Fully Device Independent Quantum Key Distribution

By Umesh Vazirani and Thomas Vidick

Abstract

Quantum cryptography promises levels of security that are impossible to attain in a classical world. Can this security be guaranteed to classical users of a quantum protocol, who may not even trust the quantum devices used to implement the protocol?

This central question dates back to the early 1990s when the challenge of achieving Device-Independent Quantum Key Distribution (DIQKD) was first formulated. We answer the challenge by rigorously proving the device-independent security of an entanglement-based protocol building on Ekert's original proposal for quantum key distribution. The proof of security builds on techniques from the classical theory of pseudo-randomness to achieve a new quantitative understanding of the non-local nature of quantum correlations.

1. INTRODUCTION

The gold standard in classical cryptography is semantic security¹⁶—given an encoded message (ciphertext), any polynomial time adversary should have a negligible chance of obtaining any information whatsoever about the plaintext (the message that was encoded). Classical cryptographers have developed encryption schemes that are able to provide this guarantee by relying on un-proven assumptions about the computational difficulty of number-theoretic problems, such as factoring large numbers,²⁸ or other more general complexity-theoretic assumptions. Such cryptographic assumptions are stronger than $P \neq NP$, and considered unlikely to be proven in the foreseeable future, so one may ask whether any cryptosystems can have security guarantees that do not rely on unproven computational assumptions.

Bennett and Brassard's seminal discovery⁷ of a Quantum Protocol for Key Distribution (QKD) in 1984 opened up the possibility for a different kind of security—"unconditional security," or security based solely on the validity of universally accepted physical laws. QKD is a protocol that allows two distant parties to collaboratively and privately generate a perfectly random string, called the users' key. Once the users have generated a private shared key they can use it to communicate with perfect security as follows. To send a message securely, the sender uses the shared key as one time pad: the encoded message is simply the bit-wise XOR (parity) of the key with the message. The semantic security of this scheme is easy to verify, provided the one-time pad (the key) is completely random and unpredictable to any adversary.

The first rigorous proofs of security of QKD,^{21,31} established over two decades ago, appeared to guarantee this ultimate notion of security could be achieved. However, the promise was short-lived. QKD researchers soon realized that practical implementations fell prey to types of attacks that were

not accounted for by the security proofs, including side-channel attacks⁸ and photon receptor blinding attacks.²⁰ More generally, these attacks pointed to a fundamental issue: given the remarkable features of quantum mechanics, such as superpositions, entanglement, and the uncertainty principle—some of which made QKD possible in the first place—how can one trust that there are not novel ways of attacking the quantum devices implementing a QKD protocol, unbeknownst to the honest, but classical, users of the protocol? The basic assumption that the user has perfect control over and trust in her quantum devices in a cryptographic protocol, on which all security proofs crucially relied, now appeared wholly unrealistic.

In a paper which appeared in the proceedings of FOCS'98, Mayers and Yao²² proposed a first principles approach to this question in the form of a new security paradigm called *device independence*.³ In this formulation, all quantum devices used in a protocol are modeled as black boxes with classical inputs and outputs. The user's confidence in a quantum device should be based only on the observed input-output behavior of the device. Concretely, in the context of QKD, a secure protocol should include a practical test that guarantees that the users' quantum devices behave according to specification, even in the scenario where the devices may have been manufactured by an adversarial party. More precisely, any execution of the protocol should either abort, or contain a proof, based only on the classical statistics observed during the execution, that no eavesdropper (a term we will use synonymously with "adversary") may have obtained any partial information about the final key (except with exponentially small probability).

Device independence may appear to set an impossibly high standard. Yet already a few years before the notion was introduced a different protocol for QKD had been proposed by Ekert,¹⁴ together with intuition suggesting that it should have strong security guarantees reminiscent of device independence. The starting point for Ekert's protocol was provided by the properties of a remarkable quantum phenomenon, entanglement—the very phenomenon which so famously puzzled the authors of the "EPR paradox".¹³ A canonical example of entanglement is the Bell state on two quantum bits, or *qubits*: $1/\sqrt{2} |00\rangle + 1/\sqrt{2} |11\rangle$. Physically, such a maximally entangled pair of qubits can be realized in the polarization of a pair of photons, or by any pair of spin- $\frac{1}{2}$ particles, such as electrons.

^a The term "device independence" was only introduced much later, in 2007.²⁴

A technical version of this paper was published in *Physical Review Letters*, Sept. 29, 2014.

In pioneering work in the 1960s (see Bell⁶) the physicist John Bell showed that suitable measurements on the two qubits in a Bell state generate correlations between their outputs that cannot be reproduced by non-communicating classical devices. Such non-local correlations provide a “test of quantumness,” one that distinguishes quantum mechanics from the kinds of hidden variable theories that Einstein championed. Non-local correlations of this form have been experimentally verified to exquisite precision in the form of Bell inequality violations.^{15, 18, 30}

The simplest example of a Bell experiment was discovered by Clauser et al.⁹ The experiment can be formulated as a game, the CHSH game, named after its inventors Clauser, Horne, Shimony, and Holt that plays an essential role in our protocol. In this game, devices D_A and D_B are given as input uniformly random bits, x and y respectively. Their goal is to return bits, a and b respectively, such that $xy = a \oplus b$. An easy argument shows that if D_A and D_B are classical, that is, each device’s output is computed as a function of the device’s input only, they cannot succeed with probability greater than $3/4$, even if the devices are allowed to access a common shared random string. By contrast, if D_A and D_B are quantum, each containing one qubit of a Bell state, then by performing suitable quantum measurements on their qubits they can achieve a success probability of $\cos^2 \pi/8 \approx 0.85 > 3/4$.

The importance of non-local correlations for device independence lies in the intuition that under certain conditions these correlations, which can be tested classically via a Bell experiment, can pin down the entangled state shared by the two devices—a phenomenon known as rigidity—as well as preclude entanglement with a third party such as the Eavesdropper—a phenomenon known as monogamy. Unfortunately, beyond its intuitive formulation monogamy is notoriously difficult to quantify, and this has been one of the main obstacles to a formal proof of security in the device-independent framework.

A significant step was taken by Barrett et al.,⁵ who were able to show security of a protocol generating a single bit of key based only on the no-signaling assumption that no information is exchanged between the users’ devices, or between the devices and the eavesdropper. This is a weaker assumption than the quantum formalism, which implies non-signaling but is generally more restrictive. Their work inspired a large number of follow-up papers^{1, 2, 29} which sought security for protocols generating multiple bits, and attempted to leverage the quantum formalism to obtain a more fine-grained security analysis, leading to potential improvements in key rates. Unfortunately all these attempts ran into a fundamental obstacle, which is that one cannot use Bayes’ rule for conditional probability in a quantum setting: the Eavesdropper’s maximum probability of guessing two bits of the key $k_1 k_2$ cannot be expressed as the product of her maximum probability of guessing k_1 times her maximum probability of guessing k_2 given that she guessed k_1 . This has to do with the quantum nature of the Eavesdropper, and the fact that different measurements on the same quantum state are not always compatible (they may not commute).

A complete proof of security of a variant of Ekert’s protocol was first given in previous work of the authors.³⁵ The goal of this paper is to highlight two key components of this security proof. The first is the use of deep properties of

randomness extractors (extremal objects from the computational complexity-based theory of pseudorandomness) and their quantum-proof extensions to bypass obstacles, such as the absence of a suitable Bayes’ rule more specifically, for analyzing quantum protocols generating multiple bits. Building upon previous work on certifiable quantum random number generation,³⁴ the proof of security uses an extraction-based proof technique called the “quantum reconstruction paradigm” to perform a reduction from global attacks of the Eavesdropper to attacks on a single round of the protocol. This component of the proof should be of broad interest to complexity theorists interested in randomness extractors and multi-player games, and the exposition tries to make these aspects accessible.

The second component of the proof rules out single-round attacks by the Eavesdropper. This involves arguing that any such attack would imply an impossibly good strategy for three players involved in a simple non-local “guessing game,” that would contradict the non-signaling principle. The proof takes the form of a reduction between guessing games, and may be particularly relevant to cryptographers in view of the recent interest in the properties of no-signaling distributions in the context of delegated computation.¹⁹

Our security proof operates within a considerable generalization of the framework of Bell experiments that encompasses the study of correlations between the outputs of three quantum devices (the third device being used to model an Eavesdropper to the protocol) involved in a complex interaction which includes some limited communication between the devices. This generalization captures new physical phenomena, such as the monogamy of entanglement mentioned earlier. In our exposition we aim to abstract as much of the quantum formalism as possible by taking the point of view of classical users in the protocol, thereby also highlighting the classically understandable elements of the proof of security while postponing as far as possible any knowledge about the quantum formalism required to provide a low-level modeling of the quantum devices used.

To achieve device independent security in an experiment, it suffices to implement devices that can carry out the honest party’s protocol with sufficient fidelity that their input-output behavior passes the statistical test specified in the protocol. This is in contrast to traditional security proofs that require e.g. a device that is guaranteed to emit a single qubit at a time, a criterion that cannot be verified based on statistical data. The recent experimental demonstrations of Bell tests reproducing the desired correlations even under stringent adversarial assumptions (so-called “loophole-free”)^{5, 18, 30} make it a near-certainty that device-independent protocols for QKD will soon be realized. An even bigger canvas for such experiments is provided by the recent demonstration of the feasibility of distributing entangled qubits from satellite to distant ground stations.³⁶ These are exciting times for quantum cryptography!

2. QUANTUM STATES, ENTANGLEMENT AND THE CHSH GAME

The qubit is the fundamental unit of quantum information. For our purposes it suffices to consider the following simplified definition: a qubit is a particle that can be in a real

superposition of two states, labeled $|0\rangle$ and $|1\rangle$. The superposition is represented by a unit vector in the two dimensional real plane, and may be specified by a parameter θ as: $(\cos \theta \sin \theta)^T$, or in ket notation as $|\psi_\theta\rangle = \cos \theta |0\rangle + \sin \theta |1\rangle$. This means that a unit vector pointing in the x direction indicates the state $|0\rangle$, and a unit vector pointing in the y direction represents the state $|1\rangle$ ^b.

An observation, or measurement, performed on a qubit is specified by a choice of basis. The distinguished basis formed by the x and y -axes is called the standard basis. The Born rule for the outcomes of quantum measurements specifies that a measurement of the state $|\psi_\theta\rangle$ in the standard basis yields the outcome 0 with probability $\cos^2 \theta$, and 1 with probability $\sin^2 \theta$. Once the measurement has been performed, the qubit collapses to the post-measurement state consistent with the outcome obtained, $|\psi_0\rangle = |0\rangle$ or $|\psi_{\pi/2}\rangle = |1\rangle$ respectively. More generally, for any angle φ the qubit can also be measured in the basis obtained by rotating the standard basis by an angle of φ . Such a measurement yields the outcome 0 with probability $\cos^2(\theta - \varphi)$, and 1 with probability $\sin^2(\theta - \varphi)$ (See Figure 1).^c The post-measurement states are $|\psi_\varphi\rangle$ or $|\psi_{\pi/2+\varphi}\rangle$ respectively.

Now we turn to a discussion of entanglement. The most fundamental entangled state is the Bell state, which is a maximally entangled state of two qubits. Using the ket notation the Bell state is written $|\phi^+\rangle = \frac{1}{\sqrt{2}}|0\rangle|0\rangle + \frac{1}{\sqrt{2}}|1\rangle|1\rangle$. To understand the remarkable properties of this state, let us start by describing the result of measuring both qubits in the same basis ($|\psi_\varphi\rangle, |\psi_{\pi/2+\varphi}\rangle$): first note that due to symmetry, the outcome of the measurement on each qubit is uniformly random; that is, 0 or 1, with probability $1/2$ each. But the symmetry of the Bell state goes deeper. Note that the Bell state is rotationally invariant and can be written as $|\phi^+\rangle = \frac{1}{\sqrt{2}} |\psi_\varphi\rangle |\psi_\varphi\rangle + \frac{1}{\sqrt{2}} |\psi_{\varphi+\pi/2}\rangle |\psi_{\varphi+\pi/2}\rangle$ (this can be easily verified by direct calculation). This means that the outcome of a measurement on the two qubits is that both outcomes are 0, or both outcomes are 1, each with probability $1/2$. Moreover the post-measurement state is $|\psi_\varphi\rangle |\psi_\varphi\rangle$, or $|\psi_{\pi/2+\varphi}\rangle |\psi_{\pi/2+\varphi}\rangle$ respectively. This conclusion

Figure 1. Measuring ψ in the standard basis yields the outcome 0 with probability $|\alpha|^2$. Measuring it in the basis $(u, u^\perp)$ yields the outcome u with probability $\cos^2 \varphi$.

^b An example of a no-signaling correlation that is neither local nor quantum is the family of distributions $p(a, b|x, y) = 1/2$ if and only if $a \oplus b = xy$, for $a, b, x, y \in \{0, 1\}$. This family gives a success probability 1 in the CHSH game with probability 1.

^c Note here we use the notation “0” to designate the outcome, even though the associated basis element is the vector $|\psi_\varphi\rangle$, not $|0\rangle$. This is a standard convention; labels for outcomes are arbitrary and have no physical meaning.

holds regardless of the spatial separation between the two qubits, a conclusion that appears very counterintuitive at first glance: how does the second qubit “know” the basis in which the first qubit has been measured, and the outcome obtained?

Such natural puzzlement, however, is an artifact of our presentation (the same given by Einstein, Podolsky and Rosen!). Indeed, a simple classical scenario recreates the same phenomenon. Consider two coins with the same random orientation, so that the normal to their surface makes the same (random) angle θ in the $x - y$ plane. The two coins are then spatially separated while leaving their orientation unchanged and measured. By measurement, we mean pick an angle φ , look at the coin along the angle φ (looking from infinity, towards the coin), and report whether one sees a heads or tails (0 or 1). Clearly the outcome is uniformly random since the orientation θ was chosen uniformly at random. And if we observe (measure) both coins from the same angle φ , the outcomes will be identical.

The fundamental distinction between the quantum and classical settings manifests itself when the two qubits are measured in different bases. If we measure the first qubit in the standard basis and the second in a φ -rotated basis, then as previously each outcome (0 or 1) is individually random, but the measurement rule indicates that the two outcomes will match with probability $\cos^2 \varphi$. This differs from the statistics obtained in the classical scenario described above, in which the chance of obtaining matching outcomes is $|1 - \varphi/\pi|$ for $\varphi \in [0, \pi)$. The difference between these two numbers, φ^2 vs. φ/π for small φ , is the foundation for Bell’s test of quantumness.

We can now see how this difference plays out in a particularly elegant way in the case of the CHSH game. The quantum strategy starts with Alice and Bob sharing a Bell state. Each of them performs a measurement of their qubit in a basis which depends on their input bit (x or y), and announces the result: Alice measures her qubit in a θ_A -rotated basis, with $\theta_A = \frac{\pi}{4} x$, and Bob measures his qubit in a θ_B -rotated basis, with $\theta_B = \frac{\pi}{8} - \frac{\pi}{4} y$ (Figure 2). In the three cases where $xy = 0$, the players measure in bases that are $\pi/8$ -rotated relative to each other, and therefore output the same bit (i.e. $a \oplus b = 0$) with probability $\cos^2 \pi/8$. In the case where $xy = 1$, they measure in bases that are $3\pi/8$ -rotated relative to each other, and therefore output different bits (i.e. $a \oplus b = 1$) with probability $\cos^2 \pi/8$ as well. Thus in each of the four cases they succeed with probability exactly $\cos^2 \pi/8$. In contrast, the classical strategy

Figure 2. Our proof of security relies on one of the most fundamental Bell inequalities, the CHSH inequality,⁹ which is pictured here as a small game. Honest devices measuring a Bell pair in the bases indicated on the right of the figure will produce outputs such that $\Pr(a \oplus b = xy) = \cos^2 \pi/8$ whenever $x, y \in \{0, 1\}$, and $a = b$ whenever $x = 2$ and $y = 0$.

based on the use of random coins described earlier, using the same angles as the quantum strategy (the angles should be doubled to account for the different way these angles are used in quantum or classical strategies), would produce valid outcomes with probability $1 - 1/4 = 3/4$ in each case.

3. EKERT'S PROTOCOL AND A PROVABLE VARIANT

We are ready to move on to the more complex task of key distribution. The goal of this task is for two trusted but distant users Alice and Bob to agree on a shared n -bit key $K \in \{0, 1\}^n$. A key distribution protocol must guarantee that if the protocol runs to completion then the users produce identical keys that are indistinguishable from a uniformly random string to any eavesdropper. The only resources available to the users are trusted local random number generators and a public, authenticated quantum communication channel.

3.1. Ekert's protocol

We first introduce a slight variant of Ekert's protocol, that carries the same intuition for security. The users, Alice and Bob, initially share a large number, N , of Bell states. These states could have been generated by Alice, with one qubit from each state transmitted to Bob. Since all communication channels are public, an eavesdropper may have interfered with the qubits. Taking an adversarial point of view, it is convenient to consider a symmetric formulation, where Alice and Bob have been provided with quantum devices, D_A and D_B , by some untrusted provider. The devices contain an arbitrary state, on which they perform arbitrary measurements when operated by the user.

In the protocol, the devices D_A and D_B are used N times in sequence. At each use, Alice's device D_A can take one of three possible inputs $x_i \in \{0, 1, 2\}$, and Bob's device D_B has two inputs $y_i \in \{0, 1\}$. These inputs are chosen uniformly at random by the users in each round. The ideal prescription for the devices is such that on two of Alice's choices of input (and both Bob's inputs) the device is supposed to perform measurements that follow the optimal quantum strategy for the CHSH game (Figure 2). This means that with probability $2/3$, Alice and Bob use their devices to play the CHSH game, in which case they check the required correlations satisfy the CHSH game condition, with sufficiently high frequency on average over the number of such game rounds. The third input for Alice's device is meant to indicate a measurement that is identical to Bob's device's measurement on input 0. This means that with probability $\frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6}$ Alice and Bob choose this matching pair of inputs, in which case they expect to obtain identical outcomes from which the key can be generated.

Ekert's intuitive argument for security of this protocol relied on a few observations. The first is that the eavesdropper cannot obtain any meaningful information while the qubits are in transit from the source, because there is no information encoded in the Bell state: as formulated in Ref. Ekert,¹⁴ the information only "comes into being" when the qubits are measured by Alice and Bob. Therefore the only reasonable attack for the Eavesdropper at this stage is to keep a quantum state that may be partially or fully entangled with the qubits that are present in Alice and Bob's devices. A second observation is

that the CHSH game test can be used to build confidence that the qubits are indeed in a Bell state. Intuitively, the Bell state is a pure state, which by definition cannot be entangled with any qubits held by the eavesdropper. Therefore, even after the users have interacted with their devices there should be no additional information that the eavesdropper may gain by performing measurements on her system. Ekert did not provide a proof of security for his QKD protocol. A proof in the weaker model for security, in which the users' quantum devices are trusted, was first achieved by reduction to the BB'84 protocol.³¹ The proof bypassed Ekert's original intuition, instead building on the nascent theory of quantum error-correction and its application to entanglement distillation, and it crucially relied on the assumption that the quantum devices are fully trusted.

3.2. Protocol E1

We will show security for a variant of the protocol described in the previous section, described in Figure 3. As previously, the users, Alice and Bob, are provided with arbitrary quantum devices D_A , which takes as input a trit $x_i \in \{0, 1, 2\}$ and D_B , which takes as input a bit $y_i \in \{0, 1\}$. The devices are used sequentially in N rounds, each time returning an output bit, $a_i, b_i \in \{0, 1\}$ respectively. The ideal specification for the devices is the same as in the previous protocol.

We divide the protocol into three phases. The first phase is the *generation phase*, in which the users execute their respective device N times in sequence by making uniformly random choices of inputs and collecting the N output bits. Then begins the *testing phase*: the users publicly select a constant fraction γ of the rounds uniformly at random and evaluate the devices' success probability in a variant of the CHSH game on average over the input-output pairs generated in those rounds. If they observe that the success probability deviates from the optimum by more than the allowed tolerance η , the users abort

Figure 3. Our DIQKD protocol, protocol E1.

1. **Inputs:** N = number of rounds, γ = testing fraction, η = noise tolerance.
2. **Generation:** For rounds $i = 1, \dots, N$: Alice selects $x_i \in \{0, 1, 2\}$, and Bob selects $y_i \in \{0, 1\}$, uniformly at random. They input x_i, y_i into their device, obtaining outputs $a_i, b_i \in \{0, 1\}$ respectively.
3. **Testing:** Alice chooses a random subset $\mathbf{B} \subseteq \{1, \dots, N\}$ of test rounds of size $\lceil \gamma N \rceil$ and shares it publicly with Bob. Alice and Bob announce their input/output pairs in all rounds in $\mathbf{B}$. They compute the fraction of pairs in $\mathbf{B}$ that satisfy the CHSH condition $a_i \oplus b_i = x_i y_i$. If this fraction is smaller than $\cos^2 \frac{\pi}{8} - \eta$ they abort the protocol. Similarly, they compute the fraction of pairs such that $a_i = b_i$, conditioned on $(x_i, y_i) = (2, 0)$, and abort if the fraction is smaller than $1 - \eta$.
4. **Extraction:** Alice and Bob publicly reveal their N choices of inputs. Let $\mathbf{C}$ be the set of check rounds: rounds i for which $(x_i, y_i) = (2, 0)$. They perform information reconciliation on the rounds in $\mathbf{C}$, followed by privacy amplification, to obtain their respective key.

the protocol. Otherwise they proceed to the *extraction phase*. For this they exchange their inputs in all rounds and discard rounds in which the inputs were not $(x_p, y_p) = (2, 0)$. Outputs from the remaining rounds, called the check rounds C , are kept private by each user and designated as the *raw key*, K_A and K_B respectively. Based on these strings the users perform post-processing steps of error reconciliation and privacy amplification to obtain the final shared key K . Information reconciliation is a procedure, based on error-correcting codes, which aims to ensure that $K_A = K_B$ after a small correction performed publicly. We will not discuss this standard procedure here. The goal of privacy amplification is to amplify the secrecy present in the raw key, from partially unknown to the Eavesdropper, to indistinguishable from uniform from the point of view of the Eavesdropper. We will describe this procedure in detail later.

The most important point of departure from Ekert's protocol lies in the small fraction γ of rounds in which the CHSH game correlations are tested. Skimping on testing may appear to weaken the security of the protocol: why not use all possible correlations for the test? On the other hand, testing involves publicly revealing the input/output pairs for those rounds, and this runs the risk of dishonest devices leaking information to the eavesdropper. Indeed, as we will see the proof of security of the protocol will require us to set a small, constant upper bound on the fraction γ of rounds selected for testing.

We formalize the intuition underlying protocol E1 in terms of two concepts, neither of which had been clearly formulated at the time of Ekert's paper. The first is known as rigidity. This is the observation that there is an essentially unique way to achieve the optimal success probability of $\cos^2 \pi/8$ in the CHSH game: any optimal strategy for the devices D_A and D_B in the game is locally equivalent to the specific strategy based on the Bell state described earlier. The second concept is the monogamy of correlations. This idea builds on the first by stating that maximally entangled qubits are necessarily pure, thus cannot share their "maximally entangled degrees of freedom" with any other qubits. It is a formalization of Ekert's intuition that, by witnessing correlations that certify that their devices share a Bell state, the users effectively preclude the possibility of entanglement between the devices and any malicious eavesdropper.

One possible approach to a formal proof of security of the protocol would be to give a quantitative formulation of both concepts. While this can be done within the mathematical formalism of quantum mechanics, there are strong impediments to the use of the resulting theory for proving security in any practically meaningful sense. First of all, achieving device independence prevents us from making any a priori assumption on the quantum systems used, such as a bound on the dimension of the underlying Hilbert space. Without such bound it is not a priori obvious how the distance to the optimal strategy scales as a function of the deviation from optimality of the statistics observed, a difficulty already faced by Mayers and Yao.²² More importantly, the interaction scenario which takes places in the QKD protocol is complex and involves multiple sources of information leaked to the eavesdropper. While a proof of security along these lines was given in Ref. Reichardt et al.,²⁷ the bounds obtained are too weak to tolerate any errors in the execution of the protocol (such as unavoidable photon losses, false detection events, etc).

4. OVERVIEW OF PROOF OF SECURITY

The proof of security can be decomposed into two parts. The first analyzes a single randomly chosen round of the protocol. It shows that if the protocol succeeds, then with high probability the devices' output in the chosen round must be at least partially unknown to the Eavesdropper. To do so, the challenge faced by the Eavesdropper is formulated in terms of a two player guessing game, and the information of the Eavesdropper is bounded by means of a reduction to a trivial guessing game. The maximum success probability in the latter game is bounded by virtue of the no-signaling principle, that is, in the absence of communication between the two players, the output distribution of one player must be independent of the other player's input. This part of the proof does not resort to the quantum formalism at all (though doing so can be used to give stronger quantitative bounds).

The second part of the proof shows that over a large number N of rounds, the protocol yields $\Omega(N)$ bits of shared random key that are secure against the eavesdropper. Analyzing multiple rounds of the protocol runs into fundamental obstacles having to do with the quantum nature of the Eavesdropper and her attack. Ideally we would like to associate the extraction of fewer than $\Omega(N)$ bits of shared random key with the failure of the guessing game argument in some round, resulting in a contradiction. The issue is that the Eavesdropper's attack could be a global measurement on her part of the quantum state based on the classical information exchanged in all rounds of the protocol, leading to a loss of control over the sequential nature of the protocol. We will get around these obstacles by appealing to powerful results from the theory of randomness extractors to directly deal with correlations between the classical information generated by Alice, Bob and the Eavesdropper. Roughly, this paradigm allows us to directly conclude that if the key generated by the protocol is distinguishable from random by the eavesdropper, then there is an (efficient) classical procedure to reconstruct Alice's raw key with non-negligible probability. This allows us to use classical information-theoretic tools to perform a reduction to the guessing game.

5. GUESSING GAMES

At its heart the security guarantees provided by a single round of protocol E1 hinge on the following guessing game (Figure 4). We start with an augmented variant of the CHSH game, where in addition to the CHSH game inputs Alice can be provided

Figure 4. The guessing game. Any devices satisfying both the CHSH condition $a \oplus b = xy$ and the guessing condition $c = a$ with high enough probability must allow signaling between D_A and D_B .

Figure 5. The proof of the guessing lemma. Using the winning condition for the CHSH game, wherever b_1 is the distance between a_0 and a_1 must be at least $2(1 - \omega)$. This implies the balls of radius $(1 - \omega)$ centered at a_0 and a_1 are disjoint.

with a third input, labeled 2 (each input is chosen with probability $1/3$). If inputs to the players are $(x, y) = (2, 0)$ then their outputs should match. Assuming the no-signaling principle and the validity of quantum mechanics (specifically, that $\omega = \cos^2 \pi/8$ is the optimal bound in the CHSH game), it is easy to see that Alice and Bob can win the augmented CHSH game with probability $\frac{2}{3}\omega + \frac{1}{3}$, which is optimal.

Now we consider adding one more rule to the game, the “guessing rule”: Bob should always produce a second output, $c \in \{0, 1\}$. The guessing rule requires that c should match Alice’s output a whenever her input is $x = 2$ (if $x \in \{0, 1\}$ then there is no requirement on c). We will show that Alice and Bob cannot achieve close to the optimal probability of $\frac{2}{3}\omega + \frac{1}{3}$ in the extended CHSH game, and simultaneously satisfy the guessing rule on Bob’s output c with probability close to 1: whatever Bob does, he cannot collaborate with Alice to succeed in the CHSH game, while simultaneously producing a valid guess for her output. This is a manifestation of the phenomenon of monogamy of quantum correlations alluded to earlier. To understand the significance of this result for the security of protocol E1, note that it also implies that Alice’s output bit must look somewhat random to the eavesdropper, even when provided with Alice and Bob’s inputs. To see this, observe that any such eavesdropper could be lumped together with Bob to derive a successful strategy in the guessing game described above.

The proof is in the form of a reduction to the following trivial guessing game. There are two cooperating players, Alice and Bob. At the start of the game, Bob receives a single bit $y \in \{0, 1\}$ chosen uniformly at random. The players are then allowed to perform arbitrary computations, but are not allowed to communicate. At the end of the game, Alice outputs a bit a , and the players win if $a = y$. Clearly, any strategy with success probability that deviates from $\frac{1}{2}$ indicates a violation of the no-signaling assumption between Alice and Bob. The idea behind the use of guessing games is to show that any undesirable outcome in the cryptographic protocol—such as a successful strategy for the Eavesdropper—can be bootstrapped into a successful strategy for Alice and Bob in the trivial guessing game, thereby contradicting the no-signaling assumption.

GUESSING LEMMA. *Suppose that Alice and Bob succeed in the augmented CHSH game with probability at least $\frac{2}{3}\omega + \frac{1}{3} - \eta$, where $\omega = \cos^2 \frac{\pi}{8}$. Then the maximum probability with which Bob can successfully guess Alice’s output on input $x = 2$ is at most $1 - C(\eta_0 - \eta)$, where C and η_0 are universal constants.*

Although in general optimization over quantum strategies is intractable, good upper bounds can often be obtained by considering convex relaxations of the set of valid quantum strategies for the players (using e.g. semidefinite programming). Here the game is sufficiently simple that an intuitive argument can be given which already provides a meaningful bound.

PROOF. For simplicity, we start by showing that when $\eta = 0$, then Bob cannot guess Alice’s output with probability 1. A more careful accounting of the parameters then provides the claimed bound. Note that $\frac{2}{3}\omega + \frac{1}{3}$ is the maximum probability with which Alice and Bob can succeed in the augmented CHSH game—for this to happen they must win the CHSH game with optimal probability ω , and on inputs $x = 2$ and $y = 0$, the outputs a and b must be equal. We assume for contradiction that Bob can guess Alice’s output (on input $x = 2$) with probability 1.

We imagine repeated sequential execution of the devices, with the same fixed inputs x and y , for some number k of executions. Let $\vec{b}_y, \vec{c}_y$ denote Bob’s outputs, and $\vec{a}_x$ denote Alice’s outputs, over these k executions. Then the condition on the augmented CHSH game implies that the players’ outputs should match when the inputs are $(x, y) = (2, 0)$: $\vec{a}_2 = \vec{b}_0$. Also by the assumption of successful guessing, on inputs $(2, y)$ for any y the players’ outputs satisfy $\vec{a}_2 = \vec{c}_y$. It follows that $\vec{b}_0 = \vec{c}_0$. The fact that Alice and Bob win the CHSH game with probability ω implies that for any $x \in \{0, 1\}$, the relative Hamming distance $d_H(\vec{a}_x, \vec{b}_0) \approx (1 - \omega)$ (by the Chernoff bound this is sharply concentrated as k grow). It follows that $d_H(\vec{a}_x, \vec{c}_0) \approx (1 - \omega)$.

Now suppose Alice knew Bob’s output $\vec{c}_y$ (we will justify this assumption later). We claim that this gives Alice an advantage in guessing Bob’s input y thus providing a contradiction with the elementary guessing game described at the start of the section. Alice’s strategy is as follows: if given her input x and output $\vec{a}_x$, she guesses $y = 0$ if $d_H(\vec{a}_x, \vec{c}_y) < (1 - \omega) + \epsilon$, and $y = 1$ otherwise, where ϵ is an arbitrarily small constant (depending on k).

First note that if $y = 0$ then she succeeds with probability close to 1. This is because as shown above, since $\vec{b}_0 = \vec{c}_0$, for any $x \in \{0, 1\}$, $d_H(\vec{a}_x, \vec{c}_0) \approx (1 - \omega)$. On the other hand, when $y = 1$, by the CHSH game condition it should be the case that $d_H(\vec{a}_0, \vec{b}_1) \approx 1 - \omega$ and $d_H(\vec{a}_1, \vec{b}_1) \approx \omega$. By the triangle inequality, $d_H(\vec{a}_0, \vec{a}_1) \gtrsim 2\omega - 1 \approx .7$. It follows that $\vec{a}_0$ and $\vec{a}_1$ cannot both be close to any fixed $\vec{c}_1$: i.e. both conditions $d_H(\vec{a}_0, \vec{c}_1) < (1 - \omega) + \epsilon$ and $d_H(\vec{a}_1, \vec{c}_1) < (1 - \omega) + \epsilon$ cannot be simultaneously satisfied. This means that in the case $y = 1$, Alice must succeed with probability about $1/2$, and therefore overall Alice can guess Bob’s output with probability close to $3/4$. Contradiction.

To justify the assumption that Alice knows Bob’s output $\vec{c}_y$, we consider the following experiment: Alice and Bob execute D_A and D_B in chunks of k executions, repeatedly. In each chunk, Alice chooses a uniformly random input $x \in \{0, 1\}$. Bob always chooses the same secret input y . Bob select a value $\vec{c}$ uniformly at random and post-selects on those chunks where $\vec{c}_y = \vec{c}$. He communicates the indices of those chunks to Alice. The information this leaks about y is very small, since the marginal distribution of $\vec{c}_y$ must match the marginal distribution $\vec{a}_2$ obtained when D_A is provided input $x = 2$, which does not depend on y ; as a consequence the indices of the

chunks sent by Bob to Alice are close to uniformly distributed, irrespective of y .

This concludes the proof for the case of $\eta = 0$. It is not difficult to reproduce the same proof to make the reasoning more quantitative in the more general case that $\eta > 0$. We leave this as an exercise to the reader. $\square$

Before continuing with the multi-round analysis, we note that the guessing lemma already gives an “in principle” proof of security for quantum key distribution, following an argument due to Barrett et al.⁴ (though they used a different, more complex Bell inequality to achieve a security parameter that scales inverse proportionally to the number of rounds). Consider the following simplified variant of Protocol E1. The users instruct their devices to sequentially play $\Theta(N)$ instances of the augmented CHSH game. Alice chooses a random time at which to stop the games, and announces it (publicly) to Bob. The users publicly exchange their inputs. Once this has been done, they determine the last round, prior to the stopping time, where they used the input pair $(2, 0)$, and should thus have matching outputs; call this the key round. They exchange their outputs in all rounds prior to the key round and estimate the (augmented) CHSH game correlations. Provided these correlations come close enough to the CHSH game bound they use their output in the key round for the final key. Based on a simple martingale argument it is not hard to show that, conditioned on the devices producing outputs that satisfy the CHSH game correlations in all rounds prior to the key round, the devices in the key round have a high probability of satisfying the CHSH game condition as well (even if it is not checked). Using the guessing lemma it must be that, in that round, the eavesdropper has a bounded probability of guessing the users’ shared bit.

Note that the above argument provides a complete proof of security (for a single bit of key), without having to resort to the quantum formalism at all: the only assumption needed for security is the no-signaling principle. Going beyond a single round, however, will require us to assume in addition that the devices’ behavior can be modeled using quantum mechanics.

6. EXTRACTING A SECURE KEY

We proceed to analyze the complete protocol^d. Our goal is to provide a reduction: we aim to show that any global strategy for the Eavesdropper that yields even partial information about the users’ final shared key implies an attack on a single round of the protocol of the form described, and ruled out, in the previous section. As mentioned earlier, analyzing multiple rounds of the protocol runs into fundamental obstacles having to do with the quantum nature of the Eavesdropper and her attack. We will get around these obstacles by appealing to powerful results from the theory of randomness extractors to directly deal with correlations between the classical information generated by Alice, Bob and the Eavesdropper.

^d The appropriate notion of entropy is the quantum conditional minentropy, which has an operational interpretation: it is simply the maximum probability with which the Eavesdropper can successfully make a correct guess for the complete string K , by performing an arbitrary measurement on her quantum side information.

The first step is to formalize exactly what constitutes a strategy for the Eavesdropper. Recall that protocol E1 includes an ultimate post-processing step called *privacy amplification*. This task converts the raw key K shared by Alice and Bob (which they derived from K_A and K_B by performing information reconciliation) into a shorter string Z . Privacy amplification guarantees that as long as K contains enough entropy from the viewpoint of the Eavesdropper, Z is indistinguishable from uniform, again from the Eavesdropper’s point of view. Privacy amplification is achieved by using a strong (quantum-proof) randomness extractor:^{12,32} a procedure Ext which takes two strings of bits as input, the source X (whose role will be played by K) and a short uniformly random seed Y (that will be generated using private randomness, and shared publicly), and combines them to produce an output $Z = \text{Ext}(X, Y)$. Think of the seed as a short tag used to select a hash function from a publicly specified family; the output is given by the evaluation of the chosen hash function on the source. The security condition for a strong extractor guarantees that no adversary, even given access to the seed, is able to distinguish the output of the extractor from a uniformly random string of the same length. This guarantee will be met, provided the source contains enough entropy from the viewpoint of the adversary.

The major challenge thus remains to establish that the source, K , contains enough entropy from the viewpoint of the Eavesdropper. Our approach to showing this is based on (the quantum generalization of) a proof technique from the theory of pseudo-randomness called the “reconstruction paradigm,” originally introduced by Trevisan³³ towards the analysis of a class of *strong randomness extractors*. Roughly, this paradigm allows us to make a stronger statement than the generic strong extractor reduction sketched above: it allows us to show directly that if the key generated by the protocol (after privacy amplification) is distinguishable from random by the eavesdropper, then there is an (efficient) classical procedure to reconstruct Alice’s raw key with non-negligible probability. This allows us to use classical information-theoretic tools to complete a reduction to the guessing game.

6.1. Global attacks

Before delving into more details about the reconstruction paradigm, we first review some of the difficulties to be overcome by the analysis.

Recall that the main difference introduced in protocol E1 compared to Ekert’s protocol is the presence of only a small fraction of test rounds. The following simple attack demonstrates that the restriction is necessary for security. Note that Alice’s device D_A is able to distinguish key rounds from test rounds, since it is given a special input, 2, for the former type of round. Suppose the eavesdropper programs D_A in a way that the device systematically outputs any bit on which its input is $x = 2$ twice: in the key round, as required, as well as in the round that immediately follows. Assuming there are comparatively few key rounds, it is likely that the second round will be a test round. The devices will probably fail the CHSH game condition in that round, but provided there are sufficiently many test rounds the effect on their average success probability will remain small. But outputs in test rounds are publicly exchanged by the users: the complete raw key is leaked to the Eavesdropper! This

attack illustrates the main difficulty faced in our scenario: the users' devices may have memory, and behave adaptively in each round, leveraging the users' public communication to covertly leak information about the final key.

Even setting the issue of adaptive devices aside, there is a second important difficulty. Assume the users had a way to enforce that their devices perform independent and identical operations in each round. It may still be the case that the Eavesdropper, on which no control is possible, benefits from a global measurement on all the side information she has collected throughout the protocol in order to implement her final attack on the key. Such an attack should be ruled out, even if its success probability is tiny, as long as it is higher than the security parameter for the protocol. So assume the Eavesdropper has performed a successful attack. Here begins the difficulty: conditioning on success of an unlikely event, which involves a measurement on the quantum side information and the public information leaked throughout the protocol, introduces correlations between the classical data observed by the users (including the devices' inputs and outputs in all rounds) that need to be taken care of by the security proof.

The confounding factor here is that Bayes' rule for conditional probabilities does not work in the presence of quantum side information. Informally, if $P_g(K|E)$ is the maximum probability with which the Eavesdropper may guess the string K , given her quantum side information E , then it is not the case that $P_g(K|E) = P_g(K_N|K_{N-1}, \dots, K_1, E) \dots P_g(K_1|E)$. In fact, it is possible for the quantities on the right-hand side to all be very close to 1, while the quantity on the left-hand side is very small, the reason being that guessing measurements for $K_1, K_2, \dots, K_N$, cannot necessarily be "stitched together" into a guessing measurement for the complete string.

In order to side-step the issue we treat the combination of the joint quantum state of the devices and the eavesdropper, and the possible effects of the global measurement, as a black box, and focus on the classical information processed by Alice and Bob and the ways in which it can correlate with classical information generated by the Eavesdropper. The main tool in the analysis is the "reconstruction paradigm" mentioned above, which we discuss next.

6.2. Security and the reconstruction paradigm

Let us first recall the main ideas behind the (classical) reconstruction paradigm.^{33, e} Consider a strong extractor Ext which takes two strings of bits as input, the source X and the seed Y , and combines them to produce an output $Z = \text{Ext}(X, Y)$. Proceeding by contradiction, assume the adversary has the ability to distinguish the output of the extractor from uniform, with non-negligible advantage ε . Observe the weakness of this starting point: the adversary's advantage could come from any kind of information; for example, she is able to predict, with small advantage ε , the parity of a small subset of the bits of Z , the location of which itself depends on other bits of Z , etc. The reconstruction paradigm uses properties of a specific construction of Ext to show that from any such adversary it is possible to construct (explicitly, and this

^e The parameters we give here are inaccurate, and are only meant to give an indication of the procedure.

will be important for us) the following stronger adversary. The stronger adversary has the ability, given as side information the information available to the original adversary (partial information about X and the seed Y) as well as a few additional "advice bits" about X , to produce a "reconstructed" guess for the complete string X that is correct with probability of the order of $(\varepsilon/m)^2$, where m is the length of Z .

Our security proof relies on an adaptation of the reconstruction paradigm to the case of quantum side information. This allows us to argue that any eavesdropper who is able to distinguish the final key Z generated by the users from a uniformly random string can be "bootstrapped" into a stronger eavesdropper who, given a few additional bits of advice, is able to guess the complete string of outputs K_A generated by Alice's device in the protocol with non-trivial probability. Note that this probability is of the same order as the security parameter ε . To see that it is a strong bound, note that it is much larger than the easy bound of inverse exponential in the extracted key length, which follows (using any strong extractor) from the correspondence between quantum conditional min-entropy and guessing probability. Formally introducing the quantum reconstruction paradigm would require notation that is beyond the scope of this article. For the expert, we mention that the main tool used in the analysis is the "pretty good measurement" of Hausladen and Wootters,¹⁷ which allows us to gain a handle on the Eavesdropper's global measurement.

6.3. Reduction to the guessing game

Recall that the raw key K_A is formed from the bits $K_A = a_1, \dots, a_{|C|}$ generated by Alice's device D_A in the key rounds $C \subseteq \{1, \dots, N\}$. Suppose for contradiction that, when an extractor Ext_c built according to the specifications of the reconstruction paradigm is applied to K_A , with a uniform choice of seed, the output Z is *not* indistinguishable from uniform: there is an attack for the adversary which uses all the quantum side information E available to the Eavesdropper at the end of protocol E1 to distinguish Z from uniform with some advantage ε . Our first step is to apply the quantum reconstruction paradigm to place ourselves in a stronger position: as argued in the previous section, it follows that there exists a "bootstrapped" adversary whom, using a combination of the side information E and a small number of additional advice bits, is able to produce a guess for the string K_A that is correct with probability of order $(\varepsilon/m)^2$, where m is the length of Z .

The second step in the analysis is to rule out such attacks. We achieve this by performing a reduction to the guessing game considered in Section 5. In order to present the reduction, it is convenient to abstract the present scenario as the following multi-round form of the guessing game. Alice gets an N -trit input x , and Bob an N -bit input y , chosen uniformly at random. They use their respective devices, D_A and D_B , sequentially to generate N -bit outputs a and b . A third player, Eve, is given x, y , and a small number of arbitrary bits of advice about Alice and Bob's outputs. (This includes the outputs in the test rounds and the bits of advice required for the reconstruction procedure.) Alice and Bob's devices are assumed to satisfy the CHSH game correlations on average, when tested on a randomly chosen fraction γ of the rounds, with non-negligible probability of order ε over an execution of the protocol. The goal is to show

that Eve is unable to produce a correct guess for the sub-string a_C , where C contains those rounds in which $(x_i, y_i) = (2, 0)$, with non-negligible probability.

To show this we perform a reduction to the single-round guessing game. We seek to identify a round $i_0 \in \{1, \dots, N\}$ such that the following holds. Let D_{A,i_0} and D_{B,i_0} denote the state of Alice and Bob's devices at the beginning of round i_0 , conditioned on all past events. This includes fixing values for all prior input and output bits, and making those values available to the three players. It should then be the case that (i) the devices D_{A,i_0} and D_{B,i_0} have a large success probability in the augmented CHSH game, when provided uniformly random inputs $(x, y) \in \{0, 1, 2\} \times \{0, 1\}$, and (ii) Eve has a strategy which allows her to produce a correct guess for the output a obtained by D_{A,i_0} , when given the input $x = 2$, with high probability. If both conditions can be shown to hold simultaneously, the single-round guessing lemma, Lemma 1, will give a contraction.

As explained earlier, the main difficulty in completing the reduction is that the adversary's attack in the original, multi-round protocol is based on a global measurement on her side information, which includes the classical information gleaned throughout the protocol, as well as her initial quantum state. The application of the quantum reconstruction paradigm dispenses with the need to deal with the adversary's quantum state at the expense of providing her with a small number of advice bits. We separate this information in two parts: first, the users' input strings (x, y) . Second, the user's outputs (a_B, b_B) in test rounds, and the advice bits $g(a_C)$. This second part of the information can be handled using a standard trick: instead of waiting for the bits to be available, we modify the adversary into one which makes a uniformly random choice for them. Using that the number of rounds used as checks is small compared to $|C|$, her success probability remains non-negligible.

The other part of the side information, the inputs (x, y) , cannot be eliminated in the same way, as it contains too many bits. We can split the inputs in two parts: those to be chosen before round i_0 , and those chosen after. The inputs chosen before round i_0 will be made publicly available as part of the specification of the devices D_{A,i_0} and D_{B,i_0} , so there is no need to worry about them. Regarding the inputs to be chosen after round i_0 , it is enough to argue that a "good" choice of inputs exist: indeed, the device's output in round i_0 cannot depend on inputs provided to the device after round i_0 ; here the sequential nature of the protocol is used in an important way.

Now that we have a single, fixed measurement for the adversary, with hard-wired values for (x, y) and the advice bits, we are finally in a position to apply the (classical!) Bayes' rule in order to identify the round i_0 satisfying conditions (i) and (ii). This provides the following straightforward implication:

$$\begin{aligned} \Pr(\text{Eve guesses } a_1 \dots a_{|C|} \text{ correctly}) &\geq \varepsilon \\ \Rightarrow \exists i_0, \Pr(\text{Eve guesses } a_{i_0} \mid \text{guesses for } & \\ a_1, \dots, a_{i_0-1} \text{ were correct}) &\geq 1 - \delta, \end{aligned} \quad (1)$$

for some small constant $\delta > 0$.

The conditioning implied by Bayes' rule, however, presents a last difficulty: it may introduce correlations between the state of the devices at the beginning of the i_0 -th round, and the inputs that the devices "expect" in that round. Indeed, the conditioning

is on the event that the adversary's guesses in rounds prior to i_0 are correct. The guesses in turn depend on the choice of inputs (x, y) that were "hard-wired" into the adversary as part of the reduction described above. Thus correctness of the adversary's guesses can bias the distribution of inputs in round i_0 , which is no longer uniform and may be jointly correlated with the state of the devices D_{A,i_0} and D_{B,i_0} . For example, it could be that the inputs in a fixed round i_0 are forced to a specific choice such as $(0, 0)$, making the guarantee (i) all but useless (if the inputs are fixed, it is easy to win the CHSH game). This difficulty is similar to one encountered in the analysis of the parallel repetition of multiplayer games. The standard approach for this problem was introduced in work of Raz²⁶ showing a parallel repetition theorem for the value of classical two-player games. The main idea consists in arguing that conditioning on an event with large enough probability cannot introduce strong correlations in all, or even most, coordinates of a distribution that is initially in product form. At a technical level, the analysis uses the chain rule for mutual information and Pinsker's inequality.

In our scenario a very similar approach can be employed to show that conditioning on the adversary's success, provided success happens with large enough probability, cannot bias the users' input distribution in most rounds by too much. This uses that initially (before the conditioning) the distribution is a product distribution, sampled independently in each of the large number of rounds of the protocol.

Unfortunately this statement is not sufficient for our purposes. We need to show, not only that the inputs to the devices in the round i_0 are close to uniformly random, but also that the post-measurement state of the devices (conditioned on the same success event) is not correlated to the users' choice of input. To show that this cannot happen we expand on Raz' technique by introducing a coding argument to deal with the quantum correlations. Assume for contradiction that the devices' state has a non-trivial correlation with their input (after a successful guess of the eavesdropper), and that this holds in most rounds. We show that such devices, including the eavesdropper's strategy, could be used to transmit classical information from Bob and the eavesdropper to Alice at a more efficient rate than is allowed by standard arguments from quantum information theory, thereby reaching a contradiction. For a further discussion of this step we refer to previous work of the authors.³⁵

7. OUTLOOK

From their origins in the EPR thought experiment, through their formulation in Bell's work and the CHSH test, to their use in device-independent cryptography, the non-local correlations of quantum mechanics have gone from undesirable paradox to operationally desirable certificates, not only of quantumness but also of randomness and privacy.^f

Our proof of security inscribes itself in a long sequence of works demonstrating progressively stronger uses of quantum

^f Note that the conditioning is performed jointly on an event involving Alice and Bob (the CHSH violation observed in previous rounds being sufficiently large) on the one hand, and Bob and Eve (Eve's guess being correct) on the other, so it can certainly introduce correlations across the whole input distribution.

correlations. Although the relevance of non-local correlations for cryptography was already pointed out by Ekert, the first concrete results considered the related task of randomness amplification, in which a single user, Alice, wishes to certify that two devices in her possession generate random bits, while using as few possible seed bits of randomness to test the device.

A protocol for randomness certification was first proposed in Colbeck,¹⁰ and the task was experimentally demonstrated in 2012.²⁵ Some of the ideas present in this paper were first formulated in the analysis of a protocol for exponential randomness expansion we introduced in previous work of the authors;³⁴ subsequently it was shown that even unbounded expansion is possible.¹¹

Subsequent to our work on quantum key distribution a second proof of security for DIQKD was put forward in Miller and Shi.²³ Neither proof achieves practical parameters, a gap recently filled by a third proof³ which establish security of a protocol with essentially optimal noise tolerance and key rate. Although the three proofs inevitably bear some similarity, they seem to rely on essentially different arguments; it remains a challenge to find a unifying framework for device-independence that would combine their respective advantages.

Recent progress in experiments^{15,18,30} place the implementation of protocols for entanglement-based quantum key distribution just around the corner. One may even start to see emerging the possibility for entanglement generation not only between single-qubit devices, as are needed for QKD, but small multi-qubit “quantum computers in the cloud” such as the one recently demonstrated by IBM. Computing devices sharing quantum entanglement may be able to implement tasks beyond quantum key distribution, from simple protocols such as, for example, quantum secret sharing or entanglement swapping, to the complex task of verifiable delegated computation. Although this remains a distant prospect, it is our hope that the techniques developed in this work will find far broader applicability in the future of quantum cryptography!

Acknowledgments

We thank the *Communications* editors and Gilles Brassard for useful feedback on an earlier version of this article.

Umesh Vazirani is supported by NSF Grant CCF-1410022, MURI Grant FA9550-18-1-0161 and a Vannevar Bush Faculty Fellowship. Thomas Vidick is supported by AFOSR YIP award number FA9550-16-1-0495, MURI Grant FA9550-18-1-0161, and a CIFAR Azrieli Global Scholar award.

References

- Acín, A., Brunner, N., Gisin, N., Massar, S., Pironio, S., Scarani, V. Device-independent security of quantum cryptography against collective attacks. *Phys. Rev. Lett.* 98, 230501 (2007).
- Acín, A., Gisin, N., Masanes, L. From Bell's theorem to secure quantum key distribution. *Phys. Rev. Lett.* 97, 120405 (2006).
- Arnon-Friedman, R., Renner, R., Vidick, T. Simple and tight device-independent security proofs. *arXiv preprint arXiv:1607.01797* (2016).
- Barrett, J., Colbeck, R., Kent, A. Unconditionally secure device-independent quantum key distribution with only two devices. *Technical report arXiv:1209.0435* (2012).
- Barrett, J., Hardy, L., Kent, A. No signaling and quantum key distribution. *Phys. Rev. Lett.* 95, 010503 (2005).
- Bell, J.S. On the Einstein-Podolsky-Rosen paradox. *Physics* 1 (1964), 195–200.
- Bennett, C., Brassard, G. Quantum cryptography: Public key distribution and coin tossing. In *Proceedings of the International Conference on Computers, Systems, and Signal Processing* (1984), 175–179.

- Brassard, G., Lütkenhaus, N., Mor, T., Sanders, B.C. Limitations on practical quantum cryptography. *Phys. Rev. Lett.* 85 (Aug 2000), 1330–1333.
- Clauser, J.F., Horne, M.A., Shimony, A., Holt, R.A. Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.* 23 (1969), 880–884.
- Colbeck, R. *Quantum and Relativistic Protocols for Secure Multi-Party Computation*. PhD thesis, Trinity College, University of Cambridge, Cambridge, UK, Nov. 2006.
- Coudron, M., Yuen, H. Infinite randomness expansion with a constant number of devices. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing* (ACM, 2014), 427–436.
- De, A., Portmann, C., Vidick, T., Renner, R. Trevisan's extractor in the presence of quantum side information. *SIAM J. Comp.* 41, 4 (2012), 915–940.
- Einstein, A., Podolsky, P., Rosen, N. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.* 47 (1935), 777–780.
- Ekert, A.K. Quantum cryptography based on Bell's theorem. *Phys. Rev. Lett.* 67 (1991), 661–663.
- Giustina, M., Versteegh, M.A., Wengerowsky, S., Handsteiner, J., Hochrainer, A., Phelan, K., Steinlechner, F., Kofler, J., Larsson, J.-A., Abellán, C., et al. Significant-loophole-free test of Bell's theorem with entangled photons. *Phys. Rev. Lett.* 115, 25 (2015), 250401.
- Goldwasser, S., Micali, S. Probabilistic encryption & how to play mental poker keeping secret all partial information. In *Proceedings of the fourteenth annual ACM symposium on Theory of computing* (ACM, 1982), 365–377.
- Hausladen, P., Wootters, W.K. A “pretty good” measurement for distinguishing quantum states. *J. Mod. Opt.* 41, 12 (1994), 2385–2390.
- Hensen, B., Bernien, H., Dréau, A.E., Reiserer, A., Kalb, N., Blok, M.S., Ruitenberg, J., Vermeulen, R.F., Schouten, R.N., Abellán, C., et al. Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres. *Nature* 526, 7575 (2015), 682–686.
- Kalai, Y.T., Raz, R., Rothblum, R.D. How to delegate computations: the power of no-signaling proofs. In *Proceedings of the forty-sixth annual ACM symposium on Theory of computing* (ACM, 2014), 485–494.
- Lydersen, L., Wiechers, C., Wittmann, C., Elser, D., Skaar, J., Makarov, V. Hacking commercial quantum cryptography systems by tailored bright illumination. *Nat. Photonics* 4, 10 (2010), 686–689.
- Mayers, D. Unconditional security in quantum cryptography. *J. ACM* 48, 3 (May 2001), 351–406.
- Mayers, D., Yao, A. Quantum cryptography with imperfect apparatus. In *Proceedings of the 39th Annual Symposium on Foundations of Computer Science, FOCS '98* (IEEE Computer Society, 1998), Washington, DC, USA, 503.
- Miller, C.A., Shi, Y. Robust protocols for securely expanding randomness and distributing keys using untrusted quantum devices. *J. ACM* 63, 4 (2016), 33.
- Pironio, S., Acín, A., Brunner, N., Gisin, N., Massar, S., Scarani, V. Device-independent quantum key distribution secure against collective attacks. *New J. Phys.* 11, 4 (2009), 045021.
- Pironio, S., Acín, A., Massar, S., De La Giroday, A.B., Matsukevich, D.N., Maunz, P., Olmschenk, S., Hayes, D., Luo, L., Manning, T.A., et al. Random numbers certified by Bell's theorem. *Nature* 464, 7291 (2010), 10.
- Raz, R. A parallel repetition theorem. *SIAM J. Comput.* 27 (1998), 763–803.
- Reichardt, B.W., Unger, F., Vazirani, U. Classical command of quantum systems. *Nature* 496, 7446 (2013), 456.
- Rivest, R.L., Shamir, A., Adleman, L. A method for obtaining digital signatures and public-key cryptosystems. *Comm. ACM* 21, 2 (1978), 120–126.
- Scarani, V., Gisin, N., Brunner, N., Masanes, L., Pironio, S., Acín, A. Secrecy extraction from no-signaling correlations. *Phys. Rev. A* 74, (Oct 2006), 042339.
- Shalm, L.K., Meyer-Scott, E., Christensen, B.G., Bierhorst, P., Wayne, M.A., Stevens, M.J., Gerrits, T., Glancy, S., Hamel, D.R., Allman, M.S., et al. Strong loophole-free test of local realism. *Phys. Rev. Lett.* 115, 25 (2015), 250402.
- Shor, P.W., Preskill, J. Simple proof of security of the BB84 quantum key distribution protocol. *Phys. Rev. Lett.* 85 (Jul 2000), 441–444.
- Tomamichel, M., Schaffner, C., Smith, A., Renner, R. Leftover hashing against quantum side information. *IEEE Transactions on Information Theory* 57, 8 (2011), 5524–5535.
- Trevisan, L. Extractors and pseudorandom generators. *J. ACM* 48 (July 2001), 860–879.
- Vazirani, U., Vidick, T. Certifiable quantum dice: Or, true random number generation secure against quantum adversaries. In *Proceedings of the 44th symposium on Theory of Computing, STOC '12* (ACM, 2011), 61–76. Also available as arXiv:1111.6054.
- Vazirani, U., Vidick, T. Fully device-independent quantum key distribution. *Phys. Rev. Lett.* 113, 14 (2014), 140501.
- Yin, J., Cao, Y., Li, Y.-H., Liao, S.-K., Zhang, L., Ren, J.-G., Cai, W.-Q., Liu, W.-Y., Li, B., Dai, H., et al. Satellite-based entanglement distribution over 1200 kilometers. *Science* 356, 6343 (2017), 1140–1144.

Umesh Vazirani (vazirani@eecs.berkeley.edu), Department of Computer Science, UC Berkeley, Berkeley, California, USA.

Thomas Vidick (vidick@cms.caltech.edu), Department of Computing and Mathematical Sciences, California Institute of Technology, Pasadena, California, USA.