

Full security of quantum key distribution from no-signaling constraints

Lluís Masanes

*H.H.Wills Physics Laboratory, University of Bristol, Tyndall Avenue,
Bristol BS8 1TL, U.K. email: ll.masanes@bristol.ac.uk*

Renato Renner and Matthias Christandl

Institute for Theoretical Physics, ETH-Zürich, CH-8093, Zürich, Switzerland

Andreas Winter

*Física Teòrica: Informació i Fenòmens Quàntics,
Universitat Autònoma de Barcelona, ES-08193 Bellaterra (Barcelona), Spain.
ICREA-Institució Catalana de Recerca i Estudis Avançats,
Pg. Lluis Companys 23, ES-08010 Barcelona, Spain.*

Jonathan Barrett

*Department of Computer Science, University of Oxford,
Parks Road, Oxford OX1 3QD, U.K.*

(Dated: October 20, 2018)

We analyze a cryptographic protocol for generating a distributed secret key from correlations that violate a Bell inequality by a sufficient amount, and prove its security against eavesdroppers, constrained only by the assumption that any information accessible to them must be compatible with the non-signaling principle. The claim holds with respect to the state-of-the-art security definition used in cryptography, known as universally-composable security. The non-signaling assumption only refers to the statistics of measurement outcomes depending on the choices of measurements; hence security is independent of the internal workings of the devices — they do not even need to follow the laws of quantum theory. This is relevant for practice as a correct and complete modeling of realistic devices is generally impossible. The techniques developed are general and can be applied to other Bell inequality-based protocols. In particular, we provide a scheme for estimating Bell-inequality violations when the samples are not independent and identically distributed.

I. INTRODUCTION

Quantum Key Distribution (QKD) [1, 2] is the task of generating a secret key such that the key gets known exclusively to two designated parties, in the following called *Alice* and *Bob*. In this work, we consider *entanglement-based QKD* [2], where Alice and Bob have access to a source of entanglement, and they can communicate over an authenticated¹ classical communication channel. Crucially, nothing is assumed about the source. In particular, the source may be fully controlled by an adversary, *Eve*, who may try to gain information about the generated key. The main idea behind entanglement-based QKD is that Alice and Bob, in a verification step, check whether the entanglement obtained from the source is sufficiently strong. If this is the case, they “distill” their key from the entanglement provided by the source. Otherwise, if the entanglement is found to be weak, they have to abort the protocol as security cannot be guaranteed.

The aim of this paper is to prove security of a class

of QKD protocols under minimal assumptions. In particular, our argument is *device-independent* [3, 4], which means that we do not make any assumptions about the internal workings of the source nor the quantum devices used by Alice and Bob. In fact, we do not even require that they work according to the laws of quantum mechanics. Instead, we only make *no-signaling assumptions*, which require that the components of the QKD scheme do not emit any undesired information [5]. This assumption can be met, for example, by perfectly isolating a large number of devices. We also note that, in order to make sense of the QKD problem, certain non-signaling assumptions are necessary: If the device processing the final key broadcasts it, secrecy obviously remains unachievable in any protocol.

In contrast to this, most standard security proofs [6–11] are device-dependent. This means that the security claim is only guaranteed to hold if Alice and Bob’s devices precisely meet a given specification. Consider, for example, an optical implementation, where a source distributes pairs of entangled photons, whose polarization is then measured by Alice and Bob. A device-dependent security proof would then require that the measurement outcomes depend in a specific way on the polarization degree of freedom, and are otherwise independent of any further properties of the photons (such

¹ An authenticated channel provides the guarantee to the receiver, say Bob, that the received messages has indeed be sent by the sender, Alice. However, such a channel does not guarantee any secrecy.

as their wave length or their arrival time) or any other parameter (such as the temperature of the device). This assumption—even if one tolerates some inaccuracies—is not only hard to meet, but also hard to verify. In fact, the assumption is not met by many existing practical realizations of QKD, as has recently been demonstrated in a series of hacking experiments (see, e.g., [12, 13]).

The mismatch between the theoretical specifications of the devices used for the security proofs and the actual practical implementations of these devices was already recognized in the late 90s. In particular, Mayers and Yao proposed the idea of *self-testing*, where the violation of Bell inequalities [14] is used to infer that the devices meet a given specification [15]. This approach has been taken further by Barrett, Hardy, and Kent, who proposed a scheme whose security is based solely on certain non-signaling assumptions [5], similarly to those used in this work. Their proof, however, only applied to an idealized setting with a noiseless source. Later, Acín, Gisin and Masanes showed that noise can be tolerated if one makes the assumption of *individual attacks*² [3]. We stress that these security claims, as well as the one presented in this work, do not rely on the correctness of quantum theory.

In a parallel line of research, device-independent security proofs have been developed which rely on the validity of quantum theory. While the first proofs of this type were restricted to a certain class of attacks, called *collective attacks*³ [16], this restriction could be relaxed recently to non-signaling constraints similar to those used here [17, 18]. Moreover, these recent results, like the one presented here, use a strong notion of security, as introduced in [19–21]. This guarantees *universal composability*, which means that the secret key generated by the protocol can safely be used in any application.

All results mentioned above rely on the no-signaling assumption. In our model, we treat each measurement as if it was carried out on a separate device, between which no signaling is allowed. In a practical setup—where Alice and Bob each have only one measurement device that is used repeatedly—this means that there should not be signaling between the individual uses of the devices. In particular, the non-signaling assumption is satisfied if the practical devices have no memory. This assumption has recently been relaxed in [22–24]. Although the protocol in [22] assumes that the adversary does not have a long term non-classical memory, [23] does not tolerate any noise, and the protocol in [24] relies on the validity of quantum theory.

Since our claims are supposed to hold independently of the correctness of quantum theory, we need a general

framework to specify the protocol, executed by Alice and Bob, as well as Eve’s attack strategy. We follow a standard approach [25–27] and describe the source of entanglement as a *random system*, which takes inputs X , Y , and Z , and produces outputs A , B , and E , accessible to Alice, Bob, and Eve, respectively. The idea is that the inputs specify the choices of measurements that can be applied to the (potentially) correlated systems, and the outputs correspond to the measurement outcomes. Crucially, we assume that this random system satisfies certain non-signaling constraints.

The paper is organized as follows. In Section II we provide an introduction for: non-signaling correlations, Bell inequalities and their relation to privacy, the notion of device-independent protocols, and the security definition that we use. In Section III we describe the protocol, state our main result, explain how to implement the protocol with quantum devices, and compare its performance with other protocols. The proofs of our results are provided in Section IV. The conclusions of our work are in Section V, and the Appendix contains a supplementary result that allows to adapt our protocol so that higher key rates can be achieved at the price of assuming the validity of quantum theory.

II. PRELIMINARIES

A. Non-signaling correlations

We use upper-case A to denote the random variable whose particular outcome is the corresponding lower-case a . We write the probability distribution for A as P_A , and the probability for $A = a$ as P_a , as well as $P_A(a)$. We use bold letters to denote strings of outcomes $\mathbf{a} = (a_1, \dots, a_N)$ or strings of random variables $\mathbf{A} = (A_1, \dots, A_N)$.

Alice and Bob share N pairs of physical systems, labeled by $n \in \{1, \dots, N\}$. Alice measures her n^{th} system with one of the M observables $X_n \in \{0, 1, \dots, M-1\}$, obtaining the outcome $A_n \in \{0, 1\}$. Analogously, Bob measures his n^{th} system with one of the $(M+1)$ observables $Y_n \in \{0, 1, \dots, M\}$ and obtains the outcome $B_n \in \{0, 1\}$. The chosen observables and their corresponding outcomes for the N pairs of systems are represented by the random variables \mathbf{A} , \mathbf{B} , \mathbf{X} , \mathbf{Y} , which are correlated according to the joint conditional probability distribution $P_{\mathbf{A}, \mathbf{B} | \mathbf{X}, \mathbf{Y}}$. The number $P_{\mathbf{a}, \mathbf{b} | \mathbf{x}, \mathbf{y}}$ is the probability of obtaining the strings of outcomes $\mathbf{a}, \mathbf{b} \in \{0, 1\}^N$ when measuring $\mathbf{x} \in \{0, \dots, M-1\}^N$ and $\mathbf{y} \in \{0, \dots, M\}^N$. The only assumption about this distribution is the following.

The non-signaling assumption: *The choice of observable for one system cannot modify the marginal distribution for the rest of the systems.*

More formally, we impose the following condition among

² An attack is called *individual* if the adversary gathers only classical information, obtained by individual measurements applied to single pairs emitted by the source of entanglement.

³ An attack is called *collective* if it is guaranteed that the individual particle pairs as received by Alice and Bob are independent and identically distributed [9].

any two sets of input/output pairs $\mathbf{I}_1, \mathbf{O}_1$ and $\mathbf{I}_2, \mathbf{O}_2$,

$$\sum_{\mathbf{o}_2} P_{\mathbf{o}_1, \mathbf{o}_2 | \mathbf{i}_1, \mathbf{i}_2} = \sum_{\mathbf{o}_2} P_{\mathbf{o}_1, \mathbf{o}_2 | \mathbf{i}_1, \mathbf{i}'_2} \quad (1)$$

for all $\mathbf{i}_2, \mathbf{i}'_2, \mathbf{o}_1, \mathbf{i}_1$. This condition could be enforced physically by using a different device for each measurement and isolating the devices from each other. In a more practical situation where the same device is used for subsequent measurements, the condition holds, for instance, if one assumes that the device has no memory.

The information available to the adversary Eve is modeled analogously: it is given by the input/output behavior of a system, which may be correlated to Alice and Bob's measurements. Specifically, Eve can choose an observable Z and obtains an outcome E . We assume that this, together with the public messages exchanged by Alice and Bob, is all information available to her. Note that we can without loss of generality assume that Eve only carries out this one measurement, for the sole assumption that we use in the security proof is that the global $(2N + 1)$ -partite distribution $P_{\mathbf{A}, \mathbf{B}, E | \mathbf{X}, \mathbf{Y}, Z}$ is a non-signaling one, but otherwise may be arbitrary.

B. Bell inequalities

A bipartite conditional distribution $P_{A, B | X, Y}$ is said to be local if it can be written as

$$P_{a, b | x, y}^{\text{local}} = \sum_v P_v P_{a | x, v} P_{b | y, v}, \quad (2)$$

for some probability distribution P_V and conditional probability distributions $P_{A | X, V}$ and $P_{B | Y, V}$. Local distributions can be generated by shared randomness (denoted V above) between the parties, plus local operations. In other words, local distributions can be generated with classical resources. A distribution $P_{A, B | X, Y}$ which cannot be written as (2) is said to be non-local. Non-local correlations are the resource consumed in our secret key distribution protocol.

By definition, Bell inequalities [14, 28, 29] are satisfied by all local distributions (2). In this paper, we concentrate on the Braunstein-Caves Bell inequality [29], or BC-inequality for short. This inequality is often stated using a different notation (not to be further used in this work), where A_x, B_y denote the random variables A, B conditioned on $X = x, Y = y$, so that it reads

$$\begin{aligned} & \langle A_1 \oplus B_1 \rangle + \langle B_1 \oplus A_2 \rangle + \langle A_2 \oplus B_2 \rangle + \dots \\ & + \langle A_M \oplus B_M \rangle + \langle B_M \oplus A_1 \oplus 1 \rangle \geq 1, \end{aligned} \quad (3)$$

where \oplus is the sum modulo 2. For our purposes it is convenient to write the BC-inequality for a given conditional distribution $P_{A, B | X, Y}$ as the expectation of the random variable

$$W = (A \oplus B \oplus \delta_X^0 \delta_Y^{M-1}) \quad (4)$$

over $P_{a, b, x, y} = P_{a, b | x, y} Q_{x, y}$, where

$$Q_{x, y} = \begin{cases} \frac{1}{2M} & \text{if } (x - y \bmod M) \in \{0, 1\} \\ 0 & \text{otherwise} \end{cases}. \quad (5)$$

The BC-inequality for M observables (3) can be written as

$$\langle W \rangle \geq \frac{1}{2M}. \quad (6)$$

As mentioned above, any local distribution (2) satisfies (6). The largest violation $\langle W \rangle = 0$ can be reached for certain non-signaling distributions, but cannot be reached within quantum theory. The largest quantum violation is obtained with the EPR state $|\phi\rangle = (|0\rangle|0\rangle + |1\rangle|1\rangle)/\sqrt{2}$ [30, 31], with the measurements specified in FIG. 1, reaching the value

$$\langle W \rangle_{|\phi\rangle} = \sin^2\left(\frac{\pi}{4M}\right). \quad (7)$$

Note that when increasing M , the quantum violation tends to zero, the maximal one.

For $M = 2$ the BC-inequality is equivalent to the famous CHSH-inequality [28], with its well-known maximal quantum violation of $\langle W \rangle = \frac{1}{2} - \frac{1}{2\sqrt{2}} \approx 0.15$ due to Tsirelson [32].

C. Guessing probability and Bell violation

Suppose that Eve is correlated with Alice and Bob through the global non-signaling distribution $P_{A, B, E | X, Y, Z}$. If Alice measures $X = 0$ and obtains the outcome A , then we can quantify the knowledge that Eve has about A by the optimal guessing probability

$$\mathcal{P}_{\text{guess}}(A|E) = \max_z \sum_e \max_a P_{A, E | X, Z}(a, e, 0, z). \quad (8)$$

If $\mathcal{P}_{\text{guess}}(A|E) = 1$ then Eve knows A with certainty. If $\mathcal{P}_{\text{guess}}(A|E) = 1/2$ then Eve is completely ignorant about the value of A . In [33] it was shown that the knowledge that Eve has about A can be bounded by the amount of non-locality present in Alice's and Bob's marginal distribution:

$$\mathcal{P}_{\text{guess}}(A|E) \leq 1/2 + M \langle W \rangle. \quad (9)$$

If the marginal for the honest parties $P_{A, B | X, Y}$ violates the BC-inequality (6), then according to (9), the probability that Eve guesses correctly is smaller than one. This is one manifestation of the monogamy of non-local correlations [26, 27]. In Appendix A, inequality (9) is generalized to the case of more than one pair of systems.

D. Device-independent QKD

Inequality (9) allows to bound Eve's knowledge about A in terms of the statistics of A, B, X, Y , regardless of

how the correlations $P_{A,B|X,Y}$ are generated. In particular, the privacy of A is independent of the functioning of the device used to generate A . Even if the devices are maliciously designed by Eve, and even if the devices violate quantum theory, the security of our protocol is not compromised.

The only assumption that we make on the devices is that they satisfy the no-signaling constraints (1). This could be enforced by performing each of the $2N$ measurements by Alice and Bob in a separate isolated device. Clearly, this approach, though theoretically possible, would be extremely costly in practice. A cheaper possibility—actually the one employed by all existing experiments—is that Alice and Bob each use one single device repeatedly for the different measurements. The constraints (1) then mean that there should be no signaling between the individual uses of the devices. This would be the case, for instance, if the devices had no memory. While such a no-memory assumption may be hard to guarantee in practice, it is still considerably weaker than the assumption that the devices can be modeled completely, which is necessary in standard (non device-independent) cryptography.

E. Security definition

Our key generation protocol starts from correlations that violate the Bell inequality (6). We model these initial correlations by a non-signaling distribution $P_{\mathbf{A},\mathbf{B},E|\mathbf{X},\mathbf{Y},Z}$. This distribution is a priori unknown and may have been chosen by the adversary. In other words, all our security claims are supposed to hold for any possible initial non-signaling distribution $P_{\mathbf{A},\mathbf{B},E|\mathbf{X},\mathbf{Y},Z}$. Furthermore, Alice and Bob have access to a public authenticated communication channel. That is, all information sent through this channel will be available to, but cannot be altered by Eve.

The key distribution protocol specifies by a sequence of instructions for Alice and Bob. In each of the protocol steps, Alice and Bob either access the correlated data, perform local calculations, or exchange messages over the public channel. In the final step of the protocol, Alice and Bob generate the keys K_A and K_B taking values on $\{0, 1\}^{N_s}$, respectively.

We say that a protocol is secure if the resulting distribution is indistinguishable from an ideal one. Suppose that at the end of the protocol all the relevant information is characterized by a distribution $P_{K_A K_B, T, E|Z}^{\text{real}}$, where T is a transcript of the communication, containing all messages exchanged between Alice and Bob through the authenticated channel (note that T is accessible to Eve). An ideal QKD protocol produces the distribution

$$P_{k_A k_B, t, e|z}^{\text{ideal}} = 2^{-N_s} \delta_{k_A}^{k_B} P_{t, e|z}^{\text{real}}, \quad (10)$$

where $P_{t, e|z}^{\text{real}}$ are the values of the marginal $P_{TE|Z}^{\text{real}}$ derived from the real distribution $P_{K_A K_B T E|Z}^{\text{real}}$. Note that

according to (10), the two versions of the secret key, K_A and K_B , are identical and uniformly distributed, independently of the values taken by T, E, Z . We say that a protocol is secure if the quantity

$$\sum_{k_A, k_B, t} \max_z \sum_e \left| P_{k_A k_B, t, e|z}^{\text{real}} - P_{k_A k_B, t, e|z}^{\text{ideal}} \right| \quad (11)$$

can be made arbitrarily small as N grows. This is the strongest notion of security, and it is called *universally composable security* [19, 20, 34, 35]. It is often the case that the secret key generated by a QKD protocol is used as an ingredient for another cryptographic task. The above security definition warrants that the composed scheme that uses a secure key distribution protocol as a component is as secure as if an ideal secret key (10) was used instead (see [21] for more details).

III. SETUP AND RESULTS

A. Description of the protocol

In what follows we describe a family of protocols parametrized by the number of settings M in the BC-inequality, and a probability $\gamma \in (0, 1)$. The role of γ is explained next. In Section III C we illustrate how to find the optimal value for these parameters. This family of protocols is similar to those introduced in [5, 38].

1. Distribution and measurements. Alice and Bob are given N pairs of systems. Alice generates the random bits $\mathbf{I} = (I_1, \dots, I_N)$ independently and with identical distribution: $P_{I_n}(0) = 1 - \gamma, P_{I_n}(1) = \gamma$. Analogously, Bob generates the random bits $\mathbf{J} = (J_1, \dots, J_N)$ independently and with identical distribution $P_{J_n} = P_{I_n}$. Pairs such that $I_n = J_n = 0$ are used to generate the raw key, and pairs such that $I_n = J_n = 1$ are used to estimate how much non-locality Alice and Bob share. For each $n \in \{1, \dots, N\}$, if $I_n = 0$ Alice measures her n^{th} system with $X_n = 0$, if $I_n = 1$ she measures it with X_n chosen uniformly on $\{0, \dots, M-1\}$, if $J_n = 0$ Bob measures his n^{th} system with $Y_n = M$, if $J_n = 1$ he measures it with Y_n chosen uniformly on $\{0, \dots, M-1\}$.

2. Estimation of non-locality. Alice and Bob announce \mathbf{I}, \mathbf{J} publicly as well as the tuples (A_n, B_n, X_n, Y_n) for the values of n where $I_n = J_n = 1$. The subset of pairs

$$\mathcal{N}_e = \left\{ n \in \{1, \dots, N\} \mid I_n = J_n = 1 \right. \\ \left. \text{and } (X_n - Y_n \bmod M) \in \{0, 1\} \right\}. \quad (12)$$

is used to compute the average value for the BC-inequality

$$\bar{W} = \frac{1}{N_e} \sum_{n \in \mathcal{N}_e} (A_n \oplus B_n \oplus \delta_{X_n}^0 \delta_{Y_n}^{M-1}), \quad (13)$$

where $N_e = |\mathcal{N}_e|$. \bar{W} thus corresponds to the average of the variables W_n defined by (4) with $n \in \mathcal{N}_e$. Note that after post-selecting on the pairs with $n \in \mathcal{N}_e$ the random variables X_n, Y_n follow the distribution $Q_{X,Y}$ defined by (5), which allows to identify W_n with the BC-inequality for the pair with index n .

The number of estimated systems is $N_e \approx 2N\gamma^2/M$ with high probability. Here and in the rest of the paper the symbol \approx denotes equality up to subleading terms. As we will see, the asymptotic efficiency of the protocol does not depend on the subleading terms. The outcomes of the pairs in the set

$$\mathcal{N}_r = \{n \in \{1, \dots, N\} \mid I_n = J_n = 0\}, \quad (14)$$

have not been published, and are denoted by $\mathbf{A}_r, \mathbf{B}_r$. These are the raw keys obtained by Alice and Bob, respectively. We denote their length by $N_r = |\mathcal{N}_r| \approx (1 - \gamma)^2 N$.

3. Error correction. Alice publishes N_c bits of information about her raw key $C = f(\mathbf{A}_r)$, which Bob uses for correcting the errors in his raw key: $(\mathbf{B}_r, C) \mapsto \mathbf{B}'_r \approx \mathbf{A}_r$. Any error-correction method, or equivalently any function $f : \{0, 1\}^{N_r} \rightarrow \{0, 1\}^{N_c}$, can be inserted here, as long as the probability that $\mathbf{B}'_r \neq \mathbf{A}_r$ vanishes as N grows. It follows from classical information theory (see [35] for more details) that error correction can be achieved asymptotically with

$$N_c \approx N_r h(\lambda), \quad (15)$$

where λ is the relative frequency of $B_n \neq A_n$ for all $n \in \mathcal{N}_r$, and

$$h(\lambda) = -\lambda \log_2 \lambda - (1 - \lambda) \log_2 (1 - \lambda) \quad (16)$$

is the binary Shannon entropy.

4. Privacy amplification. Alice chooses at random a function $G : \{0, 1\}^{N_r} \rightarrow \{0, 1\}^{N_s}$ from a set of two-universal hash functions (see Definition 5 or [41]) with output length

$$N_s(\bar{w}) = \max \left\{ 0, \max_{\theta \in [0, 1]} \left[2N_e D(\bar{w} \parallel \theta) - 2N_r \log_2 (1/2 + M\theta) \right] - N_r - N_c - 2 \log_2 (8N_e N / \epsilon) \right\}, \quad (17)$$

where the binary relative entropy is defined as

$$D(\theta_1 \parallel \theta_2) = \theta_1 \log_2 \frac{\theta_1}{\theta_2} + (1 - \theta_1) \log_2 \frac{1 - \theta_1}{1 - \theta_2}, \quad (18)$$

and \bar{w} is the observed value of the random variable (13). If $N_s(\bar{w}) > 0$ then Alice and Bob respectively compute $K_A = G(\mathbf{A}_r)$ and $K_B = G(\mathbf{B}'_r)$, which constitute their versions of the final secret key. We stress that the hash function G is chosen at random and independently of any other information. The first maximization in (17) avoids a negative length for the secret key, which obviously does not have any meaning. If $N_s(\bar{w}) = 0$ then Alice and Bob write $K_A = K_B = \perp$, which means that the protocol has not produced any secret key.

B. Main Results

The above protocol can be seen as a process which transforms process which transforms the initial distribution $P_{\mathbf{A}, \mathbf{B}, E | \mathbf{X}, \mathbf{Y}, Z}$ into the final distribution $P_{K_A, K_B, \bar{W}, T, E | Z}$, where

$$T = \left[\mathbf{I}, \mathbf{J}, C, G, (A_n, B_n, X_n, Y_n) \forall n \in \mathcal{N}_e \right] \quad (19)$$

is all the information that has been published by the honest parties. We prove that for any initial distribution $P_{\mathbf{A}, \mathbf{B}, E | \mathbf{X}, \mathbf{Y}, Z}$, the resulting distribution $P_{K_A, K_B, \bar{W}, T, E | Z}$ satisfies

$$\sum_{k_A, k_B, \bar{w}, t} \max_z \sum_e \left| P_{k_A, k_B, \bar{w}, t, e | z} - 2^{-N_s(\bar{w})} \delta_{k_A}^{k_B} P_{\bar{w}, t, e | z} \right| \leq \epsilon + 2\epsilon_{\text{erco}}, \quad (20)$$

where ϵ_{erco} is an upper-bound for the error probability of the error correction scheme. This implies that the actual key generated by the protocol has at most distance $\epsilon + 2\epsilon_{\text{erco}}$ from a perfectly secure key (see also (11) above). Note that the parameter ϵ can be controlled by the honest parties when adjusting the length of the final secret key (17).

By setting ϵ and ϵ_{erco} to sufficiently small values, the honest parties can be confident of the fact that the secret key generated by the protocol is indistinguishable from an ideal secret key (10). This implies that the protocol is secure according to the strongest notion of security, the so called universally-composable security [21, 34] (see Section II E).

The efficiency of a key distribution scheme is quantified by the asymptotic secret key rate. This is defined as the ratio N_s/N in the limit $N \rightarrow \infty$, where N_s is the number of perfect secret bits obtained and N is the number of pairs of systems consumed. Using (17) and (15) we obtain the secret key rate of our protocol:

$$\lim_{N \rightarrow \infty} \frac{N_s}{N} = -(1 - \gamma)^2 [1 + h(\lambda)] + \max_{\theta \in [0, 1]} \left[\frac{4\gamma^2}{M} D(\bar{w} \parallel \theta) - 2(1 - \gamma)^2 \log_2 \left(\frac{1}{2} + M\theta \right) \right] \quad (21)$$

It is understood that if the above quantity is negative the secret key rate is zero.

C. Implementation of the protocol with quantum devices

Here we explain how to implement the protocol with quantum-mechanical devices, i.e., if the initial correlations are generated by measurements on an entangled quantum state. We stress that this is not a part of the proof that the resulting key is secret (as the secrecy claim must hold for any possible correlations). However, it is necessary to argue that, if the adversary is passive then

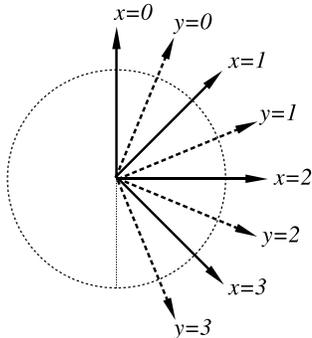


FIG. 1: Location in the equator of the Bloch sphere of the observables for $M = 4$.

the protocol actually generates a key (in particular, it should not abort), and to calculate the rate at which it does so—thereby allowing a comparison to other protocols.

Suppose Alice and Bob share many copies of the noisy EPR state

$$\rho = (1 - \xi) \Phi + \xi \frac{\mathbb{I}}{4}, \quad (22)$$

where $\xi \in [0, 1]$ is the fraction of noise, Φ is the projector onto the EPR state $|\phi\rangle = (|0\rangle|0\rangle + |1\rangle|1\rangle)/\sqrt{2}$, and $\mathbb{I}/4$ is the maximally noisy state. They perform the measurements in the following orthogonal basis. The observable $x \in \{0, \dots, M-1\}$ for Alice has eigenvectors

$$|0\rangle \pm e^{i\pi x/M} |1\rangle, \quad (23)$$

the observable $y \in \{0, \dots, M-1\}$ for Bob has eigenvectors

$$|0\rangle \pm e^{i\pi(y+1/2)/M} |1\rangle, \quad (24)$$

and the observable $y = M$ for Bob has eigenvectors

$$|0\rangle \pm |1\rangle. \quad (25)$$

Note that, while Alice has M observables, Bob has $M+1$ observables, and that $y = M$ is the same observable as Alice's $x = 0$. In the Bloch sphere, these observables correspond to the directions represented in FIG. 1. The observables $x, y \in \{0, \dots, M-1\}$ are the ones used to obtain large violations of the BC-inequality [29]. The observables $x = 0, y = M$ maximize the correlation between Alice and Bob, and hence, are used to generate the raw key. For $M = 2$, this protocol is essentially equivalent to Ekert's original protocol [2].

Using Equation (7) it is straightforward to extend the expectation of the Bell inequality violation (6) to the case where noise is added to the EPR state (22):

$$\langle W \rangle_\rho = (1 - \xi) \sin^2\left(\frac{\pi}{4M}\right) + \xi \frac{1}{2}. \quad (26)$$

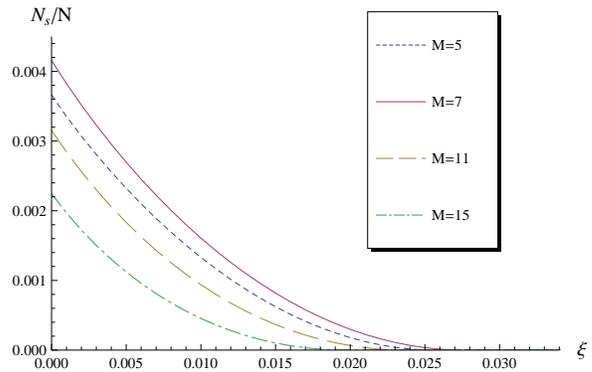


FIG. 2: The secret key rate N_s/N as a function of the noise ξ is plotted for different values of M . The parameter γ is optimized for every value of ξ . Note that in our protocol, γ tends to a non-zero value asymptotically.

This is the value taken by \bar{W} in the large N limit. Substituting this in (17) we obtain the secret key rate, which is plotted as a function of the noise ξ in Fig. 2. The value of the parameter γ is numerically optimized, such that, for each value of the noise ξ the secret key rate is maximal.

In Fig. 2 one can see that the secret key rate for $M = 2$ (and also $M = 3$) is zero, hence the security of the CHSH-protocol [3] against general attacks remains an open question. We have numerically seen that the optimal protocol is the one with $M = 7$.

D. Comparison with other results

In this section we make a survey of some proofs for device-independent security in QKD. The first division that we consider is whether the adversary is totally unrestricted [5, 17, 18, 22–24], or is constrained to perform the so called individual or collective attacks [3, 4, 16, 38]. The second division concerns whether the security relies on the validity of quantum theory [4, 16–18, 24] or not [3, 5, 22, 23, 38]. Third division, the devices are assumed to be memoryless [3, 5, 16–18, 38], or not [22–24]. Fourth division, the protocol tolerates certain degree of noise [3, 4, 16–18, 22, 24, 38], or not [5, 23]. Fifth division, the adversary is allowed to have a long-term non-classical memory [3–5, 16–18, 23, 24, 38], or not [22].

The advantages of our result are:

1. The adversary is totally unrestricted in the sense that no assumption is made about the structure of the global distribution (like in individual or collective attacks).
2. The security of our protocol does not rely on the validity of quantum theory.
3. The adversary is allowed to have a long-term non-classical (and non-quantum) memory.

The disadvantages of our result are:

1. The security of our protocol relies on the measuring devices not having an internal memory.
2. Our protocol tolerates certain degree of noise, but only a small amount $\xi = 2.7\%$ (compare this with $\xi = 11\%$ in [22]).
3. The secret key rate achieved by our protocol is very low. For example, in the noiseless case $\xi = 0$, our protocol gives $N_s/N = .004$ while the protocols of [3, 4, 16–18, 22] obtain $N_s/N = 1$.

IV. SECURITY PROOF

A. Properties of symmetric distributions

The results derived in this section provide tools for estimating properties of symmetric distributions without resorting to any de Finetti-like theorem. They are motivated by recent quantum analogues [39] (sometimes termed *postselection techniques* [40]) and may be of independent interest.

We use calligraphic letter \mathcal{V} to denote the alphabet of values for the corresponding random variable V , that is $v \in \mathcal{V}$. We use bold letters to denote strings of variables $\mathbf{v} = (v_1, \dots, v_N) \in \mathcal{V}^N$ or random variables $\mathbf{V} = (V_1, \dots, V_N)$. We say that a distribution $P_{\mathbf{V}}$ is symmetric if $P_{\mathbf{V}}(v_1, \dots, v_N) = P_{\mathbf{V}}(v_{\pi(1)}, \dots, v_{\pi(N)})$ for any permutation $\pi : \{1, \dots, N\} \rightarrow \{1, \dots, N\}$.

Definition 1 Given a string $\mathbf{v} = (v_1, \dots, v_N) \in \mathcal{V}^N$ we define its corresponding frequency distribution $q = \text{freq}(\mathbf{v})$ as

$$q(v) = \frac{|\{n : v_n = v\}|}{N}, \quad \forall v \in \mathcal{V}. \quad (27)$$

This function naturally extends to sets $\mathcal{Q} = \text{freq}(\mathcal{V}^N)$, and random variables $Q = \text{freq}(\mathbf{V})$.

For any \mathbf{v} , the frequency $q = \text{freq}(\mathbf{v})$ is a probability distribution for the random variable V , but it has the specific feature that it only takes values on the set $\{\frac{n}{N} : n = 0, \dots, N\}$. \mathcal{Q} is the set of all possible frequencies, whose cardinality can be bounded as

$$|\mathcal{Q}| \leq (N+1)^{|\mathcal{V}|-1}. \quad (28)$$

For what follows, it is convenient to define a particular kind of probability distributions for \mathbf{V} : *the distribution with well-defined frequency* $q \in \mathcal{Q}$, denoted $P_{\mathbf{V}|q}$, is the uniform distribution over all strings $\mathbf{v} \in \mathcal{V}^N$ such that $\text{freq}(\mathbf{v}) = q$. Another important kind of symmetric distributions are the *i.i.d. distributions*, representing independent and identically-distributed random variables V_1, \dots, V_N . Borrowing notation from quantum information theory we write $P_{\mathbf{V}}^{\otimes N}$ for such a distribution. If

$P_V(v) < 1$ for all v , then the i.i.d. distribution $P_{\mathbf{V}}^{\otimes N}$ does not have a well-defined frequency. However, any symmetric distribution $P_{\mathbf{V}}^{\text{sym}}$, including the i.i.d. ones, can be written as a mixture of distributions with well-defined frequency,

$$P_{\mathbf{V}}^{\text{sym}} = \sum_{q \in \mathcal{Q}} P_Q(q) P_{\mathbf{V}|q}, \quad (29)$$

$$Q = \text{freq}(\mathbf{V}). \quad (30)$$

These two equalities establish a one-to-one correspondence between Q and \mathbf{V} , for symmetric distributions. In Lemma 3 we show that, in a sense, general symmetric distributions are similar to i.i.d. distributions. But before, we need the following technical result.

Lemma 2 Let the probability distribution P_V take values on the set $\{\frac{n}{N} : n = 0, \dots, N\}$, and let $\mathbf{V} = (V_1, \dots, V_N)$ be distributed according to $P_{\mathbf{V}}^{\otimes N}$. Then the probability distribution P_Q for $Q = \text{freq}(\mathbf{V})$ takes its maximum at $Q = P_V$, that is,

$$P_Q(P_V) = \max_{q \in \mathcal{Q}} P_Q(q). \quad (31)$$

Proof We show that for any $q \in \mathcal{Q}$ with $q \neq P_V$ there exists $q' \in \mathcal{Q}$ such that $P_Q(q') > P_Q(q)$. Thus let $q \in \mathcal{Q}$ be fixed such that $q \neq P_V$. We call the *support* of q : the set of values v such that $q(v) > 0$. If the support of q is not contained in the support of P_V then $P_Q(q) = 0$. We can thus without loss of generality assume that the alphabet of V , denoted \mathcal{V} , is the support of P_V , that is, $P_V(v) > 0$ for all $v \in \mathcal{V}$. For any $v \in \mathcal{V}$ define

$$d(v) = q(v) - P_V(v).$$

Furthermore, let v_{\min} and v_{\max} be defined by

$$\begin{aligned} d(v_{\min}) &= \min_v d(v) \\ d(v_{\max}) &= \max_v d(v) \end{aligned}.$$

Because $q \neq P_V$ and the assumption of the lemma, $d(v_{\min}) \leq -1/N$ and $d(v_{\max}) \geq 1/N$. Let us define $q' \in \mathcal{Q}$ as

$$q'(v) = \begin{cases} q(v) + \frac{1}{N} & \text{if } v = v_{\min} \\ q(v) - \frac{1}{N} & \text{if } v = v_{\max} \\ q(v) & \text{otherwise.} \end{cases}$$

From the two inequalities above we have

$$\begin{aligned} q'(v_{\min}) &\leq P_V(v_{\min}) \\ q'(v_{\max}) &\geq P_V(v_{\max}) \end{aligned}. \quad (32)$$

Using the identity

$$P_Q(q) = \frac{N! \prod_v P_V(v)^{q(v)N}}{\prod_v (q(v)N)!}$$

we find

$$\frac{P_Q(q')}{P_Q(q)} = \frac{P_V(v_{\min})(q'(v_{\max}) + \frac{1}{N})}{P_V(v_{\max})q'(v_{\min})} > \frac{P_V(v_{\min}) q'(v_{\max})}{P_V(v_{\max}) q'(v_{\min})}$$

(note that the terms in the denominator cannot be zero). By (32), the right-hand side cannot be smaller than 1, which concludes the proof. \square

Lemma 3 *If there is a function $t : \mathcal{V}^N \rightarrow \mathbb{R}$ and $\epsilon > 0$ such that for any (single-copy) distribution P_V the bound*

$$\sum_{\mathbf{v}} P_V^{\otimes N}(\mathbf{v}) t(\mathbf{v}) \leq \epsilon \quad (33)$$

holds, then for any symmetric distribution $P_{\mathbf{V}}^{\text{sym}}$ we have

$$\sum_{\mathbf{v}} P_{\mathbf{V}}^{\text{sym}}(\mathbf{v}) t(\mathbf{v}) \leq |\mathcal{Q}| \epsilon. \quad (34)$$

Proof Let us first prove (34) for distributions with well-defined frequency $P_{\mathbf{V}}^{\text{sym}} = P_{\mathbf{V}|q}$, for all $q \in \mathcal{Q}$. Since any $q \in \mathcal{Q}$ is a (single-copy) distribution for V , the premise (33) applies to it:

$$\sum_{\mathbf{v}} q^{\otimes N}(\mathbf{v}) t(\mathbf{v}) \leq \epsilon \quad (35)$$

Using the decomposition (29), we know that there is a random variable Q' such that $\sum_{q' \in \mathcal{Q}} P_{Q'}(q') P_{\mathbf{V}|q'} = q^{\otimes N}$, and then

$$\sum_{\mathbf{v}} \sum_{q' \in \mathcal{Q}} P_{Q'}(q') P_{\mathbf{V}|q'}(\mathbf{v}) t(\mathbf{v}) \leq \epsilon. \quad (36)$$

In Lemma 2 it is shown that the distribution $P_{Q'}(q')$ reaches the maximum at $q' = q$, which implies $P_{Q'}(q) \geq 1/|\mathcal{Q}|$. Then

$$\begin{aligned} & \sum_{\mathbf{v}} P_{\mathbf{V}|q}(\mathbf{v}) t(\mathbf{v}) \\ & \leq |\mathcal{Q}| P_{Q'}(q) \sum_{\mathbf{v}} P_{\mathbf{V}|q}(\mathbf{v}) t(\mathbf{v}) \leq |\mathcal{Q}| \epsilon, \end{aligned} \quad (37)$$

where the last inequality follows from (36). Finally, we prove (34) by applying the bound (37) to each term in (29). \square

An equivalent way to write the above result for the case $\mathcal{V} = \{0, 1\}$, which will be useful later, is the following. For any symmetric distribution $P_{\mathbf{V}}^{\text{sym}}$ and any function t we have

$$\sum_{\mathbf{v}} P_{\mathbf{V}}^{\text{sym}}(\mathbf{v}) t(\mathbf{v}) \leq (N+1) \max_{P_V} \sum_{\mathbf{v}} P_V^{\otimes N}(\mathbf{v}) t(\mathbf{v}) \quad (38)$$

where the maximization is over single-copy distributions for V .

B. Properties of non-signaling distributions

For the following presentation it is useful to introduce some additional notation. We represent single-pair distributions $P_{A,B|X,Y}$ as vectors with components arranged

in the following way

$$P_{A,B|X,Y} = \quad (39)$$

$P(0, 0 0, 0)$	$P(0, 1 0, 0)$...	$P(0, 0 0, M-1)$
$P(1, 0 0, 0)$	$P(1, 1 0, 0)$		
\vdots		\ddots	\vdots
$P(0, 0 M-1, 0)$...		$P(0, 0 M-1, M-1)$

Define the following two vectors (which are not probability distributions)

$$\mu = \frac{1}{4M} \begin{bmatrix} 1 & 1 & & & 1 & 1 \\ 1 & 1 & & & 1 & 1 \\ 1 & 1 & 1 & 1 & & \\ 1 & 1 & 1 & 1 & & \\ & & \ddots & \ddots & & \\ & & & & 1 & 1 & 1 & 1 \\ & & & & 1 & 1 & 1 & 1 \end{bmatrix}, \quad (40)$$

$$\nu = \frac{1}{2} \begin{bmatrix} 0 & 1 & & & 1 & 0 \\ -1 & 0 & & & 0 & -1 \\ 0 & -1 & 0 & 1 & & \\ 1 & 0 & -1 & 0 & & \\ & & \ddots & \ddots & & \\ & & & & 0 & -1 & 0 & 1 \\ & & & & 1 & 0 & -1 & 0 \end{bmatrix}, \quad (41)$$

where empty boxes have to be understood as having zeros

$$\square = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad (42)$$

and ellipsis between two identical boxes have to be understood as an arbitrarily large sequence of identical boxes. From now on, the absolute value of a vector means component-wise absolute value. For example

$$|\nu| = \frac{1}{2} \begin{bmatrix} 0 & 1 & & & 1 & 0 \\ 1 & 0 & & & 0 & 1 \\ 0 & 1 & 0 & 1 & & \\ 1 & 0 & 1 & 0 & & \\ & & \ddots & \ddots & & \\ & & & & 0 & 1 & 0 & 1 \\ & & & & 1 & 0 & 1 & 0 \end{bmatrix}.$$

Also, an inequality " \preceq " between two vectors means components-wise inequality " \leq ". For example $\nu \preceq |\nu|$. Define the vectors

$$\beta_a = \mu + (-1)^a \nu, \quad (43)$$

$$\beta = \mu + |\nu|. \quad (44)$$

One can check that the Braunstein-Caves Bell inequality, defined in (4) and (6), can be written as

$$\beta \cdot P_{A,B|X,Y} = \frac{1}{2} + M\langle W \rangle \geq 1. \quad (45)$$

Above, the symbol “ \cdot ” represents the scalar product between the vectors β and $P_{A,B|X,Y}$.

Lemma 4 *If $P_{\mathbf{A},\mathbf{B}|\mathbf{X},\mathbf{Y}}$ is an arbitrary $2N$ -partite non-signaling distribution then for any $\mathbf{a} = (a_1, \dots, a_N)$ we have*

$$P_{\mathbf{A}|\mathbf{X}}(\mathbf{a}, \mathbf{0}) = \left(\bigotimes_{n=1}^N \beta_{a_n} \right) \cdot P_{\mathbf{A},\mathbf{B}|\mathbf{X},\mathbf{Y}}, \quad (46)$$

where $\mathbf{0} = (0, \dots, 0)$.

Proof: Let us first consider the bound (46) for one pair of systems ($N = 1$). The non-signaling constraint $P_{A|X,Y}(0,0,0) = P_{A|X,Y}(0,0,M-1)$ can also be expressed as the scalar product

$$\begin{array}{|c|c|c|c|} \hline -1 & -1 & & 1 & 1 \\ \hline 0 & 0 & & 0 & 0 \\ \hline & & & & \\ \hline & & \ddots & & \\ \hline & & & & \\ \hline & & & & \\ \hline \end{array} \cdot P_{A,B|X,Y} = 0$$

and the non-signaling constraint $P_{B|X,Y}(0,0,0) = P_{B|X,Y}(0,1,0)$ can be expressed as

$$\begin{array}{|c|c|c|c|} \hline -1 & 0 & & \\ \hline -1 & 0 & & \\ \hline 1 & 0 & & \\ \hline 1 & 0 & & \\ \hline & & \ddots & \\ \hline & & & \\ \hline & & & \\ \hline & & & \\ \hline \end{array} \cdot P_{A,B|X,Y} = 0.$$

The remaining non-signaling constraints can be written in an analogous fashion. A linear combination of those

equalities gives

$$\begin{array}{|c|c|c|c|} \hline 1 & 1 & & \tau & \tau \\ \hline 1 & 1 & & \tau & \tau \\ \hline 1 & 1 & 1 & 1 & \\ \hline 1 & 1 & 1 & 1 & \\ \hline & & \ddots & & \\ \hline & & & & \\ \hline & & & 1 & 1 \\ \hline & & & 1 & 1 \\ \hline \end{array} \cdot P_{A,B|X,Y} = 0, \quad (47)$$

where $\tau = 1 - 2M$. If $P_{A,B|X,Y}$ is a non-signaling distribution, the following equalities hold.

$$\begin{aligned} P_{A|X}(0,0) &= \begin{array}{|c|c|c|c|} \hline 1 & 1 & & \\ \hline 0 & 0 & & \\ \hline & & & \\ \hline & & \ddots & \\ \hline & & & \\ \hline & & & \\ \hline \end{array} \cdot P_{A,B|X,Y} \\ &= \frac{1}{2} \begin{array}{|c|c|c|c|} \hline 0 & 1 & & 1 & 1 \\ \hline -1 & 0 & & 0 & 0 \\ \hline 1 & 0 & & & \\ \hline 1 & 0 & & & \\ \hline & & \ddots & & \\ \hline & & & \ddots & \\ \hline & & & & \\ \hline \end{array} \cdot P_{A,B|X,Y} \\ &= \frac{1}{2} \begin{array}{|c|c|c|c|} \hline 0 & 1 & & 2 & 1 \\ \hline -1 & 0 & & 1 & 0 \\ \hline 0 & -1 & 0 & 1 & \\ \hline 1 & 0 & -1 & 0 & \\ \hline & & \ddots & & \\ \hline & & & \ddots & \\ \hline & & & 0 & -1 & 0 & 1 \\ \hline & & & 1 & 0 & -1 & 0 \\ \hline \end{array} \cdot P_{A,B|X,Y} \end{aligned}$$

The second and third equalities follow by adding linear combinations of non-signaling constraints. The above plus (47) times $1/4M$ gives

$$P_{A|X}(0,0) = (\mu + \nu) \cdot P_{A,B|X,Y}.$$

Under the relabeling

$$(A, B) \rightarrow (A \oplus 1, B \oplus 1),$$

we have the transformations

$$\begin{aligned} P_{A|X}(0,0) &\rightarrow P_{A|X}(1,0), \\ \mu &\rightarrow \mu, \\ \nu &\rightarrow -\nu, \end{aligned}$$

which imply $P_{A|X}(a,0) = \beta_a \cdot P_{A,B|X,Y}$. The generalization to N pairs of systems is straightforward. Each non-signaling constraint involves a linear combination of the entries of $P_{\mathbf{A},\mathbf{B}|\mathbf{X},\mathbf{Y}}$ where all indexes remain constant except the ones corresponding to one system (like for instance a_1, x_1). Hence we can apply the above argument to each of the N pairs separately, obtaining (46). \square

C. Privacy amplification

The following analysis of privacy amplification is similar to the one in [21], but has the advantage that it is valid for any choice of error correction scheme. On the other hand, it has the disadvantage that it needs a random hash function G , in particular a two-universal one [41], while the one in [21] works with a deterministic hash function.

Definition 5 A random function $G : \{0,1\}^N \rightarrow \{0,1\}^{N_s}$ is called two-universal [41] if for any pair $\mathbf{a}, \mathbf{a}' \in \{0,1\}^N$ such that $\mathbf{a} \neq \mathbf{a}'$ we have

$$\text{prob}\{G(\mathbf{a}) = G(\mathbf{a}')\} \leq 2^{-N_s}. \quad (48)$$

Lemma 6 If $G : \{0,1\}^N \rightarrow \{0,1\}^{N_s}$ is a two-universal random function, then for any subset $\mathcal{A} \subseteq \{0,1\}^N$ we have

$$\sum_{k,g} P_G(g) \left| \sum_{\mathbf{a} \in \mathcal{A}} (\delta_{g(\mathbf{a})}^k - 2^{-N_s}) \right| \leq \sqrt{2^{N_s} |\mathcal{A}|}, \quad (49)$$

where k runs over $\{0,1\}^{N_s}$.

Proof In what follows we take the square of the left-hand side of (49); use the convexity of the square function; sum over k ; partially sum over \mathbf{a}, \mathbf{a}' ; use the two-universality of G ; and a trivial bound.

$$\begin{aligned} &\left(\sum_{k,g} P_G(g) \left| \sum_{\mathbf{a} \in \mathcal{A}} (\delta_{g(\mathbf{a})}^k - 2^{-N_s}) \right| \right)^2 \\ &\leq \sum_{k,g} 2^{-N_s} P_G(g) \sum_{\mathbf{a}, \mathbf{a}' \in \mathcal{A}} \left(2^{2N_s} \delta_{g(\mathbf{a})}^k \delta_{g(\mathbf{a}')}^k + 1 - 2^{1+N_s} \delta_{g(\mathbf{a})}^k \right) \\ &= \sum_g P_G(g) \sum_{\mathbf{a}, \mathbf{a}' \in \mathcal{A}} \left(2^{N_s} \delta_{g(\mathbf{a}')}^g - 1 \right) \\ &= 2^{N_s} \sum_{\mathbf{a}, \mathbf{a}' \in \mathcal{A}: \mathbf{a} \neq \mathbf{a}'} \left(\sum_g P_G(g) \delta_{g(\mathbf{a}')}^g \right) + 2^{N_s} |\mathcal{A}| - |\mathcal{A}|^2 \\ &\leq (|\mathcal{A}|^2 - |\mathcal{A}|) + 2^{N_s} |\mathcal{A}| - |\mathcal{A}|^2 \\ &\leq 2^{N_s} |\mathcal{A}|. \end{aligned}$$

\square

Theorem 7 Let $P_{\mathbf{A},\mathbf{B},E|\mathbf{X},\mathbf{Y},Z}$ be a $(2N_r+1)$ -partite non-signaling distribution. Suppose that Alice's systems are measured with $\mathbf{X} = \mathbf{0}$, obtaining the outcomes \mathbf{A} . Let $C = f(\mathbf{A})$ where $f : \{0,1\}^{N_r} \rightarrow \{0,1\}^{N_c}$ is a given function, and $K = G(\mathbf{A})$ where $G : \{0,1\}^{N_r} \rightarrow \{0,1\}^{N_s}$ is a two-universal random function. Then

$$\begin{aligned} &\sum_{k,c,g} \max_z \sum_e \left| P_{k,c,e,g|z} - 2^{-N_s} P_{c,e,g|z} \right| \\ &\leq \sqrt{2}^{N_r+N_s+N_c+1} \left\langle \prod_{n=1}^{N_r} \left(\frac{1}{2} + MW_n \right) \right\rangle, \quad (50) \end{aligned}$$

where $W_n = (A_n \oplus B_n \oplus \delta_{X_n}^0 \delta_{Y_n}^{M-1})$ and the expectation in (50) is taken with respect to the distribution $P_{\mathbf{a},\mathbf{b}|\mathbf{x},\mathbf{y}} \prod_{n=1}^{N_r} Q_{x_n, x_n}$, where Q_{x_n, y_n} is defined in (5).

Proof For any subset $\mathcal{A} \subseteq \{0,1\}^{N_r}$ we have the following chain of component-wise inequalities.

$$\begin{aligned} &\sum_{k,g} P_G(g) \left| \sum_{\mathbf{a} \in \mathcal{A}} (\delta_{g(\mathbf{a})}^k - 2^{-N_s}) \bigotimes_{n=1}^{N_r} \beta_{a_n} \right| \\ &\preceq \sum_{k,g} P_G(g) \left(\mu^{\otimes N_r} \left| \sum_{\mathbf{a} \in \mathcal{A}} (\delta_{g(\mathbf{a})}^k - 2^{-N_s}) \right| + \right. \\ &\quad \left. + |\nu| \otimes \mu^{\otimes N_r-1} \left| \sum_{\mathbf{a} \in \mathcal{A}} (-1)^{a_1} (\delta_{g(\mathbf{a})}^k - 2^{-N_s}) \right| + \right. \\ &\quad \left. + \dots + |\nu|^{\otimes N_r} \left| \sum_{\mathbf{a} \in \mathcal{A}} (-1)^{a_1 + \dots + a_{N_r}} (\delta_{g(\mathbf{a})}^k - 2^{-N_s}) \right| \right) \\ &\preceq \mu^{\otimes N_r} \sqrt{2^{N_s} |\mathcal{A}|} + |\nu| \otimes \mu^{\otimes N_r-1} \sqrt{2^{1+N_s} |\mathcal{A}|} \\ &\quad + \dots + |\nu|^{\otimes N_r} \sqrt{2^{1+N_s} |\mathcal{A}|} \\ &\preceq \sqrt{2^{1+N_s} |\mathcal{A}|} \beta^{\otimes N_r}. \quad (51) \end{aligned}$$

In the first step we used the expansion

$$\begin{aligned} &\bigotimes_{n=1}^N \beta_{a_n} \\ &= \mu^{\otimes N_r} + (-1)^{a_1} \nu \otimes \mu^{\otimes N_r} + \dots + (-1)^{a_1 + \dots + a_{N_r}} \nu^{\otimes N_r}, \end{aligned} \quad (52)$$

as well as the component-wise triangular inequality. In the second step we used the following triangular inequality for any $\mathbf{u} \in \{0,1\}^{N_r}$

$$\begin{aligned} &\left| \sum_{\mathbf{a} \in \mathcal{A}} (-1)^{\mathbf{a} \cdot \mathbf{u}} (\delta_{g(\mathbf{a})}^k - 2^{-N_s}) \right| \\ &\leq \left| \sum_{\mathbf{a} \in \mathcal{A}: \mathbf{a} \cdot \mathbf{u} = 0 \pmod 2} (\delta_{g(\mathbf{a})}^k - 2^{-N_s}) \right| \\ &\quad + \left| \sum_{\mathbf{a} \in \mathcal{A}: \mathbf{a} \cdot \mathbf{u} = 1 \pmod 2} (\delta_{g(\mathbf{a})}^k - 2^{-N_s}) \right|, \end{aligned}$$

Lemma 6, and the concavity of the square-root function

$$\sum_{i=1}^M \sqrt{t_i} \leq \sqrt{M \sum_{i=1}^M t_i}. \quad (53)$$

For the last inequality all terms are summed up by using $\beta = \mu + |\nu|$.

In the rest of this proof the following notation is used. We denote by $P_{\mathbf{A},\mathbf{B},e|\mathbf{X},\mathbf{Y},z} = P_{\mathbf{A},\mathbf{B},E|\mathbf{X},\mathbf{Y},Z}(e,z)$ the vector with entries $P_{\mathbf{A},\mathbf{B},E|\mathbf{X},\mathbf{Y},Z}(\mathbf{a},\mathbf{b},e,\mathbf{x},\mathbf{y},z)$ for all values of $\mathbf{a},\mathbf{b},\mathbf{x},\mathbf{y}$ and fixed values of e,z . Following this notation we can write $P_{\mathbf{a}} = P_{\mathbf{A}}(\mathbf{a})$. For any subset $\mathcal{A} \subseteq \{0,1\}^{N_r}$ and any set of coefficients $\eta_{\mathbf{a}}$ we have the following chain of equalities and inequalities,

$$\begin{aligned}
& \sum_e P_{e|z} \left| \sum_{\mathbf{a} \in \mathcal{A}} \eta_{\mathbf{a}} P_{\mathbf{a}|e,z} \right| \\
&= \sum_e P_{e|z} \left| \sum_{\mathbf{a} \in \mathcal{A}} \eta_{\mathbf{a}} \bigotimes_{n=1}^{N_r} \beta_{a_n} \cdot P_{\mathbf{A},\mathbf{B}|\mathbf{X},\mathbf{Y},e,z} \right| \\
&\leq \sum_e P_{e|z} \left| \sum_{\mathbf{a} \in \mathcal{A}} \eta_{\mathbf{a}} \bigotimes_{n=1}^{N_r} \beta_{a_n} \right| \cdot P_{\mathbf{A},\mathbf{B}|\mathbf{X},\mathbf{Y},e,z} \\
&= \left| \sum_{\mathbf{a} \in \mathcal{A}} \eta_{\mathbf{a}} \bigotimes_{n=1}^{N_r} \beta_{a_n} \right| \cdot \sum_e P_{e|z} P_{\mathbf{A},\mathbf{B}|\mathbf{X},\mathbf{Y},e,z} \\
&= \left| \sum_{\mathbf{a} \in \mathcal{A}} \eta_{\mathbf{a}} \bigotimes_{n=1}^{N_r} \beta_{a_n} \right| \cdot P_{\mathbf{A},\mathbf{B}|\mathbf{X},\mathbf{Y}}, \tag{54}
\end{aligned}$$

where we have respectively used: Lemma 4, the convexity of the absolute value function, the linearity of the scalar product, and the definition of the conditional distribution. The following establishes (50).

$$\begin{aligned}
& \sum_{k,c,g} \max_z \sum_e \left| P_{k,c,g,e|z} - 2^{-N_s} P_{c,g,e|z} \right| \\
&= \sum_{k,c,g} \max_z \sum_e P_{g,e|z} \left| P_{k,c|g,e,z} - 2^{-N_s} P_{c|e,z} \right| \\
&= \sum_{k,c,g} P_g \max_z \sum_e P_{e|z} \left| \sum_{\mathbf{a} \in f^{-1}(c)} (\delta_{g(\mathbf{a})}^k - 2^{-N_s}) P_{\mathbf{a}|e,z} \right| \\
&\leq \sum_{k,c,g} P_g \left| \sum_{\mathbf{a} \in f^{-1}(c)} (\delta_{g(\mathbf{a})}^k - 2^{-N_s}) \bigotimes_{n=1}^{N_r} \beta_{a_n} \right| \cdot P_{\mathbf{A},\mathbf{B}|\mathbf{X},\mathbf{Y}} \\
&\leq \sum_c \sqrt{2^{1+N_s} |f^{-1}(c)|} \beta^{\otimes N_r} \cdot P_{\mathbf{A},\mathbf{B}|\mathbf{X},\mathbf{Y}} \\
&\leq \sqrt{2^{1+N_s+N_c+N_r}} \beta^{\otimes N_r} \cdot P_{\mathbf{A},\mathbf{B}|\mathbf{X},\mathbf{Y}}, \tag{55}
\end{aligned}$$

In the above we have respectively used: the definition of conditional distribution and the fact that C, E, Z are independent from G ; equality $P_c = \sum_{\mathbf{a} \in f^{-1}(c)} P_{\mathbf{a}}$ and the independence of \mathbf{A}, E, Z from G ; inequality (54) with $\mathcal{A} = f^{-1}(c)$; the component-wise inequality (51) together with the fact that the components of the vector $P_{\mathbf{A},\mathbf{B}|\mathbf{X},\mathbf{Y}}$ are positive; and the last inequality follows from (53) and $\sum_c |f^{-1}(c)| = 2^{N_r}$. \square

D. Security from estimated information

According to the previous theorem, the security of the secret key can be bounded in terms of the quantity

$$\beta^{\otimes N_r} \cdot P_{\mathbf{A}_r, \mathbf{B}_r | \mathbf{X}_r, \mathbf{Y}_r} = \left\langle \prod_{n=1}^{N_r} \left(\frac{1}{2} + MW_n \right) \right\rangle, \tag{56}$$

which does not depend on E at all, but only on the distribution of data held by the honest parties! This is a particular manifestation of the monogamy of non-local correlations. However, also the distribution $P_{\mathbf{A}_r, \mathbf{B}_r | \mathbf{X}_r, \mathbf{Y}_r}$ that occurs in (56) is not necessarily known. Hence, in order to be of use, we need to relate it to an observable quantity, such as \bar{W} , defined in (13). This is the purpose of Lemma 8 below.

To simplify notation, we restrict to the relevant pairs of systems, that is, the ones that are used to either estimate the amount of non-locality $n \in \mathcal{N}_e$, or the ones constituting the raw key $n \in \mathcal{N}_r$. These are the ones that are not discarded in the protocol.

Lemma 8 *Let N_r and N_e be two positive integers and $P_{\mathbf{A},\mathbf{B},E|\mathbf{X},\mathbf{Y},Z}$ a $(2N_u + 1)$ -partite non-signaling distribution, where $N_u = N_r + N_e$. Let the random variable $\mathbf{H} = (H_1, \dots, H_{N_u})$ be independent from $\mathbf{A}, \mathbf{B}, \mathbf{X}, \mathbf{Y}, E, Z$, and uniformly distributed over the strings $\{0,1\}^{N_u}$ with N_r zeroes and N_e ones. Suppose that all Alice's systems $n \in \{1, \dots, N\}$ with $H_n = 0$ are measured with $X_n = 0$ obtaining the N_r -bit outcome \mathbf{A}_r . Suppose the N_e pairs with $H_n = 1$ are measured with (X_n, Y_n) following the distribution $Q_{X,Y}$ defined in (5), obtaining the outcome $\mathbf{U} = [(A_n, B_n, X_n, Y_n) : H_n = 1]$. This is used to compute the variable*

$$\bar{W} = \frac{1}{N_e} \sum_{n: H_n=1} (A_n \oplus B_n \oplus \delta_{X_n}^0 \delta_{Y_n}^{M-1}). \tag{57}$$

Let $C = f(\mathbf{A}_r)$ where $f: \{0,1\}^{N_r} \rightarrow \{0,1\}^{N_c}$ is a given function. If $G: \{0,1\}^{N_r} \rightarrow \{0,1\}^{N_s}$ is a two-universal random function with output size

$$\begin{aligned}
N_s(\bar{w}) &= \max_{\theta \in [0,1]} [2N_e D(\bar{w} || \theta) - 2N_r \log_2(1/2 + M\theta)] \\
&\quad - N_r - N_c - 2 \log_2(8N_e N_u / \epsilon), \tag{58}
\end{aligned}$$

and $K = G(\mathbf{A}_r)$ then

$$\sum_{k,\mathbf{h},\mathbf{u},c,g} \max_z \sum_e \left| P_{k,\mathbf{h},\mathbf{u},c,e,g|z} - 2^{-N_s(\bar{w})} P_{\mathbf{h},\mathbf{u},c,e,g|z} \right| \leq \epsilon \tag{59}$$

holds for any $\epsilon > 0$.

Proof For each value of \mathbf{h} define the disjoint sets

$$\begin{aligned}
\mathcal{N}_r^{\mathbf{h}} &= \{n : h_n = 0\}, \\
\mathcal{N}_e^{\mathbf{h}} &= \{n : h_n = 1\},
\end{aligned}$$

satisfying $\mathcal{N}_r^h \cup \mathcal{N}_e^h = \{1, \dots, N_u\}$. Note that these are the same sets as (12) and (14). We also define

$$\begin{aligned} W_n &= [A_n \oplus B_n \oplus \delta_{X_n}^0 \delta_{Y_n}^{M-1}] \text{ for } n = 1, \dots, N_u, \\ \mathbf{W} &= (W_1, \dots, W_{N_u}), \\ \mathbf{W}_r &= (W_n : n \in \mathcal{N}_r^h), \\ \mathbf{W}_e &= (W_n : n \in \mathcal{N}_e^h), \\ \mathbf{U} &= [(A_n, B_n, X_n, Y_n) : n \in \mathcal{N}_e^h]. \end{aligned}$$

The distribution for \mathbf{W} is

$$P_{\mathbf{w}} = \sum_{\mathbf{a}, \mathbf{b}, \mathbf{x}, \mathbf{y}} P_{\mathbf{a}, \mathbf{b} | \mathbf{x}, \mathbf{y}} \prod_{n=1}^{N_u} Q_{x_n, y_n} \delta_{[a_n \oplus b_n \oplus \delta_{x_n}^0 \delta_{y_n}^{M-1}]}^{\delta_{w_n}},$$

where $P_{\mathbf{A}, \mathbf{B} | \mathbf{X}, \mathbf{Y}}$ is the marginal of $P_{\mathbf{A}, \mathbf{B}, E | \mathbf{X}, \mathbf{Y}, Z}$ and $Q_{X, Y}$ is defined in (5). For what follows it is also useful to define the symmetrized version of $P_{\mathbf{w}}$, namely

$$P_{w_1, \dots, w_{N_u}}^{\text{sym}} = \sum_{\pi} \frac{1}{N_u!} P_{w_{\pi(1)}, \dots, w_{\pi(N_u)}}, \quad (60)$$

where π runs over the permutations of the symbols $\{1, \dots, N_u\}$.

For each value of \mathbf{h} and \mathbf{u} the conditioned distribution $P_{\mathbf{A}, \mathbf{B}, E | \mathbf{X}, \mathbf{Y}, Z, \mathbf{h}, \mathbf{u}}$ is $(2N_r + 1)$ -partite non-signaling, hence, Theorem 7 applies to it. By definition, the random variable \mathbf{H} is independent from Z ; and by no-signaling, the random variable \mathbf{U} is independent from Z . Hence, $P_{\mathbf{h}, \mathbf{u} | z} = P_{\mathbf{h}, \mathbf{u}}$. This allows for taking the common factor $P_{\mathbf{h}, \mathbf{u}}$ out of the absolute value in (59), and applying Theorem 7 to each term, obtaining:

$$\begin{aligned} & \sum_{k, \mathbf{h}, \mathbf{u}, c, g} \max_z \sum_e \left| P_{k, \mathbf{h}, \mathbf{u}, c, e, g | z} - 2^{-N_s(\bar{w})} P_{\mathbf{h}, \mathbf{u}, c, e, g | z} \right| \\ &= \sum_{k, \mathbf{h}, \mathbf{u}, c, g} \max_z P_{\mathbf{h}, \mathbf{u} | z} \sum_e \left| P_{k, c, e, g | z} - 2^{-N_s(\bar{w})} P_{c, e, g | z} \right| \\ &= \sum_{\mathbf{h}, \mathbf{u}} P_{\mathbf{h}, \mathbf{u}} \\ & \quad \times \sum_{k, c, g} \max_z \sum_e \left| P_{k, c, e, g | z, \mathbf{h}, \mathbf{u}} - 2^{-N_s(\bar{w})} P_{c, e, g | z, \mathbf{h}, \mathbf{u}} \right| \\ &\leq \sum_{\mathbf{h}, \mathbf{u}} P_{\mathbf{h}, \mathbf{u}} \sqrt{2}^{N_r + N_s(\bar{w}) + N_c + 1} \\ & \quad \times \sum_{\mathbf{w}_r} P_{\mathbf{w}_r | \mathbf{h}, \mathbf{u}} \prod_{n \in \mathcal{N}_r^h} \left(\frac{1}{2} + M w_n \right) \\ &= \sum_{\mathbf{h}, \mathbf{w}} P_{\mathbf{h}, \mathbf{w}} \sqrt{2}^{N_r + N_s(\bar{w}) + N_c + 1} \prod_{n \in \mathcal{N}_r^h} \left(\frac{1}{2} + M w_n \right) \quad (61) \end{aligned}$$

The last equality follows from the fact that \bar{w} is a function of \mathbf{w}_e which in turn is a function of \mathbf{u} . Hence, averaging over $(\mathbf{w}_r, \mathbf{u})$ is equivalent to averaging over $(\mathbf{w}_r, \mathbf{w}_e) = \mathbf{w}$.

Now, let π_r be any permutation of the variables w_n with $n \in \mathcal{N}_r^h$, and π_e any permutation of the variables w_n with $n \in \mathcal{N}_e^h$. The fact that \bar{w} and $\prod_{n \in \mathcal{N}_r^h} (\frac{1}{2} + M w_n)$

are invariant under any permutation π_r and π_e implies that (61) is equal to

$$\begin{aligned} & \sum_{\mathbf{h}, \mathbf{w}, \pi_r, \pi_e} \frac{P_{\mathbf{h}, (\pi_r \pi_e \mathbf{w})}}{N_r! N_e!} \sqrt{2}^{N_r + N_s(\bar{w}) + N_c + 1} \\ & \quad \times \prod_{n \in \mathcal{N}_r^h} \left(\frac{1}{2} + M w_n \right) \\ &= \sum_{\mathbf{w}} P_{\mathbf{w}}^{\text{sym}} \sqrt{2}^{N_r + N_s(\bar{w}) + N_c + 1} \prod_{n \in \mathcal{N}_r^h} \left(\frac{1}{2} + M w_n \right) \quad (62) \end{aligned}$$

The above equality follows from noting that the average over all permutations π_r, π_e combined with the average over all partitions $(\mathcal{N}_r^h, \mathcal{N}_e^h)$ of $\{1, \dots, N_u\}$ (or equivalently the average over \mathbf{h}) is equivalent to the average in (60).

Since $P_{\mathbf{w}}^{\text{sym}}$ is symmetric we can apply Lemma 3 to upper-bound (62) in terms of a maximization over i.i.d. distributions $P_{w_1, \dots, w_{N_u}}^{\text{iid}} = P_{w_1} \cdots P_{w_{N_u}}$. These i.i.d. distributions are parametrized by the single number $\theta = P_W(1) \in [0, 1]$. The following is an upper bound for (62).

$$\begin{aligned} & (N_u + 1) \max_{P_{\mathbf{w}}^{\text{iid}}} \sum_{\mathbf{w}} P_{\mathbf{w}}^{\text{iid}} \sqrt{2}^{N_r + N_s(\bar{w}) + N_c + 1} \\ & \quad \times \prod_{n \in \mathcal{N}_r^h} \left(\frac{1}{2} + M w_n \right) \\ &= (N_u + 1) \max_{\theta \in [0, 1]} \sum_{\bar{w}} P_{\bar{w}}^{(\theta)} \\ & \quad \times \sqrt{2}^{N_r + N_s(\bar{w}) + N_c + 1} \left(\frac{1}{2} + M \theta \right)^{N_r} \end{aligned}$$

To obtain the above equality we use

$$\sum_{\mathbf{w}_r} P_{\mathbf{w}_r}^{\text{iid}} \prod_{n=1}^{N_r} \left(\frac{1}{2} + M w_n \right) = \left(\frac{1}{2} + M \theta \right)^{N_r},$$

and express the average over \mathbf{w}_e in terms of the distribution $P_{\bar{w}}^{(\theta)}$, which is

$$P_{\bar{w}}^{(\theta)} = \binom{N_e}{N_e \bar{w}} \theta^{N_e \bar{w}} (1 - \theta)^{N_e (1 - \bar{w})}. \quad (63)$$

It is well known [43] that the above distribution can be bounded as

$$P_{\bar{w}}^{(\theta)} \leq 2^{-N_e D(\bar{w} || \theta)}, \quad (64)$$

where $D(\bar{w} || \theta)$ is the binary relative entropy defined

in (18). Putting everything together we obtain

$$\begin{aligned}
& \sum_{k,\mathbf{h},\mathbf{u},c,g} \max_z \sum_e \left| P_{k,\mathbf{h},\mathbf{u},c,e,g|z} - 2^{-N_s(\bar{w})} P_{\mathbf{h},\mathbf{u},c,e,g|z} \right| \\
& \leq (N_u + 1) \max_{\theta \in [0,1]} \sum_{\bar{w}} 2^{(N_r + N_s(\bar{w}) + N_c + 1)/2} \\
& \quad \times 2^{-N_e D(\bar{w} \parallel \theta) + N_r \log(1/2 + M\theta)} \\
& \leq (N_u + 1) \max_{\theta \in [0,1]} \sum_{\bar{w}} 2^{1/2 - \log_2(8N_e N_u / \epsilon)} \\
& = (N_u + 1)(N_e + 1) \frac{\sqrt{2}\epsilon}{8N_e N_u} \leq \epsilon, \tag{65}
\end{aligned}$$

where we have used $\sum_{\bar{w}} 1 = N_e + 1$. \square

To avoid confusion in the following Theorem, we recall that the alphabets of the random variables $\mathbf{A}_r, \mathbf{B}'_r$ and G depend on the value of T , defined in (19) or (66). Particularly, $\mathbf{A}_r, \mathbf{B}'_r$ take values in $\{0, 1\}^{N_r}$, and G takes values in the set of functions $\{0, 1\}^{N_r} \rightarrow \{0, 1\}^{N_s}$. But the number N_r is a function of \mathbf{I}, \mathbf{J} , namely, the number of times $I_n = J_n = 0$. And as can be seen in (17), the quantity $N_s(\bar{W})$ is a function of \bar{W} , which in turn is a function of T .

Theorem 9 *At the end of the protocol described in Section III A, Alice holds K_A , Bob holds K_B , and the adversary has all the information*

$$T = \left[\mathbf{I}, \mathbf{J}, C, G, (A_n, B_n, X_n, Y_n) \forall n \in \mathcal{N}_e \right] \tag{66}$$

and the system associated to E, Z . If the error correction scheme has error probability

$$\sum_{t, \mathbf{a}_r \neq \mathbf{b}'_r} P_{t, \mathbf{a}_r, \mathbf{b}'_r} \leq \epsilon_{\text{erco}} \tag{67}$$

then we have

$$\begin{aligned}
& \sum_{k_A, k_B, t} \max_z \sum_e \left| P_{k_A, k_B, t, e|z} - 2^{-N_s(\bar{w})} \delta_{k_A}^{k_B} P_{t, e|z} \right| \\
& \leq \epsilon + 2\epsilon_{\text{erco}}. \tag{68}
\end{aligned}$$

Proof Using the triangular inequality

$$\begin{aligned}
& \left| P_{k_A, k_B, t, e|z} - 2^{-N_s(\bar{w})} \delta_{k_A}^{k_B} P_{t, e|z} \right| \\
& \leq \left| P_{k_A, k_B, t, e|z} - P_{k_A, t, e|z} \delta_{k_A}^{k_B} \right| \\
& \quad + \left| P_{k_A, t, e|z} \delta_{k_A}^{k_B} - 2^{-N_s(\bar{w})} \delta_{k_A}^{k_B} P_{t, e|z} \right| \tag{69}
\end{aligned}$$

we can bound the left-hand side of (68) with the corresponding two terms. The first term can be simplified by splitting the sum into the terms with $k_A = k_B$ and $k_A \neq k_B$, and using the no-signaling constraint

$\sum_e P_{k_A, k_B, t, e|z} = P_{k_A, k_B, t}$, that is

$$\begin{aligned}
& \sum_{k_A, k_B, t} \max_z \sum_e \left| P_{k_A, k_B, t, e|z} - P_{k_A, t, e|z} \delta_{k_A}^{k_B} \right| \\
& = \sum_{k_A \neq k_B, t} \max_z \sum_e P_{k_A, k_B, t, e|z} \\
& \quad + \sum_{k_A, t} \max_z \sum_e \left| P_{k_A, k_B = k_A, t, e|z} - P_{k_A, t, e|z} \right| \\
& \leq \sum_{k_A \neq k_B, t} \max_z P_{k_A, k_B, t} \\
& \quad + \sum_{k_A, t} \max_z \sum_e (P_{k_A, t, e|z} - P_{k_A, k_B = k_A, t, e|z}) \\
& \leq \epsilon_{\text{erco}} + \sum_{k_A, t} (P_{k_A, t} - P_{k_A, k_B = k_A, t}) \\
& \leq \epsilon_{\text{erco}} + 1 - (1 - \epsilon_{\text{erco}}) = 2\epsilon_{\text{erco}}. \tag{70}
\end{aligned}$$

In the last two inequalities we have used that

$$\begin{aligned}
& \sum_{k_A \neq k_B, t} P_{k_A, k_B, t} \\
& = \sum_{k_A \neq k_B, t, \mathbf{a}_r, \mathbf{b}'_r} P_{t, \mathbf{a}_r, \mathbf{b}'_r} \delta_{g(\mathbf{a}_r)}^{k_A} \delta_{g(\mathbf{b}'_r)}^{k_B} \\
& \leq \sum_{t, \mathbf{a}_r \neq \mathbf{b}'_r} P_{t, \mathbf{a}_r, \mathbf{b}'_r} \leq \epsilon_{\text{erco}}. \tag{71}
\end{aligned}$$

To bound the second term in (69) we invoke Lemma 8. However, in Lemma 8 the values of N_r, N_e and the set $\mathcal{N}_u = \mathcal{N}_r \cup \mathcal{N}_e$ are fixed, while here this is not the case. To overcome this problem, we can separate the sum over N_r, N_e, \mathcal{N}_u and independently apply Lemma 8 to each term with a fixed value of N_r, N_e, \mathcal{N}_u .

$$\begin{aligned}
& \sum_{k_A, k_B, t} \max_z \sum_e \delta_{k_A}^{k_B} \left| P_{k_A, t, e|z} - 2^{-N_s(\bar{w})} P_{t, e|z} \right| \\
& = \sum_{N_r, N_e, \mathcal{N}_u} P(N_r, N_e, \mathcal{N}_u) \sum_{k, \mathbf{h}, \mathbf{u}, c, g} \max_z \sum_e \\
& \quad \left| P_{k, \mathbf{h}, \mathbf{u}, c, e, g|z, N_r, N_e, \mathcal{N}_u} - 2^{-N_s(\bar{w})} P_{\mathbf{h}, \mathbf{u}, c, e, g|z, N_r, N_e, \mathcal{N}_u} \right| \\
& \leq \sum_{N_r, N_e, \mathcal{N}_u} P(N_r, N_e, \mathcal{N}_u) \epsilon = \epsilon. \tag{72}
\end{aligned}$$

To understand the above equality note that, except for the discarded pairs (with $I_n \neq J_n$) which are not considered in Lemma 8, all the information contained in t is also contained in $[N_r, N_e, \mathcal{N}_u, \mathbf{h}, \mathbf{u}, c, g]$. More specifically, the set \mathcal{N}_u tells us which of the pairs satisfy $I_n = J_n$, and among those, \mathbf{h} designates the ones with $I_n = 0$ and $I_n = 1$. Additionally, the information $(A_n, B_n, X_n, Y_n) \forall n \in \mathcal{N}_e$ is contained in \mathbf{u} .

Finally, the combination of (70) and (72) gives (68). \square

V. CONCLUSIONS

We have showed that it is possible to generate secret key from correlations that violate the Braunstein-Caves inequality [29] by a sufficient amount. We proved this according to the strongest notion of security, the so-called universally-composable security [19, 20]. The only assumption used in the security proof is that, when measuring a system, the outcome does not depend on the choice of observables measured on other systems. One clean (although expensive) way to achieve this within our device-independent scenario would be to use a separate isolated measurement device for each of the measurements. A more practical (but not fully device-independent) variant is to use a single device per party, whose design is chosen such that it can reasonably be assumed that subsequent measurements are independent of each other (i.e., the device should not have any internal memory).

On the technical level, we introduced a scheme for estimating symmetric properties of general probability distributions. Applied to our setting, this allows Alice and Bob to treat arbitrary and unknown correlations as if they were independent and identically-distributed samples. This may be more generally useful to quantify Bell-inequality violations without the i.i.d. assumption.

This work is inspired by, but goes beyond the philosophy of [2] in which the validity of quantum mechanics, in particular, Tsirelson's bound [32], is still assumed. In contrast, *all* we assume is no-signaling. This idea, proposed in [5], is conceptually simpler. Nevertheless, we hope that our results also contribute to the understanding of the more practical scenario of device-independent quantum cryptography where quantum theory is assumed to hold, but where the honest users still do not have complete control of their quantum apparatuses, or distrust them [3, 15, 16]. In this case, our techniques may still be applied. Furthermore, with appropriate modifications (cf. Appendix A), it is possible to obtain higher key rates, compared to the pure non-signaling scenario.

VI. ACKNOWLEDGEMENTS

LM acknowledges support from CatalunyaCaixa, the EU ERC Advanced Grant NLST (PHYS RQ8784), EU Qessence project and the Templeton Foundation. RR acknowledges support by the Swiss National Science Foundation (SNF) through the National Centre of Competence in Research "Quantum Science and Technology (QSIT)" and through grant No. 200020-135048, the European Research Council (grant 258932), and the CHIST-ERA project DIQIP. MC acknowledges financial support by the German Science Foundation (grant CH 843/2-1), the Swiss National Science Foundation (grants PP00P2-128455, 20CH21-138799 (CHIST-ERA project CQC)), the Swiss National Center of Competence in Research "QSIT", the Swiss State Secretariat for Edu-

cation and Research supporting COST action MP1006, and the European Research Council under the European Union's Seventh Framework Program (FP/2007-2013) (grant 337603). AW acknowledges support by the EC STREP QCS, the ERC (Advanced Grant IRQUAT), and the Philip Leverhulme Trust. JB is supported by the EPSRC, and the CHIST-ERA DIQIP projects.

Appendix A: Monogamy of non-local correlations

The following lemma is not necessary for the security proof, but we include it for the following two reasons. First, it provides insight on phenomenon of monogamy, that is, the trade-off between Bell-inequality violation and correlation with a third party. Second, it allows to highly improve the secret key rate of our protocol if we additionally assume that the global initial distribution $P_{\mathbf{A},\mathbf{B},E|\mathbf{X},\mathbf{Y},Z}$ is compatible with quantum theory [17]. Note that this extra assumption does not invalidate the fact that the security is device independent.

In order to incorporate the validity of quantum theory as an extra assumption, we can use the results in [42], which provide a bound for the security of the secret key in terms of the guessing probability of the raw key. And this is the quantity addressed by the following lemma.

Lemma 10 *Let $P_{\mathbf{A},\mathbf{B},E|\mathbf{X},\mathbf{Y},Z}$ be an arbitrary $(2N+1)$ -partite non-signaling distribution and define*

$$\mathcal{P}_{\text{guess}}(\mathbf{A}|E, \mathbf{x}) = \max_z \sum_e \max_{\mathbf{a}} P_{\mathbf{A},E|\mathbf{X},Z}(\mathbf{a}, e, \mathbf{x}, z). \quad (\text{A1})$$

For any \mathbf{x} we have

$$\mathcal{P}_{\text{guess}}(\mathbf{A}|E, \mathbf{x}) \leq \left\langle \prod_{n=1}^N \left(\frac{1}{2} + MW_n \right) \right\rangle, \quad (\text{A2})$$

where the expectation of $W_n = (A_n \oplus B_n \oplus \delta_{X_n}^0 \delta_{Y_n}^{M-1})$ is taken with the distribution Q_{X_n, Y_n} defined in (5).

Proof First, note that by using the non-signaling condition we can write

$$P_{\mathbf{A},\mathbf{B}|\mathbf{X},\mathbf{Y}} = \sum_e P_{E|Z}(e, z) P_{\mathbf{A},\mathbf{B}|\mathbf{X},\mathbf{Y},E,Z}(e, z). \quad (\text{A3})$$

Now, let us show that

$$\beta_{a_1} \otimes \dots \otimes \beta_{a_n} \preceq \beta^{\otimes n}, \quad (\text{A4})$$

for any n and any $(a_1, \dots, a_n) \in \{0, 1\}^n$. First, expand each side of this inequality according to definitions (43) and (44); second, note that $\mp \nu^{\otimes n} \preceq |\nu^{\otimes n}| = |\nu|^{\otimes n}$; and finally, use this to show that each term in the left is component-wise bounded by the corresponding term in the right. Let us show (A2) for the case $\mathbf{x} = (0, \dots, 0)$. In the following chain of equalities and inequalities we use, respectively: the definition of $\mathcal{P}_{\text{guess}}$ in

(A1); Lemma 4; inequality (A4) and positivity of the vectors $P_{\mathbf{A},\mathbf{B}|\mathbf{X},\mathbf{Y},E,Z}(e,z)$; the linearity of the scalar product; decomposition (A3); and the identity (45).

$$\begin{aligned}
& \mathcal{P}_{\text{guess}}(\mathbf{A}|E, \mathbf{x}) \\
&= \max_z \sum_e P_{E|Z}(e,z) \max_{\mathbf{a}} P_{\mathbf{A}|\mathbf{X},E,Z}(\mathbf{a}, \mathbf{x}, e, z) \\
&= \max_z \sum_e P_{E|Z}(e,z) \max_{\mathbf{a}} \left(\bigotimes_{n=1}^N \beta_{a_n} \right) \cdot P_{\mathbf{A},\mathbf{B}|\mathbf{X},\mathbf{Y},E,Z}(e,z) \\
&\leq \max_z \sum_e P_{E|Z}(e,z) \beta^{\otimes N} \cdot P_{\mathbf{A},\mathbf{B}|\mathbf{X},\mathbf{Y},E,Z}(e,z) \\
&= \max_z \beta^{\otimes N} \cdot \left(\sum_e P_{E|Z}(e,z) P_{\mathbf{A},\mathbf{B}|\mathbf{X},\mathbf{Y},E,Z}(e,z) \right) \\
&= \beta^{\otimes N} \cdot P_{\mathbf{A},\mathbf{B}|\mathbf{X},\mathbf{Y}} \\
&= \left\langle \prod_{n=1}^N \left(\frac{1}{2} + MW_n \right) \right\rangle
\end{aligned}$$

In order to extend this inequality to all values of \mathbf{x} , consider the relabeling. For any $m \in \{0, \dots, M-1\}$

$$\begin{aligned}
X &\rightarrow X + m \bmod M \\
Y &\rightarrow Y + m \bmod M \\
A &\rightarrow A \oplus I\{M-m \leq X \leq M-1\} \\
B &\rightarrow B \oplus I\{M-m \leq Y \leq M-1\}
\end{aligned} \tag{A5}$$

This relabeling corresponds to a permutation of the entries of the vectors (39) such that

$$P_{A|X}(a, 0) \rightarrow P_{A|X}(a, m) .$$

This relabeling leaves the vector β invariant. Hence, performing the relabeling to each pair with $m = x_n$, the above inequality for $\mathbf{x} = (0, \dots, 0)$ is generalized to any value of \mathbf{x} . \square

-
- [1] C. H. Bennett, G. Brassard; In Proceedings of International Conference on Computer Systems and Signal Processing, page 175 (1984).
- [2] A. Ekert; Phys. Rev. Lett. **67**, 661 (1991).
- [3] A. Acín, N. Gisin, Ll. Masanes; Phys. Rev. Lett. **97**, 120405 (2006).
- [4] A. Acin, N. Brunner, N. Gisin, S. Massar, S. Pironio, V. Scarani; Phys. Rev. Lett. **98**, 230501 (2007).
- [5] J. Barrett, L. Hardy, A. Kent; Phys. Rev. Lett. **95**, 010503 (2005).
- [6] H.-K. Lo, H. F. Chau; Science, 283:20502056 (1999).
- [7] D. Mayers; Journal of the ACM, 48(3):351406 (2001).
- [8] P. Shor, J. Preskill; Phys. Rev. Lett. 85:441 (2000).
- [9] E. Biham, T. Mor; Phys. Rev. Lett. 78(11):22562259 (1997).
- [10] M. Christandl, R. Renner, A. Ekert; <http://arxiv.org/abs/quant-ph/0402131> (2004).
- [11] R. Renner, N. Gisin, B. Kraus; Phys. Rev. A 72:012332 (2005).
- [12] F. Xu, B. Qi, H.-K. Lo; New J. Phys. 12, 113026 (2010).
- [13] L. Lydersen, C. Wiechers, C. Wittmann, D. Elser, J. Skaar, V. Makarov; Nat. Photonics 4, 686 (2010).
- [14] J. S. Bell; Physics **1**(3), 195 (1964).
- [15] D. Mayers, A. Yao; Quantum Inf. Comput. 4, 273 (2004).
- [16] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, V. Scarani; Phys. Rev. Lett. **98**, 230501 (2007).
- [17] L. Masanes, S. Pironio, A. Acín; Nat. Commun. 2, 238 (2011)
- [18] E. Hänggi, R. Renner; arXiv:1009.1833 (2010).
- [19] R. Renner, R. König; Proc. of TCC 2005, LNCS, Springer, vol. 3378 (2005).
- [20] M. Ben-Or, M. Horodecki, D. W. Leung, D. Mayers, J. Oppenheim; Theory of Cryptography: Second Theory of Cryptography Conference, TCC 2005, J. Kilian (ed.) Springer Verlag 2005, vol. 3378 of Lecture Notes in Computer Science, pp. 386-406.
- [21] Ll. Masanes; Phys. Rev. Lett. **102**, 140501 (2009).
- [22] S. Pironio, L. Masanes, A. Leverrier, A. Acin; arXiv:1211.1402
- [23] J. Barrett, R. Colbeck, A. Kent; Phys. Rev. A **86**, 062326 (2012).
- [24] U. Vazirani, T. Vidick; arXiv:1210.1810.
- [25] S. Popescu, D. Rohrlich; Found. Phys. **24**, 379 (1994).
- [26] J. Barrett, N. Linden, S. Massar, S. Pironio, S. Popescu, D. Roberts; Phys. Rev. A **71**, 022101 (2005).
- [27] Ll. Masanes, A. Acín, N. Gisin; Phys. Rev. A **73**, 012112 (2006).
- [28] J. F. Clauser, M. A. Horne, A. Shimony, R. A. Holt; Phys. Rev. Lett. **23**, 880 (1969).
- [29] S. Braunstein, C. Caves; Ann. Phys. **202**, p. 22 (1990).
- [30] A. Einstein, B. Podolsky, N. Rosen; Phys. Rev. **47**, 777 (1935).
- [31] S. Wehner; Phys. Rev. A **73**, 022110 (2006).
- [32] B. S. Tsirelson; Hadronic J. Suppl. **8**, 329 (1993).
- [33] J. Barrett, A. Kent, S. Pironio; Phys. Rev. Lett. **97** 170409 (2006).
- [34] J. Mueller-Quade, R. Renner; New Journal of Physics **11**, 085006 (2009).
- [35] R. Renner; *Security of Quantum Key Distribution*, PhD thesis, ETH Zurich, arXiv:quant-ph/0512258 (2005).
- [36] U. Leonhardt; Phys. Rev. Lett. **74**, 4101 (1995).
- [37] Vogel, Risken; Phys. Rev. A **40** 7113 (1989).
- [38] A. Acín, S. Massar, S. Pironio; New J. Phys. **8**, 126 (2006).
- [39] M. Christandl, R. König, R. Renner; Phys. Rev. Lett. **102**, 020504 (2009).
- [40] R. Renner, NATO Advanced Research Workshop Quantum Cryptography and Computing: Theory and Implementation, vol. 26, 66-75 (2010). <http://ebooks.iospress.nl/publication/23842>
- [41] C. H. Bennett, G. Brassard, C. Crépeau, U. M. Maurer; IEEE Trans. Inf. Theory, vol. 41, no. 6 (1995).
- [42] R. König, R. Renner, C. Schaffner; IEEE Trans. Inf. Th., vol. 55, no. 9 (2009).
- [43] T. M. Cover, J. A. Thomas; *Elements of Information Theory*, John Wiley & Sons, Inc.