Device-independent quantum key distribution based on measurement inputs

Ramij Rahaman,[1,2,*] Matthew G. Parker,[3,†] Piotr Mironowicz,[4,5,‡] and Marcin Pawłowski[2,§]
[1]*Department of Mathematics, University of Allahabad, Allahabad 211002, U.P., India*
[2]*Institute of Theoretical Physics & Astrophysics, University of Gdańsk, 80-952 Gdańsk, Poland*
[3]*Department of Informatics, University of Bergen, Post Box-7803, 5020, Bergen, Norway*
[4]*Department of Algorithms and System Modelling, Faculty of Electronics, Telecommunications and Informatics, Gdańsk University of Technology, Gdańsk 80-233, Poland*
[5]*National Quantum Information Centre in Gdańsk, Sopot 81-824, Poland*
(Received 23 September 2013; revised manuscript received 27 April 2015; published 1 December 2015)

We provide an analysis of a family of device-independent quantum key distribution (QKD) protocols that has the following 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. (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. The main advantage of these protocols is a smaller vulnerability to imperfect random number generators made possible by feature (a). We prove the security and the robustness of such protocols. We show that using our method it is possible to construct a QKD protocol which retains its security even if the source of randomness used by communicating parties is strongly biased. As a proof of principle, an explicit example of a protocol based on the Hardy's paradox is presented. Moreover, in the noiseless case, the protocol is secure in a natural way against any type of memory attack, and thus allows one to reuse the device in subsequent rounds. We also analyze the robustness of the protocol using semidefinite programming methods. Finally, we present a postprocessing method, and observe a paradoxical property that rejecting some random part of the private data can increase the key rate of the protocol.

DOI: 10.1103/PhysRevA.92.062304 PACS number(s): 03.67.Dd, 03.65.Ud, 03.67.Ac, 03.67.Mn

I. INTRODUCTION

Developments in quantum mechanics lead to emergence of many new research areas including quantum cryptography [1] and quantum computation [2]. 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 [3], the most secure known way for sending secret messages, is a significant achievement in the field of cryptography. It's techniques allow Alice and Bob to 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 relies on some other assumptions. The foremost among them, 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 use and they also know the exact operation of their measuring devices, etc. The goal of the device-independent [4] analysis of quantum protocols is to eliminate the latter assumption, viz. players can distrust the source of particles and they can also distrust their measuring apparatuses as they might have been prepared by a malicious party.

In 2007, Acín *et al.* [5] 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 [6–8]. In 2011, Masanes *et al.* [9] provided a more general security scheme based on causally independent measurement processes. The security of all these protocols is undermined as the measurement at any step 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 [10,11].

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 eavesdropper's point of view and, if indeed they are, use them as their secret key. In this manuscript we present a family of protocols which are significantly different. 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 prepare-and-measure Scarani-Acín-Ribordy-Gisin 2004 (SARG04) protocol [12].

The potential benefit of flipping the roles of outcomes and settings is that the latter are chosen by the parties using their random number generators (RNGs), which are typically assumed to be perfect, while the former are obtained from measurements on the systems supplied by the eavesdropper. It was shown that even small imperfectness in RNGs are a big threat to QKD [13,14]. We demonstrate that they are a much bigger threat to the protocols where the key is obtained from the outcomes than from the settings. More precisely, we take a standard device-independent QKD and show that it cannot

*ramijrahaman@gmail.com
†Matthew.Parker@ii.uib.no
‡piotr.mironowicz@gmail.com
§dokmpa@univ.gda.pl

 ©2015 American Physical Society

be secure if the bias of RNGs is greater than 0.1 while our protocol allows for positive key rates far beyond this point. This is the main motivation of our approach.

Obviously, the parties need a way to convince themselves that the correlations they share cannot be classical. Checking for a violation of a Bell inequality is only one possible way of doing so. Another option is to, e.g., verify the so-called Hardy's paradox [15,16]. There, more than a single security parameter is estimated, which gives the parties more knowledge about the correlations they share. In [17–19] this approach has been used to improve the rate of certified randomness.

The main result of this paper is the presentation of a protocol which remains secure, even if the source of randomness used by the parties is strongly biased. Besides that, the protocol serves as a proof of principle that one can use the bits from private random number generators as a key for the device-independent QKD. It is shown how a protocol with these properties can be constructed and how its security can be proven. These features may be exploited by some future protocols.

We generalize the results of [9] stating that a condition imposed on a single Bell inequality may certify the randomness of the outcomes. Here we consider the case when there are many parameters used, and the key is formed from the measurement settings with the outcomes made public. What is more, we show that this intrinsically many-valued estimation can be as simple to conduct experimentally as the standard Bell scenario.

Apart from proving the security of such protocols, we provide a way of using semidefinite programming relaxations to evaluate their key rates. We give explicit numerical results for a protocol basing on the original Hardy's paradox.

A. Organization of the paper

The organization of our paper is as follows.

We start in Sec. II with recapitulation of the original Hardy's paradox, show the uniqueness of the Hardy state (Sec. II A), then we describe a QKD protocol and show that in the perfect case it allows for reusing the devices (Sec. II B).

In Sec. III we present the main motivation of the paper by discussing the case when the distribution of settings is biased, and compare the presented protocol with other [9] QKD schemes.

We develop methods that allow one to analyze the introduced family of protocols when the measurements are causally independent in Sec. IV. In Sec. IV A we describe the notation used in the analysis, and in Sec. IV B the system configuration. Section IV C discusses the definition of the *guessing probability of a measurement setting*.

The following Sec. V presents a method of evaluation of the guessing probability of a measurement setting using semidefinite programming.

Section VI discusses the methods of evaluating robustness of the protocol and describes several strategies of postprocessing that allow one to increase the privacy.

The key rates obtained for these strategies are evaluated, again using semidefinite programming, for the case of Hardy's paradox, in Sec. VII.

II. HARDY'S PARADOX AND QUANTUM KEY DISTRIBUTION

In this section we introduce Hardy's paradox [15], and describe a quantum key distribution protocol based on it.

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.

A. Hardy's state

The Hardy-type argument starts with the following set of four joint probability conditions for two two-level systems:

$$\begin{aligned} P(a=0,b=0|A=0,B=0) &\equiv q >0,\\ P(a=0,b=0|A=1,B=0) &=0,\\ P(a=0,b=0|A=0,B=1) &=0,\\ P(a=1,b=1|A=1,B=1) &=0. \end{aligned} \tag{1}$$

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

$$\begin{aligned} |\phi_3\rangle &= |0\rangle_A|0\rangle_B,\\ |\phi_2\rangle &= |0'\rangle_A|0\rangle_B,\\ |\phi_1\rangle &= |0\rangle_A|0'\rangle_B,\\ |\phi_0\rangle &= |1'\rangle_A|1'\rangle_B. \end{aligned} \tag{2}$$

Let

$$\begin{aligned} |0'\rangle_P &\equiv \alpha_P|0\rangle_P + \beta_P|1\rangle_P, \text{ and}\\ |1'\rangle_P &\equiv \beta_P^*|0\rangle_P - \alpha_P^*|1\rangle_P, \end{aligned} \tag{3}$$

where $|\alpha_P|^2+|\beta_P|^2=1$ and $0<|\alpha_P|<1$ for $P=A,B$. The last condition is due to the noncommutativity of $X=0$ and $X=1$.

Let $\mathcal{S}$ 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 (1), ρ has to be confined to a subspace $\mathcal{S}^\perp$ of $\mathcal{C}^2\otimes\mathcal{C}^2$, which is orthogonal to $\mathcal{S}$ but not orthogonal to $|\phi_3\rangle$. The dimension of $\mathcal{S}^\perp$ is one. Therefore, ρ must be an unique (up to a local unitary) 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 [20]. It can also be shown that no two maximally entangled qubit states satisfy Hardy's argument [15].

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

MOST WIEDZY Downloaded from mostwiedzy.pl

$|\phi_3'\rangle = |\psi^H\rangle$ is its last member:

$$|\phi_0'\rangle = |\phi_0\rangle,$$
$$|\phi_i'\rangle = \frac{|\phi_i\rangle - \sum_{j=0}^{i-1}\langle\phi_j'|\phi_i\rangle|\phi_j'\rangle}{\sqrt{1-\sum_{j=0}^{i-1}|\langle\phi_j'|\phi_i\rangle|^2}}, \text{ for } i = 1,2,3. \tag{4}$$

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

$$q = |\langle\psi^H|\phi_3\rangle|^2 = 1 - \sum_{i=0}^{2}|\langle\phi_i'|\phi_3\rangle|^2 = \frac{|\alpha_A\alpha_B|^2|\beta_A\beta_B|^2}{1-|\alpha_A\alpha_B|^2}.$$

Its maximum is $\frac{5\sqrt{5}-11}{2}$ for $|\alpha_A| = |\alpha_B| = \sqrt{\frac{\sqrt{5}-1}{2}}$ [21].

B. Protocol

We consider a scenario in which two distant parties, Alice and Bob, want to generate a secure key. They are allowed to use public classical communication. The QKD protocol proceeds as follows.

S1. In the *initial phase* of the protocol, the two parties obtain pairs of entangled qubits. In each round one of the qubits from each pair is given to Alice, and the other to Bob. Each pair is called a subsystem.

S2. Alice randomly chooses whether to measure $A = 0$, or $A = 1$ on each of her qubits. Bob does the same by choosing randomly between measurements of $B = 0$ and $B = 1$. Parties repeat such measurements on all subsystems, and collect statistics. In each run, labeled by i, they write down the chosen observables, A_i and B_i, respectively, together with the obtained results, a_i and b_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). Alice and Bob publicly compare their announced measurement choices in order to estimate security parameters. For this reason this phase is called the *estimation phase*.

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 got outcome 0. (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.)

S5. For each run with outcomes 0, they assign a bit value according to their settings.

If the pairs of entangled qubits emitted by the shared source are all perfect copies of the two-qubit "Hardy" states $|\psi^H\rangle$ given by Eqs. (4) the assigned bit values will be perfectly correlated due to (1). That is, in the ideal case they generate the same key. In the noisy case, *key reconciliation* is required.

Device-independent approach allows one to quantify all possible interventions of the eavesdropper. 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, by emitting a different state, or by using measurement settings different from those specified by the protocol.

As mentioned earlier, the ideal Hardy test (1) for maximum probability of success $q = \frac{5\sqrt{5}-11}{2}$ is fully device independent [22]—there is a unique quantum probability distribution associated with this value. The conditions (1) assure that both parties got outcome 0 only if they have chosen the same measurement basis. Then we have

$$0 < P(a=0,b=0|A=0,B=0) = \frac{5\sqrt{5}-11}{2}$$
$$< P(a=0,b=0|A=1,B=1) = \sqrt{5}-2$$

for a given set of observables and the choice of observable on each side is fully random. The protocol is secure against the most general form of collective memory attacks. Unfortunately, this case requires perfect experimental data which is not possible to obtain in practice. The remaining part of this paper analyzes the noisy case.

III. BIASED SOURCES OF RANDOMNESS

Before we move to the detailed analysis of the protocol in the case with imperfect experimental data let us present the main motivation of this approach. To this end we will compare the robustness to compromised random number generators in our protocol and the standard one based on CHSH inequality. In both cases we assume that the observed data corresponds to what an experimenter would expect from perfect states and devices.

A common assumption in QKD states that the source of randomness is perfect, meaning that settings are i.i.d. with a probability distribution defined by numbers

$$\mathbb{P}_{\text{perfect}}(A,B) = [P(0,0) = p_A p_B, P(0,1) = p_A(1-p_B),$$
$$P(1,0) = (1-p_A)p_B, P(1,1) = (1-p_A)(1-p_B)], \tag{5}$$

with $p_A = p_B = \frac{1}{2}$ for the uniform probability distribution. In Sec. VII B 2 we introduce a *nonuniform* probability distribution with $p_A = p_B = \frac{1}{2}(\sqrt{5}-1)$.

In this section we consider the case in which the average probability distribution of the source of randomness is given by Eq. (5), but in particular runs, the probability distribution is biased in a way known to eavesdropper. For the sake of simplicity we consider biases modeled by changing the parameters p_A and p_B to $p_A \pm \epsilon$ and $p_B \pm \epsilon$, respectively, for given ϵ, which gives four possible biased distributions, $[\mathbb{P}_{\text{biased},i}(A,B)]_{i=1,2,3,4}$.

It is easy to see that the average distribution (5) can be obtained only if the proportions of all biased distributions in the total number of runs are equal.

Note that if we know only the average distribution given by Eq. (5), then for runs with a particular biased distribution $\mathbb{P}_{\text{biased},i}(A,B)$, the observed conditional probabilities are under- or overestimated, viz.

$$P_{\text{observed}}(a,b|A,B) = P_{\text{actual}}(a,b|A,B)\frac{P_{\text{biased},i}(A,B)}{P_{\text{perfect}}(A,B)}. \tag{6}$$

Let us consider the case without noise described by Eq. (1) with

$$P(a=0,b=0|A=0,B=0) = q = \frac{5\sqrt{5}-11}{2},$$

MOST WIEDZY Downloaded from mostwiedzy.pl

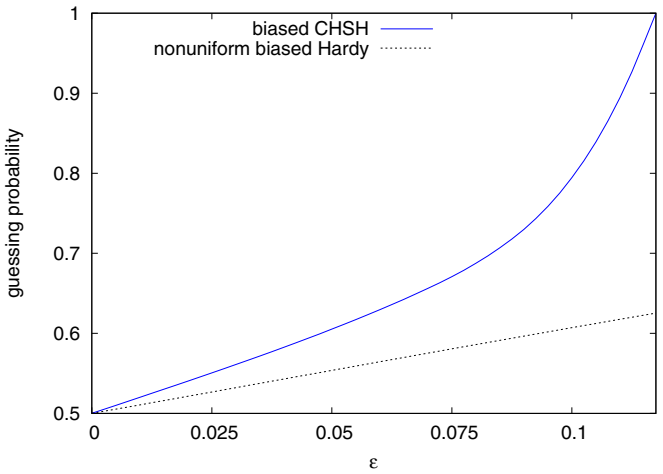

FIG. 1. (Color online) Comparison of guessing probabilities of key values certified with the protocols using Hardy paradox (solid blue line) and CHSH inequality (dotted black line) in the case when the distribution of settings is biased. ϵ refers to the bias defined in Sec. III.

and thus with

$$P(a=0,b=0|A=1,B=1)=\tilde{q}=\sqrt{5}-2, \qquad (7)$$

and with nonuniform distribution of settings. Then, for a given biased distribution $\mathbb{P}_{\text{biased},i}(A,B)$, the probability that the generated key is 0, $P_{i,\text{key}=0}$ and is given by [cf. Eq. (14a)]

$$P_{i,\text{key}=0} \equiv \frac{q\,P_{\text{biased},i}(0,0)}{q\,P_{\text{biased},i}(0,0)+\tilde{q}\,P_{\text{biased},i}(1,1)}.$$

The guessing probability is given by $P_{\text{guess},i} = \max(P_{i,\text{key}=0}, 1-P_{i,\text{key}=0})$, since the eavesdropper tries to guess the more probable key value. To obtain the average guessing probability this expression has to be averaged over all four possible biased probability distributions, namely

$$\sum_{i=1}^{4}\frac{1}{4}P_{\text{guess},i}. \qquad (8)$$

Guessing probabilities for different ϵ with nonuniform distribution of settings are shown in Fig. 1.

In order to compare the efficiency of the presented protocol with other QKD protocols, we use the method of [23,24] to evaluate the guessing probability of the outcomes of Alice certified by the maximal violation of the CHSH inequality with uniform, but biased, distribution of settings,[1] We consider the bound on the guessing probability implied by the observed value of $2\sqrt{2}$ of the following expression:

$$4[P_{\text{biased},i}(0,0)C(0,0)+P_{\text{biased},i}(0,1)C(0,1) + P_{\text{biased},i}(1,0)C(1,0)-P_{\text{biased},i}(1,1)C(1,1)],$$

where 4 is the inverse of the uniform probability of each pair of settings if the distribution were unbiased, and $C(A,B)$ is the correlation between outcomes, when the pair of settings A and B is chosen,

$$C(A,B) \equiv P(0,0|A,B)-P(0,1|A,B) - P(1,0|A,B)+P(1,1|A,B).$$

Similarly, as in the case of the Hardy protocol, the guessing probability has to be averaged over cases with different biases. The results are shown in Fig. 1. We see that for $\epsilon \approx \frac{1}{10}$ the Hardy protocol is still able to work, whereas the CHSH protocol gives zero key rate.

IV. METHODS FOR ANALYSIS OF QUANTUM KEY DISTRIBUTION PROTOCOLS BASED ON HARDY-LIKE PARADOXES

Below we describe the notation used within this paper, the arrangement used in the analysis of the QKD, and give more details about the phases of the QKD protocol.

For the sake of simplicity we consider here a case with perfect RNGs. This can be extended in a natural way to the case with biased probability distributions.

A. Notation and arrangement

In the perfect case, we can use the uniqueness of the Hardy state to protect against the collective memory attacks, whereas if the noise occurs we need to assume that the measurements are causally independent, meaning that their operators commute. 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.

We treat successively emitted pairs of particles as separated subsystems. These subsystems together with the subsystem of Eve form one system. We assume that the order of the subsystems is irrelevant.

Let L_0 be a set of labels of pairs of Alice's and Bob's subsystems. For $l \in L_0$ we denote the Hilbert space of the relevant subsystems of Alice and Bob by $\mathcal{H}_l^{\mathcal{A}}$ and $\mathcal{H}_l^{\mathcal{B}}$, respectively. The subsystem of Eve lives on a Hilbert space $\mathcal{H}^{\mathcal{E}}$. We assume that all spaces are finite dimensional. The Hilbert space of the whole system is

$$\mathcal{H} \equiv \mathcal{H}^{\mathcal{E}} \otimes \prod_{l\in L_0}\mathcal{H}_l^{\mathcal{A}} \otimes \mathcal{H}_l^{\mathcal{B}}. \qquad (9)$$

Vectors on $\mathcal{H}^{\mathcal{E}}$ are denoted by $|e\rangle^{\mathcal{E}}$, and on $\mathcal{H}_l^{\mathcal{P}}$ for $\mathcal{P}\in\{\mathcal{A},\mathcal{B}\}$ by $|e\rangle_l^{\mathcal{P}}$.

For every pair of subsystems both Alice and Bob perform one of the two measurements, each labeled by either 0 or 1. The measurements are binary POVMs denoted by $\tilde{M}_{l,X,x}^{\mathcal{P}}$, where $\mathcal{P}\in\{\mathcal{A},\mathcal{B}\}$ denotes the party, $l\in L_0$ denotes the pair of subsystems, $X\in\{0,1\}$ denotes the party's setting, and $x\in$

[1]The phrase *uniform, but biased distribution* may sound a little bit paradoxical. We mean by this phrase the following situation in which we get a series of random values. Each element of the series is generated with some probability distribution, which does not have to be uniform; thus it may be biased. Nonetheless, we assume that the distribution averaged over the whole series is uniform; $p_A = p_B = \frac{1}{2}$ in Eq. (5). Similarly the nonuniform biased distribution refers to a situation with a series of distributions biased in some way, but on average giving the desired distribution.

MOST WIEDZY Downloaded from mostwiedzy.pl

$\{0,1\}$ denotes the outcome. The measuring operator $\tilde{M}^{\mathcal{P}}_{l,X,x}$ acts on $\mathcal{H}^{\mathcal{P}}_l$.

The natural extension of the operator $\tilde{M}^{\mathcal{P}}_{l,X,x}$ to the space $\mathcal{H}$ is denoted by $M^{\mathcal{P}}_{l,X,x}$, and acts with identity operators on spaces different to $\mathcal{H}^{\mathcal{P}}_l$. From the construction, $M^{\mathcal{P}_1}_{l_1,X_1,x_1}$ commutes with $M^{\mathcal{P}_2}_{l_2,X_2,x_2}$ if $\mathcal{P}_1 \neq \mathcal{P}_2$ or $l_1 \neq l_2$. Recall that $l \in L_0$, and we denote by a_l and b_l the outcomes, and by A_l and B_l the settings, of Alice and Bob, respectively.

Let $S_{\mathcal{E}}$ be an arbitrary set, and $\{|e\rangle^{\mathcal{E}}\}_{e\in S_{\mathcal{E}}}$ be a set of orthogonal states on $\mathcal{H}^{\mathcal{E}}$. Without any loss of generality, we assume the concerning state in the device-independent scenario is of the following form:

$$|\Phi\rangle^{\mathcal{AB}\mathcal{E}} \equiv \sum_{e\in S_{\mathcal{E}}} c_e |\phi_e\rangle, \tag{10}$$

with $c_e \in \mathbb{C}$, $\sum_{e\in S_{\mathcal{E}}} |c_e|^2 = 1$, where

$$|\phi_e\rangle \equiv |e\rangle^{\mathcal{E}} \otimes \left(\bigotimes_{l\in L_0} |e\rangle^{\mathcal{A}}_l \otimes |e\rangle^{\mathcal{B}}_l\right),$$

$\mathcal{P} \in \{\mathcal{A},\mathcal{B}\}, l \in L_0$, and $|e\rangle^{\mathcal{P}}_l$ is a state on $\mathcal{H}^{\mathcal{P}}_l$. Eve is allowed to choose the state $|\Phi\rangle^{\mathcal{AB}\mathcal{E}}$, and the measuring operators $\{M^{\mathcal{P}}_{l,X,x}\}$.

For a subsystem l a conditional probability distribution $\mathbb{P}_l(a,b|A,B) = [P_l(a,b|A,B)]_{a,b,A,B}$ can be defined by

$$P_l(a,b|A,B) \equiv \mathrm{Tr}\left(\tilde{M}^{\mathcal{A}}_{l,a,A}\tilde{M}^{\mathcal{B}}_{l,b,B}\rho_l\right), \tag{11}$$

where ρ_l is the state obtained by tracing all other subsystems in (10).

Let us consider a family of sets of N_B functionals, $\{H_k\}_{k=1,\ldots,N_B}$, acting on conditional probability distributions of the form $\mathbb{P}(a,b|A,B) = [P(a,b|A,B)]_{a,b,A,B}$ (thus $\mathbb{P}_l$ fits this form). These functionals are defined by a set of values $\{\alpha_{k,a,b,A,B}\}$, with $k = 1,\ldots,N_B$, $a,b,A,B \in \{0,1\}$, and are linear combinations of conditional probabilities of the form

$$H_k[\mathbb{P}] = \sum_{a,b,A,B} \alpha_{k,a,b,A,B} P(a,b|A,B). \tag{12}$$

From Eq. (1), it follows that, in the case of the protocol using the original Hardy's paradox, $\alpha_{1,0,0,0,0} = 1$, $\alpha_{2,0,0,1,0} = -1$, $\alpha_{3,0,0,0,1} = -1$, $\alpha_{4,1,1,1,1} = -1$, and the remaining $\alpha_{k,a,b,A,B}$s are equal to 0.

B. Setups of interest

The protocol presented in this paper is device independent, since it relies only on the observed statistics. The main aim is to show that it is possible to prove the security of a key generated out of settings.

In order to illustrate what orders of key rates can be expected to occur in real experiments, we refer to the setup of each subsystem with Hardy's measurements and the following state:

$$\rho(\eta) \equiv (1-\eta)\frac{\mathbb{1}}{4} + \eta|\psi^H\rangle\langle\psi^H|. \tag{13}$$

The observed statistics $\mathbb{P}(a,b|A,B)$ do not depend on the distribution of settings, $\mathbb{P}(A,B)$; nevertheless, the key rate does. We consider two distributions: uniform and the one described in Sec. VII B 2, referred to further as *nonuniform*.

An additional benefit of using nonuniform distribution is the fact that it requires less randomness.

C. Guessing probability of a setting

Let us consider a particular subsystem l. We are interested in conditional probabilities of Alice's *settings* A, when we know that both Alice and Bob got the outcome 0, namely $P(A|a=0,b=0)$. These probabilities may be expressed in terms of $\mathbb{P}(a,b|A,B)$ with use of the Bayes rule, as

$$P(A=0|a=0,b=0) = \frac{\sigma}{\sigma+\nu}, \text{ and} \tag{14a}$$

$$P(A=1|a=0,b=0) = \frac{\nu}{\sigma+\nu}, \tag{14b}$$

where

$$\begin{aligned}\sigma \equiv\ & P(a=0,b=0|A=0,B=0)P(A=0,B=0)\\ &+ P(a{=}0,b{=}0|A=0,B=1)P(A=0,B=1),\end{aligned} \tag{15a}$$

$$\begin{aligned}\nu \equiv\ & P(a=0,b=0|A=1,B=0)P(A=1,B=0)\\ &+ P(a{=}0,b{=}0|A=1,B=1)P(A=1,B=1).\end{aligned} \tag{15b}$$

Let $\mathbf{h} = (h_1,\ldots,h_{N_B})$ denote the values of functionals defined by Eq. (12) over the probability distribution $\mathbb{P}_l$ defined by Eq. (11), so that $h_k = H_k[\mathbb{P}_l]$. For the setup of interest, (13), $\mathbf{h}$ is given by the following [cf. (1)]:

$$\begin{aligned} h_1 &= \eta q + \frac{1-\eta}{4},\\ h_2 = h_3 = h_4 &= \frac{1-\eta}{4}. \end{aligned} \tag{16}$$

Now, let us ignore the full knowledge about $\mathbb{P}_l$, and consider only the vector $\mathbf{h}$. We introduce two functions, $\Gamma_0(\mathbf{h})$ and $\Gamma_1(\mathbf{h})$, that give upper bounds for values of $P_l(A=0|a=0,b=0)$ and $P_l(A=1|a=0,b=0)$, respectively, allowed by quantum mechanics. Note that these functions do not depend on l, since they do not make any assumptions about the state and the measurements, so they give device-independent bounds. Examples of these functions for $\mathbf{h}$ given by Eq. (16) are shown in Fig. 2.

V. SEMIDEFINITE PROGRAMMING RELAXATION OF THE GUESSING PROBABILITY

This section describes how to use semidefinite programming [25] methods to evaluate upper bounds for functions $\Gamma_0(\mathbf{h})$ and $\Gamma_1(\mathbf{h})$. Expressing them as a semidefinite problem is desired, since such programs may be efficiently treated numerically using the primal-dual interior point algorithm [26,27].

In [9] the authors have been able to use a hierarchy of semidefinite programs from [23,24], called NPA, to find upper bounds for their case. It was possible because they were interested in the probability of guessing the outcome if the setting is known, $P(a|A)$. These probabilities appear directly in the semidefinite programs as problem variables. In our case there is no variable corresponding to $P(A|a,b)$, and therefore we have to find it another way.

MOST WIEDZY Downloaded from mostwiedzy.pl

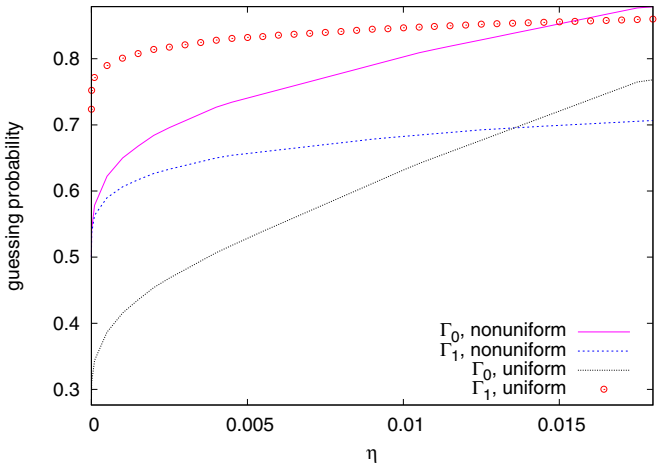

FIG. 2. (Color online) Functions $\Gamma_0(\mathbf{h})$ and $\Gamma_1(\mathbf{h})$ for uniform and nonuniform (see Sec. VII B 2) distribution of settings. These functions give upper bounds for values of $P(A=0|a=0,b=0)$ and $P(A=1|a=0,b=0)$ for $\mathbf{h}$ given by (16).

Let us consider a subsystem l. Without loss of generality, using no signaling principle, we may assume that the eavesdropper performs her measurement with result e before Alice and Bob start the protocol. This does not reduce the generality of the attacks available to the eavesdropper [9]. Moreover, to consider probability distributions allowed for a particular subsystem l, we may trace out other subsystems and perform optimization over bipartite states.

To use the NPA method, we introduce functions $\tilde{\Gamma}_0(\mathbf{h})$, and $\tilde{\Gamma}_1(\mathbf{h})$, which give the relevant upper bounds on $P(A|a=0,b=0)$ assuming that the state under consideration is pure. Then $\Gamma_0(\mathbf{h})$, and $\Gamma_1(\mathbf{h})$, are concave hulls of $\tilde{\Gamma}_0(\mathbf{h})$, and $\tilde{\Gamma}_1(\mathbf{h})$, respectively.

We are interested in using the NPA hierarchy in order to calculate $\tilde{\Gamma}_0(\mathbf{h})$. Since the expressions in (14a) and (14b) are not linear in variables occurring in NPA, they cannot by used either as target, or as constraint.

To overcome this difficulty, we notice that both σ and ν in (15a) and (15b) are linear in NPA variables. In the general case we need to perform the optimization in two stages. In the first stage we impose the constraints given by the vector $\mathbf{h}$, which can be easily done in the NPA hierarchy, and calculate the scope of the allowed values of σ for given $\mathbf{h}$. In the second stage we calculate the scope of the allowed values of ν for given $\mathbf{h}$ and given value of σ, for some grid of values. This way we obtain a boundary of some region for which it is possible to evaluate the bounds on both (14a) and (14b).

For Hardy's paradox the optimization is much simpler. In this case σ is a function of $\mathbf{h}$. It is easy to see that the expressions (14a) and (14b) achieve their maximal values, if ν gets it's minimal or maximal value, respectively.

Obviously calculating a function for all possible values of a continuous parameter is impossible. Instead we calculate it only for some grid of values. Now, if we represent the function with a set of vectors, each containing the coordinates of a single point together with the value of the function, then the problem of linearly constrained optimization over this function can be solved with linear programming. Examples of such problems are programs stated in Eqs. (17) and (18) further in this paper.

VI. EVALUATING KEY RATES AND POSTPROCESSING STRATEGIES

Here we describe the method of evaluating the key rate achieved by protocols based on Hardy-like paradoxes. We also discuss some postprocessing strategies that allow one to increase the key rate.

A. Basic case

In the simplest case described in Sec. II B, the best thing the eavesdropper may do is to maximize his guessing probability, namely $P^{(1)}_{\text{guess}}(\mathbf{h})$. Subsystems can be divided into two groups. For states within the first group, the eavesdropper makes a guess that the key value is 0, and for states from the second group, she guesses the key value 1.

The probability that a subsystem belongs to the first group is p_0. The average values of the Bell observables (12) (or *statistics*) from this group is given by $\mathbf{h}_0$, which allows for guessing 0 by the eavesdropper with the probability upper bounded by $\Gamma_0(\mathbf{h}_0)$. The remaining part of subsystems (with probability $p_1 = 1 - p_0$) has the statistics given by $\mathbf{h}_1$, and the eavesdropper guesses correctly the key value 1 with probability not exceeding $\Gamma_1(\mathbf{h}_1)$.

The solution of the following program gives the relevant upper bound for the average guessing probability:

$$\begin{aligned} &\text{maximize } p_0\Gamma_0(\mathbf{h}_0) + p_1\Gamma_1(\mathbf{h}_1), \\ &\text{subject to } p_0\mathbf{h}_0 + p_1\mathbf{h}_1 = \mathbf{h}, \\ &\qquad p_0 + p_1 = 1, \\ &\qquad p_0, p_1 \geqslant 0. \end{aligned} \tag{17}$$

Solutions of this program for $\mathbf{h}$ given by (16) are shown in Fig. 3.

The key rate is given by the following formula:

$$K_1 = P(a=0,b=0)\left\{ -\log_2\left[P^{(1)}_{\text{guess}}(\mathbf{h})\right] - H(A|B)\right\},$$

where both expressions, $P(a=0,b=0)$ and the conditional entropy $H(A|B)$, can be evaluated from the setup. Examples

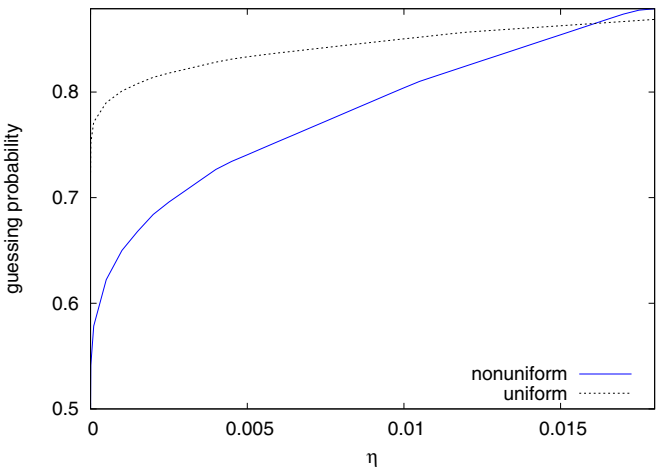

FIG. 3. (Color online) Solutions of the program (17) for cases with uniform and nonuniform (see Sec. VII B 2) distribution of settings.

MOST WIEDZY Downloaded from mostwiedzy.pl

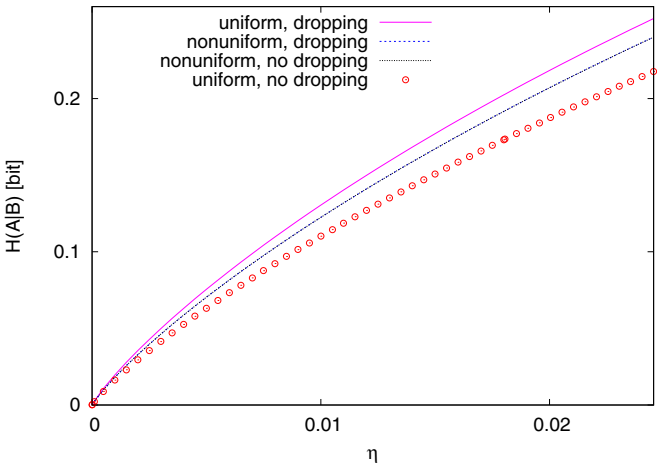

FIG. 4. (Color online) Values of conditional entropies of the setting of Alice given the setting of Bob, $H(A|B)$, if outcomes on both sides were equal to 0. The cases with uniform and nonuniform distribution of settings, and with and without dropping strategy, are considered. The η parameter refers to the state given by Eq. (13). In the case with nonuniform distribution, the line referring to use of dropping strategy is slightly above the one without dropping.

of conditional entropies for different setups from Sec. IV B and postprocessing strategies are shown in Fig. 4.

B. Dropping strategy

In a long sequence of N runs, the number of runs with both outcomes equal to 0 is

$$n \approx P(a=0,b=0)N.$$

Eve in p_0 fraction of all runs tries to guess that the key value is 0, and in $p_1 \equiv 1-p_0$ part of the runs, that it is 1. The former part of runs gives statistics $\mathbf{h}_0$, and the latter $\mathbf{h}_1$. Since the observed statistics are given by $\mathbf{h}$, we have $\mathbf{h} = p_0\mathbf{h}_0 + p_1\mathbf{h}_1$.

Let her success probability in each of these two cases be denoted by P_0 and P_1, respectively.

Now, let us consider only those runs in which both published outcomes were 0. Among them the number of runs with the setting of Alice equal to 0 is

$$[p_0P_0 + p_1(1-P_1)]n \equiv p_0^A n,$$

and equal to 1 is

$$[p_0(1-P_0) + p_1P_1]n \equiv p_1^A n.$$

If $p_0^A < p_1^A$, then Alice discards (or *drops*) $(p_1^A - p_0^A)n$ runs with the value 1. After dropping she has equal number of both values, namely $p_0^A n$ of each. In this situation Eve correctly guesses p_0P_0n of runs with the value 0, and $\frac{p_0^A}{p_1^A}p_1P_1n$ of runs with the value 1, so her guessing probability (among those runs that were not dropped) is given by

$$P^{(2)}_{\text{guess}} \equiv \frac{1}{2p_0^A n}\left(p_0P_0n + \frac{p_0^A}{p_1^A}p_1P_1n\right) = \frac{1}{2}\left(\frac{p_0P_0}{p_0^A} + \frac{p_1P_1}{p_1^A}\right).$$

The case with $p_0^A > p_1^A$ gives exactly the same formula.

To calculate $P^{(2)}_{\text{guess}}(\mathbf{h})$ (as a function of $\mathbf{h}$) *via* a linear optimization, the values $p_0^A = P(A=0|a=0,b=0)$ and $p_1^A = P(A=1|a=0,b=0)$ have to be calculated from the setup. Then the bound on the guessing probability is given by the solution of the following program:

$$\begin{aligned} \text{maximize } & \frac{1}{2}\left(\frac{1}{p_0^A}p_0\Gamma_0(\mathbf{h}_0) + \frac{1}{p_1^A}p_1\Gamma_1(\mathbf{h}_1)\right) \\ \text{subject to } & p_0\mathbf{h}_0 + p_1\mathbf{h}_1 = \mathbf{h}, \\ & p_0 + p_1 = 1, \\ & p_0, p_1 \geqslant 0. \end{aligned} \tag{18}$$

The key rate is now given by the following formula:

$$K_2 = P(a=0,b=0)[2\ \min(p_0^A, p_1^A)] \times \left\{-\log_2\left[P^{(2)}_{\text{guess}}(\mathbf{h})\right] - H(A|B,\text{dropping})\right\}.$$

Both expressions $P(a=0,b=0)$ and the conditional entropy $H(A|B,\text{dropping})$ can be evaluated directly from the setup.

VII. ROBUSTNESS OF QUANTUM KEY DISTRIBUTION PROTOCOLS BASING ON HARDY'S PARADOX

In this section the method described in Secs. V and VI is applied to the experimental setup described in Sec. IV B. The resulting key rates are shown in Fig. 5.

A. Results

The numerical results concerning the obtained key rates in different situations are shown in Fig. 5. The optimal choice of the distribution of settings depends on the value of the noise parameter η. Although the nonuniform distribution gives better key rates with lower noise, the uniform distribution can be more robust.

In the case of nonuniform distribution, the role of the dropping strategy seems to be marginal. This is not surprising, since the aim of that distribution is to make the number of

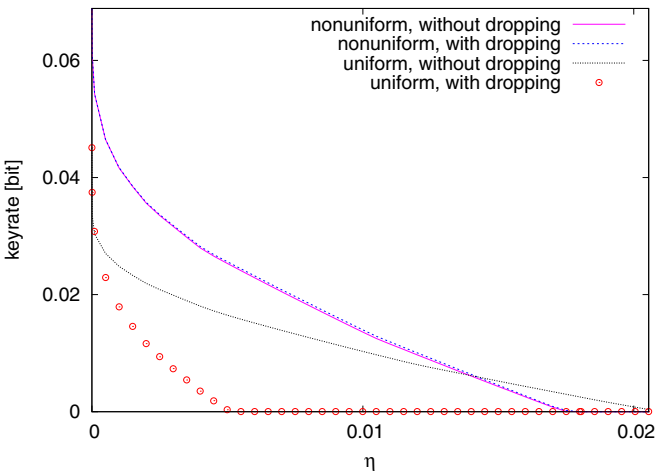

FIG. 5. (Color online) Comparison of key rates in different scenarios.

MOST WIEDZY Downloaded from mostwiedzy.pl

values 0 and 1 more or less equal. Similar results refer to conditional entropies (cf. Fig. 4).

A characteristic property of these protocols is the fact that the use of nonuniform distribution of settings not only requires less randomness, but also in some cases improves the key rate.

B. Case without noise

The analysis of the case without noise gives an insight to the reason why the use of the dropping strategy can increase the key rate. It also explains the role of nonuniform distribution of settings.

1. Dropping strategy

In the perfect case with uniform distribution we have $P(a=0,b=0|A=0,B=0)=\frac{5\sqrt{5}-11}{2}\approx 0.090167$ and $P(a=0,b=0|A=1,B=1)=\sqrt{5}-2\approx 0.236068$, so $P(a=0,b=0|A=0,B{=}0){<}P(a=0,b=0|A{=}1,B{=}1)$. Hence the guessing probability for Eve is higher than $\frac{1}{2}$. The following dropping strategy makes the guessing probability equal to $\frac{1}{2}$.

After performing her measurements, Alice randomly selects only $\frac{P(a=0,b=0|A=0,B=0)}{P(a=0,b=0|A=1,B=1)}$ runs from the total runs with $a=b=0$, where her measurement settings were $A=1$ (in the perfect case, then also $B=1$). Alice sends the list of selected runs to Bob via a public channel.

For this reduced list (from which the key is generated) of runs Alice has an equal number of 0's and 1's. (In the perfect case they correspond to the same values on the side of Bob.) Hence the guessing probability is now exactly equal to $\frac{1}{2}$.

We have

$$\begin{aligned}&P_{\text{not dropped}}(a=0,b=0|A=0,B=0)\\&\quad+P_{\text{not dropped}}(a=0,b=0|A=1,B=1)\\&\quad=5\sqrt{5}-11\approx 0.180334.\end{aligned}$$

To get the actual ratio of runs that are contained in the key, this should be multiplied by

$$P(A=B=0)=P(A=B=1)=\tfrac{1}{4}.$$

Thus the key rate is approximately $\frac{5\sqrt{5}-11}{4}\approx 0.04508$.

2. Nonuniform distribution of settings

Instead of choosing measurement settings with equal probabilities, both Alice and Bob may choose the observables $A=0$ ($B=0$) and $A=1$ ($B=1$) with a ratio $r:1-r$, for some r.

Let us denote for simplicity $P(a=0,b=0|A=0,B=0)=x$ and $P(a=0,b=0|A=1,B=1)=y$. Then to obtain guessing probability equal to $\frac{1}{2}$, the condition for r is

$$xr^2=y(1-r)^2, \text{ or equivalently} \tag{19a}$$

$$r=\frac{\sqrt{y}}{\sqrt{x}+\sqrt{y}}. \tag{19b}$$

The key rate is thus $2xr^2=2x\frac{y}{(\sqrt{x}+\sqrt{y})^2}$.

In the perfect case $x=\frac{5\sqrt{5}-11}{2}$ and $y=\sqrt{5}-2$, so the key rate is $2xr^2\approx 0.06888$. The ratio is $r=\frac{1}{2}(\sqrt{5}-1)\approx 0.61803$.

VIII. CONCLUSIONS

Our paper provides an example of an entirely different class of QKD protocols and provides tools for their analysis.

We have presented a QKD protocol based on Hardy's paradox and analyzed its security in both ideal and noisy scenarios. It has two 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 more difficult for an eavesdropper to influence. (b) Instead of a single security parameter a set of them is used to estimate the level of trust in the secrecy of the key, or to construct a certifying observable.

We have shown that these two properties were not chosen by accident. They both make the eavesdropping harder, leading to protocols which can produce a positive amount of shared key even if the biases of the source of randomness are strong. 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.

ACKNOWLEDGMENTS

This work is supported by the Foundation for Polish Science TEAM project (TEAM/2011-8/9/styp7), cofinanced by EU European Regional Development Fund and ERC grant QOLAPS (No. 291348), a grant of Ministry of Science and Higher Education of the Republic of Poland IDEAS PLUS (No. IdP2011 000361), a National Science Centre (NCN) grant No. 2013/08/M/ST2/00626, and a National Science Centre project Maestro No. DEC-2011/02/A/ST2/00305. R.R. also acknowledges partial support by the UGC (University Grants Commission, Govt. of India) Start-Up Grant. S.D.P. was implemented in the free software package GNU OCTAVE [28] using the SeDuMi [26,27] toolbox.

[1] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dusek, N. Lutkenhaus, and M. Peev, Rev. Mod. Phys. **81**, 1301 (2009).

[2] D. DiVincenzo, Science **270**, 255 (1995).

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

[4] D. Mayers and 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.

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

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

MOST WIEDZY Downloaded from mostwiedzy.pl

[7] A. Acín, N. Gisin, and L. Masanes, Phys. Rev. Lett. **97**, 120405 (2006).
[8] V. Scarani, N. Gisin, N. Brunner, L. Masanes, S. Pino, and A. Acín, Phys. Rev. A **74**, 042339 (2006).
[9] L. Masanes, S. Pironio, and A. Acín, Nat. Commun. **2**, 238 (2011).
[10] J. Barrett, R. Colbeck, and A. Kent, Phys. Rev. A **86**, 062326 (2012).
[11] U. Vazirani and T. Vidick, Phys. Rev. Lett. **113**, 140501 (2014).
[12] V. Scarani, A. Acín, G. Ribordy, and N. Gisin, Phys. Rev. Lett. **92**, 057901 (2004).
[13] J. Bouda, M. Pivoluska, M. Plesch, and C. Wilmott, Phys. Rev. A **86**, 062308 (2012).
[14] M. Huber and M. Pawłowski, Phys. Rev. A **88**, 032309 (2013).
[15] L. Hardy, Phys. Rev. Lett. **68**, 2981 (1992).
[16] S. Mansfield and T. Fritz, Found. Phys. **42**, 709 (2012).
[17] P. Mironowicz and M. Pawłowski, Phys. Rev. A **88**, 032319 (2013).
[18] J.-D. Bancal, L. Sheridan, and V. Scarani, New J. Phys. **16**, 033011 (2014).
[19] O. Nieto-Silleras, S. Pironio, and J. Silman, New J. Phys. **16**, 013035 (2014).
[20] G. Kar, Phys. Lett. A **228**, 119 (1997).
[21] T. F. Jordan, Phys. Rev. A **50**, 62 (1994).
[22] R. Rabelo, L. Y. Zhi, and V. Scarani, Phys. Rev. Lett. **109**, 180401 (2012).
[23] M. Navascues, S. Pironio, and A. Acín, Phys. Rev. Lett. **98**, 010401 (2007).
[24] M. Navascues, S. Pironio, and A. Acín, New J. Phys. **10**, 073013 (2008).
[25] L. Vandenberghe and S. Boyd, SIAM Rev. **38**, 49 (1996).
[26] J. F. Sturm, Optim. Methods Softw. **11**, 625 (1999).
[27] J. F. Sturm, Optim. Methods Softw. **17**, 6 (2002).
[28] J. W. Eaton, D. Bateman, and S. Hauberg, GNU Octave version 3.0.1 manual: a high-level interactive language for numerical computations (CreateSpace Independent Publishing Platform, 2009).

MOST WIEDZY Downloaded from mostwiedzy.pl