Simple and practical DIQKD security analysis via BB84-type uncertainty relations and Pauli correlation constraints

Michele Masini, Stefano Pironio, and Erik Woodhead

Laboratoire d'Information Quantique, Université libre de Bruxelles (ULB), Belgium

According to the entropy accumulation theorem, proving the unconditional security of a device-independent quantum key distribution protocol reduces to deriving tradeoff functions, i.e., bounds on the single-round von Neumann entropy of the raw key as a function of Bell linear functionals, conditioned on an eavesdropper's quantum side information. In this work, we describe how the conditional entropy can be bounded in the 2-input/2-output setting, where the analysis can be reduced to qubit systems, by combining entropy bounds for variants of the well-known BB84 protocol with quantum constraints on qubit operators on the bipartite system shared by Alice and Bob. The approach gives analytic bounds on the entropy, or semi-analytic ones in reasonable computation time, which are typically close to optimal. We illustrate the approach on a variant of the device-independent CHSH QKD protocol where both bases are used to generate the key as well as on a more refined analysis of the original single-basis variant with respect to losses. We obtain in particular a detection efficiency threshold slightly below 80.26%.

1 Introduction

Based on Bell's theorem [1, 2], device-independent quantum key distribution (DIQKD) aims to allow cryptographic keys to be generated and proved secure based on minimal assumptions about the quantum devices [3]. Following its proposal fifteen years ago, realizing a working DIQKD protocol has long presented a significant challenge both to theorists, due to the mathematical difficulty of devising practical and rigorous security proofs, and to experimental researchers, due to the difficulty of distributing entangled quantum systems with low noise and high detection rates over long distances. Recent advances paved the way to three successful proof-of-principle experiments demonstrating the feasibility of this technology [4–6]. However, there is still a long way from these proof-of-principle experiments to practical DIQKD implementations, with the necessity to improve the distance and the rate at which the keys are distributed.

Michele Masini: michele.masini@ulb.be
Stefano Pironio: stefano.pironio@ulb.be
Erik Woodhead: erik.woodhead@ulb.be

One major theoretical advance introduced a few years ago is the entropy accumulation theorem [7], and the related technique of quantum probability estimation [8], which reduces proving the unconditional security of a generic DIQKD protocol in the finite-key regime to the problem of obtaining a lower bound (called a *min-tradeoff function* in [7]) on the conditional von Neumann entropy $H(K_A|E)$ of Alice's raw key variable K_A conditioned on an eavesdropper's possible quantum side information E, as a function of the expected value of a Bell expression. For instance the security of the simplest DIQKD protocol based on the CHSH inequality follows from the following lower bound on the conditional von Neumann entropy of Alice's measurement outcome A_1

$$H(A_1|E) \geq 1 - \phi\left(\sqrt{S^2/4 - 1}\right), \tag{1}$$

where $\phi(x) = h\left(\frac{1}{2} + \frac{1}{2}x\right)$, $h(x)$ is the binary entropy, and $S = \langle A_1B_1\rangle + \langle A_1B_2\rangle + \langle A_2B_1\rangle - \langle A_2B_2\rangle$ is the expected value of the CHSH Bell expression [3].

The basic CHSH protocol based on the above lower bound is, however, not optimal in a number of respects. There has thus been in the last few years a search for ways to bound the conditional entropy for more general DIQKD protocols, either focusing on the 2-input/2-output setting [9–11], or finding numerical methods to tackle the problem in a more general way [12, 13]. Despite these efforts, bounding the entropy can be a numerically-intensive problem, with one recent approach [11] notably requiring thousands of processor core-hours of computing time to numerically bound the average entropy for a two-basis variant [10] of the CHSH-based DIQKD protocol. This has significant drawbacks, reducing confidence in the results (as they are harder for others to reproduce), increasing the difficulty to optimize over parameters in simulations, and generally increasing the time and computing resources necessary just to calculate a key rate.

In this work, we present a new and versatile approach to bound the conditional entropy in the 2-input/2-output device-independent setting that is conceptually and technically relatively simple. It is a generalization of the approach in [14] that was used to derive an analytical bound on the conditional en-

tropy for a family of asymmetric CHSH inequalities. As we explain here, the main conceptual steps of this security analysis are not specific to the protocol considered in [14] but can actually be easily adapted to other 2-input/2-output device-independent protocols.

The starting point is, as usual in the 2-input/2-output scenario, to use Jordan's lemma to reduce the analysis to convex combinations of qubit strategies. From there, our approach is based on three steps. First, as in a standard qubit QKD protocol like BB84, we bound the conditional entropy of Alice's key generating measurement, say, A_1 through an uncertainty relation involving the correlations $\langle \bar{A}_1 \otimes B\rangle$ between an *orthogonal* measurement $\bar{A}_1$ on Alice's subsystem and a binary observable B on Bob's system. In a device-independent setting, though, and contrarily to, e.g., BB84, we cannot have direct access to the correlations $\langle \bar{A}_1 \otimes B\rangle$ as we cannot assume that Alice's measurement devices perform measurements in two orthogonal bases A_1, $\bar{A}_1$. The second step is then to establish a device-independent qubit constraint on $\langle \bar{A}_1 \otimes B\rangle$ which is based on correlations between Alice and Bob that are actually observed in the protocol, e.g., the CHSH expectation value or some other Bell expression. Combining the first and second step, we obtain a bound on the conditional entropy which is device-independent, apart from the assumptions that Alice and Bob are measuring qubits. The third step then involves a convexity analysis: either the resulting bound happens to be convex or, if this is not the case, we convexify it. In this way, we get a lower bound that is valid for convex combination of qubit strategies, and thus by Jordan's lemma, for arbitrary, dimension-free strategies.

We illustrate this new approach in detail on two variants of the CHSH-based DIQKD protocol: the two-basis variant [10] and a new variant that incorporates, in addition to the CHSH value, information about the bias in the key generating measurement A_1. This last feature is particularly relevant for photonic implementations of DIQKD where no-click outcomes $\emptyset$ are mapped to a given key bit value, say $\emptyset \mapsto +1$, resulting in highly biased outcomes. The bounds that we obtain are optimal or close to optimal and significantly simpler technically and less computationally demanding than other approaches. We show in particular that a qubit DIQKD protocol can tolerate detector efficiencies as low as 80.26%.

We first provide in Section 2 a high-level description of our approach to bounding the conditional entropy in 2-input/2-output scenarios and then illustrate it in detail on the two-basis variant of the CHSH DIQKD protocol in Section 3.1 and on the variant optimized for losses in Section 3.2.

2 Description of our approach

We start by specifying the class of problems that we aim to solve. We consider a tripartite setup involving a state ρ_{ABE} shared among Alice, Bob, and the eavesdropper Eve. We assume that Alice can measure one of two ±1-valued observables A_1 or A_2 on her system, and similarly Bob can measure one of two ±1-valued observables B_1 or B_2. We refer to the tuple $\mathcal{Q} \equiv (\rho_{ABE}, A_1, A_2, B_1, B_2)$ as a *strategy*.

A strategy $\mathcal{Q}$ can be seen as describing a single round of a multi-round DIQKD protocol. The measurements by Alice and Bob serve two purposes: generating some random variable K_A on Alice's side (which will constitute Alice's copy of the *raw key* in the DIQKD protocol) and establishing some correlations between Alice and Bob (which will be estimated in a *parameter estimation* step of the DIQKD protocol). Any strategy $\mathcal{Q}$ implies some tradeoff between how random K_A is to Eve and how correlated Alice's and Bob's measurement outcomes are. This tradeoff can be formalized as follows.

Eve's information on the raw key K_A. Let us assume that Alice uses the following general procedure to generate a random key value K_A: she first selects a measurement choice $X = 1, 2$ according to a probability distribution μ_X, she measures the corresponding observable A_1 or A_2, she gets the classical output $A = \pm1$, and finally she applies to A a (possibly stochastic) map $\$_x : \{\pm1\} \to \mathcal{K}_A : A \mapsto K_A$ to obtain a value K_A in some finite alphabet $\mathcal{K}_A$. A measure of how random K_A is to Eve, given knowledge of the measurement choice X, is the conditional von Neumann entropy

$$H(K_A|XE) = H(\rho_{K_AXE}) - H(\rho_{XE}) \tag{2}$$

where $H(\rho) = -\operatorname{Tr}[\rho \log_2(\rho)]$ is the von Neumann entropy and $\rho_{XE} = \operatorname{Tr}_{K_A}[\rho_{K_AXE}]$ where

$$\rho_{K_AXE} = \sum_{k_A,x} \mu(x)|k_A, x\rangle\langle k_A, x| \otimes \rho_E^{k_A,x} \tag{3}$$

is the classical-quantum state describing the correlations between K_A, X, and E. In the above expression, the reduced states of Eve are given by

$$\rho_E^{k_A,x} = \sum_{a=\pm1} p_x(k_A|a) \operatorname{Tr}_{AB}\left[\rho_{ABE} \frac{\mathbb{1} + aA_x}{2} \otimes \mathbb{1}_B \otimes \mathbb{1}_E\right] \tag{4}$$

where $p_x(k|a)$ are the transition probabilities of the map $\$_x$.

In this paper, we will often be interested in the case where K_A is simply obtained as the outcome of one of Alice's measurement, e.g., A_1 (i.e., there is no random input choice X and no classical preprocessing.) By a

slight abuse of notation, we write A_1 both for the random variable denoting the measurement outcome of A_1 and for the measurement A_1 itself. We thus write in such cases $K_A = A_1$ and $H(K_A|XE) = H(A_1|E)$. We will also consider noisy preprocessing [15, 16], where Alice's raw key bit K_A is again the outcome of the measurement A_1, but with probability q she flips it and with probability $1-q$ she keeps it as it is. We write $K_A = A_1^q$ for the corresponding random variable and thus $H(K_A|XE) = H(A_1^q|E)$ for the conditional entropy. Finally, the last case we will consider is one where K_A is obtained by choosing the observables A_1 and A_2 with probabilities p and $\bar{p} = 1-p$, respectively, and applying noisy preprocessing with flip probability q to the measurement output. We then write $K_A = A_X^q$ and $H(K_A|XE) = H(A_X^q|XE)$.

Alice-Bob correlations. In a device-independent setting, the correlations between Alice and Bob can be characterized through *Bell linear functionals*, which are linear functions of 1-body and 2-body correlators. In the 2-input/2-output scenario, 1-body and 2-body correlators can all be written in the common form

$$\langle A_x \otimes B_y \rangle = \mathrm{Tr}\big[\rho_{AB}\, A_x \otimes B_y\big] \quad \text{for } x = 0,1,2 \tag{5}$$

if we define $A_0 = \mathbb{1}_A$ and $B_0 = \mathbb{1}_B$. A Bell linear functional S is then specified by 9 real coefficients $\{S_{xy}\}_{x,y=0,1,2}$ $(x, y = 0, 1, 2)$ and its value on a given set of correlators $\{\langle A_x \otimes B_y\rangle\}$ is given by

$$S = \sum_{x,y=0}^{2} S_{xy} \langle A_x \otimes B_y \rangle \,. \tag{6}$$

We refer to S as a *Bell expectation*. We will particularly be interested in the following in the CHSH functional

$$S = \langle A_1 B_1 \rangle + \langle A_1 B_2 \rangle + \langle A_2 B_1 \rangle - \langle A_2 B_2 \rangle \,. \tag{7}$$

Tradeoff between Eve's information on the raw key and Alice-Bob correlations. Assume that a procedure for generating a raw key value (as specified by a measurement probability distribution μ_X and preprocessing maps $\$_x$) and a series of $m \geq 1$ Bell expectation values $\boldsymbol{S} = (S_1, \ldots, S_m)$[1] are fixed. Our objective is to establish a lower bound

$$H(K_A|XE) \geq f(\boldsymbol{S}) \tag{8}$$

that is device independent, in the sense that it is satisfied by every quantum strategy $\mathcal{Q}$. For technical reasons, we require f to be a convex function of its arguments[2].

[1] This can range from a single Bell functional, such as CHSH, to the entire set of correlators $\{\langle A_x \otimes B_y\rangle\}$, or anything in between.

[2] This is required for application of the entropy accumulation theorem, and follows naturally when reducing the analysis to qubits. Furthermore, if f defines a bound on $H(K|XE)$ that is tight, it must necessarily be convex by concavity of the conditional entropy and because any convex mixture of two strategies defines a valid strategy.

Relation to the security of DIQKD protocols. In a typical DIQKD protocol, Alice's and Bob's devices are successively used for n rounds. Some of the rounds are used to generate raw key values K_A on Alice's side and K_B on Bob's side. Some of the rounds are used to gather statistical data to decide, based on whether one or several Bell statistics are above some thresholds, if the protocol should be aborted or if it can proceed. In the latter case, error correction and privacy amplification are applied to the final raw key string. Following the application of the entropy accumulation theorem [7], the security of such a generic multi-round protocol can be reduced to deriving a tradeoff bound (8), which can be understood as characterizing the behavior of a single round[3] in expectation. In particular a tradeoff bound allows one to compute the key rate in the finite-key regime and in the asymptotic one, where it simply reduces to the Devetak-Winter formula [17]

$$r = H(K_A|XE) - H(K_A|K_B) \,, \tag{9}$$

where $H(K_A|K_B)$ is the conditional Shannon entropy of the classical random variables K_A and K_B.

2.1 Reduction to qubits

The lower bounds (8) we aim to derive must be proven valid for any quantum strategy $\mathcal{Q} = (\rho_{ABE}, A_1, A_2, B_1, B_2)$, defined a priori on Hilbert spaces of arbitrary dimension. However, because the strategies we consider involve only two binary measurements for Alice and for Bob, it is well-known that it is sufficient, thanks to Jordan's lemma, to consider pure qubit strategies [18].

More specifically, suppose that we have derived a lower bound $H(K_A|XE) \geq f(\boldsymbol{S})$, that is valid for any strategy $\mathcal{Q} = (|\Psi\rangle_{ABE}, A_1, A_2, B_1, B_2)$ where *i)* Alice's and Bob's systems are two-dimensional, *ii)* $|\Psi\rangle_{ABE}$ is a pure state, *iii)* A_1, A_2, B_1, B_2 are qubit, non-degenerate Pauli observables constrained to the Z–X plane on the Bloch sphere, and where *iv)* the function f is convex. Then this lower bound is valid for arbitrary strategies. For details, see for instance [14].

Note that the "2-input/2-output" restriction, which allows to make this qubit simplification, only applies to Alice's measurements and to those measurements of Bob that are involved in the definition of the Bell functionals $\boldsymbol{S}$, as these are the only measurements involved in the relation (8). The raw key generation procedure on Bob's side leading to the raw key

[3] The raw key generation procedure and the set of Bell statistics to be used in the single-round bound (8) should obviously coincide with those of the multi-round protocol.

value K_B can, however, involve further measurement choices with more outputs, see examples in the Section 3.

We now assume the above simplification and present our approach to deriving tradeoff bounds, which follows three technical steps described in the next three subsections.

2.2 BB84-type uncertainty relations

The first non-trivial step in our approach is *device-dependent* and consists in deriving a qubit uncertainty relation akin to those used in the analysis of the standard entanglement-based BB84 protocol and variants of it. Let us illustrate this on several examples. In the following, $\phi(x) = h\big(\frac{1}{2} + \frac{1}{2}x\big)$, where $h(x)$ is the binary entropy.

Consider first the simple situation where Alice's raw key bit $K_A = A_1$ is simply obtained as the outcome of the measurement A_1, i.e., there is no random input choice X and no classical preprocessing. We then have the following bound.

Entropy bound 1 (BB84).

$$H(A_1|E) \geq 1 - \phi\big(|\langle \bar{A}_1 \otimes B\rangle|\big)\,, \tag{10}$$

where $\bar{A}_1$ is a Pauli observable orthogonal to A_1 on the Bloch sphere and B any given ±1-valued observable on Bob's subsystem.

This bound is simply a reexpression of the one-sided device-independent entropy bound $H(\mathrm{Z}|E) \geq 1 - \phi\big(|\langle \mathrm{X} \otimes B\rangle|\big)$ for the BB84 protocol [19] that relates the information Eve has about the outcome of a Z measurement by how much Bob is correlated to the complementary X measurement. The bound (10) directly follows from the fact that A_1 and $\bar{A}_1$ are Pauli operators, which we can identify with the Z and X operators.

As a second example, let us add noisy preprocessing [15, 16] to the raw key procedure: Alice's raw key bit $K_A = A_1^q$ is again the outcome of the measurement A_1, but with probability q she flips it and with probability $1-q$ she keeps it as it is.

Entropy bound 2 (BB84 bound with noisy preprocessing).

$$H(A_1^q|E) \geq f_q(|\langle \bar{A}_1 \otimes B\rangle|)\,, \tag{11}$$

where

$$f_q(x) = 1 + \phi\Big(\sqrt{(1-2q)^2 + 4q(1-q)x^2}\Big) - \phi(x)\,, \tag{12}$$

and $\bar{A}_1$ is a Pauli observable orthogonal to A_1 on the Bloch sphere and B any given ±1-valued observable on Bob's subsystem.

This again follows by identifying A_1 and $\bar{A}_1$ with the Z and X operators and reusing a one-sided device-independent bound known for BB84 with noisy preprocessing [14, 20].

The two above bounds were used in [14] to analyze the security of a family of CHSH-based DIQKD protocols. But more generally, it is also possible to obtain other bounds, such as the two ones below, which we will apply to other variants of CHSH-based DIQKD protocols in Section 3.

Entropy bound 3 (BB84 with noisy preprocessing and bias).

$$H(A_1^q|E) \geq g_q\big(|\langle A_1\rangle|, |\langle \bar{A}_1 \otimes B\rangle|\big)\,, \tag{13}$$

where

$$g_q(z,x) = \phi\big(\tfrac{1}{2}(R_+ + R_-)\big) + \phi\big(\tfrac{1}{2}(R_+ - R_-)\big) - \phi\big(\sqrt{z^2+x^2}\big)\,, \tag{14}$$

with

$$R_\pm = \sqrt{(1-2q\pm z)^2 + 4q(1-q)x^2}\,, \tag{15}$$

and $\bar{A}_1$ is a Pauli observable orthogonal to A_1 on the Bloch sphere and B any given ±1-valued observable on Bob's subsystem.

This bound represents a refinement of the bound 2, as it depends not only on $\langle \bar{A}_1 \otimes B\rangle$, but also on the value of the 1-body correlator $\langle A_1\rangle$ measuring how much Alice's raw output is biased.

Our last example is one where Alice's raw key bit $K_A = A_X^q$ is obtained by choosing the observables A_1 and A_2 with probability p and $\bar{p} = 1-p$, respectively, and applying noisy preprocessing with flip probability q to the measurement output. The conditional entropy is then

$$H(A_X^q|XE) = pH(A_1^q|E) + \bar{p}H(A_2^q|E)\,, \tag{16}$$

and one has the following bound.

Entropy bound 4 (Two-basis bound).

$$H(A_X^q|XE) \geq f_q\Big(\sqrt{p\langle \bar{A}_1 \otimes B\rangle^2 + \bar{p}\langle \bar{A}_2 \otimes B'\rangle^2}\Big) \tag{17}$$

where $\bar{A}_1$ and $\bar{A}_2$ are observables orthogonal to A_1, A_2, respectively and $f_q(x)$ is the function defined in (12).

The above bounds are essentially similar to those used in the analysis of standard entanglement-based QKD. They are valid for arbitrary entangled states $|\Psi\rangle_{ABE}$ where Alice's and Bob's systems are two dimensional and are expressed in terms of correlators $\langle A \otimes B\rangle$ between Alice and Bob that involve (contrarily to the device-independent case) *specific, fixed* observables, such as $\bar{A}_1$ on Alice's side. As such they can be derived using existing techniques.

We remark that all of these bounds can be derived from bound 3, which we derive in detail in Appendix A. In particular, bound 2 is a special case of bound 3 evaluated with $\langle A_1\rangle = 0$[4], while bound 1 is obtained by further setting $q = 0$. Bound 4 follows from bounding both contributions to the average entropy separately using bound 2,

$$\begin{aligned} H(A_X^q|XE) &= pH_q(A_1|E) + \bar{p}H_q(A_2|E) \\ &\geq pf_q\big(|\langle \bar{A}_1 \otimes B\rangle|\big) + \bar{p}f_q\big(|\langle \bar{A}_2 \otimes B'\rangle|\big)\,, \end{aligned} \tag{18}$$

and then using that the function $x \mapsto f_q(\sqrt{x})$ is convex (see Appendix B of [14] for a proof of this property).

Importantly, we also show in Appendix A that all the above bounds satisfy a type of monotonicity property. We say that a bound $H(K_A|XE) \geq f(x)$ is *monotone* in x if the bound $H(K_A|XE) \geq f(x_-)$ holds for all $x_- \leq x$ and similarly in the multivariate case for each variable independently, e.g., $H(K_A|XE) \geq f(x, y)$ is *monotone* in x and y if the bound $H(K_A|XE) \geq f(x_-, y_-)$ hold for all $x_- \leq x$ and $y_- \leq y$. Note that this monotonicity property is weaker than monotonicity of the function f itself: if the function f is monotonically increasing then the bound $H(K_A|XE) \geq f(x)$ is monotone, but the converse does not necessarily hold.

Monotonicity property. *The entropy bounds* (10) *and* (11) *are monotone in* $|\langle \bar{A}_1 \otimes B\rangle|$*, the bound* (13) *is monotone in* $|\langle A\rangle_1|$ *and* $|\langle \bar{A}_1 \otimes B\rangle|$*, and the bound* (17) *is monotone in* $p\langle \bar{A}_1 \otimes B\rangle^2 + \bar{p}\langle \bar{A}_2 \otimes B'\rangle^2$*.*

The monotonicity of the bound (13) is established in Appendix A from which the monotonicity of the other bounds follows[5]. This property will be important in Section 2.3 as it allows replacing in the entropy bounds the correlators on which they depend in the right-hand side by a lower bound on these correlators and in Section 2.4 where it allows the systematic computation of a convex envelope based on a discrete set of points.

2.3 Pauli correlation constraints

The bounds on the conditional entropy $H(K_A|XE)$ that we have given in the previous subsection are expressed in terms of correlators involving observables which are not necessarily accessible through the devices, e.g., the correlator $\langle \bar{A}_1 \otimes B\rangle$ involving the observable $\bar{A}_1$. The second step of our approach consists in deriving a constraint on these correlators in terms of correlators involving only the observables A_1, A_2, B_1, B_2 *actually measured* by the devices.

For instance, it is a straightforward exercise, see [14], to show the following bound.

Correlation bound 1 (CHSH).

$$|\langle \bar{A}_1 \otimes B\rangle| \geq \sqrt{S^2/4 - 1}\,, \tag{19}$$

where $S = \langle A_1B_1\rangle + \langle A_1B_2\rangle + \langle A_2B_1\rangle - \langle A_2B_2\rangle$ *is the expected value of the CHSH statistic and* $B \propto B_1 - B_2$*.*

More generally, one can also consider a family of asymmetric versions of the CHSH statistic for which the following bounds are shown in [14].

Correlation bound 2 (asymmetric CHSH). *Let* $S_\alpha = \alpha\langle A_1B_1\rangle + \alpha\langle A_1B_2\rangle + \langle A_2B_1\rangle - \langle A_2B_2\rangle$ *be a variant of CHSH depending on a given parameter* $\alpha \in \mathbb{R}$*. Then for some appropriate choice of a* ± 1*-valued observable* B*,*

$$|\langle \bar{A}_1 \otimes B\rangle| \geq E_\alpha(S_\alpha)\,, \tag{20}$$

where

$$E_\alpha(S_\alpha) = \sqrt{{S_\alpha}^2/4 - \alpha^2} \tag{21}$$

if $|\alpha| \geq 1$ *or* $|S_\alpha| \geq 2\sqrt{1 + \alpha^2 - \alpha^4}$ *and*

$$E_\alpha(S_\alpha) = \sqrt{1 - \Big(1 - \tfrac{1}{|\alpha|}\sqrt{(1-\alpha^2)({S_\alpha}^2/4 - 1)}\Big)^2} \tag{22}$$

otherwise.

The correlation bounds (19) and (20) can be derived analytically. But more generically, one can derive numerical lower bounds on polynomial functions of arbitrary qubit correlators, such as $\langle \bar{A}_1 \otimes B\rangle$ or $\langle \bar{A}_2 \otimes B'\rangle$, in terms of Bell functionals involving only the accessible correlators $\langle A_x \otimes B_y\rangle$ $(x, y = 0, 1, 2)$, using the Lasserre hierarchy of semidefinite programming relaxations for polynomial optimization [21, 22]. This can be done by parameterizing explicitly all qubit operators in the Z–X plane.

We illustrate this general idea on the specific problem of deriving a lower bound for the expression

$$p\langle \bar{A}_1 \otimes B\rangle^2 + \bar{p}\langle \bar{A}_2 \otimes B'\rangle^2 \tag{23}$$

appearing on the right-hand side of (17) in terms of the CHSH expectation value S.

We first recall that we can use any ± 1-valued observables B and B' in (17). Taking these to be of the form

$$B^{(\prime)} = \cos\big(\varphi_B^{(\prime)}\big)\mathrm{Z} + \sin\big(\varphi_B^{(\prime)}\big)\mathrm{X} \tag{24}$$

and then choosing the angles φ_B and φ'_B that maximize (23) we obtain

$$\begin{aligned} &p\langle \bar{A}_1 \otimes B\rangle^2 + \bar{p}\langle \bar{A}_2 \otimes B'\rangle^2 \\ &\quad = p\big(\langle \bar{A}_1 \otimes \mathrm{Z}\rangle^2 + \langle \bar{A}_1 \otimes \mathrm{X}\rangle^2\big) \\ &\qquad + \bar{p}\big(\langle \bar{A}_2 \otimes \mathrm{Z}\rangle^2 + \langle \bar{A}_2 \otimes \mathrm{X}\rangle^2\big)\,. \end{aligned} \tag{25}$$

[4] The resulting bound holds independently of the actual value of $\langle A_1\rangle$ thanks to the monotonicity property discussed below: if we make in bound 3 the replacement $|\langle A_1\rangle| \mapsto 0$ we obtain a bound that remains valid.

[5] In the case of bounds (10), (11), (17), it also follows from the stronger property that the function $f_q(x)$ is monotonically increasing in x, as shown in Appendix B. of [14].

We then choose Alice's basis such that

$$A_1 = \cos\left(\tfrac{\varphi_A}{2}\right)\mathrm{Z} - \sin\left(\tfrac{\varphi_A}{2}\right)\mathrm{X}\,, \tag{26}$$
$$A_2 = \cos\left(\tfrac{\varphi_A}{2}\right)\mathrm{Z} + \sin\left(\tfrac{\varphi_A}{2}\right)\mathrm{X} \tag{27}$$

and the complementary operators are

$$\bar{A}_1 = \sin\left(\tfrac{\varphi_A}{2}\right)\mathrm{Z} + \cos\left(\tfrac{\varphi_A}{2}\right)\mathrm{X}\,, \tag{28}$$
$$\bar{A}_2 = -\sin\left(\tfrac{\varphi_A}{2}\right)\mathrm{Z} + \cos\left(\tfrac{\varphi_A}{2}\right)\mathrm{X} \tag{29}$$

for some unknown angle φ_A. Using these in the above expression we obtain, explicitly,

$$\begin{aligned} &p\langle \bar{A}_1 \otimes B\rangle^2 + \bar{p}\langle \bar{A}_2 \otimes B'\rangle^2 \\ &\quad = \sin\left(\tfrac{\varphi_A}{2}\right)^2 ({E_{\mathrm{zz}}}^2 + {E_{\mathrm{zx}}}^2) + \cos\left(\tfrac{\varphi_A}{2}\right)^2 ({E_{\mathrm{xz}}}^2 + {E_{\mathrm{xx}}}^2) \\ &\qquad + 2(2p-1)\sin\left(\tfrac{\varphi_A}{2}\right)\cos\left(\tfrac{\varphi_A}{2}\right)(E_{\mathrm{zz}}E_{\mathrm{xz}} + E_{\mathrm{zx}}E_{\mathrm{xx}})\,, \end{aligned} \tag{30}$$

where we note the expectation values of products of Pauli operators $E_{\mathrm{xx}} = \langle \mathrm{X}\otimes\mathrm{X}\rangle$ and similarly for E_{xz}, E_{zx}, and E_{zz}.

We wish to constrain (30) for a given value of the CHSH expectation value which, in the choice of basis made above, takes the form

$$\begin{aligned} S &= \langle (A_1 + A_2)\otimes B_1\rangle + \langle (A_1 - A_2)\otimes B_2\rangle \\ &= 2\cos\left(\tfrac{\varphi_A}{2}\right)\langle \mathrm{Z}\otimes B_1\rangle - 2\sin\left(\tfrac{\varphi_A}{2}\right)\langle \mathrm{X}\otimes B_2\rangle\,. \end{aligned} \tag{31}$$

Maximizing the second line over (nondegenerate) ±1-valued observables B_1 and B_2 in the Z–X plane gives

$$\begin{aligned} S/2 \leq\ & \left|\cos\left(\tfrac{\varphi_A}{2}\right)\right| \sqrt{{E_{\mathrm{zz}}}^2 + {E_{\mathrm{zx}}}^2} \\ &+ \left|\sin\left(\tfrac{\varphi_A}{2}\right)\right| \sqrt{{E_{\mathrm{xz}}}^2 + {E_{\mathrm{xx}}}^2}\,, \end{aligned} \tag{32}$$

which can be read as a constraint on the unknown angle φ_A and Pauli correlations E_{xx}, E_{xz}, E_{zx}, and E_{zz} appearing in (30).

To complete the problem, we finally remark that E_{xx}, E_{xz}, E_{zx}, and E_{zz} can be interpreted as expectations of products of the Z and X Pauli operators for some underlying state only if they satisfy

$${E_{\mathrm{zz}}}^2 + {E_{\mathrm{zx}}}^2 \leq 1\,, \tag{33}$$
$${E_{\mathrm{xz}}}^2 + {E_{\mathrm{xx}}}^2 \leq 1\,, \tag{34}$$

and

$$\begin{aligned} &\left(1 - {E_{\mathrm{zz}}}^2 - {E_{\mathrm{zx}}}^2\right)\left(1 - {E_{\mathrm{xz}}}^2 - {E_{\mathrm{xx}}}^2\right) \\ &\quad \geq \left(E_{\mathrm{zz}}E_{\mathrm{xz}} + E_{\mathrm{zx}}E_{\mathrm{xx}}\right)^2 \end{aligned} \tag{35}$$

as shown in Section 4.3 of [14].

To get a valid lower bound on (40), it is thus sufficient to minimize the left-hand side of (30) given the constraints (32)–(35). The problem can be simplified by introducing the new variables

$$E_{\mathrm{zz}} = \lambda\cos(z)\,, \qquad E_{\mathrm{zx}} = \lambda\sin(z)\,, \tag{36}$$
$$E_{\mathrm{xz}} = \mu\cos(x)\,, \qquad E_{\mathrm{xx}} = \mu\sin(x)\,, \tag{37}$$
$$s = \sin\left(\tfrac{\varphi_A}{2}\right)\,, \qquad c = \cos\left(\tfrac{\varphi_A}{2}\right)\,, \tag{38}$$
$$\Delta = \cos(x - z)\,. \tag{39}$$

Using the trigonometric identity $\cos\left(\tfrac{\varphi_A}{2}\right)^2 + \sin\left(\tfrac{\varphi_A}{2}\right)^2 = 1$ and that we can drop the absolute values from (32) without substantially changing the problem, we arrive at the following.

Correlation bound 3 (two-basis). *There exist ±1-valued qubit operators B and B' acting on Bob's subsystem such that*

$$p\langle \bar{A}_1 \otimes B\rangle^2 + \bar{p}\langle \bar{A}_2 \otimes B'\rangle^2 \geq E_p(S)^2\,, \tag{40}$$

where $E_p(S)^2$ is the solution to the minimization problem

$$\begin{aligned} E_p(S)^2 = \min \quad & s^2\lambda^2 + c^2\mu^2 + 2(2p-1)sc\lambda\mu\Delta \\ \text{s.t.} \quad & c\lambda + s\mu \geq S/2 \\ & \lambda^2 \leq 1 \\ & \mu^2 \leq 1 \\ & (1-\lambda^2)(1-\mu^2) \geq \lambda^2\mu^2\Delta^2 \\ & c^2 + s^2 = 1 \\ & \Delta^2 \leq 1 \end{aligned} \tag{41}$$

in the five variables $\lambda, \mu, c, s, \Delta \in \mathbb{R}$.

As the above is a polynomial optimization problem, it can be reduced to a sequence of semidefinite programs using the Lasserre hierarchy [21, 22]. Importantly, every SDP relaxation at a given order in the hierarchy provides a valid lower bound to the optimization problem and consequently a valid lower bound of the form (40). At level 3 of the Lasserre hierarchy, the problem takes less than a second to solve and appears to already give the optimal solution.

In the case in which $p = 1/2$, the above problem can actually be solved analytically, as shown in Appendix B. The result in that case is

$$E_{\frac{1}{2}}(S)^2 = \frac{1 + {x_*}^2}{1 - x_*} + \frac{S^2}{4}\frac{1 + x_*}{1 - x_*} - \frac{S}{\sqrt{2}}\frac{(1 + x_*)^{3/2}}{1 - x_*}\,, \tag{42}$$

where the variable x_* is the solution of

$$4x(2 - x) + 2(S^2 + 2) + S(x - 5)\sqrt{2(1 + x)} = 0 \tag{43}$$

in the range

$$-\frac{S}{4}\sqrt{8 - S^2} \leq x \leq \frac{S}{4}\sqrt{8 - S^2}\,. \tag{44}$$

Eq. (43) can be rearranged to a root-finding problem for a degree 4 polynomial in x and can thus be solved analytically, though the solution is quite lengthy and we do not explicitly report it here.

2.4 Convexity and fully device-independent bounds

Combining the above correlation bounds and the entropy bounds of the previous section, one obtains

bounds on the conditional entropy that are device independent modulo the qubit reduction. For instance, using the CHSH correlation bound (19) in the BB84 entropy bound (10), where the substitution of (19) in (10) is possible thanks to the monotonicity property of the BB84 entropy bound, we recover the CHSH entropy bound

$$H(A_1|E) \geq 1 - \phi\big(\sqrt{S^2/4 - 1}\big) \tag{45}$$

given in the introduction and originally derived in [3]. Using (20) in the BB84 bound with noisy preprocessing (11), one obtains the more general qubit bound

$$H(A_1^q|E) \geq f_q\big(E_\alpha(S_\alpha)\big) \tag{46}$$

derived in [14].

But other combinations are also possible, such as the two original following ones, which we are going to consider in more detail in Section 3.

The first, which gives a bound on the entropy in terms of $\langle A_1 \rangle$ in addition to CHSH, is simply obtained by combining (19) and (13):

$$H(A_1^q|E) \geq g_q\big(|\langle A_1 \rangle|, \sqrt{S^2/4 - 1}\big)\,. \tag{47}$$

For the second, let $\tilde{E}_p(S)^2$ denote any lower bound to $E_p(S)^2$ obtained by solving analytically or numerically the polynomial optimization problem (41) or any of its relaxations in the Lasserre hierarchy. Then using such a bound in (17), we obtain

$$H(A_X^q|XE) \geq f_q\big(\tilde{E}_p(S)\big) \tag{48}$$

with $\tilde{E}_p(S) \equiv \sqrt{\tilde{E}_p(S)^2}$.

2.4.1 Convexity analysis

Regardless of the combination used, the result is a bound on the conditional entropy valid for two-qubit systems, which can only be extended to give a fully device-independent bound, valid in arbitrary dimension, if it is convex. The third and final step thus consists of a convexity analysis.

If we obtain a qubit bound on the conditional entropy with a reasonably simple analytic expression then it may be feasible to study its properties directly. Either we simply prove it is convex, as can be done for (45), or more generally as was done in [14] for (46) for $|\alpha| \geq 1$. Or we analytically establish that it is not convex and determine its convex envelope, as was done in [14] for (46) for $|\alpha| < 1$.

More generally, however, the qubit bound may be obtained numerically or it may be analytic but of a form that does not easily lend itself to an analytic convexity analysis, as is the case for the bounds (47) and (48). In such cases, we need a way of constructing a convex lower bound on whatever qubit bound we obtain.

2.4.2 Convex lower bounds through linear programming

A simple solution that we can use, provided our entropy bounds satisfy the monotonicity property introduced in subsection 2.2, is based on a discretization of the qubit bound, similar to the approach used in [10]. In the following, let us generically write the bound valid for two-qubit systems as

$$H(K_A|XE) \geq f(\boldsymbol{S})\,, \tag{49}$$

where $f\colon \mathcal{D} \to \mathbb{R}$ is a function, defined on some domain $\mathcal{D}$, that we either know analytically or can compute numerically, of one or more Bell expectation values $\boldsymbol{S} = (S_1, S_2, \ldots, S_n) \in \mathcal{D}$.

Let us introduce a covering $\mathcal{K} = \{K\}$ of the domain $\mathcal{D}$ by polytopes $\{K\}$, such that every $\boldsymbol{S} \in \mathcal{D}$ is contained in at least one of the polytopes K. In practice, we would typically use a grid partition in terms of hyperrectangles where each point (outside of vertices and shared edges) is contained in only one hyperrectangle K (but this is not strictly necessary for the method to work).

Let us suppose, furthermore, that for every K we have a way of identifying a value $f[K]$ that we can use as a lower qubit bound on the conditional entropy valid for the entire polytope, i.e., such that

$$H(K_A|XE) \geq f[K]\,, \qquad \forall \boldsymbol{S} \in K\,. \tag{50}$$

We can then define a discretized qubit bound,

$$H(K_A|XE) \geq f_{\mathcal{K}}(\boldsymbol{S}) \tag{51}$$

where $f_{\mathcal{K}}$ is defined as

$$f_{\mathcal{K}}(\boldsymbol{S}) = \min_{K \ni \boldsymbol{S}} f[K]\,, \tag{52}$$

where the minimization is taken over all polytopes K that contain $\boldsymbol{S}$. This, in particular, associates unique values $f_{\mathcal{K}}(\boldsymbol{S}_j)$ to the vertices $\boldsymbol{S}_j$ of the polytopes. The convex envelope of the discretized function $f_{\mathcal{K}}$, finally, is readily given by the solution to the following linear programming problem,

$$\begin{aligned} \bar{f}_{\mathcal{K}}(\boldsymbol{S}) = \ & \text{minimize} \sum_j \theta_j\, f_{\mathcal{K}}(\boldsymbol{S}_j) \\ & \text{subject to} \sum_j \theta_j \boldsymbol{S}_j = \boldsymbol{S} \\ & \sum_j \theta_j = 1 \\ & \theta_j \geq 0\,, \end{aligned} \tag{53}$$

where the $\boldsymbol{S}_j$ are the combined vertices of all the polytopes K in $\mathcal{K}$. We thus obtain a bound

$$H(K_A|XE) \geq \bar{f}_{\mathcal{K}}(\boldsymbol{S}) \tag{54}$$

on the conditional entropy that is convex and extends to the fully device-independent setting.

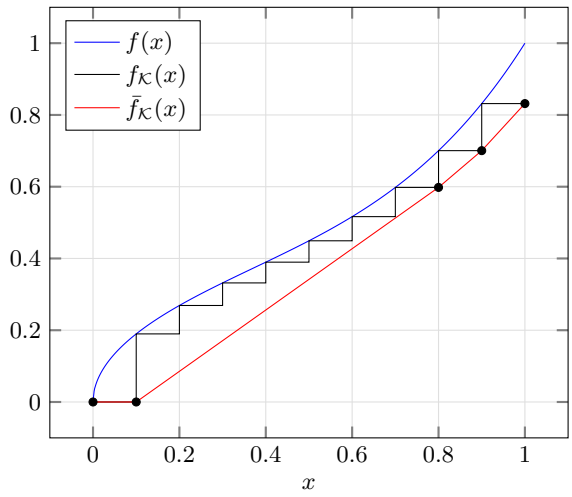

Figure 1: Convex lower bound $\bar{f}_{\mathcal{K}}$ of a function f constructed on n (in the figure $n = 10$) equally spaced subdivisions of its domain, i.e., the polytopes K are here n consecutive line segments between $x = 0$ and $x = 1$. We actually used this method on the qubit bound (48), but the function $f_q(\tilde{E}_p(S))$ is too close to convex to make a visually interesting example. The construction is thus illustrated on the figure for the visibly non-convex function $f(x) = 0.6\sqrt{x} + 0.4x^4$.

We have not explained, however, how one can identify in (50) the lower-bound values $f[K]$ for each polytope K, which is crucial to define a discretized qubit bound. This can be done if the bound (49) is monotone in $|\boldsymbol{S}| = (|S_1|, |S_2|, \ldots, |S_n|)$, i.e., if the bound still holds if we replace in (49) any of the n Bell expectation values S_i by a value s_i that is smaller in absolute value, $|s_i| \leq |S_i|$. This is in particular the case for all the bounds (45)–(48) presented above since they are obtained by combining the monotone entropy bounds of subsection 2.2 with the monotonically increasing correlation bounds of subsection 2.3. Using this monotonicity property, we can now simply divide the domain $\mathcal{D}$ into hyperrectangles K and use as the lower-bound value $f[K]$ for each hyperrectangle K, the value of the qubit bound evaluated at the corner that is closest to the origin.

Finally, in the special case that we are working with a qubit entropy bound $H(K_A|XE) \geq f(S)$ of a single variable S, we remark that one can avoid the linear program and compute $f_{\mathcal{K}}(S)$ very rapidly essentially by eliminating the redundant vertices and interpolating between the remaining ones, as illustrated in Figure 1. This can be done in linear time in the number of vertices [23, 24]. We in particular applied this technique to the two-basis bound (48) to compute the key-rate bounds obtained in Section 3.1 below.

2.4.3 Certifying an affine tradeoff bound

While we can always use the above approach when we have a qubit entropy bound satisfying the monotonicity property, it is not always necessary to solve the linear programming problem to obtain a valid convex lower bound on the conditional entropy. An alternative approach, which would ultimately lend itself to more direct use in the entropy accumulation theorem, is to certify a linear or affine lower bound on the entropy.

Here, let us suppose we believe that the conditional entropy respects an affine lower bound

$$H(K_A|XE) \geq \beta + \boldsymbol{\alpha} \cdot \boldsymbol{S} - \varepsilon\,, \tag{55}$$

that we wish to certify up to some precision ε. Such a bound may be obtained, for example, by computing at a particular point the tangent of a function $\bar{f}(\boldsymbol{S})$ that we believe to be the convex hull of a known qubit bound $f(\boldsymbol{S})$. As above, we introduce a covering $\mathcal{K} = \{K\}$ of the domain $\mathcal{D}$ with polytopes K and assume for every K a lower bound $f[K]$ on the conditional entropy, as defined in (50). We also define

$$\begin{aligned} \alpha[K] &= \max_{\boldsymbol{S} \in K} \boldsymbol{\alpha} \cdot \boldsymbol{S} \\ &= \max_{\boldsymbol{S} \in \mathrm{Vert}(K)} \boldsymbol{\alpha} \cdot \boldsymbol{S} \end{aligned} \tag{56}$$

where $\mathrm{Vert}(K)$ are the vertices of K. To check that (55) holds, we then only need to verify that

$$\beta + \alpha[K] - f[K] \leq \varepsilon \tag{57}$$

holds for all polytopes K in the covering $\mathcal{K}$, which is now a finite problem. Alternatively, we can compute the maximal value over $\mathcal{K}$ of $\beta + \alpha[K] - f[K]$ to determine the best possible precision ε we can achieve given our covering choice.

An important difference with the linear programming approach above is that we do not necessarily have to decide on a covering $\mathcal{K}$ in advance. In fact, this is often very wasteful as, to obtain a good bound with a small tolerance, we would typically find we need a fine discretization of the domain only close to where the bound coincides with its tangent. Finding a suitable discretization can then be done naturally, and in practice often very rapidly, by starting by testing (57) for the polytopes K in an initially coarse covering (which could consist of just one polytope containing the entire domain) and then, for each K for which the test fails, subdividing K into smaller polytopes and recursively applying the test to each of those (see illustration in Figure 2).

Application to the bound (47) including the bias $\langle A_1 \rangle$. We used this recursive certification method, coupled with a guess on the optimal linear tradeoff functions, for the qubit bound (47) which depends on the two variables $\langle A_1 \rangle$ and S. The function $\tilde{g}_q(\langle A_1 \rangle, S) \equiv g_q\big(|\langle A_1 \rangle|, \sqrt{S^2/4 - 1}\big)$ defining this bound is not convex as its Hessian matrix is not positive semidefinite everywhere. It appears, though, to be convex in each of the parameters $\langle A_1 \rangle$

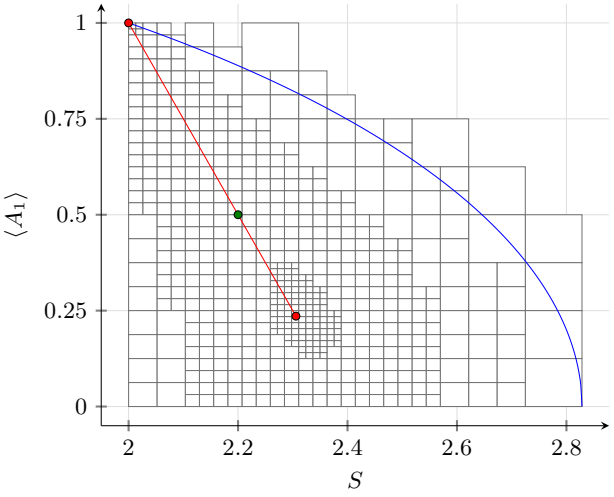

Figure 2: Certification of an affine lower entropy bound based on the qubit bound (47) depending on the CHSH expectation value S and the one-body correlator $\langle A_1 \rangle$. The blue curve represents the boundary of the domain $\mathcal{D} \subset [0,1] \times [2, 2\sqrt{2}]$ where the values of $(\langle A_1 \rangle, S)$ are consistent with quantum theory. We conjecture that the convex envelope of the function $\tilde{g}_q(\langle A_1 \rangle, S) = g_q\big(|\langle A_1 \rangle|, \sqrt{S^2/4 - 1}\big)$ in $\mathcal{D}$ is obtained by taking a convex decomposition of the point $(1, 2)$ and a point on the line from $(1, 2)$ to $(\langle A_1 \rangle, S)$. The figure illustrates such a convex decomposition (red points) for the point $(0.5, 2.2)$ (green point). From this, we can compute a candidate affine function (55) that optimally certifies the entropy of the point $(0.5, 2.2)$. Setting a value for ε, we then run a recursive algorithm to find a rectangle covering, depicted in the figure, that certifies the candidate affine function. We chose a value $\varepsilon = 0.025$ such that the resultant covering is coarse enough that it can be visualized, but much smaller values, e.g., $\varepsilon \approx 10^{-8}$ or less can readily be used.

and S individually, and more generally in any direction passing through the positive orthant in the plane $\langle A_1 \rangle$–S. This implies that the convex envelope of $\tilde{g}_q(\langle A_1 \rangle, S)$ can be constructed by considering at most convex combinations of *two* points in the plane, instead of three points as follows by Carathéodory's theorem. Indeed, any non-trivial convex combination of three points in the plane $\langle A_1 \rangle$–S would have at least two of those points joined by a segment aligned in the direction of the positive orthant. But since the function is convex in that direction, one can advantageously replace the two points by a mixture of those.

Furthermore, if we are interested in computing a valid entropy bound for a point with $\langle A_1 \rangle$ positive, it is sufficient to consider convex combinations in the domain $\mathcal{D} \subset [0,1] \times [2, 2\sqrt{2}]$ of the plane $\langle A_1 \rangle$–S, i.e., points with negative values of $\langle A_1 \rangle$ can be neglected. To see this, consider a convex combination

$$(\langle A_1 \rangle, S) = t\,(\langle A_1 \rangle', S') + (1-t)\,(\langle A_1 \rangle'', S'') \tag{58}$$

where $\langle A_1 \rangle' < 0$ is negative for the point $(\langle A_1 \rangle, S)$ yielding a corresponding value for the entropy function

$$t\,\tilde{g}_q\big(\langle A_1 \rangle', S'\big) + (1-t)\,\tilde{g}_q\big(\langle A_1 \rangle'', S''\big) \tag{59}$$

that is a valid lower bound for $H(A_1^q|E)$. Replace now this convex strategy by the (valid) convex combination

$$(\langle A_1 \rangle, S) = t\,(0, S') + (1-t)\left(\frac{\langle A_1 \rangle}{1-t},\, S''\right). \tag{60}$$

The corresponding value for the entropy function is

$$t\,\tilde{g}_q(0, S') + (1-t)\,\tilde{g}_q\left(\frac{\langle A_1 \rangle}{1-t},\, S''\right), \tag{61}$$

which is still a valid lower bound for $H(A_1^q|E)$ because of the monotonicity property of the bound and the fact that $\frac{\langle A_1 \rangle}{1-t} \leq \langle A_1 \rangle''$ (since $\langle A_1 \rangle' < 0$).

Finally, we numerically observed that the convex envelope of $\tilde{g}_q(\langle A_1 \rangle, S)$ in the domain $[0,1] \times [2, 2\sqrt{2}]$ was always obtained by taking a convex decomposition of two particular points: the point $(1, 2)$ and a point on the line from $(1, 2)$ to $(\langle A_1 \rangle, S)$. This observation gives a conjecture for the convex envelope of the qubit bound (47), from which candidate linear tradeoff functions of the form (55) can readily be computed as tangents to this envelope. We can then attempt to certify that such candidates are indeed proper tradeoff functions through a rectangle covering and the recursive procedure described above, as illustrated in Figure 2. We can in principle perform such certification to arbitrary precision ε, though, in practice, we may be limited by the number of rectangles required to reach a very small ε and by the limited precision of hardware floating-point arithmetic on typical computers. The key rates and results presented in Section 3.2 have been computed using this procedure. From our results, it appears that our conjecture on the convex envelope of $\tilde{g}_q(\langle A_1 \rangle, S)$ is correct as we are always able to certify the resultant linear tradeoff functions up to a precision of the order of $\varepsilon \approx 10^{-6}$ or better.

3 Applications

Here, we apply our method to bound the asymptotic one-way key rate, given by the Devetak-Winter rate

$$r = H(K_A|XE) - H(K_A|K_B)\,, \tag{62}$$

for DIQKD in two situations of interest: white noise, where we assume that Alice and Bob share an attenuated version,

$$\rho = v\phi^+ + (1-v)\mathbb{1}/4\,, \tag{63}$$

depending on some visibility v, of the ideal maximally-entangled state

$$|\phi^+\rangle = \frac{1}{\sqrt{2}}\big(|00\rangle + |11\rangle\big)\,, \tag{64}$$

and limited detection efficiency, where we assume that Alice's and Bob's devices return one of the expected outcomes ± 1 with a probability η less than one.

The qubit bound (45) (which is already convex) was used in [3] to compute the key rate of the standard CHSH DIQKD protocol and the convexification of (46) was used in [14] to generalize the analysis in terms of the asymmetric CHSH expressions S_α and incorporating noisy preprocessing. We will now illustrate the use of the two other qubit bounds (47) and (48) given in the preceding section, in subsections 3.2 and 3.1, respectively.

In [14], the asymmetric CHSH expressions were chosen for parameter estimation because they retain the same symmetries as the version of the DIQKD protocol where only one of Alice's measurements, A_1, is used to generate the key and they can be used to derive the optimal one-way key rate for that protocol with respect to white noise. There is no analogous connection between the asymmetric CHSH expressions and losses and, in fact, the lowest threshold, $\eta \approx 82.57\%$, on the global detection efficiency reported in [14] was obtained using CHSH (the special case of S_α with $\alpha = 1$).

In the following, we reanalyze these correlation models using different setups. In particular, as [14] already does an optimal analysis for white noise using one measurement basis for key generation and with noisy preprocessing, the only remaining way to improve the noise robustness is to use a different protocol. For that case, we apply our approach to a variant of the protocol based on CHSH, proposed recently in [10], in which both of Alice's measurements A_1 and A_2 are used to generate the key. For losses, by contrast, as remarked in [14] the analysis performed there was likely not optimal as the treatment of losses introduced biases in the probabilities of Alice's and Bob's measurement outcomes, while the analytic bound on the entropy used there was optimized for the case that Alice's outcomes are obtained equiprobably. For losses, therefore, we concentrate on bounding the key rate using the expectation value $\langle A_1 \rangle$ of Alice's key-generation measurement in addition to the Bell violation.

3.1 White noise analysis for the two-basis protocol

In the two-basis protocol of [10], Alice and Bob ideally share a maximally-entangled state $|\phi^+\rangle$ and have devices that, for Alice, ideally perform the two measurements

$$A_1 = \mathrm{Z}\,, \qquad A_2 = \mathrm{X}\,, \tag{65}$$

and, for Bob, the four measurements

$$B_1 = \frac{\mathrm{Z}+\mathrm{X}}{\sqrt{2}}\,, \qquad B_3 = \mathrm{Z}\,, \tag{66}$$

$$B_2 = \frac{\mathrm{Z}-\mathrm{X}}{\sqrt{2}}\,, \qquad B_4 = \mathrm{X}\,. \tag{67}$$

This ideal realization is designed so that the measurements A_1, A_2, B_1, and B_2 yield a maximal violation of the CHSH Bell inequality while Bob's measurements B_3 and B_4 yield outcomes that are perfectly correlated with Alice's when she measures, respectively, A_1 and A_2, i.e., $\langle A_1 B_3 \rangle = \langle A_2 B_4 \rangle = 1$.

In the protocol, Alice and Bob use rounds where Bob measures B_1 or B_2 to estimate CHSH; they use a small fraction of the rounds where Bob measures B_3 and B_4 to estimate how correlated the outcomes are with A_1 and A_2, and use the results of the remaining rounds where Alice and Bob measured A_1 and B_3 or A_2 and B_4 as their raw key. We also assume in the following that Alice flips her outcomes in the key generation rounds (i.e., applies noisy preprocessing) with some probability q.

Let us suppose that Alice uses the measurements A_1 and A_2 with probabilities p' and $\bar{p}' = 1 - p'$ and that Bob uses the measurements B_3 and B_4 with the same relative probabilities. Then, out of the rounds not used for parameter estimation, the asymptotic key rate, taking into account the effect of sifting[6], is

$$\begin{aligned} r &= p'^2\, r_{13} + \bar{p}'^2\, r_{24} \\ &= (p'^2 + \bar{p}'^2)(p\, r_{13} + \bar{p}\, r_{24})\,, \end{aligned} \tag{68}$$

where

$$r_{xy} = H(A_x^q|E) - H(A_x^q|B_y) \tag{69}$$

and we introduced $p = p'^2/(p'^2 + \bar{p}'^2)$ and $\bar{p} = 1 - p$ in the second line. Here, $H(A_1^q|B_3)$ and $H(A_2^q|B_4)$ depend only on the correlations between Alice's and Bob's measurement outcomes, which they know from parameter estimation. Assuming Alice and Bob perform the ideal measurements on an attenuated state (63), the entropies of Alice's outcomes conditioned on Bob are

$$H(A_1^q|B_3) = H(A_2^q|B_4) = h\big(q + \delta(1-2q)\big)\,, \tag{70}$$

where the channel error rate δ is related to the visibility v in (63) by $v = 1 - 2\delta$, while the CHSH expectation value is

$$S = 2\sqrt{2}(1-2\delta)\,. \tag{71}$$

Bounding the key rate thus amounts to establishing a lower bound on the weighted average conditional entropy

$$pH(A_1^q|E) + \bar{p}H(A_2^q|E) = H(A_X^q|XE) \tag{72}$$

depending on the CHSH violation. A valid *qubit* bound in terms of the CHSH expectation value S is given by (48), from which a valid, fully device-independent, convex lower bound can be obtained using the techniques discussed in Section 2.4.2.

[6] In particular, the key rate is attenuated by the probability $p'^2 + \bar{p}'^2$ that Alice and Bob use matching bases. It has been pointed out in [11] that this can be avoided, but this requires the parties to either possess quantum memories or to use a very long preshared key to coordinate the measurement choices.

We can thus express the bound we obtain on the key rate, via CHSH, in terms of δ using our approach as

$$r \geq (p'^2 + \bar{p}'^2)\Big[\tilde{f}_q\big(2\sqrt{2}(1-2\delta)\big) - h\big(q+\delta(1-2q)\big)\Big], \tag{73}$$

where $\tilde{f}_q(S)$ is the convex lower bound we obtain for the entropy, evaluated at $S = 2\sqrt{2}(1-2\delta)$.

We remark here that we could, in principle, bound the average entropy in terms of any correlation Bell inequality. We use only the CHSH expectation value here both for simplicity and because, in the most interesting case where the bases are used equiprobably (i.e., $p = 1/2$), we can infer from the symmetries of the protocol that CHSH is already the optimal measure of nonlocality for white noise (see Appendix C for details).

The key rate we obtain using our approach for $p = 0.5$ and $p = 0.75$ are illustrated, and compared with the known analytical bounds for $p = 1$, without noisy preprocessing (i.e., $q = 0$) and with the optimal amount of noisy preprocessing applied in Figures 3 and 4. The threshold noise rates up to which we obtain a positive key rate are reported for different values of q in Table 1. For $q = 0$ and q close to $1/2$, the results essentially rigorously confirm the thresholds of 8.36% and 9.24% that were anticipated could be obtained in the conclusion of [14]. For $0 < p < 1/2$, similar to [10], we did not see any improvement to the key rate; the highest rate appeared to always be obtained with either $p = 1$ or $p = 1/2$, depending on the value of S. However, as it may not be realistic to be sure that the measurements are used *exactly* equiprobably in a real implementation, we note that it is important to be able to bound the entropy for values of p that may deviate a little from 0.5. The key rate is in fact very robust against deviations of p from 0.5, as can be seen comparing the results for $p = 0.5$ and $p = 0.75$ in Figures 3 and 4.

The best threshold of 9.24% obtained for q close to $1/2$ using our method is close to the best threshold of 9.33% recently reported in [11] and obtained for $q = 0.3$, although the method we have used allows the key rate to be bounded much more rapidly[7]. Without noisy preprocessing, the threshold of 8.36% we obtain is slightly better than the threshold around 8.24% found in [10] and the same as the threshold that would be obtained using the "conjectured alternative proof" (after taking the convex envelope of the result) proposed in section I.H of the supplementary information to the same paper[8].

[7]Ref. [11] reports requiring ~ 5000 processor-core hours to obtain a numerical bound on the average conditional entropy. For comparison, using our method we could generate a plot of the conditional entropy with 500 points in a minute or two on a regular laptop using the Lasserre hierarchy or almost instantaneously using the analytic method for $p = 1/2$ described in Appendix B.

[8]This is not a coincidence. The section in question proposes to bound the key rate using a lower bound on the conditional entropy in terms of the fidelity of Eve's marginal states. This is very closely related to the BB84 bound [25] and, in fact, all of the lower bounds we derive on the correlation terms $|\langle \bar{A}_x \otimes B\rangle|$ appearing in the BB84 bounds we use are also (typically tight) lower bounds on the fidelity of Eve's marginals following the qubit reduction.

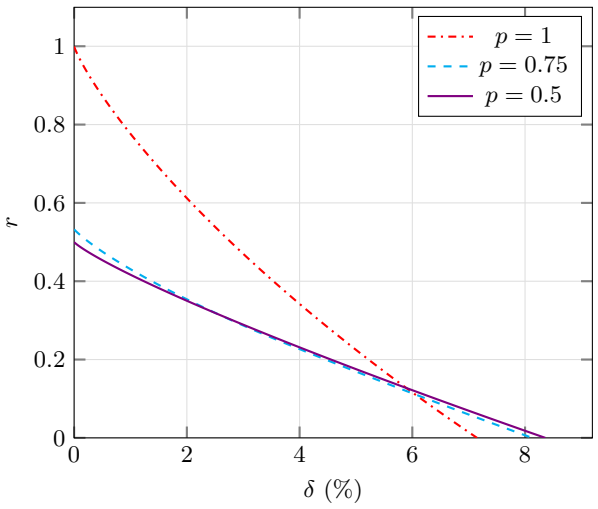

Figure 3: Lower bound on the Devetak-Winter rate as a function of the channel error rate δ, assuming $q = 0$.

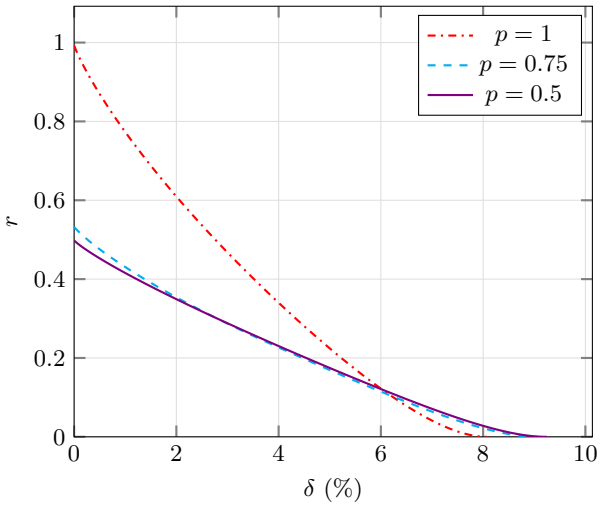

Figure 4: Lower bound on the Devetak-Winter rate as a function of the channel error rate δ, using an optimal noisy preprocessing.

p	$q = 0$	$q = 0.2$	$q = 0.3$	$q = 0.49$	$q \to 1/2$
1	7.1492	7.9503	8.0321	8.0848	8.0848
0.5	8.3599	9.1130	9.1923	9.2434	9.2435

Table 1: Threshold error rates (%) obtained for different probabilities p of measuring A_1 after sifting non-matching basis.

We provide an indication of how close the key-rate bound we obtain in the case $p = 1/2$ is to being optimal by comparing with a specific strategy, which was already identified as a likely candidate for the optimal collective attack for $q = 0$ in [14], and described in Appendix D. This attack yields the following value for the average entropy

$$\tfrac{1}{2}H(A_1^q|E) + \tfrac{1}{2}H(A_2^q|E) = \bar{f}_q\big(S/\sqrt{8}\big)\,, \tag{74}$$

where

$$\bar{f}_q(x) = \begin{cases} f_q(x) & \text{if } x \geq x_* \\ h(q) + f_q'(x_*)(x - 1/\sqrt{2}) & \text{if } x \leq x_* \end{cases} \tag{75}$$

with x_* (dependent on q) such that

$$h(q) + f_q'(x_*)(x - 1/\sqrt{2}) = f_q(x_*)\,, \tag{76}$$

and where $f_q(x)$ is defined in Eq. (12).

The results of numerical tests done without noisy preprocessing in [14] and [26] strongly suggest that (74) actually gives the optimal bound on the average entropy for $q = 0$. Additional tests we did for this work did not find a counterexample for $q \neq 0$. But even without a proof of optimality, as (74) is obtained with a known collective attack it gives an upper bound on the one-way asymptotic key rate with noisy preprocessing. A comparison of the key rates, optimized over q, using our numerical lower bound (already given in Figure 4) and using (74) is given in Figure 5 and shows the two to be very close. The threshold error rate obtained using (74) ranges from $\delta \approx 8.4447\%$ for $q = 0$ up to $\delta \approx 9.4756\%$ for $q \to 1/2$, and is compared with the threshold obtained using our numerical method in Figure 6.

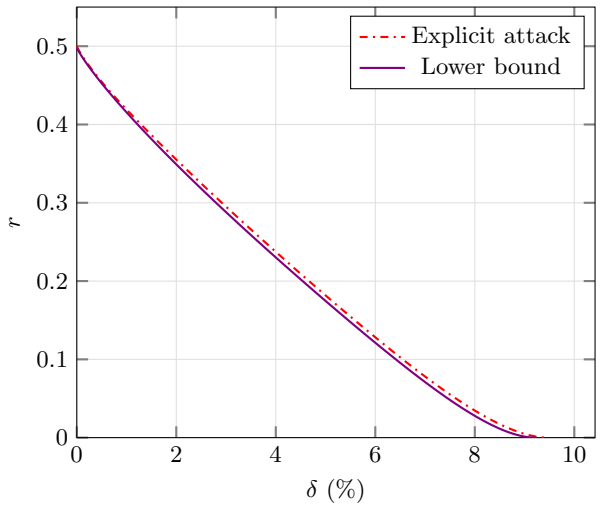

Figure 5: Comparison between the conjectured optimal attack and the lower bound on the Devetak-Winter rate as a function of the channel error rate δ, using an optimal noisy preprocessing.

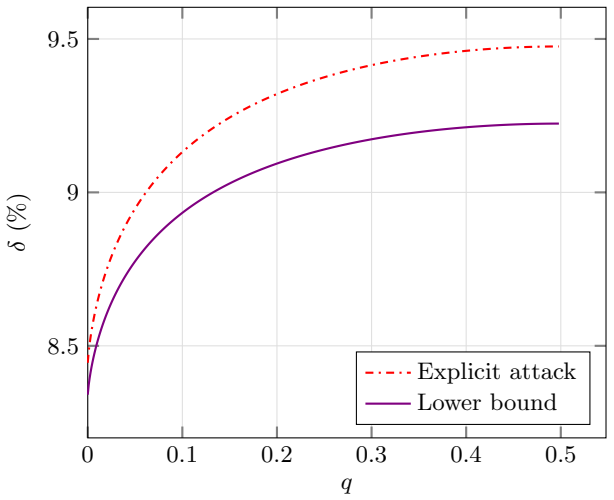

Figure 6: Thresholds for the channel error rate as a function of the noisy preprocessing computed using the conjectured optimal attack and our lower bound on the conditional entropy.

3.2 More refined loss analysis exploiting bias

Here, we consider a setup where we suppose that the main imperfection is that Alice's and Bob's devices have a detection efficiency that is less than perfect, i.e., we suppose that, in each protocol round, each of their devices outputs one of the regular outcomes ± 1 with probability η and outputs nothing, or a "nondetection" outcome $\emptyset$, with probability $1 - \eta$. In order to use our approach, which strictly applies to protocols in which the measurements in the Bell test have binary outcomes, we map nondetection events resulting from the measurements A_1, A_2, B_1, and B_2 used to perform the Bell test to $+1$.

In this case we consider the usual, single-basis, version of the DIQKD protocol, but with different states and measurements. Similar to the Eberhard scheme [27], we suppose that Alice and Bob (ideally) share a partially-entangled two-qubit state

$$|\psi_\theta\rangle = \cos\big(\tfrac{\theta}{2}\big)|00\rangle + \sin\big(\tfrac{\theta}{2}\big)|11\rangle\,, \tag{77}$$

and that Alice and Bob (ideally) perform, respectively, two and three measurements

$$A_x = \cos(\varphi_{A,x})\mathrm{Z} + \sin(\varphi_{A,x})\mathrm{X}\,, \quad x = 1, 2 \tag{78}$$

$$B_y = \cos(\varphi_{B,y})\mathrm{Z} + \sin(\varphi_{B,y})\mathrm{X}\,, \quad y = 1, 2, 3\,, \tag{79}$$

determined by angles $\varphi_{A,x}$ and $\varphi_{B,y}$ that we will optimize over when bounding the key rate[9]. Alice and Bob use the measurements A_1, A_2, B_1, and B_2 to estimate the CHSH expectation value and use A_1 and B_3 to generate the key.

As we are only considering the usual single-basis version of the protocol, the asymptotic key rate is

$$r = H(A_1^q|E) - H(A_1^q|B_3) \tag{80}$$

[9]Note that this is a slight generalization with respect to [14], which fixed A_1 and B_3 to Z.

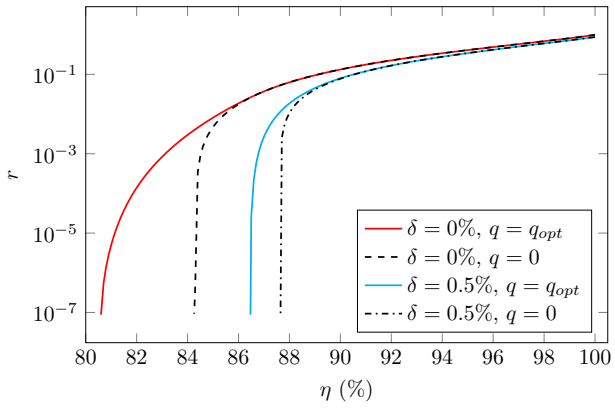

Figure 7: Key rate as a function of the detection efficiency with no channel error rate and with a little error rate.

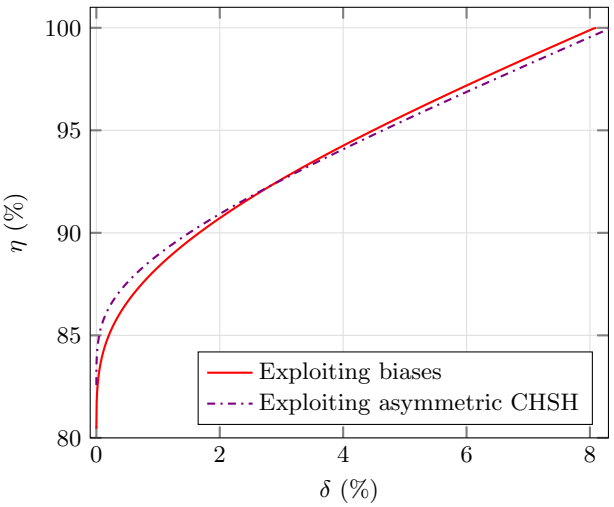

Figure 8: Threshold detection efficiency η as a function of the channel error rate δ.

where the Shannon entropy of Alice's outcome conditioned on Bob,

$$H(A_1^q|B_3) = -\sum_{a,b} p(a,b)\log_2\big(p(a|b)\big)\,, \tag{81}$$

depends on the joint probability $p(a,b)$ that Alice obtains the outcome $a \in \{+1,-1\}$ from measuring A_1 after mapping nondetection events to $+1$ and flipping the result with probability q, and Bob obtains the outcome $b \in \{+1,-1,\emptyset\}$ from measuring B_3 and possibly obtaining the loss outcome $\emptyset$ with probability $1-\eta$.

To bound the key rate we need to bound $H(A_1^q|E)$. As mentioned above, mapping nondetection events deterministically to $+1$ and deliberately using a partially-entangled state bias Alice's and Bob's measurements to giving one of the outcomes more frequently than the other. We can exploit this by taking into account the expectation value $\langle A_1\rangle$ of Alice's key generation measurement, in addition to the CHSH expectation value S, to derive a better lower bound on the entropy.

The expectation value $\langle A_1\rangle$ can be taken into account using the qubit bound (47) and the convexification procedure discussed at the end of Section 2.4.3 and illustrated in Figure 2. Using this approach, we optimized the key rate numerically over the angles φ_{A_j}, φ_{B_k}, and θ. The optimized key rates, both assuming no noise and a white noise rate of $\delta = 0.5\%$ are represented both for $q = 0$ and with optimized q in Figure 7.

As one can see in the figure, the highest key rate is very small for a significant range of global detection efficiencies close to the threshold as a result of being obtained for values of q close to $1/2$ and very weakly entangled states. Due to this, the threshold detector efficiency above which a positive key rate can be certified is very sensitive and, for example, significantly worsened by the addition of even a small amount of depolarizing noise. To illustrate this, we plot the threshold global detection efficiency as a function of the error rate δ in Figure 8, where a comparison is provided

	$q=0$	$q=0.2$	$q=0.3$	$q=0.49$
Certified	84.2149	80.4642	80.3411	80.2593
Conjectured	84.2147	80.4362	80.3046	80.2283

Table 2: Threshold detection efficiencies (%) for different probabilities q of flipping Alice's outcome assuming no channel noise. For q between 0.49 and 0.5, we did not observe an improvement of the threshold up to the precision reported in the table.

with the earlier results of [14] using the analytic entropy bound for the asymmetric CHSH expressions.

Table 2 gives the thresholds on the detection efficiency that we find using our approach for different values of q assuming no additional noise. We include in the table both the thresholds for which we can certify a positive key rate and the ones obtained using our conjecture regarding the convex envelope of the qubit bound. The small discrepancy between the two values, particularly for larger values of q, is due to the difficulty of numerically certifying the key rate accurately when the key rate becomes very small (the key rate for the last column of Table 2 is of $O(10^{-12})$). Indeed to certify the entropy to a very high precision using a discretized qubit bound requires using a very dense covering, which at some point becomes too time-consuming computationally.

This issue however only affects the certification of extremely small asymptotic key rates, such as the long tail observed in Figure 7, which are probably too low to be of practical value and likely to be dwarfed by the difference made by even small amounts of noise or corrections due to finite-key effects. To illustrate this, in Table 3 we report the detection efficiency thresholds in the presence of a channel noise rate of $\delta = 0.5$. In this case, the thresholds using the conjectured convex envelope and those that can be properly certified are the same up to the precision to which we report the results.

	$q = 0$	$q = 0.2$	$q = 0.3$	$q = 0.49$
$\delta = 0.5\%$	87.6017	86.5842	86.5013	86.4490

Table 3: Certified threshold detection efficiencies (%) obtained for different probabilities q of flipping Alice's outcome and with $\delta = 0.5\%$ of channel error rate. We do not observe a difference with the conjectured case up to the precision reported in the table.

Finally, we remark that the qubit bound (47) is tight in $\langle A_1 \rangle$ and S for all q as there is an explicit attack, described in Appendix E, that saturates it. This means that our conjecture regarding the convex envelope of the qubit bound represents a valid attack yielding upper bounds on the key rate (as it corresponds to an explicit mixture of two-qubit strategies). This means that the certified bounds that we report in Table 3 are, up to the precision we use, optimal in terms of $\langle A_1 \rangle$ and S, and that the second line of Table 2 corresponds to the minimal detection thresholds one can hope to attain using only information about $\langle A_1 \rangle$ and S.

4 Discussion

Building on [14], we have introduced a flexible approach to derive practical and fully device-independent bounds on the key rate for DIQKD in the 2-input/2-output setting. We have illustrated it on to the two-basis variant of the CHSH DIQKD protocol as well as to undertake a more optimized analysis of the single-basis variant when the main anticipated experimental imperfection is losses. Contrarily to [14], we used numerical methods to solve part of the problem in both cases and obtain optimal or close to optimal bounds on the conditional entropy within a very low amount of computation time. The results may be used to derive bounds on the key rate in the asymptotic limit or in the finite-key regime via the entropy accumulation theorem. They may also be useful as a point of comparison with different numerical approaches used to bound the conditional entropy in the device-independent setting.

When considering losses we found that the global detection efficiency can be brought under 80.26%. This is notably below the detection efficiency of 87.49% attained in the recent experimental demonstration of device-independent quantum key distribution based on a photonic setup [6]. As we remarked in the previous section, however, our threshold is attained using a very weakly entangled state and increases significantly if any realistic amount of noise is added to the model we studied. (Separately, a finite-key analysis would likely have the same effect.)

While writing this manuscript, a new promising numerical method to bound the conditional entropy in general DI scenarios was proposed [13]. Our detection threshold, derived using only the expectation value $\langle A_1 \rangle$ of Alice's key-generation measurement in addition to CHSH, is slightly lower than the threshold of 80.5% reported in [13] using full statistics. This is not a limitation of the method of [13], but rather a matter of using a suboptimal state and measurement implementation parameters in that work. Indeed, running their method on the correlations achieving the threshold of 80.2593% in Table 2, the authors of [13] confirmed to us that they also find a positive key rate [28] (though, again, using full statistics instead of only $\langle A_1 \rangle$ and S). This illustrates the interest of having complementary methods. While [13] can in principle be used to tackle very general problems, our method specializing on the 2-input/2-output scenario allows us to rapidly explore the parameter space to find a good implementation. Moreover, there exist scenarios in which our analysis can provide slightly better bounds compared to the numerical method as one can observe from [13, Figure 6b].

A recent result [29] obtained lower bounds on the key rate for the finite-size case without the use of the entropy accumulation theorem in the two-input/two-output scenario. It might be interesting to investigate whether our results involving different parameters to bound the conditional Von Neumann entropy can be used in combination with their technique.

Finally, although we discussed in detail two specific examples illustrating our approach to bounding the conditional von Neumann entropy, we point out that other bounds can be derived. For instance, we could combine the BB84-type bound (13) using bias with the correlation bound (20) in terms of the asymmetric CHSH expectations. As suggested by Figure 8, this should slightly improve the analysis presented here (are least for larger amounts of noise δ). One could also, much more generally, use numerical techniques [30] to derive device-dependent bounds on the conditional von Neumann entropy that are more stringent and combine them with correlation bounds involving full-statistics obtained through relaxations of the Lasserre hierarchy. Our method can also in principle be applied to the n-partite setting, e.g., to derive entropy bounds based on Mermin-type Bell inequalities [31, 32].

The code used to obtain the numerical results in this paper is available on GitHub [33].

Acknowledgments

This work was supported by the EU Quantum Flagship project QRANGE and the F.R.S-FNRS through the grant PDR T.0171.22. S.P. is a Senior Research Associate of the Fonds de la Recherche Scientifique – FNRS.

References

[1] John S. Bell. On the Einstein Podolsky Rosen paradox. *Physics*, 1(3):195–200, 1964. URL http://cds.cern.ch/record/111654/.

[2] Nicolas Brunner, Daniel Cavalcanti, Stefano Pironio, Valerio Scarani, and Stephanie Wehner. Bell nonlocality. *Rev. Mod. Phys.*, 86:419–478, Apr 2014. DOI: 10.1103/RevModPhys.86.419.

[3] Antonio Acín, Nicolas Brunner, Nicolas Gisin, Serge Massar, Stefano Pironio, and Valerio Scarani. Device-independent security of quantum cryptography against collective attacks. *Phys. Rev. Lett.*, 98:230501, Jun 2007. DOI: 10.1103/PhysRevLett.98.230501.

[4] DP Nadlinger, P Drmota, BC Nichol, G Araneda, D Main, R Srinivas, DM Lucas, CJ Ballance, K Ivanov, EY-Z Tan, et al. Experimental quantum key distribution certified by bell's theorem. *Nature*, 607(7920):682–686, 2022. DOI: 10.1038/s41586-022-04941-5.

[5] Wei Zhang, Tim van Leent, Kai Redeker, Robert Garthoff, René Schwonnek, Florian Fertig, Sebastian Eppelt, Wenjamin Rosenfeld, Valerio Scarani, Charles C-W Lim, et al. A device-independent quantum key distribution system for distant users. *Nature*, 607(7920):687–691, 2022. DOI: 10.1038/s41586-022-04891-y.

[6] Wen-Zhao Liu, Yu-Zhe Zhang, Yi-Zheng Zhen, Ming-Han Li, Yang Liu, Jingyun Fan, Feihu Xu, Qiang Zhang, and Jian-Wei Pan. Toward a photonic demonstration of device-independent quantum key distribution. *Phys. Rev. Lett.*, 129(5):050502, 2022. DOI: 10.1103/PhysRevLett.129.050502.

[7] Rotem Arnon-Friedman, Frédéric Dupuis, Omar Fawzi, Renato Renner, and Thomas Vidick. Practical device-independent quantum cryptography via entropy accumulation. *Nat. Commun.*, 9:459, Jan 2018. DOI: 10.1038/s41467-017-02307-4.

[8] Yanbao Zhang, Honghao Fu, and Emanuel Knill. Efficient randomness certification by quantum probability estimation. *Phys. Rev. Research*, 2: 013016, Jan 2020. DOI: 10.1103/PhysRevResearch.2.013016.

[9] Ernest Y-Z Tan, René Schwonnek, Koon Tong Goh, Ignatius William Primaatmaja, and Charles C-W Lim. Computing secure key rates for quantum cryptography with untrusted devices. *npj Quantum Information*, 7(1):1–6, 2021. DOI: 10.1038/s41534-021-00494-z.

[10] René Schwonnek, Koon Tong Goh, Ignatius W. Primaatmaja, Ernest Y.-Z. Tan, Ramona Wolf, Valerio Scarani, and Charles C.-W. Lim. Device-independent quantum key distribution with random key basis. *Nat. Commun.*, May 2021. DOI: 10.1038/s41467-021-23147-3.

[11] Ernest Y.-Z. Tan, Pavel Sekatski, Jean-Daniel Bancal, René Schwonnek, Renato Renner, Nicolas Sangouard, and Charles C.-W. Lim. Improved DIQKD protocols with finite-size analysis. Dec 2020. URL https://doi.org/10.48550/arXiv.2012.08714.

[12] Peter Brown, Hamza Fawzi, and Omar Fawzi. Computing conditional entropies for quantum correlations. *Nat. Commun.*, 12:575, Jan 2021. DOI: 10.1038/s41467-020-20018-1.

[13] Peter Brown, Hamza Fawzi, and Omar Fawzi. Device-independent lower bounds on the conditional von neumann entropy. Jun 2021. URL https://doi.org/10.48550/arXiv.2106.13692.

[14] Erik Woodhead, Antonio Acín, and Stefano Pironio. Device-independent quantum key distribution with asymmetric CHSH inequalities. *Quantum*, 5:443, Apr 2021. DOI: 10.22331/q-2021-04-26-443.

[15] Renato Renner, Nicolas Gisin, and Barbara Kraus. Information-theoretic security proof for quantum-key-distribution protocols. *Phys. Rev. A*, 72:012332, Jul 2005. DOI: 10.1103/PhysRevA.72.012332.

[16] Oliver Kern and Joseph M. Renes. Improved one-way rates for BB84 and 6-state protocols. *Quantum Inf. Comput.*, 8(8,9):0756–0772, Sep 2008. DOI: 10.26421/QIC8.8-9-6.

[17] Igor Devetak and Andreas Winter. Distillation of secret key and entanglement from quantum states. *Proc. R. Soc. A*, 461(2053):207–235, Jan 2005. DOI: 10.1098/rspa.2004.1372.

[18] Stefano Pironio, Antonio Acín, Nicolas Brunner, Nicolas Gisin, Serge Massar, and Valerio Scarani. Device-independent quantum key distribution secure against collective attacks. *New J. Phys.*, 11(4):045021, Apr 2009. DOI: 10.1088/1367-2630/11/4/045021.

[19] Mario Berta, Matthias Christandl, Roger Colbeck, Joseph M. Renes, and Renato Renner. The uncertainty principle in the presence of quantum memory. *Nature Phys.*, 6:659–662, Jul 2010. DOI: 10.1038/nphys1734.

[20] Erik Woodhead. Tight asymptotic key rate for the Bennett-Brassard 1984 protocol with local randomization and device imprecisions. *Phys. Rev. A*, 90:022306, Aug 2014. DOI: 10.1103/PhysRevA.90.022306.

[21] Jean B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM J. Comput.*, 11:796–817, 2001. DOI: 10.1137/S1052623400366802.

[22] D. Henrion and J.-B. Lasserre. Convergent relaxations of polynomial matrix inequalities and static output feedback. *IEEE Trans. Autom. Control*, 51(2):192–202, 2006. DOI: 10.1109/TAC.2005.863494.

[23] Duncan McCallum and David Avis. A linear

algorithm for finding the convex hull of a simple polygon. *Information Processing Letters*, 9 (5):201–206, Dec 1979. ISSN 0020-0190. DOI: 10.1016/0020-0190(79)90069-3.

[24] Alejandro A. Schäffer and Christopher J. Van Wyk. Convex hulls of piecewise-smooth Jordan curves. *J. Algorithms*, 8(1):66–94, Mar 1987. ISSN 0196-6774. DOI: 10.1016/0196-6774(87)90028-9.

[25] Erik Woodhead. Quantum cloning bound and application to quantum key distribution. *Phys. Rev. A*, 88:012331, Jul 2013. DOI: 10.1103/PhysRevA.88.012331.

[26] Rutvij Bhavsar, Sammy Ragy, and Roger Colbeck. Calculation and application of various von Neumann entropies in CHSH-based device-independent randomness expansion. Mar 2021. URL https://doi.org/10.48550/arXiv.2103.07504.

[27] Philippe H. Eberhard. Background level and counter efficiencies required for a loophole-free Einstein-Podolsky-Rosen experiment. *Phys. Rev. A*, 47:R747–R750, Feb 1993. DOI: 10.1103/PhysRevA.47.R747.

[28] Peter Brown. private communication.

[29] Xingjian Zhang, Pei Zeng, Tian Ye, Hoi-Kwong Lo, and Xiongfeng Ma. Quantum complementarity approach to device-independent security. Nov 2021. URL https://doi.org/10.48550/arXiv.2111.13855.

[30] Adam Winick, Norbert Lütkenhaus, and Patrick J. Coles. Reliable numerical key rates for quantum key distribution. *Quantum*, 2:77, Jul 2018. DOI: 10.22331/q-2018-07-26-77.

[31] N. David Mermin. Extreme quantum entanglement in a superposition of macroscopically distinct states. *Phys. Rev. Lett.*, 65:1838–1840, Oct 1990. DOI: 10.1103/PhysRevLett.65.1838.

[32] Federico Grasselli, Gláucia Murta, Hermann Kampermann, and Dagmar Bruß. Entropy bounds for multiparty device-independent cryptography. *PRX Quantum*, 2:010308, Jan 2021. DOI: 10.1103/PRXQuantum.2.010308.

[33] https://github.com/MicheleMasini1996/diqkd-2input2output.

[34] Stefano Pironio, Antonio Acín, Serge Massar, Antoine Boyer de La Giroday, Dzmitry N. Matsukevich, Peter Maunz, Steven Olmschenk, David Hayes, Le Luo, T. Andrew Manning, and Christopher Monroe. Random numbers certified by Bell's theorem. *Nature*, 464:1021–1024, Apr 2010. DOI: 10.1038/nature09008.

[35] Antonio Acín, Serge Massar, and Stefano Pironio. Randomness versus nonlocality and entanglement. *Phys. Rev. Lett.*, 108:100402, Mar 2012. DOI: 10.1103/PhysRevLett.108.100402.

A Derivation of BB84 bound with bias

The BB84 entropy bound (13) is a generalization of the two bounds (10) and (11), which give the special cases of (13) with $\langle A_1\rangle = 0$ and both with $\langle A_1\rangle = 0$ and no noisy preprocessing ($q = 0$). It can be derived, in a way that also confirms the monotonicity property, essentially by modifying the symmetrization step in the derivation done in section 4.2 of the paper [14]. We do this in detail here.

As in the derivation of [14], we suppose that Alice, Bob, and Eve share a pure tripartite state

$$|\psi\rangle_{ABE} = |0\rangle_A|\psi_0\rangle_{BE} + |1\rangle_A|\psi_1\rangle_{BE}\,, \tag{82}$$

where $|0\rangle$ and $|1\rangle$ are the eigenstates of A_1, which we identify here with Z, and $|\psi_0\rangle$ and $|\psi_1\rangle$ are arbitrary (and not necessarily orthogonal) states shared by Bob and Eve normalized so that

$$\langle\psi_0|\psi_0\rangle + \langle\psi_1|\psi_1\rangle = 1\,. \tag{83}$$

After Alice measures $A_1 = \mathrm{Z}$ and flips the outcome with probability q, the correlations between Alice and Eve are described by the classical-quantum state

$$\tau_{AE} = [0]_A\otimes(\bar q\psi_0^E + q\psi_1^E) + [1]_A\otimes(q\psi_0^E + \bar q\psi_1^E)\,, \tag{84}$$

where $\bar q = 1 - q$ and $\psi_a^E = \mathrm{Tr}_B[\psi_a]$ are the partial traces of the states $|\psi_a\rangle$ accessible to Eve.

Now, since renaming the outcomes does not change the entropy, the conditional entropy $H(\mathrm{Z}|E) = H(\mathrm{Z}E) - H(E)$ computed on the above state is the same as the conditional entropy computed on

$$\tau'_{AE} = [1]_A\otimes(\bar q\psi_0^E + q\psi_1^E) + [0]_A\otimes(q\psi_0^E + \bar q\psi_1^E)\,, \tag{85}$$

which is the same state as above except that we have swapped $[0]_A$ and $[1]_A$. They in addition have the same entropy as a partly symmetrized state,

$$\bar\tau_{AEF} = \bar p\,\tau_{AE}\otimes[0]_F + p\,\tau'_{AE}\otimes[1]_F\,, \tag{86}$$

for any probability p and $\bar p = 1 - p$, since

$$H(\mathrm{Z}|EF)_{\bar\tau} = \bar p\,H(\mathrm{Z}|E)_\tau + p\,H(\mathrm{Z}|E)_{\tau'} = H(\mathrm{Z}|E)_\tau\,. \tag{87}$$

The above state, written out explicitly, is

$$\begin{aligned}\bar\tau_{AEF} = [0]_A\otimes\Big[&\bar p(\bar q\psi_0^E + q\psi_1^E)\otimes[0]_F \\ &+ p(q\psi_0^E + \bar q\psi_1^E)\otimes[1]_F\Big] \\ + [1]_A\otimes\Big[&\bar p(q\psi_0^E + \bar q\psi_1^E)\otimes[0]_F \\ &+ p(\bar q\psi_0^E + q\psi_1^E)\otimes[1]_F\Big]\,.\end{aligned} \tag{88}$$

We rewrite this as

$$\bar\tau_{AEF} = [0]_A\otimes(\bar q\sigma_= + q\sigma_{\neq}) + [1]_A\otimes(q\sigma_= + \bar q\sigma_{\neq}) \tag{89}$$

with the (unnormalized) states

$$\sigma_= = \bar p\,\psi_0^E\otimes[0]_F + p\,\psi_1^E\otimes[1]_F\,, \tag{90}$$
$$\sigma_{\neq} = \bar p\,\psi_1^E\otimes[0]_F + p\,\psi_0^E\otimes[1]_F\,. \tag{91}$$

The state can be obtained as the marginal of an extended one,

$$\begin{aligned}\bar\tau_{ABEE'FF'} = &[0]_A\otimes(\bar q\chi_= + q\chi_{\neq}) \\ &+ [1]_A\otimes(q\chi_= + \bar q\chi_{\neq})\,,\end{aligned} \tag{92}$$

where $|\chi_=\rangle, |\chi_{\neq}\rangle \in \mathcal{H}_B\otimes\mathcal{H}_E\otimes\mathcal{H}_{E'}\otimes\mathcal{H}_F\otimes\mathcal{H}_{F'}$ are unnormalized pure states

$$|\chi_=\rangle = \sqrt{\bar p}|\psi_0\rangle|\phi_0\rangle|00\rangle + \sqrt{p}|\psi_1'\rangle|\phi_1\rangle|11\rangle\,, \tag{93}$$
$$|\chi_{\neq}\rangle = \sqrt{\bar p}|\psi_1'\rangle|\phi_1\rangle|00\rangle + \sqrt{p}|\psi_0\rangle|\phi_0\rangle|11\rangle\,, \tag{94}$$

in which

$$|\psi_1'\rangle = e^{i\varphi}B\otimes\mathbb{1}_E|\psi_1\rangle \in \mathcal{H}_B\otimes\mathcal{H}_E\,, \tag{95}$$

where B is a Hermitian unitary operator (thus satisfying $B^2 = \mathbb{1}_B$) acting on $\mathcal{H}_B$ and φ is a phase chosen such that $\langle\psi_0|\psi_1'\rangle$ is real and nonnegative, and

$$|\phi_0\rangle, |\phi_1\rangle \in \mathcal{H}_{E'} \tag{96}$$

are normalized states chosen to have some nonnegative real overlap $\langle\phi_0|\phi_1\rangle = \lambda_{\mathrm{X}} \in [0, 1]$.

Using that the conditional entropy cannot increase if we extend the Hilbert space being conditioned on, direct calculation of the conditional entropy on the state (92) gives

$$\begin{aligned}H(\mathrm{Z}|E)_\tau &= H(\mathrm{Z}|EF)_{\bar\tau} \\ &\geq H(\mathrm{Z}|BEE'FF')_{\bar\tau} \\ &= S(\bar\tau_{ABEE'FF'}) - S(\chi_= + \chi_{\neq}) \\ &= H(\boldsymbol{\lambda}) - \phi\Big(\sqrt{Z'^2 + X'^2}\Big)\,,\end{aligned} \tag{97}$$

where

$$Z' = \|\chi_=\| - \|\chi_{\neq}\| \equiv \langle\chi_=|\chi_=\rangle - \langle\chi_{\neq}|\chi_{\neq}\rangle\,, \tag{98}$$
$$X' = 2|\langle\chi_=|\chi_{\neq}\rangle|\,, \tag{99}$$

and $H(\boldsymbol{\lambda}) = -\sum_{jk}\lambda_{jk}\log_2(\lambda_{jk})$ is the Shannon entropy associated to the four eigenvalues of (92),

$$\lambda_{11} = \frac{1}{4}\Big[1 + QZ' + \sqrt{R'^2 + 2QZ'}\Big]\,, \tag{100}$$
$$\lambda_{12} = \frac{1}{4}\Big[1 - QZ' + \sqrt{R'^2 - 2QZ'}\Big]\,, \tag{101}$$
$$\lambda_{21} = \frac{1}{4}\Big[1 - QZ' - \sqrt{R'^2 - 2QZ'}\Big]\,, \tag{102}$$
$$\lambda_{22} = \frac{1}{4}\Big[1 + QZ' - \sqrt{R'^2 + 2QZ'}\Big]\,, \tag{103}$$

where Q is related to the amount of noisy preprocessing applied by

$$Q = \bar q - q = 1 - 2q \tag{104}$$

and

$$R' = \sqrt{Z'^2 + Q^2 + (1 - Q^2)X'^2}\,. \tag{105}$$

We can factorize the four eigenvalues above as $\lambda_{jk} = p_j p'_k$ with

$$p_1 = \frac{1}{2} + \frac{1}{4}\big(R'_+ + R'_-\big)\,, \tag{106}$$

$$p_2 = \frac{1}{2} - \frac{1}{4}\big(R'_+ + R'_-\big)\,, \tag{107}$$

$$p'_1 = \frac{1}{2} + \frac{1}{4}\big(R'_+ - R'_-\big)\,, \tag{108}$$

$$p'_2 = \frac{1}{2} - \frac{1}{4}\big(R'_+ - R'_-\big)\,, \tag{109}$$

and

$$R'_\pm = \sqrt{R'^2 \pm 2QZ'}\,, \tag{110}$$

so that $H(\boldsymbol{\lambda}) = H(\boldsymbol{p}) + H(\boldsymbol{p}')$. This allows us to express the qubit entropy bound more concisely as

$$H(\mathrm{Z}|E) \geq g_q(Z', X') \tag{111}$$

with

$$g_q(Z', X') = \phi\big(\tfrac{1}{2}(R'_+ + R'_-)\big) + \phi\big(\tfrac{1}{2}(R'_+ - R'_-)\big) - \phi\big(\sqrt{Z'^2 + X'^2}\big) \tag{112}$$

and

$$R'_\pm = \sqrt{(Q \pm Z')^2 + (1-Q^2)X'^2}\,. \tag{113}$$

At this point, we have recovered the form of the function g_q defined in section 2. To complete the derivation note that, from the definitions of $|\chi_=\rangle$ and $|\chi_\neq\rangle$ we have

$$\begin{aligned} Z' &= \|\chi_0\| - \|\chi_1\| \\ &= \bar{p}\|\psi_0\| + p\|\psi_1\| - \bar{p}\|\psi_1\| - p\|\psi_0\| \\ &= \lambda_{\mathrm{Z}}\big(\|\psi_0\| - \|\psi_1\|\big) \\ &= \lambda_{\mathrm{Z}}\langle A_1\rangle\,, \end{aligned} \tag{114}$$

where $\lambda_{\mathrm{Z}} \in [-1, 1]$ is related to the symmetrization-step probability by $\lambda_{\mathrm{Z}} = \bar{p} - p$, and that

$$\begin{aligned} \langle\chi_=|\chi_\neq\rangle &= \bar{p}\langle\psi_0|\psi'_1\rangle\langle\phi_0|\phi_1\rangle + p\langle\psi'_1|\psi_0\rangle\langle\phi_1|\phi_0\rangle \\ &= \lambda_{\mathrm{X}}\, e^{i\varphi}\langle\psi_0|B\otimes\mathbb{1}_E|\psi_1\rangle \\ &= \lambda_{\mathrm{X}}\,\big|\mathrm{Re}\big[\langle\psi_0|B\otimes\mathbb{1}_E|\psi_1\rangle\big]\big|\,, \end{aligned} \tag{115}$$

where we recall that we set $\langle\phi_0|\phi_1\rangle = \lambda_{\mathrm{X}} \in [0, 1]$, while

$$\langle \mathrm{X}\otimes B\rangle = 2\,\mathrm{Re}\big[\langle\phi_0|B\otimes\mathbb{1}_E|\phi_1\rangle\big]\,, \tag{116}$$

so that

$$2\langle\chi_=|\chi_\neq\rangle = \lambda_{\mathrm{X}}\,|\langle \mathrm{X}\otimes B\rangle|\,. \tag{117}$$

Putting all this together and recalling that we identify A_1 with Z, and can choose $\bar{A}_1 = \mathrm{X}$, means that we finally get

$$H(A_1|E) \geq g_q\big(\lambda_{\mathrm{Z}}\langle A_1\rangle,\, \lambda_{\mathrm{X}}|\langle\bar{A}_1\otimes B\rangle|\big) \tag{118}$$

for all $-1 \leq \lambda_{\mathrm{Z}}, \lambda_{\mathrm{X}} \leq 1$ (as the derivation we have given applies for any values of the symmetrization probability p and overlap $\langle\phi_0|\phi_1\rangle$ we may wish to use). This confirms that the inequality

$$H(A_1|E) \geq g_q(Z, X) \tag{119}$$

holds for any (real) numbers satisfying

$$|Z| \leq |\langle A_1\rangle| \quad \text{and} \quad |X| \leq |\langle\bar{A}_1\otimes B\rangle|\,. \tag{120}$$

B Analytic solution for $p = 1/2$

Here we derive in detail the average entropy bound for the two-basis protocol in the case that Alice's measurements are used equiprobably. When $p = 1/2$, the minimization problem (41) in Section 2.3 simplifies to

$$\begin{aligned} \text{minimize} \quad & f(\lambda, \mu, \varphi_A) = \sin\big(\tfrac{\varphi_A}{2}\big)^2\lambda^2 + \cos\big(\tfrac{\varphi_A}{2}\big)^2\mu^2 \\ \text{subject to} \quad & \big|\cos\big(\tfrac{\varphi_A}{2}\big)\big||\lambda| + \big|\sin\big(\tfrac{\varphi_A}{2}\big)\big||\mu| \geq S/2 \\ & \lambda^2 \leq 1 \\ & \mu^2 \leq 1\,, \end{aligned} \tag{121}$$

where we have reintroduced the angle φ_A from earlier in the section explicitly and used that the single constraint involving the variable Δ becomes irrelevant. As we stated in Section 2.3 and show here, the above problem can be solved analytically subject to finding the root of a degree four polynomial.

In the following, we will assume that $S > 2$, since the solution to the classical case $S = 2$ is trivially ${E_{\frac{1}{2}}}^2 = 0$.

First, we note that, as our problem is invariant under the transformations $\lambda \mapsto -\lambda$ and $\mu \mapsto -\mu$ and that, for $S > 2$, the points $\mu = 0$ or $\lambda = 0$ do not satisfy the first constraint

$$\big|\cos\big(\tfrac{\varphi_A}{2}\big)\big||\lambda| + \big|\sin\big(\tfrac{\varphi_A}{2}\big)\big||\mu| \geq S/2\,, \tag{122}$$

we can replace the constraints $\lambda^2 \leq 1$ and $\mu^2 \leq 1$ with $0 < \lambda \leq 1$ and $0 < \mu \leq 1$.

Moreover, the problem is also invariant under the transformation $\varphi_A \mapsto 2\pi - \varphi_A$, meaning that for all solutions such that $\varphi_A \in [0, \pi]$, there exists an equivalent solution in $[\pi, 2\pi]$. Thus, we can restrict the domain of φ_A to be $0 < \varphi_A < \pi$, where we excluded the boundaries since the cases $\varphi_A = 0, \pi$ are not in agreement with $S > 2$.

The function that we need to minimize can be rewritten as

$$f(\lambda, \mu, \varphi_A) = \frac{\lambda^2}{2}\big(1 - \cos(\varphi_A)\big) + \frac{\mu^2}{2}\big(1 + \cos(\varphi_A)\big)\,. \tag{123}$$

Let us look for a minimum for our function by checking where its derivatives are zero. We start with

$$\frac{d}{d\mu} f(\lambda, \mu, \varphi_A) = \mu\big(1 + \cos(\varphi_A)\big)\,. \tag{124}$$

Here, $\frac{d}{d\mu} f(\lambda, \mu, \varphi_A) = 0$ if and only if $\mu = 0$ or $\varphi_A = \pi$. These points are not part of the restricted domain that we are considering. We conclude that the minimum must be at the boundaries of our domain. From now on, we will analyze this case.

Case 1: We consider the boundary $\lambda = 1$. We have

$$f(1, \mu, \varphi_A) = \frac{1 + \mu^2}{2} + \frac{\cos(\varphi_A)}{2}(\mu^2 - 1) \tag{125}$$

and

$$\frac{d}{d\mu}f(1,\mu,\varphi_A) = \mu\big(1+\cos(\varphi_A)\big)\,, \tag{126}$$

thus $\frac{d}{d\mu}f(\lambda,\mu,\varphi_A)=0$ if and only if $\mu=0$ or $\varphi_A=\pi$. Such solutions are not in the domain.

Case 2: We consider the boundary $\mu=1$. Analogously, we obtain non-feasible solutions.

Case 3: We consider the boundary $\cos\big(\frac{\varphi_A}{2}\big)\lambda + \sin\big(\frac{\varphi_A}{2}\big)\mu = S/2$. This region is the one in which

$$\begin{aligned}\mu_* &= \lambda\frac{\sin(\varphi_A)}{\cos(\varphi_A)-1} - S\frac{\sin\big(\frac{\varphi_A}{2}\big)}{\cos(\varphi_A)-1}\\ &= \lambda\frac{\sqrt{1-x^2}}{x-1} - \frac{S}{\sqrt{2}}\frac{\sqrt{1-x}}{x-1}\,,\end{aligned} \tag{127}$$

where we made the change of variable $x=\cos(\varphi_A)$. The domain of x is $-1<x<1$.

We have

$$f(\lambda,\mu_*,x) = \lambda^2\frac{1+x^2}{1-x} - \lambda\frac{S(1+x)^{3/2}}{\sqrt{2}(1-x)} + \frac{S^2(1+x)}{4(1-x)} \tag{128}$$

and

$$\frac{d}{d\lambda}f(\lambda,\mu_*,x) = 2\lambda\frac{1+x^2}{1-x} - \frac{S(1+x)^{3/2}}{\sqrt{2}(1-x)}\,. \tag{129}$$

Now, recalling that we assumed $x\neq 1$, we have that $\frac{d}{d\lambda}f(\lambda,\mu_*,x)=0$ iff

$$\lambda = \frac{S(x+1)^{3/2}}{2\sqrt{2}(x^2+1)} = \lambda_*\,. \tag{130}$$

Thus,

$$f(\lambda_*,\mu_*,x) = \frac{S^2}{8}\frac{1-x^2}{1+x^2}\,, \tag{131}$$

which is a concave function of x, meaning the minimum is at the intersection between boundaries.

Case 3+1: We intersect the boundary of case 3 with $\lambda=1$. We get

$$\mu_* = \frac{\sqrt{1-x^2}}{x-1} - \frac{S}{\sqrt{2}}\frac{\sqrt{1-x}}{x-1}\,. \tag{132}$$

Here, requiring $\mu_*\leq 1$, we obtain the condition

$$-\frac{S}{4}\sqrt{8-S^2} \leq x \leq \frac{S}{4}\sqrt{8-S^2}\,. \tag{133}$$

We have

$$f(1,\mu_*,x) = \frac{x^2+1}{1-x} - \frac{S(x+1)^{3/2}}{\sqrt{2}(1-x)} + \frac{S^2(x+1)}{4(1-x)} \tag{134}$$

and

$$\frac{d}{dx}f(1,\mu_*,x) = \frac{4x(2-x)+2(S^2+2)+S(x-5)\sqrt{2(1+x)}}{4(x-1)^2}\,, \tag{135}$$

hence, since $x\neq 1$, $\frac{d}{dx}f(1,\mu_*,x)=0$ iff

$$4x(2-x)+2(S^2+2)+S(x-5)\sqrt{2(1+x)} = 0\,. \tag{136}$$

Case 3+2: We intersect the boundary of case 3 with $\mu=1$. Here, one can check that we obtain the same result as in case 3+1.

Case 1+2: We consider $\lambda=\mu=1$. With this choice we have ${E_{\frac{1}{2}}}^2 = 1\ \forall\varphi_A$. This region of parameters does not contain in general the absolute minimum.

We conclude that the solution to the optimization problem must be the one of case 3+1 (or equivalently 3+2). If there is more than one solution to Eq. (136) satisfying the constraints (133), then we take the smallest one.

We used Mathematica to find the roots of Eq. (136) analytically. Moreover, imposing the constraints (133) and $S>2$, we found a single solution. We used the resulting expression for the computations for $p=1/2$ done in Section 3.1.

C Optimality of CHSH for the two-basis protocol

In the case that the bases are used equiprobably, i.e., $p=1/2$, the symmetries of the two-basis DIQKD protocol studied in section 3.1 imply that the CHSH Bell expectation value alone already gives the optimal bound on the average conditional entropy

$$H(A_x|XE) \propto \tfrac{1}{2}H(A_1|E) + \tfrac{1}{2}H(A_2|E) \tag{137}$$

for the optimal CHSH-violating correlations attenuated by white noise. The reason for this is that, given any quantum strategy giving a particular value of the average entropy and CHSH expectation value, one can construct a new symmetrized strategy giving the same entropy and CHSH expectation value.

To see this, let us suppose we have a particular quantum strategy $\mathcal{Q}=(\rho_{ABE},A_1,A_2,B_1,B_2)$. We note first that both conditional entropies $H(A_x|E)$ and the CHSH expectation value $S=\langle A_1B_1\rangle + \langle A_1B_2\rangle + \langle A_2B_1\rangle - \langle A_2B_2\rangle$ are unchanged if we flip all the measurements, i.e., do $A_x\mapsto -A_x$ and $B_y\mapsto -B_y$. By randomly and equiprobably using these two strategies we can force Alice's and Bob's local outcomes to become equiprobable. This corresponds to using a

new strategy $\mathcal{Q}' = (\rho'_{ABE}, A'_1, A'_2, B'_1, B'_2)$ with

$$A'_x = A_x \oplus -A_x\,, \tag{138}$$
$$B'_y = B_y \oplus -B_y\,, \tag{139}$$
$$\rho'_{ABE} = \tfrac{1}{2}\rho_{ABE} \oplus \tfrac{1}{2}\rho_{ABE}\,, \tag{140}$$

for which the CHSH expectation value and the values of the entropies are unchanged, but for which $\langle A'_x\rangle = \langle B'_y\rangle = 0$.

Next, we use that the average entropy and CHSH both remain unchanged under the two transformations

$$T_1 : \begin{cases} A_1 \mapsto A_1 \\ A_2 \mapsto -A_2 \\ B_1 \mapsto B_2 \\ B_2 \mapsto B_1 \end{cases} \qquad T_2 : \begin{cases} A_1 \mapsto A_2 \\ A_2 \mapsto A_1 \\ B_1 \mapsto B_1 \\ B_2 \mapsto -B_2 \end{cases}, \tag{141}$$

as well as their composition $T_2 \circ T_1$. By randomly using the strategy $\mathcal{Q}'$ with neither, either one, or both transformations applied, we construct a new strategy $\mathcal{Q}'' = (\rho''_{ABE}, A''_1, A''_2, B''_1, B''_2)$ with

$$A''_1 = A'_1 \oplus A'_1 \oplus A'_2 \oplus A'_2\,, \tag{142}$$
$$A''_2 = A'_2 \oplus -A'_2 \oplus A'_1 \oplus -A'_1\,, \tag{143}$$
$$B''_1 = B'_1 \oplus B'_2 \oplus B'_1 \oplus -B'_2\,, \tag{144}$$
$$B''_2 = B'_2 \oplus B'_1 \oplus -B'_2 \oplus B'_1\,, \tag{145}$$
$$\rho''_{ABE} = \tfrac{1}{4}\rho'_{ABE} \oplus \tfrac{1}{4}\rho'_{ABE} \oplus \tfrac{1}{4}\rho'_{ABE} \oplus \tfrac{1}{4}\rho'_{ABE}\,, \tag{146}$$

for which

$$\langle A''_1 B''_1\rangle = \langle A''_1 B''_2\rangle = \langle A''_2 B''_1\rangle = -\langle A''_2 B''_2\rangle = S/4\,. \tag{147}$$

As, given any strategy $\mathcal{Q}$, we can in this way always construct a strategy $\mathcal{Q}''$ with the same average entropy and CHSH expectation value, but satisfying $\langle A''_x\rangle = \langle B''_y\rangle = 0$ and $\langle A''_1 B''_1\rangle = \langle A''_1 B''_2\rangle = \langle A''_2 B''_1\rangle = -\langle A''_2 B''_2\rangle$, we can infer that these constraints, if they are satisfied for real correlations, do not contain any information other than the CHSH expectation value that can be used to improve the entropy bound.

D Explicit attack for the two-basis protocol

We describe here an explicit attack for the two-basis protocol in the case $p = 1/2$, which we conjecture to be optimal.

Suppose that Alice, Bob, and Eve share the optimal symmetric BB84 attack state

$$\begin{aligned} |\Psi\rangle_{ABE} = \frac{1}{2}\Big[&(1+E)|\phi^+\rangle_{AB}|{+}{+}\rangle_E \\ &+ \sqrt{1-E^2}|\phi^-\rangle_{AB}|{+}{-}\rangle_E \\ &+ \sqrt{1-E^2}|\psi^+\rangle_{AB}|{-}{+}\rangle_E \\ &+ (1-E)|\psi^-\rangle_{AB}|{-}{-}\rangle_E \Big]\,, \end{aligned} \tag{148}$$

where $|\phi^\pm\rangle$ and $|\psi^\pm\rangle$ are the four Bell states, depending on some number $0 \le E \le 1$. Its marginal once Eve is traced out is

$$\Psi_{AB} = \frac{1}{4}\Big[\mathbb{1}\otimes\mathbb{1} + E\,\mathrm{X}\otimes\mathrm{X} - E^2\,\mathrm{Y}\otimes\mathrm{Y} + E\,\mathrm{Z}\otimes\mathrm{Z}\Big]\,. \tag{149}$$

By measuring $A_1 = \mathrm{Z}$, $A_2 = \mathrm{X}$, and $B_{1,2} = (\mathrm{Z} \pm \mathrm{X})/\sqrt{2}$, the highest possible CHSH expectation value of $S = 2\sqrt{2}E$ with this state is obtained. Direct computation of the conditional entropies after Alice measures Z and X on this state gives

$$\tfrac{1}{2}H(A_1^q|E) + \tfrac{1}{2}H(A_2^q|E) = f_q\big(S/\sqrt{8}\big) \tag{150}$$

where f_q is the same BB84 bound with noisy preprocessing used earlier and given by Eq. (12). This is too high to be the optimal bound on the average entropy for all S, as the correct bound must attain $h(q)$ at $S = 2$. But we can construct a plausible strategy by taking a convex mixture (similar to the construction in Section 2 of [14]) of the strategy just described with a deterministic one giving $\big(H(A_X^q|XE),\, S\big) = (h(q),\, 2)$. This gives

$$\tfrac{1}{2}H(A_1^q|E) + \tfrac{1}{2}H(A_2^q|E) = \bar{f}_q\big(S/\sqrt{8}\big)\,, \tag{151}$$

where

$$\bar{f}_q(x) = \begin{cases} f_q(x) & \text{if } x \ge x_* \\ h(q) + f'_q(x_*)(x - 1/\sqrt{2}) & \text{if } x \le x_* \end{cases} \tag{152}$$

with x_* (dependent on q) such that

$$h(q) + f'_q(x_*)(x - 1/\sqrt{2}) = f_q(x_*)\,. \tag{153}$$

E Explicit attack saturating the qubit entropy bound with bias (47)

One can verify that the qubit bound (47) is attained with measurements and an initial state of the form

$$A_1 = \mathrm{Z}\,, \tag{154}$$
$$A_2 = \mathrm{X}\,, \tag{155}$$
$$B_1 = \cos\big(\tfrac{\varphi_B}{2}\big)\mathrm{Z} + \sin\big(\tfrac{\varphi_B}{2}\big)\mathrm{Z}\,, \tag{156}$$
$$B_2 = \cos\big(\tfrac{\varphi_B}{2}\big)\mathrm{Z} - \sin\big(\tfrac{\varphi_B}{2}\big)\mathrm{Z}\,, \tag{157}$$

and

$$|\Psi\rangle_{ABE} = \cos\big(\tfrac{\theta}{2}\big)|00\rangle_{AB}|\psi_0\rangle_E + \sin\big(\tfrac{\theta}{2}\big)|11\rangle_{AB}|\psi_1\rangle_E \tag{158}$$

with

$$\cos(\theta) = \langle A_1\rangle\,, \tag{159}$$
$$\sin(\theta)\langle\psi_0|\psi_1\rangle = \sqrt{S^2/4 - 1}\,, \tag{160}$$
$$\cos\big(\tfrac{\varphi_B}{2}\big) = 2/S\,, \tag{161}$$
$$\sin\big(\tfrac{\varphi_B}{2}\big) = \sqrt{1 - 4/S^2}\,. \tag{162}$$

Note that, because $\cos(\theta)^2 + \sin(\theta)^2 |\langle\psi_0|\psi_1\rangle|^2 \leq 1$, (159) and (160) are only consistent with each other if

$$\langle A_1\rangle^2 + S^2/4 \leq 2\,, \tag{163}$$

but this is a known boundary of the quantum set [34, 35].