the algorithm is supposed to compute or approximate some function of the input, having only space of size S available. See for instance [2, 24].

There is a well-established connection between one-way communication complexity and the streaming model: if we view the input as consisting of two consecutive parts x and y, then the content of the memory after x has been processed, together with y, contains enough information to compute $f(x, y)$. Hence, a space-S streaming algorithm for f implies a one-way protocol for f of communication S with the same success probability. The classical lower bound for our Boolean communication complexity problem, together with the observation that our quantum protocol can be implemented in the streaming model, implies a separation between the quantum and classical streaming model. Namely, there is a partial Boolean function f that can be computed in the streaming model with small error probability using quantum space of $O(\log n)$ qubits, but requires $\Omega(\sqrt{n})$ bits if the space is classical.

Le Gall [12] constructed a problem that can be solved in the streaming model using $O(\log n)$ qubits of space, while any classical algorithm needs $\Omega(n^{1/3})$ classical bits. His $\log n$-vs-$n^{1/3}$ separation is a bit smaller than our $\log n$-vs-$\sqrt{n}$, but his separation is for a *total* Boolean function while ours is only partial (i.e. requires some promise on the input). Le Gall's result predates ours, though we only learned about it after finishing the conference version of our paper. We remark also that Le Gall's separation holds only in the streaming model variant where the bits arrive in order, while ours holds in the more general model where we allow the different pieces of the input to arrive in any order.

The algorithm for solving our problem in the streaming model starts out with a $\log n$-qubit superposition $\frac{1}{\sqrt{n}}\sum_{i=1}^{n}|i\rangle$. Whenever a bit x_i streams by in the input, the algorithm unitarily multiplies basis state $|i\rangle$ with a phase $(-1)^{x_i}$. Whenever an edge (i_ℓ, j_ℓ) streams by, the algorithm measures with operators $E_1 = |i_\ell\rangle\langle i_\ell| + |j_\ell\rangle\langle j_\ell|$ and $E_0 = I - E_1$; in case of outcome E_1, the algorithm records the values i_ℓ and j_ℓ (note that E_1 can be obtained at most once, as the edges are pairwise disjoint). And whenever a bit (i_ℓ, j_ℓ, w_ℓ) streams by, the algorithm unitarily multiplies basis state $|\min(i_\ell, j_\ell)\rangle$ with a phase $(-1)^{w_\ell}$. At the end, with probability 2α the algorithm is left with a classical record of $(i_\ell, j_\ell) \in M$ and the corresponding quantum state $\frac{1}{\sqrt{2}}((-1)^{x_{i_\ell} \oplus w_\ell}|i_\ell\rangle + (-1)^{x_{j_\ell}}|j_\ell\rangle)$. The algorithm can learn the function value $x_{i_\ell} \oplus x_{j_\ell} \oplus w_\ell$ from this by a final measurement.

2. The problem and its quantum and classical upper bounds. We assume basic knowledge of quantum computation [26] and (quantum) communication complexity [21, 34].

Before giving the definition of our variant of the Boolean Hidden Matching Problem, we fix some notation. Part of Bob's input will be a sequence M of αn disjoint edges $(i_1, j_1), \ldots, (i_{\alpha n}, j_{\alpha n})$ over $[n]$, which we call an α-matching. We use $\mathcal{M}_{\alpha n}$ to denote the set of all such matchings. If $\alpha = 1/2$ then the matching is *perfect*, if $\alpha < 1/2$ then the matching is *partial*. We can view M as an $\alpha n \times n$ matrix over $GF(2)$, where the ℓ-th row has exactly two 1s, at positions i_ℓ and j_ℓ. Let $x \in \{0,1\}^n$. Then the matrix-vector product Mx is an αn-bit string $z = z_1, \ldots, z_\ell, \ldots z_{\alpha n}$ where $z_\ell = x_{i_\ell} \oplus x_{j_\ell}$. Using this notation, we define the following *α-Partial Matching (αPM)* problem, whose one-way communication complexity we will study.

Alice: $x \in \{0,1\}^n$
Bob: an α-matching M and a string $w \in \{0,1\}^{\alpha n}$
Promise on the input: there is a bit b such that $w = Mx \oplus b^{\alpha n}$ (equivalently, $w = z$ or $w = \overline{z}$)
Function value: b

Actually, most of our analysis will not be concerned with Bob's second input w. Rather, we will show that given only a short message about x, Bob will know essentially nothing about $z = Mx$. Note that to compute b, it suffices that Bob learns one bit z_ℓ of the string z, since $b = z_\ell \oplus w_\ell$. We will first give quantum and classical upper bounds on the message length needed for this.

Quantum upper bound:. Suppose Alice sends a uniform superposition of her bits to Bob:

$$|\psi\rangle = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (-1)^{x_i} |i\rangle.$$

Bob completes his αn edges to a perfect matching in an arbitrary way, and measures with the corresponding set of $n/2$ 2-dimensional projectors. With probability 2α he will get one of the edges (i_ℓ, j_ℓ) of his input M. The state then collapses to

$$\frac{1}{\sqrt{2}} \left((-1)^{x_{i_\ell}} |i_\ell\rangle + (-1)^{x_{j_\ell}} |j_\ell\rangle \right),$$

from which Bob can obtain the bit $z_\ell = x_{i_\ell} \oplus x_{j_\ell}$ by measuring in the corresponding $|\pm\rangle$-basis. Note that this protocol has so-called "zero-sided error": Bob knows when he didn't learn any bit z_ℓ. If Bob is given $O(k/\alpha)$ copies of $|\psi\rangle$, then with high probability (at least while $k \ll \alpha n$) he can learn k distinct bits of z.

Remark. This protocol can be modified to a protocol in the simultaneous message passing model in a standard way, first suggested by Buhrman (see [14]). Alice and Bob share the maximally entangled state $\frac{1}{\sqrt{n}}\sum_i |i,i\rangle$. Alice implements the transformation $|i\rangle \to (-1)^{x_i}|i\rangle$ on her half. Bob performs the measurement with his projectors on his half. If he gets one of the edges of his input, he sends the resulting (i_ℓ, j_ℓ) and w_ℓ to the referee. Now Alice and Bob perform a Hadamard transform on their halves, measure and send the result to the referee, who has enough information to reconstruct z_ℓ.

Classical upper bound:. We sketch an $O(\sqrt{n/\alpha})$ *classical* upper bound. Suppose Alice uniformly picks a subset of $d \approx \sqrt{n/\alpha}$ bits of x to send to Bob. By the Birthday Paradox, with high probability Bob will have both endpoints of at least one of his αn edges and so he can compute a bit of z (and hence the function value b) with good probability. In this protocol Alice would need to send about $d\log n$ bits to Bob, since she needs to describe the d indices as well as their bit values. However, by Newman's Theorem [25], Alice can actually restrict her random choice to picking one out of $O(n)$ possible d-bit subsets, instead of one out of all $\binom{n}{d}$ possible subsets. Hence $d + O(\log n)$ bits of communication suffice. This matches our lower bound up to constant factors.

3. Main proof.

In this section we prove our main technical result (Theorem 1.1), which shows that Bob knows hardly anything about the string $z = Mx$ unless Alice sends him a long message.

3.1. Preliminaries.

We begin by providing a few standard definitions from Fourier analysis on the Boolean cube. For functions $f, g : \{0,1\}^n \to \mathbb{R}$ we define their inner product and ℓ_2-norm by

$$\langle f, g\rangle = \frac{1}{2^n}\sum_{x\in\{0,1\}^n} f(x)g(x) \quad , \quad \| f \|_2^2 = \langle f, f\rangle = \frac{1}{2^n}\sum_{x\in\{0,1\}^n} |f(x)|^2. \tag{3.1}$$

The Fourier transform of f is a function $\widehat{f} : \{0,1\}^n \to \mathbb{R}$ defined by

$$\widehat{f}(s) = \langle f, \chi_s\rangle = \frac{1}{2^n}\sum_{y\in\{0,1\}^n} f(y)\chi_s(y),$$

where $\chi_s : \{0,1\}^n \to \mathbb{R}$ is the character $\chi_s(y) = (-1)^{y\cdot s}$ with "$\cdot$" being the scalar product; $\widehat{f}(s)$ is the Fourier coefficient of f corresponding to s. We have the following relation between f and $\widehat{f}$:

$$f = \sum_{s\in\{0,1\}^n} \widehat{f}(s)\chi_s.$$

We will use two tools in our analysis, Parseval's identity and the KKL lemma.

Lemma 3.1 (Parseval). *For every function $f : \{0,1\}^n \to \mathbb{R}$ we have $\| f \|_2^2 = \sum_{s\in\{0,1\}^n} \widehat{f}(s)^2$.*

Note in particular that if f is an arbitrary probability distribution on $\{0,1\}^n$ and U is the uniform distribution on $\{0,1\}^n$, then $\widehat{f}(0^n) = \widehat{U}(0^n) = 1/2^n$ and $\widehat{U}(s) = 0$ for nonzero s, hence

$$\| f - U \|_2^2 = \sum_{s\in\{0,1\}^n} (\widehat{f}(s) - \widehat{U}(s))^2 = \sum_{s\in\{0,1\}^n\setminus\{0^n\}} \widehat{f}(s)^2. \tag{3.2}$$

LEMMA 3.2 ([17]). *Let f be a function $f : \{0,1\}^n \to \{-1,0,1\}$. Let $A = \{x \mid f(x) \neq 0\}$, and let $|s|$ denote the Hamming weight of $s \in \{0,1\}^n$. Then for every $\delta \in [0,1]$ we have*

$$\sum_{s\in\{0,1\}^n} \delta^{|s|}\widehat{f}(s)^2 \leq \left(\frac{|A|}{2^n}\right)^{\frac{2}{1+\delta}}.$$

We also need the following combinatorial lemma about uniformly chosen matchings.[4]

LEMMA 3.3. *Let $v \in \{0,1\}^n$. If $|v| = k$ for even k, then*

$$\Pr_M[\exists\, s \in \{0,1\}^{\alpha n} s.t.\ M^T s = v] = \frac{\binom{\alpha n}{k/2}}{\binom{n}{k}},$$

where the probability is taken uniformly over all α-matchings M.

Proof. We can assume without loss of generality that $v = 1^k 0^{n-k}$. We will compute the fraction of matchings M for which there exists such an s. The total number of matchings M of αn edges is $n!/(2^{\alpha n}(\alpha n)!(n-2\alpha n)!)$. This can be seen as follows: pick a permutation of n, view the first αn pairs as αn edges, and ignore the ordering within each edge, the ordering of the αn edges, and the ordering of the last $n - 2\alpha n$ vertices. Note that $\exists\ s\ s.t.\ M^T s = v$ if and only if M has exactly $k/2$ edges in $[k]$ and $\alpha n - k/2$ edges in $[n]\backslash[k]$. The number of ways to pick $k/2$ edges in $[k]$ (i.e. a perfect matching) is $k!/(2^{k/2}(k/2)!)$. The number of ways to pick $\alpha n - k/2$ edges in $[n] - [k]$ is $(n-k)!/(2^{\alpha n - k/2}(\alpha n - k/2)!(n-2\alpha n)!)$. Hence the probability in the lemma equals

$$\frac{k!/(2^{k/2}(k/2)!)\ \cdot\ (n-k)!/(2^{\alpha n - k/2}(\alpha n - k/2)!(n-2\alpha n)!)}{n!/(2^{\alpha n}(\alpha n)!(n-2\alpha n)!)} = \frac{\binom{\alpha n}{k/2}}{\binom{n}{k}}.$$

$\square$

This probability is exponentially small in k if $\alpha < 1/2$, but it equals 1 if $\alpha = 1/2$ and $v = 1^n$.

3.2. Proof of Theorem 1.1. In order to prove Theorem 1.1, consider any set $A \subseteq \{0,1\}^n$ with $|A| \geq 2^{n-c}$ and let $f : \{0,1\}^n \to \{0,1\}$ be its characteristic function (i.e. $f(x) = 1$ if and only if $x \in A$). Let $\varepsilon \in (0,2]$, $\alpha \in (0,1/4]$, and $1 \leq c \leq \gamma\varepsilon\sqrt{n/\alpha}$ for some γ to be determined later.

With x uniformly distributed over A, we can write down Bob's induced distribution on z as

$$p_M(z) = \frac{|\{x \in A \mid Mx = z\}|}{|A|}.$$

We want to show that p_M is close to uniform, for most M. By Eq. (3.2), we can achieve this by bounding the Fourier coefficients of p_M. These are closely related to

[4]We use the standard convention $\binom{a}{b} = 0$ whenever $b > a$.

the Fourier coefficients of f:

$$\begin{aligned}
\widehat{p_M}(s) &= \frac{1}{2^{\alpha n}} \sum_{z\in\{0,1\}^{\alpha n}} p_M(z)(-1)^{z\cdot s} \\
&= \frac{1}{|A|2^{\alpha n}} \left(|\{x \in A \mid (Mx)\cdot s = 0\}| - |\{x \in A \mid (Mx)\cdot s = 1\}|\right) \\
&= \frac{1}{|A|2^{\alpha n}} \left(|\{x \in A \mid x\cdot (M^T s) = 0\}| - |\{x \in A \mid x\cdot (M^T s) = 1\}|\right) \\
&= \frac{1}{|A|2^{\alpha n}} \sum_{x\in\{0,1\}^{n}} f(x)(-1)^{x\cdot(M^T s)} \\
&= \frac{2^n}{|A|2^{\alpha n}} \cdot \widehat{f}(M^T s). \qquad (3.3)
\end{aligned}$$

Note that the Hamming weight of $v = M^T s \in \{0,1\}^n$ is twice the Hamming weight of $s \in \{0,1\}^{\alpha n}$.

Using KKL, we get the following bound on the level sets of the Fourier transform of f:

LEMMA 3.4. *For every $k \in \{1,\dots,4c\}$ we have* $\displaystyle\frac{2^{2n}}{|A|^2}\sum_{v:|v|=k} \widehat{f}(v)^2 \le \left(\frac{4\sqrt{2}c}{k}\right)^k.$

Proof. By the KKL inequality (Lemma 3.2), for every $\delta \in [0,1]$ we have

$$\frac{2^{2n}}{|A|^2}\sum_{v:|v|=k} \widehat{f}(v)^2 \le \frac{2^{2n}}{|A|^2}\frac{1}{\delta^k}\left(\frac{|A|}{2^n}\right)^{2/(1+\delta)} = \frac{1}{\delta^k}\left(\frac{2^n}{|A|}\right)^{2\delta/(1+\delta)} \le \frac{1}{\delta^k}\left(\frac{2^n}{|A|}\right)^{2\delta} \le \frac{2^{2\delta c}}{\delta^k}.$$

Plugging in $\delta = k/4c$ (which is in $[0,1]$ by our assumption on the value of k) gives the lemma. □

We bound the expected *squared* total variation distance between p_M and U as follows:

$$\begin{aligned}
\mathbb{E}_M[\| p_M - U \|_{tvd}^2] &\le 2^{2\alpha n}\, \mathbb{E}_M\left[\| p_M - U \|_2^2\right] \\
&= 2^{2\alpha n}\, \mathbb{E}_M\left[\sum_{s\in\{0,1\}^{\alpha n}\setminus\{0^{\alpha n}\}} \widehat{p_M}(s)^2\right] \\
&= \frac{2^{2n}}{|A|^2}\mathbb{E}_M\left[\sum_{s\in\{0,1\}^{\alpha n}\setminus\{0^{\alpha n}\}} \widehat{f}(M^T s)^2\right]
\end{aligned}$$

where we used, respectively, the Cauchy-Schwarz inequality (recall that our definition of $\|\cdot\|_2^2$ in Eq. (3.1) already includes a factor $1/2^{\alpha n}$), Eq. (3.2), and Eq. (3.3). Note that for each $v \in \{0,1\}^n$, there is at most one $s \in \{0,1\}^{\alpha n}$ for which $M^T s = v$ (and the only s that makes $M^T s = 0^n$, is $s = 0^{\alpha n}$). This allows us to change the expectation over M into a probability and use Lemma 3.3:

$$
\begin{aligned}
&= \frac{2^{2n}}{|A|^2}\mathbb{E}_M\left[\sum_{v\in\{0,1\}^n\setminus\{0^n\}} |\{s\in\{0,1\}^{\alpha n} \mid M^T s = v\}|\cdot \widehat{f}(v)^2\right]\\
&= \frac{2^{2n}}{|A|^2}\sum_{v\in\{0,1\}^n\setminus\{0^n\}} \Pr_M\left[\exists\, s\in\{0,1\}^{\alpha n} s.t.\ M^T s = v\right]\cdot \widehat{f}(v)^2\\
&= \frac{2^{2n}}{|A|^2}\sum_{\text{even } k=2}^{2\alpha n} \frac{\binom{\alpha n}{k/2}}{\binom{n}{k}} \sum_{v:|v|=k} \widehat{f}(v)^2.
\end{aligned}
$$

We first upper bound the part of this sum with $k < 4c$. Applying Lemma 3.4 for each k, using the standard estimates $(n/k)^k \leq \binom{n}{k} \leq (en/k)^k$, and our upper bound $c \leq \gamma\varepsilon\sqrt{n/\alpha}$, we get:

$$
\frac{2^{2n}}{|A|^2}\sum_{\text{even } k=2}^{4c-2} \frac{\binom{\alpha n}{k/2}}{\binom{n}{k}} \sum_{v:|v|=k} \widehat{f}(v)^2 \leq \sum_{\text{even } k=2}^{4c-2} \frac{(2e\alpha n/k)^{k/2}}{(n/k)^k}\left(\frac{4\sqrt{2}c}{k}\right)^k \leq \sum_{\text{even } k=2}^{4c-2}\left(\frac{64e\gamma^2\varepsilon^2}{k}\right)^{k/2}.
$$

Picking γ a sufficiently small constant, this is at most $\varepsilon^2/2$ (note that the sum starts at $k=2$).

In order to bound the part of the sum with $k \geq 4c$, note that the function $g(k) := \binom{\alpha n}{k/2}/\binom{n}{k}$ is decreasing for the range of even k up to $2\alpha n$ (which is $\leq n/2$ because $\alpha \leq 1/4$):

$$
\begin{aligned}
\frac{g(k-2)}{g(k)} &= \frac{\binom{\alpha n}{k/2-1}/\binom{n}{k-2}}{\binom{\alpha n}{k/2}/\binom{n}{k}}\\
&= \frac{(n-k+2)(n-k+1)k/2}{(\alpha n - k/2+1)(k-1)k}\\
&= \frac{(n-k+2)(n-k+1)}{(2\alpha n - k+2)(k-1)}\\
&\geq \frac{n-k+1}{k-1} \geq 1.
\end{aligned}
$$

We also have $\sum_{v\in\{0,1\}^n} \widehat{f}(v)^2 = \frac{|A|}{2^n}$ by Parseval (Lemma 3.1), and $\frac{2^n}{|A|} \leq 2^c$ by assumption. Hence

$$
\frac{2^{2n}}{|A|^2}\sum_{\text{even } k=4c}^{2\alpha n} g(k)\sum_{v:|v|=k}\widehat{f}(v)^2 \leq 2^c g(4c) \leq \left(\frac{8\sqrt{2}e\alpha c}{n}\right)^{2c} \leq \left(8\sqrt{2}e\gamma\varepsilon\sqrt{\frac{\alpha}{n}}\right)^{2c} \leq \varepsilon^2/2,
$$

where in the last step we used $\alpha/n \leq 1$ and $c \geq 1$, and picked γ a sufficiently small constant.

Hence we have shown $\mathbb{E}_M[\| p_M - U \|_{tvd}^2] \leq \varepsilon^2$. By Jensen's inequality we have

$$
\mathbb{E}_M[\| p_M - U \|_{tvd}] \leq \sqrt{\mathbb{E}_M[\| p_M - U \|_{tvd}^2]} \leq \varepsilon.
$$

This concludes the proof of

Theorem 1.1. *Let x be uniformly distributed over a set $A \subseteq \{0,1\}^n$ of size $|A| \geq 2^{n-c}$ for some $c \geq 1$, and let M be uniformly distributed over the set $\mathcal{M}_{\alpha n}$ of all α-matchings, for some $\alpha \in (0, 1/4]$. There exists a universal constant $\gamma > 0$ (independent of n, c, and α), such that for all $\varepsilon \in (0,2]$: if $c \leq \gamma\varepsilon\sqrt{n/\alpha}$ then*

$$\mathbb{E}_M \left[\| \, p_M - U \, \|_{tvd}\right] \leq \varepsilon.$$

The ε^2 upper bound on $\mathbb{E}_M[\| \, p_M - U \, \|_{tvd}^2]$ is essentially tight. This can be seen in the communication setting as follows. With probability $\Omega(\varepsilon^2)$ over the choice of M, at least one edge of M will have both endpoints in the first $c = \varepsilon\sqrt{n/\alpha}$ bits. Then if Alice just sends the first c bits of x to Bob, she gives him a bit of z. This makes $\| \, p_M - U \, \|_{tvd}$ at least 1, hence $\mathbb{E}_M[\| \, p_M - U \, \|_{tvd}^2] = \Omega(\varepsilon^2)$.

3.3. Proof of Theorem 1.2. Our Theorem 1.2, stated in the introduction, easily follows from Theorem 1.1. By the Yao principle [35], it suffices to analyze *deterministic* protocols under some "hard" input distribution. Our input distribution will be uniform over $x \in \{0,1\}^n$ and $M \in \mathcal{M}_{\alpha n}$. The inputs x and M together determine the αn-bit string $z = Mx$. To complete the input distribution, with probability 1/2 we set $w = z$ and with probability 1/2 we set w to z's complement $\overline{z}$.

Fix $\varepsilon > 0$ to a small constant, say 1/1000. Let $c = \gamma\varepsilon\sqrt{n/\alpha}$, and consider any classical deterministic protocol that communicates at most $C = c - \log(1/\varepsilon)$ bits. This protocol partitions the set of 2^n x's into 2^C sets $A_1, \ldots, A_{2^C}$, one for each possible message. On average, these sets have size 2^{n-C}. Moreover, by a simple counting argument, at most a $2^{-\ell}$-fraction of all $x \in \{0,1\}^n$ can sit in sets of size $\leq 2^{n-C-\ell}$. Hence with probability at least $1 - \varepsilon$, the message that Alice sends corresponds to a set $A \subseteq \{0,1\}^n$ of size at least $2^{n-C-\log(1/\varepsilon)} = 2^{n-c}$. In that case, by Theorem 1.1 and Markov's inequality, for at least a $(1 - \sqrt{\varepsilon})$-fraction of all M, the random variable $Z = MX$ (with X uniformly distributed over A) is $\sqrt{\varepsilon}$-close to the uniform distribution U. Given w, Bob needs to decide whether $w = Z$ or $w = \overline{Z}$. In other words, he is given one sample w, and needs to decide whether it came from distribution Z or $\overline{Z}$. As we mentioned after Eq. (1.1), he can only do this if the distributions of Z and $\overline{Z}$ have large total variation distance. But by the triangle inequality

$$\| \, Z - \overline{Z} \, \|_{tvd} \leq \| \, Z - U \, \|_{tvd} + \| \, \overline{Z} - U \, \|_{tvd} = 2\| \, Z - U \, \|_{tvd} \leq 2\sqrt{\varepsilon}.$$

Hence Bob's advantage over randomly guessing the function value will be at most ε (for the unlikely event that A is very small) plus $\sqrt{\varepsilon}$ (for the unlikely event that M is such that MX is more than $\sqrt{\varepsilon}$ away from uniform) plus $\sqrt{\varepsilon}/2$ (for the advantage over random guessing when $\| \, Z - U \, \| \leq \sqrt{\varepsilon}$). To sum up: if the communication is much less than $\sqrt{n/\alpha}$ bits, then Bob cannot decide the function value with probability significantly better than 1/2.

4. Viewing our construction as an extractor. So far, we have proved that if the n-bit string X is uniformly distributed over a set A with $|A| \geq 2^{n-c}$ (i.e., a flat distribution on A), and Y is uniformly distributed over all α-matchings, then $(Y, Z(X,Y))$ is close to uniform. In order to conclude the result about extractors mentioned in Section 1.2, we need to prove the same result in the more general situation when X has min-entropy at least $n - c$ (instead of just being uniform on a set of size at least 2^{n-c}). However, a result by Chor and Goldreich [10, Lemma 5] based on the fact that any distribution can be thought of as a convex combination of flat distributions, shows that the second statement follows from the first: flat distributions are the "worst distributions" for extractors.

5. Conclusion. In this paper we presented an extractor that is reasonably good when some small amount of *classical* information is known about the random source X (technically: $H_{min}(X) \geq n - O(\sqrt{n/\alpha})$), but that fails miserably if even a very small (logarithmic) amount of *quantum* information is known about X. We gave the following applications of this:

1. An exponential quantum-classical separation for one-way communication complexity for a Boolean function (which, in particular, implies near-optimality of Aaronson's classical simulations of quantum one-way protocols).
2. A classically-secure privacy amplification scheme that is insecure against a quantum adversary.
3. A key-expansion scheme that is secure against memory-bounded classical adversaries, but not against memory-bounded quantum adversaries.
4. An exponential quantum-classical separation in the streaming model of computation.

They all can be viewed as examples where *quantum memory is much more powerful than classical.* This contrasts, for instance, with the results about privacy amplification based on two-universal hashing [19, 30], where quantum memory is not significantly more powerful than classical memory.

Acknowledgments. Many thanks to Oded Regev for his ideas on simplifying the two proofs in the conference version of our paper to the current proof. Also many thanks to Renato Renner for explaining why min-entropy determines the key length when doing privacy amplification based on two-universal hashing (both in the classical and in the quantum case), and him, Yevgeniy Dodis, Christian Schaffner, and Barbara Terhal for discussions about the bounded-storage model. Thanks to Shengyu Zhang for bringing the streaming model to our attention, and to Scott Aaronson for prodding us to make explicit the dependence on α in Theorem 1.1 in light of the results mentioned at the end of Section 1.1. Thanks to Guy Kindler and Jaikumar Radhakrishnan for various references that were needed for the earlier proofs in our conference paper (but that are not needed anymore now). Finally, thanks to the anonymous Sicomp referees for some comments that improved the presentation of the paper.

REFERENCES

[1] S. Aaronson, *The learnability of quantum states*, Proceedings of the Royal Society of London, A463 (2007). quant-ph/0608142.

[2] N. Alon, Y. Matias, and M. Szegedy, *The space complexity of approximating the frequency moments*, Journal of Computer and System Sciences, 58 (1999), pp. 137–147. Earlier version in STOC'96.

[3] Y. Aumann, Y. Z. Ding, and M.O. Rabin, *Everlasting security in the bounded storage model*, IEEE Transactions of Information Theory, 48 (2002), pp. 1668–1680.

[4] Z. Bar-Yossef, T. S. Jayram, and I. Kerenidis, *Exponential separation of quantum and classical one-way communication complexity*, in Proceedings of 36th ACM STOC, 2004, pp. 128–137.

[5] W. Beckner, *Inequalities in Fourier analysis*, Annals of Mathematics, 102 (1975), pp. 159–182.

[6] C. H. Bennett, G. Brassard, and J.-M. Robert, *Privacy amplification by public discussion*, SIAM Journal on Computing, 17 (1988), pp. 210–229.

[7] A. Bonami, *Étude des coefficients de Fourier des fonctions de $L^p(G)$*, Annales de l'Institut Fourier, 20 (1970), pp. 335–402.

[8] H. Buhrman, R. Cleve, J. Watrous, and R. de Wolf, *Quantum fingerprinting*, Physical Review Letters, 87 (September 26, 2001). quant-ph/0102001.
[9] H. Buhrman, R. Cleve, and A. Wigderson, *Quantum vs. classical communication and computation*, in Proceedings of 30th ACM STOC, 1998, pp. 63–68. quant-ph/9802040.
[10] B. Chor and O. Goldreich, *Unbiased bits from sources of weak randomness and probabilistic communication complexity*, SIAM Journal on Computing, 17 (1988), pp. 230–261. Earlier version in FOCS'85.
[11] S. Dziembowski and U. Maurer, *Optimal randomizer efficiency in the bounded-storage model*, Journal of Cryptology, 17 (2004), pp. 5–26.
[12] F. Le Gall, *Exponential separation of quantum and classical online space complexity*, in Proceedings of 18th ACM SPAA, 2006, pp. 67–73. quant-ph/0606066.
[13] D. Gavinsky, J. Kempe, I. Kerenidis, R. Raz, and R. de Wolf, *Exponential separations for one-way quantum communication complexity, with applications to cryptography*, in Proceedings of 39th ACM STOC, 2007, pp. 516–525. quant-ph/0611209.
[14] D. Gavinsky, J. Kempe, O. Regev, and R. de Wolf, *Bounded-error quantum state identification and exponential separations in communication complexity*, in Proceedings of 38th ACM STOC, 2006, pp. 594–603. quant-ph/0511013.
[15] A. S. Holevo, *Bounds for the quantity of information transmitted by a quantum communication channel*, Problemy Peredachi Informatsii, 9 (1973), pp. 3–11. English translation in *Problems of Information Transmission*, 9:177–183, 1973.
[16] R. Impagliazzo, L. A. Levin, and M. Luby, *Pseudo-random generation from one-way functions*, in Proceedings of 21st ACM STOC, 1989, pp. 12–24.
[17] J. Kahn, G. Kalai, and N. Linial, *The influence of variables on Boolean functions*, in Proceedings of 29th IEEE FOCS, 1988, pp. 68–80.
[18] H. Klauck, *Lower bounds for quantum communication complexity*, in Proceedings of 42nd IEEE FOCS, 2001, pp. 288–297. quant-ph/0106160.
[19] R. König, U. Maurer, and R. Renner, *On the power of quantum memory*, IEEE Transactions on Information Theory, 51 (2005), pp. 2391–2401.
[20] R. König and B. M. Terhal, *The bounded storage model in the presence of a quantum adversary*, 11 Aug 2006. quant-ph/0608101.
[21] E. Kushilevitz and N. Nisan, *Communication Complexity*, Cambridge University Press, 1997.
[22] C.-J. Lu, *Encryption against storage-bounded adversaries from on-line strong extractors.*, Journal of Cryptology, 17 (2004), pp. 27–42.
[23] U. Maurer, *Conditionally-perfect secrecy and a provably-secure randomized cipher*, Journal of Cryptology, 5 (1992), pp. 53–66.
[24] M. Muthukrishnan, *Data Streams: Algorithms and Applications*, Now Publishers, 2005.
[25] I. Newman, *Private vs. common random bits in communication complexity*, Information Processing Letters, 39 (1991), pp. 67–71.
[26] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, 2000.
[27] R. Raz, *Fourier analysis for probabilistic communication complexity*, Computational Complexity, 5 (1995), pp. 205–221.
[28] ———, *Exponential separation of quantum and classical communication complexity*, in Proceedings of 31st ACM STOC, 1999, pp. 358–367.
[29] R. Renner, *Security of Quantum Key Distribution*, PhD thesis, ETH Zürich, 2005. quant-ph/0512258.
[30] R. Renner and R. König, *Universally composable privacy amplification against quantum adversaries*, in Proceedings of 2nd TCC, vol. 3378 of Lecture Notes in Computer Science, Springer, 2005, pp. 407–425.
[31] R. Shaltiel, *Recent developments in explicit constructions of extractors*, Bulletin of the EATCS, 77 (2002), pp. 67–95.
[32] P. W. Shor, *Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer*, SIAM Journal on Computing, 26 (1997), pp. 1484–1509. Earlier version in FOCS'94. quant-ph/9508027.
[33] S. P. Vadhan, *Constructing locally computable extractors and cryptosystems in the bounded-storage model*, Journal of Cryptology, 17 (2004), pp. 43–77.
[34] R. de Wolf, *Quantum communication and complexity*, Theoretical Computer Science, 287 (2002), pp. 337–353.
[35] A. C-C. Yao, *Probabilistic computations: Toward a unified measure of complexity*, in Proceedings of 18th IEEE FOCS, 1977, pp. 222–227.
[36] ———, *Some complexity questions related to distributive computing*, in Proceedings of 11th ACM STOC, 1979, pp. 209–213.