Device-independent quantum key distribution based on Hardy’s paradox

Ramij Rahaman,1,2 Matthew G. Parker,3 Piotr Mironowicz,4,5 and Marcin Pawlowski6,2,†

1Department of Mathematics, University of Allahabad, Allahabad 211002, U.P., India
2Institute of Theoretical Physics & Astrophysics, University of Gdańsk, 80-952 Gdańsk, Poland
3Department of Informatics, University of Bergen, Post Box-7803, 5020, Bergen, Norway
4Department of Algorithms and System Modelling, Faculty of Electronics, Telecommunications and Informatics, Gdańsk University of Technology, Gdańsk 80-233, Poland
5National Quantum Information Centre in Gdańsk, Sopot 81-824, Poland
6Department of Mathematics, University of Bristol, Bristol BS8 1TW, United Kingdom

(Dated: May 11, 2014)

We present a secure device-independent quantum key distribution scheme based on Hardy’s paradox. In comparison with protocols based on Bell inequalities, it has several novel features: (a) The bits used for the secret key do not come from the results of the measurements on an entangled state but from the choices of settings which are harder for an eavesdropper to influence; (b) Instead of a single security parameter (a violation of some Bell inequality) a set of them is used to estimate the level of trust in the secrecy of the key. This further restricts the eavesdropper’s options. We prove the security of our protocol for both ideal and noisy cases.

PACS numbers: 03.67.Ac, 03.67.Dd, 03.67.Mn, 03.65.Ud
Keywords: Hardy’s paradox, quantum key distribution, min-entropy, NPA, semi-definite programming

I. INTRODUCTION

The goal of quantum information theory is to develop new technologies for information processing that will take us from the traditional classical information age into the age of quantum information. Quantum key distribution (QKD) [1], the most secure known way for sending secret messages, is a significant achievement in this context: if Alice and Bob have access to a preshared key for authentication, they can establish a shared secret key using an insecure quantum channel and public communication.

Besides the validity of the laws of quantum physics, the security of all QKD schemes rely on some other assumptions. The first and foremost assumption, always present in any such protocol, is that all parties concerned have secure laboratories, i.e., at no stage should there be a leakage of secure classical data from any laboratory. This assumption is crucial and cannot be removed. Another basic assumption is that all players have complete control over their own physical devices, i.e., they have full knowledge over what quantum system their apparatuses operate and they also know the exact operation of their measuring devices, etc. The goal of the device-independent analysis of quantum protocols is to eliminate the last assumption, namely players can distrust the source of particles and the communication channel, but they could also distrust their measuring apparatuses as they might have been fabricated by a malicious party. In 2007, Acín et al. [3], introduced a device-independent QKD protocol secure against collective attacks. Earlier questions of a similar type were also addressed by several researchers in different contexts [4–6]. In 2011, Masanes et al. [7] provided a more general security scheme based on causally independent measurement processes. But the security of all these protocols is undermined as the measurement at step $k$ may depend on the classical or quantum memory of all previous inputs and outputs. Recently secure protocols where device re-use is allowed were introduced [8, 9].

In all protocols mentioned above the parties make measurements on entangled subsystems, check for a violation of some Bell inequality to see if their outcomes are random from the eavesdroppers point of view and, if indeed they are, use them as their secret key. In this manuscript we present a protocol which is significantly different. Obviously, the parties need a way to convince themselves that the correlations they share cannot be classical. But checking for a violation of a Bell inequality is only one possible way of doing so. Another option is to verify Hardy’s paradox [10]. There, instead of estimating a single parameter, the estimation of four is made, which gives the parties more knowledge about the correlations they share. Moreover, the structure of this paradox implies that the parties announce their outcomes and use their choices of measurement settings for key generation. Our protocol shares this property with the non-device-independent protocol of SARG04 [11]. The benefit of flipping the roles of outcomes and settings is that the latter are chosen by the parties using their private random number generators over which Eve has no influence (this is a standard assumption in QKD). The former, on the other hand, come from measurements on the systems supplied by the eavesdropper.

The main messages of this paper are that one can:
use other things, besides Bell inequalities, for device-independent cryptography; check more than one parameter to know more about the correlations [12, 14], and use the bits from private random number generators for the secret key to make the eavesdropper’s life harder. The main result is to show that a protocol with these properties can be constructed and how to prove its security.

We organize our paper as follows: we start with a description of the so-called Hardy paradox [10]. Next we present a QKD protocol based on the paradox and discuss why it is secure against collective attack in the case when perfect correlations are observed. Then we move to a more realistic noisy case and prove its security when the measurements are causally independent. We end with a discussion of our results.

II. HARDY’S PARADOX

Consider a physical system consisting of two subsystems shared between two distant parties. The two observers (Alice and Bob) have access to one subsystem each. Both can choose one of two binary measurement settings labeled 0 and 1, with outcomes 0 and 1. The settings are chosen at random in subsequent runs of the experiment. Settings are denoted by letters A and B while outcomes by a and b, for Alice and Bob respectively. The Hardy-type argument starts with the following set of four joint probability conditions for two two-level systems:

\[
P(a = 0, b = 0|A = 0, B = 0) = q > 0, \\
P(a = 0, b = 0|A = 1, B = 0) = 0, \\
P(a = 0, b = 0|A = 0, B = 1) = 0, \\
P(a = 1, b = 1|A = 1, B = 1) = 0.
\]

(1)

This set of conditions cannot be satisfied by any local-realistic theory (LRT) [10, 15].

Let us find the set of states \( \rho \) for which the conditions for the Hardy-type argument given in (1) are satisfied for a given pair of observables. Let us denote the eigenstates of the observable \( X = 0(1) \) by \( |0\rangle_X (|1\rangle_X) \) and \( |1\rangle_X (|0\rangle_X) \) for the outcome 0 and 1 respectively. We now associate a product state with every condition in the test of the Hardy-type argument given in (1), say:

\[
|\phi_0\rangle = |1\rangle_A |1\rangle_B; \\
|\phi_2\rangle = |0\rangle_A |0\rangle_B; \\
|\phi_1\rangle = |0\rangle_A |1\rangle_B; \\
|\phi_3\rangle = |1\rangle_A |0\rangle_B.
\]

(2)

Let

\[
|0\rangle_X = a_X |0\rangle_X + b_X |1\rangle_X,
\]

and

\[
|1\rangle_X = a_X |0\rangle_X + b_X |1\rangle_X,
\]

where \( |a_X|^2 + |b_X|^2 = 1 \) and \( 0 < |a_X| < 1 \) for \( A = 0, B \). The last condition is due to the non-commutativity of \( X = 0 \) and \( X = 1 \).

Let \( S \) be the subspace spanned by the three linearly independent states \( |\phi_0\rangle, |\phi_1\rangle \) and \( |\phi_2\rangle \) given in (2). To satisfy the conditions given in Eqs. (1), a state \( \rho \) has to be confined to a subspace of \( \mathcal{C}^2 \otimes \mathcal{C}^2 \), which is orthogonal to the subspace \( S \) but not orthogonal to \( |\phi_3\rangle \). The dimension of the subspace \( (\mathcal{S}^2) \), orthogonal to the subspace \( S \) is one. Therefore, \( \rho \) must be unique (up to a local unitary) and a pure two-qubit entangled state, which we denote \( |\psi^H\rangle \). Thus, no mixed state of two spin-1/2 particles will satisfy Hardy’s argument [10]. It was also shown that no two-qubit maximally entangled states satisfy Hardy’s argument [10].

The four product states \( \{|\phi_i\rangle\}_{i=0}^3 \) are linearly independent, hence, by the Gram-Schmidt orthogonalization procedure, one can find an orthonormal basis \( \{|\phi'_i\rangle\}_{i=0}^3 \) in which state \( |\phi'_0\rangle = |\psi^H\rangle \) is its last member:

\[
|\phi'_0\rangle = |\phi_0\rangle, \\
|\phi'_i\rangle = \frac{|\phi_i\rangle - \sum_{j=0}^{i-1} \langle \phi_j | \phi_i \rangle |\phi'_j\rangle}{\sqrt{1 - \sum_{j=0}^{i-1} |\langle \phi'_j | \phi_i \rangle|^2}}, \quad \text{for } i = 1, 2, 3.
\]

(3)

The probability \( q \) in conditions (1), for the Hardy state \( |\psi^H\rangle \), reads

\[
q = |\langle \psi^H | \phi_0 \rangle|^2 = 1 - \sum_{i=0}^{2} |\langle \phi'_i | \phi_3 \rangle|^2 = \frac{|\alpha_A \alpha_B|^2 - |\beta_A \beta_B|^2}{|\alpha_A \beta_B|^2}
\]

Its maximum is \( \frac{\sqrt{5} - 1}{2} \) for \( |\alpha_A| = |\alpha_B| = \sqrt{\frac{\sqrt{5} - 1}{2}} \).

III. QUANTUM KEY DISTRIBUTION PROTOCOL BASED ON HARDY’S PARADOX

Imagine that two distant parties, Alice and Bob, want to generate a secure key. They have been provided with the facility of public classical communication (which may be insecure but authenticated) as well as transportation of a physical system like a spin-\( \frac{1}{2} \) system or a polarized photon. The quantum key distribution (QKD) protocol can run as follows:

- **S1.** The two parties share a source which emits two-qubit non-maximally entangled ‘Hardy’ states \( |\psi^H\rangle \) given by Eq. (3). It sends one of the qubits from each pair \( |\psi^H\rangle \) to Alice and the other to Bob.

- **S2.** Alice randomly chooses whether to measure \( A = 0 \), or \( A = 1 \) on her qubit. Bob does the same by choosing randomly between measurements of \( B = 0 \) and \( B = 1 \). The observables and the state are chosen such that Hardy’s conditions are met by the measurements. Alice and Bob repeat

---

1 If a state \( |\psi^H\rangle \) attains the maximum probability of success for conditions (1), then \( |\psi^H\rangle = |\eta\rangle \otimes |\phi^H\rangle \), where \( |\eta\rangle \) is an arbitrary bipartite state [16], which allows for a device-independent self test of the state, where \( |\phi^H\rangle \) is the unique two-qubit Hardy state that provides the maximum probability of success for a given set of observables.
such measurements on consecutively emitted pairs of qubits, until they collect sufficiently large statistics. In each run, \(i\), they write down not only the chosen observables, \(A_i\) and \(B_i\) respectively, but also the obtained results, \(a_i\) and \(b_i\), and the timing of the measurement (which allows to link Alice’s and Bob’s measurement data with single acts of emission, or with the emitted pairs - in a one-to-one manner, just like in a standard Bell experiment, and thus it will be denoted just by \(i\)).

- **S3. Check for eavesdropping:** For some randomly selected runs, Alice and Bob both announce their measurement choices (\(A_i\) and \(B_i\)) and the corresponding outcomes (\(a_i\) and \(b_i\)). A state other than \(|\psi^H\rangle\) cannot satisfy all Hardy conditions, \(1\), for the given pairs of local observables. All possible interventions of the eavesdropper can be found out. These may include influencing the internal working of the devices used by Alice and Bob, e.g. by establishing any type of correlation, by coupling to the state \(|\psi^H\rangle\), or by emitting a different state, or by using measurement settings different from those specified by the protocol. Alice and Bob simply publicly compare their announced measurement choices and the corresponding outcomes with the Hardy conditions. Thus we are working in a device-independent scenario.

- **S4.** For the remaining runs, Alice and Bob announce only their measurement outcomes not their bases. Next, to generate their key, they select only those runs for which both of them get 0 outcomes. The conditions \(1\) assure that both of them receive a 0 outcome only if they have chosen the same measurement basis. Let’s call the list \(L\). In a successful Hardy experiment, \(0 < P(a = 0, b = 0|A = 0, B = 0) \neq P(a = 0, b = 0|A = 1, B = 1) > 0\) for a given set of observables and the choice of observable on each side is fully random.

- **S5.** For each run of the list \(L\), they assign a bit value (0, 1) according to their choice of observable. As their observable settings for these runs are correlated, they will generate the same key. Note that the probability of obtaining values 0 and 1 may not be equal and depend on the probability distribution of the randomly chosen settings. We may further drop some part of the results to get an equal number of 0s and 1s in the key. The role of both changing the distribution of settings and the dropping strategy is described in Supplementary Material (Section: The Key Rate).

As mentioned earlier, the ideal Hardy test \(1\) for maximum probability of success \(q = \frac{5\sqrt{5} - 11}{2}\) is, fully device-independent \(15\) - there is only one quantum probability distribution associated with this value. The protocol is secure against general collective memory attacks, but requires perfect experimental data which is not possible to obtain in practice. Therefore we now move to a more realistic, noisy case.

### IV. NOISY CASE

So far we have studied the ideal case in which the conditions \(1\) are always satisfied with \(q\) exactly equal to the maximal quantum value of \(\frac{5\sqrt{5} - 11}{2}\). Now we move to a more practical scenario with some imperfections. Let us assume that the observed probabilities instead of \(1\) satisfy

\[
\begin{align*}
P(a = 0, b = 0|A = 0, B = 0) &\geq q - \epsilon, \\
P(a = 0, b = 0|A = 1, B = 0) &\leq \epsilon, \\
P(a = 0, b = 0|A = 0, B = 1) &\leq \epsilon, \\
P(a = 1, b = 1|A = 1, B = 1) &\leq \epsilon.
\end{align*}
\]

A simple calculation shows that it is only relevant to consider \(0 \leq \epsilon < \frac{1}{3}\) \(15\). We can now adopt the methods of \(6\) to prove that for \(\epsilon\) small enough the protocol remains secure against general attacks if the devices are causally independent. To this end we first need to bound the probability of guessing the setting of one party, say Alice, if the outcomes of both are announced and they happen to be the same. In other words, we find the function \(G\) such that

\[
P_{\text{guess}}(A|a = b) = \max_{A} P(A|a = b) \leq G(\epsilon). \quad (5)
\]

In \(6\) the authors have been able to use a hierarchy of semi-definite programs \(14\) from \(20, 21\) to find the function \(G\) for their case.\(^3\) This was possible because they were interested in the probability, \(P(a|A)\), of guessing the outcome if the setting is known, which is equal \(\text{Tr}(M^a_\rho)\) where \(M^a\) is the projector in basis \(A\) corresponding to the outcome \(a\), where \(\rho\) is the state. Operators \(M^a\) appear directly in the semi-definite programs. In our case there is no linear operator corresponding to \(P(A|a, b)\) and therefore we have to find another way.

To find \(G\) we first use Bayes rule to express the conditional probabilities \(P(A = a|0, b = 0)\) and \(P(A = 1|a = 0, b = 0)\) as

\[
\begin{align*}
P(A = 0|a = 0, b = 0) &= \frac{x}{x + y} \quad \text{and} \\
P(A = 1|a = 0, b = 0) &= \frac{y}{x + y},
\end{align*}
\]

\(^3\) The function \(G\) is concave and upper bounds the guessing probability. Expressing it as a semi-definite problem is desired, since such programs may be efficiently treated numerically using the primal-dual interior point algorithm \(23, 24\).

\(2\) In the noisy case the key is generated from the setting of Alice, which is then reconciled with Bob.
Here, where
\[ x \equiv P(a = 0, b = 0|A = 0, B = 0)P(A = 0, B = 0) + P(a = 0, b = 1|A = 0, B = 1)P(A = 0, B = 1), \]
\[ y \equiv P(a = 0, b = 0|A = 1, B = 0)P(A = 1, B = 0) + P(a = 0, b = 1|A = 1, B = 1)P(A = 1, B = 1). \]

Both expressions \( x \) and \( y \) are variables in the methods of [20, 21], and may be used either as constraints or as the target function, but still their quotient cannot. To overcome this difficulty we first find the allowed range of values of \( x \) with constraints \([3]\) for some \( \epsilon \). Then, using the same procedure, we find the range for \( y \) with \([3]\) and the previously established range of \( x \) as constraints. This allows us to find an area of values for \( x \) and \( y \) for any given \( \epsilon \). Examples of such areas are given in the Supplementary Material.

This gives us the function \( G(\epsilon) \) which we can use in the security proof. The security proof is given in the Supplementary Material. It states that the constraints imposed on an arbitrary set of Bell operators may be used to upper-bound the probability of guessing the measurement settings by Eve, when she knows the outcomes. In particular it applies to the case of Hardy’s paradox with constraints on 4 Bell operators. For the proof we need to assume that the measurements performed by Alice and Bob are independent, so security is limited to collective attacks with causal independence constraints \([22]\).

As mentioned in S5, the probabilities of obtaining 0 and 1 as bits in the key depends on the distribution of the settings and we may apply a strategy where we drop some of the bits. We consider here the case with a uniform distribution of settings and the case with the probability of setting 0 equal to 0.6180. The latter gives \( P(A = 0|a = 0, b = 0) = P(A = 1|a = 0, b = 0) = \frac{1}{2} \) for \( \epsilon = 0 \), so the guessing probability in this case is \( \frac{1}{7} \), giving one bit of min-entropy and the key rate is 0.0689. We admit that our protocol does not provide a better key rate than certain other DIQKD solutions, but still it has other pleasant characteristics. Namely: it uses the settings instead of outcomes as key bits, has more than one security parameter and does not rely on the Bell inequality. These properties may be exploited by some future protocols. In the supplementary material, we provide the tools to prove that this protocol achieves a very strong level of security. Our methods are general and can be used for other protocols. Moreover, all these characteristics are desirable because they: 1. allow more flexibility in the design of QKD protocols; 2. make it more difficult for the eavesdropper to influence settings in comparison to outcomes; 3. provide stronger constraints on the probability of guessing more than one parameter.

V. CONCLUSIONS

We have presented a QKD protocol based on Hardy’s paradox and analyzed its security in both ideal and noisy scenarios. It has two novel features: (a) The bits used for the secret key do not come from the results of the measurements on an entangled state but from the choices of settings which are harder for an eavesdropper to influence; (b) Instead of a single security parameter (a violation of some Bell inequality) a set of them is used to estimate the level of trust in the secrecy of the key. We have also shown how to adopt the existing security proofs for a protocol with these features.

These two properties were not chosen by accident. They both make the eavesdropping harder. The benefit of flipping the roles of outcomes and settings is that the latter are chosen by the parties using their private random number generators over which Eve has no influence. The former, on the other hand, come from the measurements on the systems supplied by the eavesdropper. Using more than a single parameter for security provides more information to the parties about the correlations that they share and puts more limits on the eavesdropping strategies. Note that the parameters related to the security proof of the protocol are joint probabilities so their estimation is not more difficult than the estimation of any Bell inequality violation.

The results presented here show that using two independent private random number generators and checking for Bell inequality violation is not the only option for device independent quantum key distribution. We have also presented a security proof for our protocol.

![Key rates with respect to the noise parameter η](image-url)

FIG. 1. (Color online) Key rates with respect to the noise parameter \( \eta \) given in [3] for different variants of the protocol.
Our paper provides an example of an entirely new class of QKD protocols and provides tools for their analysis.

Acknowledgments:- This work is supported by the Foundation for Polish Science TEAM project (TEAM/2011-8/9/styp7), co-financed by EU European Regional Development Fund and ERC grant QOLAPS (291348), UK EPSRC and NCN grant 2013/08/M/ST2/00626. SDP was implemented in OCTAVE using the SeDuMi [23, 24] toolbox.

VI. SUPPLEMENTARY MATERIAL

In this Supplementary Material we generalise the proof given in Ref. 16 to the case of many Bell operators and replace outcomes with settings. In particular our proof may be applied to the protocol that uses the Hardy paradox. Its key rate is also evaluated using semi-definite programming.

The original theorem states that a condition imposed on a single Bell inequality may certify the randomness of the outcomes. In our proof below we generalise to many Bell inequalities and certify the randomness of the settings, given the outcomes. The latter is more problematic, since the certified entity is non-linear in the associated quantum states.

In the first section, we describe the notation used in the security proof and briefly recapitulate the stages of our QKD protocol. The following section contains a proof of a theorem, that allows to upper-bound the probability of guessing by an eavesdropper of a part of the key. The proof is divided into three parts and summarized in statements 13. The next section relates this upper-bound to the data that may be accessed in the device independent scenario. Then a method of evaluating the guessing probability, using semi-definite programming, is described. Finally, we describe several strategies of post-processing, that allow to increase the privacy and evaluate the key rates obtained for these strategies.

A. Notation

We treat successively emitted pairs of particles as separated subsystems which, with the subsystem of Eve, form one system. In the perfect case, we can use the uniqueness of the Hardy state to protect against the general collective memory attacks, whereas when the noise occurs we need to assume that the measurements are causally independent. This is justified by the no-signaling principle if we use many spatially separated measuring devices and perform the measurements on all emitted pairs simultaneously, or if we use a single measuring device that does not have a memory.

Further, we assume that the order of the subsystems is irrelevant. Since the measurements commute, this may be achieved if after all rounds of the experiment, Alice chooses randomly a permutation of the sequence of rounds, publishes this sequence, and then both Alice and Bob change the ordering of their data according to this permutation. This assures that the results are symmetrically-distributed, which is crucial for deriving the equation (13) below.

Let us consider a family of $N_B$ operators \{$H^i_k\}$, with $k = 1, \ldots, N_B$ for certain $N_B$, where $i$ indexes the subsystem these operators act on. These operators are a linear combination of conditional probabilities, of the form

$$H^i_k = \sum_{a,b,A,B} \alpha_{k,a,b,A,B} P (a_i = a, b_i = b | A_i = A, B_i = B),$$

where $a_i$ and $b_i$ refer to the outcomes, $A_i$ and $B_i$ to the settings of the measurements, for Alice and Bob, respectively, and for a subsystem given by the index $k$. Let us fix the set of $H^i_k$ operators.

In the protocol, Alice and Bob perform experiments with all possible randomized settings and publish all their outcomes. In order to estimate the privacy, Alice and Bob publish a subset of their measurement settings of size $N_{ext}$ in the estimation phase.

After estimation, Alice and Bob ignore those unrevealed pairs that do not have outcomes on both sides equal to 0. So they are working on some subset of the states. Let $N$ denote the number of retained pairs, and $\rho_{ABE} = \sum_{e} P (e|z) \rho_{AB|e|z} \otimes |e|z\rangle \langle e|z|$ the $(2N + 1)$-partite system of Alice, Bob and Eve, with $\text{Tr}_E (\rho_{ABE}) = \sum_{e} P (e|z) \rho_{AB|e|z} \equiv \rho_{AB}$. And let $i$ index the $N$ pairs of subsystems. If Eve chooses (for her subsystem) a measurement setting $z$, she gets a result $e$ with probability $P (e|z)$ and this leaves the Alice-Bob subsystem in a state $\rho_{AB|e|z}$. Let $E$ denote Eve’s knowledge of the pair $z$ and $e$ of her measurement, her influence on the measurements performed and on the preparation of the state $\rho_{ABE}$.

Finally Alice and Bob perform the reconciliation, that requires Alice to publish $N_{cor}$ bits. Afterward both parties obtain a shared set of $N - N_{cor}$ secure bits.

We denote

$$h^{(i)}_{\text{exp}} (\rho_{AB}) \equiv (\text{Tr} (H^i_1 \rho_{AB}), \ldots, \text{Tr} (H^i_{N_B} \rho_{AB})).$$

The probability that Eve guesses correctly $A_i$, given $h^{(i)}_{\text{exp}} (\rho_{AB})$, is

$$P_{\text{guess}} (A_i \parallel h^{(i)}_{\text{exp}} (\rho_{AB})) = \max_{A_i} \text{P} \left(A_i = A_i | E, h_{\text{exp}}^{(i)} (\rho_{AB})\right).$$

The value of $P (A_i = A_i | E, h)$ depends on $A_i$ and $h$, but does not depend directly on the quantum state $\rho_{AB}$, and is the same for all $i$. Thus, let us define

$$\gamma (A, h) \equiv \text{Concave} \left( \max_{E} \text{P} (A_i = A_i | E, h) \right), \quad (7)$$

---

5 From the equation (12), it follows that, in the case of the protocol described in the main text, $\alpha_{1,0,0,0,0} = 0$, $\alpha_{2,0,0,1,0} = -1$, $\alpha_{3,0,0,0,1} = -1$, $\alpha_{4,1,1,1,1} = -1$ and the remaining $\alpha_{k,a,b,A,B}$ are equal to 0.
where Concave(·) means an arbitrarily non-decreasing concave upper-bounding function. Defining
\[ \Gamma(h) \equiv \text{Concave} \left( \max_A \gamma(A, h) \right), \tag{8} \]
we have
\[ P_{\text{guess}}(A_i|h) \leq \max_A \gamma(A, h) \leq \Gamma(h). \]

The reason to use a concave upper-bound is twofold. Eve can use a mixed strategy by emitting states that on average gives the value h for the relevant subsystem. Since Γ is concave it also upper-bounds the guessing probability with this strategy. Moreover, from the definition of concavity, this function is upper-bounded by a hyperplane tangent to an arbitrary point, which is used below in equation (10).

For two operators \( O_1 \) and \( O_2 \) acting on the same Hilbert space \( \mathbb{H} \), we define \( O_1 \succeq O_2 \) if, and only if \( \text{Tr}(O_1(\rho)) \geq \text{Tr}(O_2(\rho)) \) for all states \( \rho \) on \( \mathbb{H} \). It is easy to see that this relation is a partial ordering.

\[ \text{B. Upper-bounding theorem} \]

**Theorem 1.** There exists a linear operator \( G \) in a product form, such that for any \( \rho_{AB} \) we have
\[ \prod_{i=1}^N P_{\text{guess}}(A_i|h_{\text{exp}}^{(i)}(\rho_{AB})) \leq \text{Tr}(G\rho_{AB}). \]

In the first step of the proof of the theorem we need the following

**Statement 1.** For any \( i \), there exist operators \( G_i \) and \( G_i^A \) such that
\[ G_i(\cdot) \succeq \max_{A \in \{0,1\}} \gamma(A, h_{\text{exp}}^{(i)}(\cdot)) \geq 0 \]
\[ G_i^A(\cdot) \succeq \gamma(A, h_{\text{exp}}^{(i)}(\cdot)) \geq 0, \text{ for } A \in \{0,1\} \]
and, for \( i \neq j \), the operator \( \gamma(A, h_{\text{exp}}^{(j)}(\cdot)) \) commutes with operators \( G_j(\cdot) \) and \( G_j^A(\cdot) \).

**Proof.** For any state \( \rho_{AB} \) and \( A \in \{0,1\} \), we have
\[ \text{Tr} \left( \gamma(A, h_{\text{exp}}^{(i)}(\rho_{AB})) \right) \]
\[ \geq \max_E \text{Tr} \left( P \left( A_i = A | E, h_{\text{exp}}^{(i)}(\rho_{AB}) \right) \rho_{AB} \right) \]
\[ = \max_E P \left( A_i = A | E, h_{\text{exp}}^{(i)}(\rho_{AB}) \right) \geq 0, \]
so we obtain \( \gamma(A, h_{\text{exp}}^{(i)}(\cdot)) \geq 0 \).

From the concavity of \( \Gamma(\cdot) \), the following inequality holds for all \( h = (h_1, \ldots, h_N) \) and for any \( h_0 \)
\[ \Gamma(h) \leq \Gamma(h_0) + (h - h_0)^T \cdot \nabla \Gamma(h_0) \]
\[ = \mu(h_0) + \sum_{k=1}^{N_B} \nu_k(h_0) h_k, \tag{10} \]
where \( \mu(h_0) = \Gamma(h_0) - h_0^T \cdot \nabla \Gamma(h_0) \) and \( \nu_k(h_0) = \partial_k \Gamma(h_0). \) Moreover
\[ \Gamma(h) = \min_{h_0} \left( \mu(h_0) + \sum_{k=1}^{N_B} \nu_k(h_0) h_k \right). \tag{11} \]

Thus, for all states \( \rho_{AB} \), for all \( h, h_0 \in \mathbb{R}^{N_B} \) and any \( i \), we have
\[ 0 \leq P_{\text{guess}}(A_i|h) \leq \max_A \gamma(A, h) \]
\[ \leq \Gamma(h) = \min_{h_0} \left( \mu(h_0) + \sum_{k=1}^{N_B} \nu_k(h_0) h_k \right). \tag{12} \]

Let us define an operator
\[ G_i(h_0) = \mu(h_0) \cdot 1 + \sum_{k=1}^{N_B} \nu_k(h_0) H_{ki}^i. \tag{13} \]

Since \( h_{\text{exp}}^{(i)} \) and \( G_j \) act on \( i \)-th and \( j \)-th subsystems, respectively, it is easy to see that \( \gamma(A, h_{\text{exp}}^{(i)}) \) and \( G_j \) commute for \( i \neq j \).

From equation (12), for any \( \rho_{AB} \) and any \( h_0 \), we have
\[ \text{Tr} \left( \gamma(A, h_{\text{exp}}^{(i)}(\rho_{AB})) \right) \]
\[ \leq \text{Tr} \left( \mu(h_0) + \sum_{k=1}^{N_B} \nu_k(h_0) h_{\text{exp},k}^{(i)}(\rho_{AB}) \right) \]
\[ = \text{Tr} \left( G_i(h_0) \rho_{AB} \right). \]

This proves that \( G_i \succeq \gamma(A, h_{\text{exp}}^{(i)}(\cdot)) \).

Similar reasoning leads to the construction of operators \( G_i^A(\cdot) \). Instead of functions \( \mu \) and \( \{\nu\} \), we use the related functions \( \mu^A \) and \( \{\nu^A\} \), which can be constructed because the \( \gamma(A, h) \)'s are concave.

The proof is completed by invoking the transitivity of the relation \( \succeq \).

**Statement 1** points to a difference between our proof and the one given in Ref. 7. We do not assume that
\[ \text{We differ from } 6 \text{ in that we upper-bound a concave function of many variables.} \]
\( \gamma \left( A, h^{(i)}_{\exp}(\rho_{AB}) \right) \) is linear in \( \rho_{AB} \), only that it is semi-positive in the sense \( \gamma \left( A, h^{(i)}_{\exp}(\cdot) \right) \geq 0 \). This is a weaker condition than the standard semi-positive definiteness (see Ref. [18]).

Let us denote \( A_{\text{guess}} \equiv (A_{\text{guess},1}, \ldots, A_{\text{guess},N}) \). In the second step of the proof of theorem [1] we need the following

Statement 2. For any state \( \rho_{AB} \), any \( A_{\text{guess}} \) and any \( h_0 \), we have

\[
\text{Tr} \left( \sum_{i=1}^{N} \gamma \left( A_{\text{guess},i}, h^{(i)}_{\exp}(\cdot) \right) (\rho_{AB}) \right) \leq \text{Tr} \left( \prod_{i=1}^{N} G_i(h_0) \rho_{AB} \right),
\]

where \( \gamma \) is defined in equation [7] and \( G_i \) in equation [13].

Proof. Let \( Z_i(\cdot) = G_i(h_0)(\cdot) - \gamma \left( A_{\text{guess},i}, h^{(i)}_{\exp}(\cdot) \right) \).

From statement [1] we have \( Z_i(\cdot) \geq 0 \). For any state \( \rho_{AB} \), the following holds

\[
\text{Tr} \left( \sum_{i=1}^{N} \left( Z_i(\cdot) + \gamma \left( A_{\text{guess},i}, h^{(i)}_{\exp}(\cdot) \right) \right) (\rho_{AB}) \right) = \text{Tr} \left( \prod_{i=1}^{N} \gamma \left( A_{\text{guess},i}, h^{(i)}_{\exp}(\cdot) \right) (\rho_{AB}) \right).
\]

From the definition of \( \gamma \) (see equation [7]) we get, that this is lower-bounded by

\[
\text{Tr} \left( \prod_{i=2}^{N} Z_i(\cdot) \prod_{i=1}^{N} P \left( A_i = A|E, h^{(i)}_{\exp}(\cdot) \right) + \cdots + \prod_{i=1}^{N} P \left( A_i = A|E, h^{(i)}_{\exp}(\cdot) \right) Z_N(\cdot) \right) (\rho_{AB}) \geq 0
\]

for any \( E \).

This statement is similar to the commonly known relation between semi-definite operators [10], but here one of the involved operators is not linear.

Now we define

\[
G(h_0) \equiv \prod_{i=1}^{N} G_i(h_0).
\]

Let \( P_{\text{guess}}(A|E) \) denote the probability of guessing the vector of settings by Eve \( A \equiv (A_1, \ldots, A_N) \). To prove theorem [1] we also need the following

\[
\text{Statement 3.} \quad \text{Let } G(h_0) \text{ be defined by equation [15] with arbitrary } h_0. \text{ Then for all states } \rho_{AB} \text{ we have } P_{\text{guess}}(A|E) \leq \text{Tr} \left( G(h_0) \rho_{AB} \right).
\]

Proof. Eve prepares some state \( \rho_{ABE} \) such that it optimizes her guessing probability. The value of \( P_{\text{guess}}(A|E) \) is upper-bounded by

\[
\max_{z \neq 0} \sum_{e} P(e|z) \max_{A_{\text{guess}}} \prod_{i=1}^{N} \left( \gamma \left( A_{\text{guess},i}, h^{(i)}_{\exp}(\rho_{AB}|e) \right) \right).
\]

From this, using statement [2] we get

\[
P_{\text{guess}}(A|E) \leq \max_{h_0} \text{Tr} \left( \sum_{e} P(e|z) G(h_0) \rho_{AB}|e \right) = \max_{h_0} \text{Tr} \left( G(h_0) \rho_{AB} \right) = \min_{h_0} \text{Tr} \left( G(h_0) \rho_{AB} \right). \]

We conclude the proof of theorem [1] by taking \( G \equiv \min_{\rho_{AB}} G(h_0) \), which is in a product form, and notice that

\[
\prod_{i=1}^{N} P_{\text{guess}} \left( A_i, h^{(i)}_{\exp}(\rho_{AB}) \right) \leq \max_{E} P_{\text{guess}}(A|E).
\]

C. The estimation phase

Above we have worked with the expectation values, but we cannot access them directly, since we have only one result of the experiment for each subsystem. Below we use a method to estimate these values.

In Ref. [22] the following lemma has been proved:

Lemma 1. If \( V_1, \ldots, V_{N+N_{\text{est}}} \) are symmetrically distributed random variables with a finite set of possible values and \( V_{\text{est}} \equiv \text{Avg}_i \in \{N+1, \ldots, N+N_{\text{est}}\} V_i \), then

\[
P \left( \left( V_{\text{est}} + N_{\text{est}} \right)^{\frac{1}{N}} \leq E \prod_{i=1}^{N} V_i \right) \leq C_1 \exp(-C_2 \sqrt{N_{\text{est}}}),
\]

where \( C_1, C_2 \in \mathbb{R}_+ \), and \( E \) means the expectation value.

We define \( h^{(i)} = (h_1^{(i)}, \ldots, h_N^{(i)}) \) as the vector of the values of the operators estimated in the experiment on the \( i \)-th pair, \( i = 1, \ldots, N+N_{\text{est}} \). Indices \( i \in \{N+1, \ldots, N+N_{\text{est}}\} \) refer to experiments performed in the estimation phase. For a given \( h_0 \), consider a set of random variables

\[
\left\{ V(h_0, h^{(i)}) \equiv V_i \equiv E \left[ \mu(h_0) + \sum_{k=1}^{N_p} \nu_k(h_0) h_k^{(i)} \right] \right\}.
\]

Since the value of each \( V_i \) depends only on a finite set of outcomes, it has a finite set of possible values. Since all the experiments occur in a randomized order, their results are symmetrically distributed, and so are the \( V_i \)s.
We use lemma 1 to obtain that for any $h_0$, and
$V_{est}(h_0) = \text{Avg}_{i \in \{N+1, \ldots, N+N_{est}\}} V(h_0, h^{(i)})$, the following holds with probability $1 - \delta$.

$$\text{Tr} (G(h_0) \rho_{AB})$$

$$= \text{Tr} \left( \prod_{i=1}^{N} \left( \mu(h_0) \cdot I_i + \sum_{k=1}^{N_h} \nu_k(h_0) H_k^i \right) \rho_{AB} \right)$$

$$= E \prod_{i=1}^{N} V_i(h_0) \leq \left( V_{est}(h_0) + N_{est} \right)^N,$$

where $\delta$ decreases exponentially with $N_{est}$. From this, using statement 3 we infer that, for any $\rho_{ABE}$,

$$P_{\text{guess}}(A|\rho_{ABE}) \leq \min_{h_0} \text{Tr} (G(h_0) \rho_{AB})$$

$$\leq \min_{h_0} \left( V_{est}(h_0) + N_{est} \right) = \left( V_{est} + N_{est} \right)$$

holds with probability $1 - \delta$, where $V_{est} \equiv \min_{h_0} V_{est}(h_0)$.

From this, following the reasoning in 7, we get that the certified asymptotic secret key rate is

$$R = - \log_2 V_{est} - H(A|B),$$

where $H(A|B)$ quantifies the amount of communication in the error correction phase (i.e. for reconciliation). The way of calculating it, in the case of the protocol described in this paper, is described below.

Similarly, it is possible to estimate the guessing probability of all values 0, or 1, on a given subset of $N$ bits. To this end we redefine the $V_i$s with $\mu^A$ and $\{\nu^A\}$, instead of $\mu$ and $\{\nu\}$, with $A = 0$ or $A = 1$ respectively.

D. The guessing probability

This section applies the security proof to the QKD protocol demonstrated in the main text, and describes how to use semi-definite programming relaxations 20,21 to evaluate upper-bounds for functions from equations 7 and 5.

In the case of the Hardy paradox we have $N_B = 4$ with

$$H_1^i = P(a_i = 0, b_i = 0|A_i = 0, B_i = 0),$$

$$H_2^i = -P(a_i = 0, b_i = 0|A_i = 1, B_i = 0),$$

$$H_3^i = -P(a_i = 0, b_i = 0|A_i = 0, B_i = 1),$$

$$H_4^i = -P(a_i = 1, b_i = 1|A_i = 1, B_i = 1).$$

In principle, it is possible to evaluate values of $\gamma(A, h)$ and $\Gamma(h)$ (see equations 7 and 8) for an arbitrary vector $h$. However, this is computationally expensive, since it requires a calculation on a many-dimensional grid. Therefore, instead of using a vector of values, we take a single value that gives us a reasonable upper-bound. To this end, we define

$$\epsilon_{est}(h) \equiv \min \{ \epsilon : h(\epsilon) \leq h \},$$

$$h(\epsilon) = \left( \frac{5\sqrt{5} - 11}{2} \right) - \epsilon, -\epsilon, -\epsilon, -\epsilon,$$

and

$$g_0(\epsilon) \equiv \text{Concave} (\gamma(0, h(\epsilon))) \geq \max_{h \leq h(\epsilon)} \gamma(0, h),$$

$$g_1(\epsilon) \equiv \text{Concave} (\gamma(1, h(\epsilon))) \geq \max_{h \leq h(\epsilon)} \gamma(1, h).$$

Here, the relation $\leq$ for vectors is element-wise.

In Figs. 2a and 2b, examples of these are shown for a uniform distribution and in Fig. 3a for a non-uniform distribution. Each area is included in areas with higher $\epsilon$.

Considering the case of uniform settings, we define

$$P(a = 0, b = 0|A = 0) \equiv$$

$$\frac{1}{2} \left( P(a = 0, b = 0|A = 0, B = 0) + P(a = 0, b = 0|A = 0, B = 1) \right).$$

Performing two tiers of optimization, as described in the main text, we obtain areas with allowed $x$s and $y$s for given $\epsilon$, denoted $\text{Area}(\epsilon)$. Examples of these are shown in Fig. 2. If $\epsilon_1 \leq \epsilon_2$, then $\text{Area}(\epsilon_1) \subseteq \text{Area}(\epsilon_2)$.

Functions

$$g_0(\epsilon) = \max_{(x,y) \in \text{Area}(\epsilon)} \left\{ \frac{x}{x+y} \right\}$$

and

$$g_1(\epsilon) = \max_{(x,y) \in \text{Area}(\epsilon)} \left\{ \frac{y}{x+y} \right\}$$

fulfill definitions from equations 19a and 19b, respectively. They are plotted in Fig. 2 for a uniform distribution and in Fig. 3b for a non-uniform distribution.

What remains to be done is to give an experimental method to estimate the value of $\epsilon$. Taking $V^{(k)}_i = h^{(i)}_k$,
and using lemma \[11\] for each \(k = 1, \ldots, N_B\), we get that for \(i \in \{1, \ldots, N\}\), the probability that \(EV_{i}^{(k)} \geq V_{est}^{(k)} + N_{est}^{-\frac{1}{2}}\) for any \(k\), decreases exponentially with \(N_{est}\). Now, defining \(h_{exp} \equiv (V_{est}^{(1)}, \ldots, V_{est}^{(N)})\), we may take 

\[\epsilon = \epsilon_{est}(h_{exp})\]

### E. The key rate

In this section, the key rate of the described protocol for different variants is obtained. We recall that we consider uniform and non-uniform distributions of measurement settings, and usage or not of a dropping strategy. Evaluation of the key rate consists of several steps.

We assume that Alice and Bob tolerate imperfections that, on average, do not exceed a certain value, denoted \(\epsilon\).

First we consider the situation where all settings of Alice with both outcomes equal to 0 are taken as the bits of the key (i.e. no dropping strategy is used). In this case an upper-bound to the guessing probability may be taken directly from the two-tier semi-definite program, by computing the function

\[G(\epsilon) \equiv \text{Concave}(\max\{g_0(\epsilon), g_1(\epsilon)\}) = \text{Concave}\left(\max_{(x,y) \in \text{Area}(\epsilon)} \frac{x}{x+y}, \frac{y}{x+y}\right)\]

for a given distribution of the measurement settings.

We now consider the situation where Alice uses the following dropping strategy. After performing all Hardy experiments, creating the list of \(N\) runs with both outcomes equal to 0, and collecting the measurement settings for these runs, she checks how many bits with values 0 and 1 occur. If these numbers are not equal, she randomly selects some part of the more frequent values and announces them via a public channel to be dropped and not used as the key. Finally, Alice has a key with an equal number of 0s and 1s. The reason for dropping is to increase the average min-entropy of the key\[9\].

We consider the following scenario. For each round of the Hardy experiment, Eve generates a mixed state \(\rho_0\) (here index 0 means that she tries to guess 0 for that round, so she maximizes the probability of obtaining this result) with probability \(p_0\), or a mixed state \(\rho_1\) (meaning she guesses 1) with probability \(p_1 \equiv 1 - p_0\). If the outcomes were both 0s, the state \(\rho_0\) gives the value 0 of the key with probability \(P_0\) and the state \(\rho_1\) gives the value 1 with probability \(P_1\). Upper-bounds for both of these values are obtained from the described two-tier semi-definite program, as functions \(g_0(\epsilon)\) and \(g_1(\epsilon)\), defined by equations \[17a\] and \[19a\], respectively. Fig. 3 shows these bounds for two distributions of the settings.

Applying the formula \[18\] for the expectation values of operators \[17\] in states \(\rho_0\) and \(\rho_1\), we get values \(\epsilon_0\) and \(\epsilon_1\), respectively. Since the order of rounds is not important, Alice and Bob observe the mixed state \(\rho \equiv p_0 \rho_0 + p_1 \rho_1\) and

\[\epsilon = p_0 \epsilon_0 + p_1 \epsilon_1.\] 

\[9\] Obviously after dropping some bits the total min-entropy decreases, but this lessens the communication needed in the reconciliation phase.
for the state $\rho^\Omega$. 

Thus, Eve chooses the values of $\epsilon_0$ and $\epsilon_1$, being constrained by $\min(\epsilon_0, \epsilon_1) \leq \epsilon \leq \max(\epsilon_0, \epsilon_1)$. Then, $P_0 = g_0(\epsilon)$ and $P_1 = g_1(\epsilon)$. When $\epsilon_0 \neq \epsilon_1$, then, from equation (20) and, and from the fact that $p_1 = 1 - p_0$, we get 

$$p_0 = \frac{\epsilon - \epsilon_1}{\epsilon_0 - \epsilon_1} \in [0, 1].$$

(21)

When $\epsilon_0 = \epsilon_1 = \epsilon$, Eve may choose the probability $p_0 \in [0, 1]$ arbitrarily.

Let us define 

$$p_0^A \equiv p_0 P_0 + p_1 (1 - P_1)$$

and

$$p_1^A \equiv p_0 (1 - P_0) + p_1 P_1.$$ 

Before dropping bits, Alice has on average $p_0^A N$ values 0 and $p_1^A N$ values 1. When $p_0^A < p_1^A$ Alice randomly drops $(p_1^A - p_0^A) N$ values 1, so that she has $2p_0^A N$ values with half being equal to 0 and half being equal to 1. Eve will correctly guess $p_0 P_0 N$ values 0 and $p_0^A p_1 P_1 N$ values 1, so her guessing probability is 

$$P_N(p, P_0, P_1) \equiv \frac{1}{2p_0^A N} \left( p_0 P_0 N + \frac{p_0^A}{p_1^A} p_1 P_1 N \right)$$

$$= \frac{1}{2} \left( \frac{p_0 P_0}{pP_0 + (1 - p)(1 - P_1)} + \frac{(1 - p) P_1}{p(1 - P_0) + (1 - p) P_1} \right),$$

where $p_0 = p$, and $p_1 = 1 - p$. We get the same result if $p_0^A > p_1^A$.

This way we arrive at the following formula for the average probability of guessing a setting for a given $\epsilon$:

$$P_{\text{guess}}(\epsilon) = \max \left\{ \max_{\epsilon_0, \epsilon_1} P_G \left( \frac{\epsilon - \epsilon_1}{\epsilon_0 - \epsilon_1}, g_0(\epsilon), g_1(\epsilon) \right), \max_{\epsilon} P_G (p, g_0(\epsilon), g_1(\epsilon)) \right\}.$$ 

From equation (15), we have

$$R \geq - \log_2 P_{\text{guess}}(\epsilon) - H(\mathbb{A}|\mathbb{B})_{a=0,b=0},$$

where $P_{\text{guess}}$ is evaluated using data from Fig. 3 and $H(\mathbb{A}|\mathbb{B})_{a=0,b=0}$ equals the conditional entropy of the setting $\mathbb{A}$ given the setting $\mathbb{B}$, under condition that both outcomes $a$ and $b$ are equal to 0. This expresses the uncertainty of the key established on the side of Alice from Bob’s perspective. It equals

$$H(\mathbb{A}|\mathbb{B})_{a=0,b=0} = - \sum (P(A,B) \log_2 P(A,B))_{a=0,b=0}.$$ 

To evaluate the conditional entropy $H(\mathbb{A}|\mathbb{B})_{a=0,b=0}$ under noise, we consider the state

$$\rho(\eta) \equiv (1 - \eta) \frac{1}{4} + \eta |\psi^+\rangle \langle \psi^+|,$$

and the measurements given in the main text. In this case, for a considered region $\eta \in [0, 1]$, we have $\epsilon = \frac{1 - \eta}{4}$. The certified asymptotic key rate is

$$R(\epsilon) = P_{00} (H_\infty(\epsilon) - H(\mathbb{A}|\mathbb{B})_{a=0,b=0}),$$

where $H_\infty(\epsilon) \equiv - \log_2 P_{\text{guess}}(\epsilon)$ and $P_{00}$ is the probability that the result of a given run of the experiment is accepted as a key. This probability is the probability of getting both results equal to 0 times the probability of dropping a single result (if the dropping strategy is used). The key rate is shown in Fig. 1 in the main text.

---

[1] C. H. Bennett, G. Brassard, Quantum Cryptography: Public key distribution and coin tossing, in Proceedings of the IEEE International Conference on Computers, Systems, and Signal Processing, Bangalore, p. 175 (1984).

[2] D. Mayers, A. Yao, in Proceedings of the 39th Annual Symposium on Foundations of Computer Science (FOCS-98) (IEEE Computer Society, Los Alamitos, CA, 1998), pp. 503-509.

[3] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, V. Scarani, Phys. Rev. Lett. 98, 230501 (2007).

[4] J. Barrett, L. Hardy, A. Kent, Phys. Rev. Lett. 95, 010503 (2005).

[5] A. Acín, N. Gisin, L. Masanes, Phys. Rev. Lett. 97, 120405 (2006).

[6] V. Scarani, N. Gisin, N. Brunner, L. Masanes, S. Pino, A. Acín, Phys. Rev. A 74, 042339 (2006).

[7] L. Masanes, S. Pironio, A. Acín, Nat. Commun. 2, 238 (2011).

[8] J. Barrett, R. Colbeck, A. Kent, Phys. Rev. A, 86, 062326 (2012).

[9] U. Vazirani, T. Vidick, arXiv:1210.1810 (2012).

[10] L. Hardy, Phys. Rev. Lett. 68, 2981 (1992).

[11] V. Scarani, A. Acín, G. Ribordy, N. Gisin, Phys. Rev. Lett. 92, 057901 (2004).

[12] P. Mironowicz, M. Pawłowski, Phys. Rev. A 88, 032319 (2013).

[13] J.-D. Bancal, L. Sheridan, V. Scarani, New J. Phys. 16, 033011 (2014).

[14] O. Nieto-Silleras, S. Pironio, J. Silman, New J. Phys. 16, 013035 (2014).

[15] A. Einstein, B. Podolsky, N. Rosen, Phys. Rev. 47, 777.
[16] G. Kar, *Phys. Lett. A* **228**, 119 (1997).
[17] T. F. Jordan, *Phys. Rev. A* **50**, 62 (1994).
[18] R. Rabelo, L. Y. Zhi, V. Scarani, *Phys. Rev. Lett.* **109**, 180401 (2012).
[19] L. Vandenberghe, S. Boyd, *SIAM Review*, **38**(1): 49-95, March 1996.
[20] M. Navascues, S. Pironio, A. Acín, *Phys. Rev. Lett.* **98**, 010401 (2007).
[21] M. Navascues, S. Pironio, A. Acín, *New J. Phys.* **10**, 073013 (2008).
[22] L. Masanes, R. Renner, M. Christandl, A. Winter, J. Barrett, *Unconditional security of key distribution from causality constraints*, arXiv:quant-ph/0606049.
[23] J. F. Sturm, *Optimization Methods and Software* **11**, 625 (1999).
[24] J. F. Sturm, *Optimization Methods and Software* **17**, 6 (2002).