A Device-Independent Protocol for XOR Oblivious Transfer

Srijita Kundu
Centre for Quantum Technologies, National University of Singapore, Singapore
srijita.kundu@u.nus.edu

Jamie Sikora
Perimeter Institute for Theoretical Physics, Waterloo, Ontario, Canada
jsikora@perimeterinstitute.ca

Ernest Y.-Z. Tan
Institute for Theoretical Physics, ETH Zürich, Switzerland
ernestt@ethz.ch

Abstract

Oblivious transfer is a cryptographic primitive where Alice has two bits and Bob wishes to learn some function of them. Ideally, Alice should not learn Bob's desired function choice and Bob should not learn any more than logically implied by the function value. While decent quantum protocols for this task are known, many quickly become insecure if an adversary were to control the quantum devices used in the implementation of the protocol. Here we present how some existing protocols fail in this device-independent framework, and give a fully-device independent quantum protocol for XOR oblivious transfer which is provably more secure than any classical protocol.

2012 ACM Subject Classification Security and privacy → Cryptography; Theory of computation → Cryptographic primitives

Keywords and phrases Quantum cryptography, device independence, oblivious transfer, semidefinite programming, security analysis

Digital Object Identifier 10.4230/LIPIcs.TQC.2020.12

Related Version A full version of the paper is available at `https://arxiv.org/abs/2006.06671`.

Supplementary Material The code used for the SDP computations is available at `https://github.com/ernesttyz/dixot`.

Funding *Srijita Kundu*: Supported by the Singapore Ministry of Education and the National Research Foundation, Prime Minister's Office, Singapore. Part of this work was done when S. K. was visiting the Institute for Quantum Computing at the University of Waterloo, Canada.
Jamie Sikora: Supported by Government of Canada through the Department of Innovation, Science and Economic Development Canada, and by the Province of Ontario through the Ministry of Economic Development, Job Creation and Trade.
Ernest Y.-Z. Tan: Supported by the Swiss National Science Foundation via the National Center for Competence in Research for Quantum Science and Technology (QSIT), the Air Force Office of Scientific Research (AFOSR) via grant FA9550-19-1-0202, and the QuantERA project eDICT.

Acknowledgements We thank Jean-Daniel Bancal, Andrea Coladangelo, Lídia del Rio, Honghao Fu, Anand Natarajan, Christopher Portmann, Xingyao Wu and Vilasini Venkatesh for helpful discussions.

1 Introduction

Oblivious transfer is an important cryptographic primitive in two-party computation as it can be used as a universal building block for constructing more elaborate protocols [10]. Indeed, some quantum protocols for this task are known [22, 18, 6, 5]. It can be shown that there do not exist classical protocols with any level of information-theoretic security, and there do not exist quantum protocols with perfect security [6, 11].

In this paper, we consider a variant of oblivious transfer called *XOR oblivious transfer (XOT)*. This is the two-party cryptographic primitive in which two spatially separated parties, Alice and Bob, wish to do the following task: Alice outputs two bits (x_0, x_1), which are uniformly random in $\{0,1\}^2$, and Bob outputs b which is uniformly random in $\{0,1,2\}$, as well as x_b where we define $x_2 = x_0 \oplus x_1$. In other words, Alice and Bob communicate and Bob learns one bit of information from Alice's two bits (either the first bit, second bit, or their XOR). When designing XOT protocols, the security goals are:

1. *Completeness:* If both parties are honest, then their outcomes are consistent (i.e., x_b is the correct value), uniformly random, and neither party aborts.
2. *Soundness against cheating Bob:* If Alice is honest, then a dishonest (i.e., cheating) Bob cannot learn both x_0 and x_1 by digressing from protocol.
3. *Soundness against cheating Alice:* If Bob is honest, then a dishonest (i.e., cheating) Alice cannot learn b by digressing from protocol.

▶ Remark 1. One could imagine a situation where Alice already has a fixed choice of (x_0, x_1) that she wishes to input into a XOT protocol (perhaps from the result of an earlier computation). However, we can use the outcomes of an XOT protocol as described above as a one-time pad to convey the information to Bob. For more details, see [6].

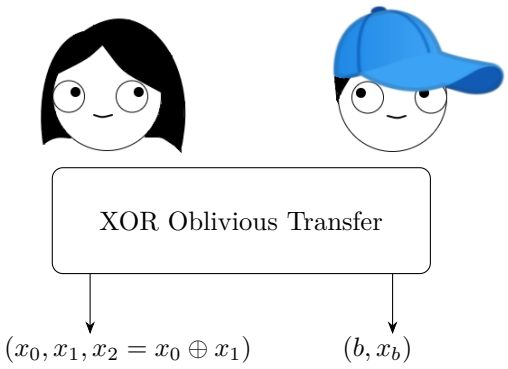

Figure 1 Desired outputs for XOR oblivious transfer.

In this paper we are concerned with *information-theoretic security*, meaning that Alice and Bob are only bounded by the laws of quantum mechanics. In other words, Alice and Bob can perform arbitrarily complicated computations, have arbitrarily large quantum memories, and so on. We shall have occasion to change how much control Alice and Bob have over the protocol, but precisely what actions are allowed to be performed by dishonest parties should be clear from context, and will be described shortly.

We focus on studying XOT protocols from the perspective of assuming perfect completeness and trying to make them as sound as possible.[1] To this end, we choose to quantify the soundness via *cheating probabilities*, which we define as follows:

$\mathrm{P}_B^{\mathrm{XOT}}$: The maximum probability with which a dishonest Bob can learn both of honest Alice's outcome (x_0, x_1) and the protocol does not abort.

$\mathrm{P}_A^{\mathrm{XOT}}$: The maximum probability with which a dishonest Alice can learn honest Bob's choice outcome b and the protocol does not abort.

Any XOT protocol with perfect completeness necessarily has $\mathrm{P}_A^{\mathrm{XOT}} \geq \frac{1}{3}$ and $\mathrm{P}_B^{\mathrm{XOT}} \geq \frac{1}{2}$, since a dishonest Alice could always guess one of three choices for Bob's outcome b uniformly at random; similarly, dishonest Bob can follow the honest protocol to gain perfect knowledge of x_0, x_1, or $x_0 \oplus x_1$, and then randomly guess the unknown bit in Alice's outcome (x_0, x_1).

As mentioned earlier, XOT is a variation on the more well-known cryptographic primitive, oblivious transfer (OT). In an oblivious transfer protocol, Alice and Bob wish to do the following task: Alice outputs two bits (x_0, x_1), which are uniformly random in $\{0,1\}^2$, and Bob outputs (b, x_b) where b is uniformly random in $\{0,1\}$. The completeness and soundness against cheating Alice and Bob for OT are the same as in XOT, the only difference being that here Alice is trying to learn which one of two possible values of b Bob has. Any OT protocol with perfect completeness has $P_A^{\mathrm{OT}} \geq 1/2, P_B^{\mathrm{OT}} \geq \frac{1}{2}$.

▶ Remark 2. In this work, we chose to quantify soundness via cheating probabilities, but we note that such a measure of security is not necessarily *composable* [23, 21]. Unfortunately, it can be very challenging to prove that a protocol is composably secure, and in some settings such protocols are in fact impossible [21]. As a first analysis of the protocol proposed in this work, we will restrict ourselves to studying the cheating probabilities only.

Since the lowest possible bounds on $\mathrm{P}_A^{\mathrm{XOT}}$ and $\mathrm{P}_B^{\mathrm{XOT}}$ for perfectly complete XOT protocols are asymmetric, we shall consider them in pairs and will not concern ourselves with finding an "optimal protocol". That is, we consider the security of XOT protocols a partial ordering. Instead, we motivate our work by asking the following question:

"Is is possible to create quantum protocols where both $\mathrm{P}_A^{\mathrm{XOT}}, \mathrm{P}_B^{\mathrm{XOT}} < 1$ *when Alice and Bob do not even trust their own quantum devices?"*

Taken literally, this statement cannot be true, since arbitrarily malicious devices could simply broadcast all desired information to a dishonest party. However, it turns out that there exist quantum protocols that can be proven secure using almost no assumptions other than ruling out this extreme scenario (which seems a rather necessary assumption in any case). This is the notion of *device-independent* security, which typically exploits *nonlocal games* played using entangled states. In a fully device-independent model, one only assumes that the parties' devices do not directly broadcast certain information to the dishonest party and/or each other (we shall explain this in more detail in Section 1.4). In particular, one does not assume that the states and/or measurements implemented by the devices are known, and even the dimensions of the quantum systems are not specified. Device-independent security analyses exist for other cryptographic tasks such as quantum key distribution [17, 3], bit commitment [20, 2], and coin-flipping [1].

[1] To contrast, the task of finding protocols with perfect soundness and the best possible completeness was considered in [19].

While we have described the fully device-independent framework above, one can instead choose to trust some subset of the properties described, leading to various levels of semi-device-independent security. For instance, Alice and Bob could trust state preparation devices, measurement devices, quantum operations, or any combination of the above.

In this work, we examine the security of XOT quantum protocols in semi-device-independent and device-independent scenarios. By a slight abuse of notation, we use the same notation $\mathrm{P}_A^{\mathrm{XOT}}$ and $\mathrm{P}_B^{\mathrm{XOT}}$ to denote the cheating probabilities of Alice and Bob in all the different scenarios, corresponding to differently defined cheating capabilities of the dishonest party. For example, if we were to allow a dishonest Alice to control Bob's measurements, it may lead to a different value of $\mathrm{P}_A^{\mathrm{XOT}}$. The cheating capabilities of cheating parties should be clear when we discuss $\mathrm{P}_A^{\mathrm{XOT}}$ and $\mathrm{P}_B^{\mathrm{XOT}}$.

1.1 Trivial protocols

For readers new to oblivious transfer, we present two bad classical protocols and one bad quantum protocol.

▶ **Protocol 1** (Bad XOT Protocol 1)**.**

1. *Alice chooses* (x_0, x_1) *uniformly at random and sends* (x_0, x_1) *to Bob.*
2. *Bob chooses* b *uniformly at random.*
3. *Alice outputs* (x_0, x_1) *and Bob outputs* (b, x_b)*.*

A moment's thought shows that Bob has full information (he clearly learns (x_0, x_1)) while Alice has no information. Therefore, we have

$$\mathrm{P}_A^{\mathrm{XOT}} = 1/3 \quad \text{and} \quad \mathrm{P}_B^{\mathrm{XOT}} = 1 \tag{1}$$

which is as insecure concerning cheating Bob as possible.

▶ **Protocol 2** (Bad XOT Protocol 2)**.**

1. *Bob chooses* b *uniformly at random and tells Alice his choice.*
2. *Alice chooses and outputs* (x_0, x_1) *uniformly at random and sends* x_b *to Bob.*
3. *Alice outputs* (x_0, x_1) *and Bob outputs* (b, x_b)*.*

Here Alice has full information while Bob has none. Therefore, here we have

$$\mathrm{P}_A^{\mathrm{XOT}} = 1 \quad \text{and} \quad \mathrm{P}_B^{\mathrm{XOT}} = 1/2. \tag{2}$$

▶ Remark 3. Surprisingly, these protocols can be useful for protocol design. For instance, suppose Alice wishes to test Bob to see if he has been cheating, and aborts if and only if the test fails. Then if the test passes, the parties need a way to finish executing the protocol, which they could do using Protocol 1 (which is independent of previous steps in the tested protocol).

1.2 A quantum protocol for XOR oblivious transfer with no device-independent security

It is tricky creating a protocol in which Alice and Bob both cannot cheat with high probability. To this end, we often enlist the aid of quantum mechanics. Here we present the oblivious transfer protocol from [6] adapted to the XOT setting.

For $b \in \{0,1,2\}$, let $|\psi_b^{\pm}\rangle \in \mathcal{XY}$ denote the following two-qutrit state:

$$|\psi_b^{\pm}\rangle = \begin{cases} \frac{1}{\sqrt{2}}(|00\rangle_{\mathcal{XY}} \pm |22\rangle_{\mathcal{XY}}) & \text{if } b = 0, \\ \frac{1}{\sqrt{2}}(|11\rangle_{\mathcal{XY}} \pm |22\rangle_{\mathcal{XY}}) & \text{if } b = 1, \\ \frac{1}{\sqrt{2}}(|00\rangle_{\mathcal{XY}} \pm |11\rangle_{\mathcal{XY}}) & \text{if } b = 2. \end{cases} \tag{3}$$

Note that for every $b \in \{0,1,2\}$, we have that $|\psi_b^+\rangle$ and $|\psi_b^-\rangle$ are orthogonal. We are now ready to state the protocol.

▶ **Protocol 3.**

1. *Bob chooses $b \in \{0,1,2\}$ uniformly at random, prepares the state $|\psi_b^+\rangle$ in registers $\mathcal{XY}$, and sends the register $\mathcal{X}$ to Alice.*
2. *Alice chooses (x_0, x_1) uniformly at random, performs the unitary*

$$U_{(x_0,x_1)} = (-1)^{x_0}|0\rangle\langle 0| + (-1)^{x_1}|1\rangle\langle 1| + |2\rangle\langle 2| \tag{4}$$

   on $\mathcal{X}$, and then sends it back to Bob.
3. *Bob performs the 2-outcome measurement $\{|\psi_b^+\rangle\langle\psi_b^+|, \mathbb{1} - |\psi_b^+\rangle\langle\psi_b^+|\}$ and records his outcome as $c = 0$ if he gets $|\psi_b^+\rangle\langle\psi_b^+|$ and $c = 1$ otherwise.*
4. *Alice outputs (x_0, x_1), Bob outputs (b, c).*

Protocol 3 can be checked to be complete (i.e., Bob gets the correct outcome). The cheating probabilities in this protocol in the cases of trusted and untrusted devices are given by Theorem 4 below. We give a proof for the trusted case Section 3, and the relatively simple proof for the untrusted case is given here.

▶ **Theorem 4.** *In Protocol 3, the cheating probabilities are as listed in the following table. (In the untrusted setting, Alice controls Bob's state preparation and Bob controls Alice's unitary.)*

	$\mathrm{P}_A^{\mathrm{XOT}}$	$\mathrm{P}_B^{\mathrm{XOT}}$
Trusted devices	1/2	3/4
Untrusted devices	1	1

We show here that this protocol breaks down when they do not trust their own devices. For example, assume Bob has full control over Alice's unitary. It could be a unitary which implements a superdense coding protocol:

$$U_{(x_0,x_1)}^{cheat} = \begin{cases} \mathbb{1}_{\mathcal{X}} & \text{if } (x_0, x_1) = (0,0), \\ \sigma_X & \text{if } (x_0, x_1) = (0,1), \\ \sigma_Y & \text{if } (x_0, x_1) = (1,0), \\ \sigma_Z & \text{if } (x_0, x_1) = (1,1). \end{cases} \tag{5}$$

Note that this unitary acts on a qubit (which we can assume Bob sends if he wishes), or just define it such that it acts trivially on the $|2\rangle$ subspace. Now, if Bob creates $|\psi_2^+\rangle$, he is left with a state from the Bell basis. In other words, he can perfectly learn (x_0, x_1).

In the case of cheating Alice, she can simply control Bob's state preparation device to prepare the state $|b\rangle$ on input b, and have Bob send that. From this state, Alice simply measures it to learn b. Thus, we have

$$\mathrm{P}_A^{\mathrm{XOT}} = 1 \quad \text{and} \quad \mathrm{P}_B^{\mathrm{XOT}} = 1 \quad \text{[Devices are NOT trusted]}. \tag{6}$$

Thus, we need a clever way to design protocols where Alice and Bob cannot cheat in the above way. This motivates the need for device-independent XOT protocols and some protocol design ideas that should be avoided.

1.3 XOR oblivious transfer from the magic square game – A case of untrusted sources

Similar to device-independent protocols which exist for other cryptographic tasks, we design our protocols using nonlocal games. In this work, we shall make use of the nonlocal game known as the Mermin-Peres magic square game. In the magic square game,

- Alice and Bob receive respective inputs $a \in \{0,1,2\}$ and $b \in \{0,1,2\}$ independently and uniformly at random.
- Alice outputs $(x_0, x_1, x_2) \in \{0,1\}^3$ such that $x_0 \oplus x_1 \oplus x_2 = 0$ and Bob outputs $(y_0, y_1, y_2) \in \{0,1\}^3$ such that $y_0 \oplus y_1 \oplus y_2 = 1$.
- Alice and Bob win the game if $x_b = y_a$.

If Alice and Bob are allowed only classical strategies (using e.g. shared randomness), the magic square game cannot be won with probability greater than 8/9. However, there exists a quantum strategy where Alice and Bob share prior entanglement, with which the magic square game can be won with probability 1. We shall refer to this strategy as the *magic square strategy* which is detailed in Section 2.1.

Now, suppose Alice and Bob play the magic square game according to the above description. Notice that x_2 will always be equal to $x_0 \oplus x_1$, similar to the definition of XOT, and that (x_0, x_1) is uniformly distributed (see Section 2.1). Also, for each of Bob's input choices, he learns either x_0, x_1, or $x_2 = x_0 \oplus x_1$ depending on the choice of input a for Alice. Since a is chosen uniformly at random, this is almost a proper XOT protocol (putting aside soundness for now). The only missing ingredient is that Bob knows he has x_b, but does not know which of the bits of (y_0, y_1, y_2) it is. To fix this small issue, we simply have Alice tell Bob which bit it is.

We formalize this protocol below and add in a *test* step that helps to prevent cheating. Strictly speaking, Protocol 4 should be thought of as a protocol framework, as we have not specified who creates the entangled state that Alice and Bob share – either party can. We consider different security analyses of Protocol 4, corresponding to each of these different cases. In the trusted state analysis, the honest party (whomever that may be) creates the state, and in the untrusted state analysis it is the cheating party who does so.

▶ **Protocol 4** (XOR oblivious transfer from the magic square game)**.**

1. *Alice and Bob share the bipartite state used in the magic square strategy.*
2. *Bob chooses $b \in \{0,1,2\}$ uniformly at random, performs the measurements corresponding to b in the magic square strategy to his state to get the outcome (y_0, y_1, y_2), and sends (y_0, y_1, y_2) to Alice.*
3. *Alice chooses $a \in \{0,1,2\}$ uniformly at random and sends a to Bob.*
4. *Alice performs the measurement corresponding to a in her magic square strategy on her state to get outcome (x_0, x_1, x_2).*
5. ***Test:** If $(x_0, x_1) = (0,0)$ and Bob has sent (y_0, y_1) such that $y_a = 1$, then Alice aborts.*
6. *Alice outputs (x_0, x_1) and Bob outputs (b, y_a).*

Intuitively, the test step in Protocol 4 serves as a weak test that the magic square winning condition is fulfilled (though the test only occurs with somewhat small probability). This provides a way to partially certify that Bob has measured his share of the state before

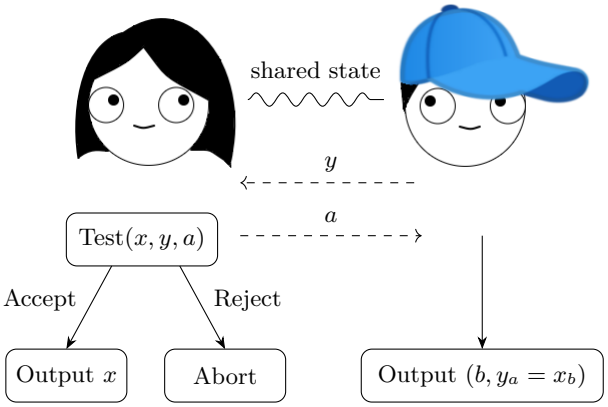

Figure 2 Schematic depiction of the messages sent in Protocol 4.

learning Alice's input. (Note that if Bob has not measured his state yet, then even when the states are trusted, he has the potential to perfectly learn Alice's output after learning her input, by simply performing the same measurement as Alice on his state. However, forcing him to perform the magic square measurement "deletes" some of this information; a notion referred to as *certified deletion* in [8].)

It can be checked that Protocol 4 accomplishes what XOT wants and that it never aborts in the honest case; in other words, the protocol is complete. To prove its soundness, we bound the cheating probabilities using appropriate SDPs, as described in the full version. Our numerical results are shown as Theorem 5 below.

▶ **Theorem 5.** *We assume Alice and Bob play the "canonical" strategy for the magic square game (see Section 2.1). Then the cheating probabilities for Alice and Bob in Protocol 4 are bounded by the values in the table below, rounded to 5 decimal places. The bounds for cheating Alice are tight.*

Upper bounds	P_A^{XOT}	P_B^{XOT}
Trusted state	0.83333	0.93628
Untrusted state	0.87268	0.94096
Untrusted measurement	1	1

As the last row of the table in Theorem 5 indicates, Protocol 4 is not fully device-independent. To see this, note that if Bob were to control Alice's measurements, he can force $(x_0, x_1) = (0, 1)$ to always occur (regardless of the state, by performing a trivial measurement), and then he will never be tested. Moreover, he will learn (x_0, x_1) perfectly. Conversely, if Alice controls Bob's measurement, she can force the output to be such that $y_0 + y_1 = b$ (note the sum is not modulo 2). Then, Bob's message will reveal b to Alice.

1.4 A device-independent protocol

We now aim to find an XOT protocol based on the magic square game that is fully device-independent. We shall first clarify the premises and assumptions in such a setting. Specifically, we shall suppose that Alice and Bob each possess one of a pair of black boxes, each of which will accept a classical input in $\{0, 1, 2\}$ and return a classical output in $\{0, 1\}^3$. We shall

only require a single use of these boxes. In the honest scenario, the boxes will simply be implementing the ideal magic square states and measurements. If either party is dishonest, however, we shall suppose only that the boxes' behaviour can be modelled as follows: the boxes share some entangled state between them, and when one of the boxes receives an input, it returns the output of some measurement (conditioned on the input) performed on its share of the quantum state, *without broadcasting either its input or output* to any party other than the one holding the box[2]. While the honest party can only interact with the box as specified, the dishonest party is able to "open" any box they possess and perform arbitrary quantum operations or measurements on the share of the state held by that box. However, the dishonest party cannot interact with or change the behaviour of a box while it is in the honest party's possession.

This describes the general device-independent setting. For the purposes of this work, we shall impose a small additional assumption on the states and measurements the boxes implement, namely that they are described by a tensor product of Hilbert spaces, one for each box. More general scenarios could be considered (for instance, one could require only that the two boxes' measurements commute), but these are outside the scope of this work.

It would seem difficult to design a secure protocol under such weak assumptions. However, we can exploit the fact that many nonlocal games (including the magic square game) exhibit the important property of *self-testing* or *rigidity*: if the boxes win the game with probability equal to the maximum quantum value, then they must be implementing the ideal state and measurements (up to trivial isometries). A robust version of this statement is formally expressed as Lemma 9 in the next section.

This suggests the following idea to make Protocol 4 fully device-independent: we introduce an initial step where with some probability, either party may ask the other to send over their box, so they can perform a single-shot test of whether the boxes win the magic square game. To prevent a dishonest party from always calling for a test, we shall enforce that a party calling for a test must then cede all control if the test is passed, performing a XOT protocol that is perfectly secure against them. If a test is not called, the parties simply perform Protocol 4. We describe this idea more formally as Protocol 5 below.

Qualitatively, Protocol 5 imposes a "tradeoff" for the cheating party between passing the test (if it is called) and the extent to which they deviate from the ideal implementation of Protocol 4. More specifically, if (say) Bob is dishonest, he could cheat perfectly if Alice decides to test, by having both boxes implement the honest behaviour, and he could also cheat perfectly if Alice decides not to test, but to do so he needs to modify Alice's box's behaviour at least (since our device-dependent arguments show that perfect cheating is impossible when Alice's box is honest). Since Alice's box must behaves differently in the two scenarios and Bob cannot change how that box behaves once the protocol starts, Alice can constrain his cheating probability by randomly choosing between testing and not testing. A similar argument applies to cheating Alice.

Note that for this reasoning to be valid, it is important that the honest party's box cannot be allowed to detect whether it is being subjected to a magic square test or whether it is being used for Protocol 4 (we are implicitly assuming that the honest party's box behaves the same way in both situations). An assumption of this nature is typically required in device-independent protocols that involve performing a test with some probability, e.g. [1, 3].

[2] To take a slightly different perspective (used in e.g. [3]), we could suppose that the honest party is able to "shield" their lab in a way such that signals cannot be broadcast out of it once they have supplied the input to their box.

In particular, as observed in [1], we note that if the behaviour of the boxes could be time-dependent, then the honest party must ensure they provide the input to their box at a fixed pre-determined time, regardless of whether the box is being tested or used for Protocol 4.

▶ **Protocol 5** (XOR oblivious transfer from the magic square game with an extra test step).

1. *Alice flips a coin whose outcome is 0 with probability* $1-q^A$*, to obtain* $c^A \in \{0,1\}$*, which she sends to Bob.*
2. **a.** *If* $c^A = 0$*, Bob flips a coin whose outcome is 0 with probability* $1-q^B$*, to obtain* $c^B \in \{0,1\}$*, which he sends to Alice.*
   b. *If* $c^A = 1$*, Bob sends his box to Alice.*
3. **a.** *If* $c^A = 0, c^B = 0$*, Alice and Bob perform Protocol 4 henceforth.*
   b. *If* $c^A = 0, c^B = 1$*, Alice sends her box to Bob.*
   c. *If* $c^A = 1$*, Alice receives Bob's box, picks* $a^A, b^A \in \{0,1,2\}$ *uniformly at random to input into her and Bob's boxes, and checks if the outputs* x^A, y^A *satisfy* $x^A_{b^A} = y^A_{a^A}$*. If not, she aborts.*
4. *If* $c^A = 0, c^B = 1$*, Bob receives Alice's box, picks* $a^B, b^B \in \{0,1,2\}$ *uniformly at random to input into his and Alice's boxes, and checks if the outputs* x^B, y^B *satisfy* $x^B_{b^B} = y^B_{a^B}$*. If not, he aborts.*
5. **a.** *If* $c^A = 1$ *and Alice has not aborted, Alice and Bob perform Protocol 1 henceforth.*
   b. *If* $c^A = 0, c^B = 1$ *and Bob has not aborted, Alice and Bob perform Protocol 2 henceforth.*

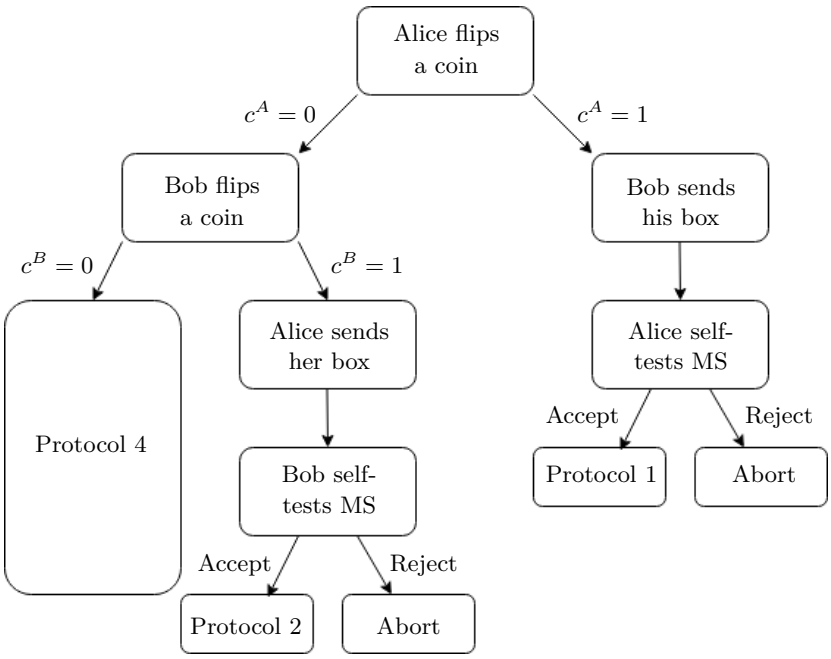

Figure 3 Flowchart for Protocol 5.

We have required that when either party calls for a test, the tested party must send over their box so that the testing party supplies an input to both boxes themselves, rather than having the tested party self-report an input-output pair for their box. This is to ensure that the inputs to the boxes are indeed uniformly chosen. Also, while it would be convenient if Protocol 4 did not involve Alice sending her input to Bob (thereby more closely resembling a standard nonlocal game), it would seem this step is necessary to allow an honest Bob to know which bit of his output he should use, as previously mentioned regarding Protocol 4.

We give two soundness arguments for Protocol 5. The first consists of explicit numerical bounds on the cheating probabilities, based on the family of SDPs known as the Navascués-Pironio-Acín (NPA) hierarchy [16]. We state the results as Theorem 6 below, and give the proof in the full version. The second is an analytic proof that the cheating probabilities are bounded away from 1, based on the robust self-testing bounds for the magic square game [24, 7]. We state this result formally as Theorem 8 below, and give the proof in the full version.

▶ **Theorem 6.** *Upper bounds on Alice and Bob's cheating probabilities for Protocol 5 (the fully device-independent scenario) with $q^A = 0.6$, $q^B = 0.6$ are given below, rounded to 5 decimal places.*

Upper bounds	$\mathrm{P}_A^{\mathrm{XOT}}$	$\mathrm{P}_B^{\mathrm{XOT}}$
Fully Device-Independent	0.96440	0.99204

▶ Remark 7. The choices for q^A and q^B in Theorem 6 were made by computing the bounds for different choices of q^A and q^B in intervals of 0.1, then simply taking the value that yields the best bounds on the cheating probabilities. We note that the result obtained for $\mathrm{P}_B^{\mathrm{XOT}}$ is rather close to 1; however, the significant figures shown here are within the tolerance levels of the solver.

▶ **Theorem 8.** *For any $q^A, q^B > 0$ in Protocol 5, there exists some $\delta > 0$ such that $P_A^{\mathrm{XOT}}, P_B^{\mathrm{XOT}} \leq 1 - \delta$.*

2 Background

In this section, we give the necessary background to prove the results mentioned in the introduction.

2.1 The magic square game

The optimal quantum strategy for magic square can be described as follows. Alice and Bob share the state

$$|\Psi^{\mathrm{MS}}\rangle = \frac{1}{\sqrt{2}}(|00\rangle_{\mathcal{X}_0\mathcal{Y}_0} + |11\rangle_{\mathcal{X}_0\mathcal{Y}_0})\frac{1}{\sqrt{2}}(|00\rangle_{\mathcal{X}_1\mathcal{Y}_1} + |11\rangle_{\mathcal{X}_1\mathcal{Y}_1}) \tag{7}$$

with Alice holding the registers $\mathcal{X}_0\mathcal{X}_1$, and Bob holding the registers $\mathcal{Y}_0\mathcal{Y}_1$. The measurements of Alice and Bob are given by the following table.

On input a, Alice simultaneously performs the three 2-outcome measurements $\{(\Pi^0_{ab},\Pi^1_{ab})\}_b$ in the row indexed by a in Table 1 (it can be checked that the three measurements in every row are compatible, so they can be performed simultaneously) on her registers $\mathcal{X}_0\mathcal{X}_1$. Her output (x_0, x_1, x_2) is the outputs of the three measurements (in order). Similarly, on input b, Bob simultaneously performs the three 2-outcome measurements $\{(\Pi^0_{ab}, \Pi^1_{ab})\}_a$ in the column indexed by b (the three measurements in every column are also compatible, so they can be performed simultaneously) on his registers $\mathcal{Y}_0\mathcal{Y}_1$, and gives the outcomes of the three measurements as his output (y_0, y_1, y_2).

Clearly, the measurement Alice performs to output x_b is the same as the measurement Bob performs to output y_a. Since these measurements are performed on maximally entangled states, one can show Alice and Bob always get the same outcome for x_b and y_a. Also, it can

Table 1 Possible measurements for either party in the quantum strategy for magic square.

a \ b	0	1	2
0	$\Pi^0_{00} = \|0\rangle\langle 0\| \otimes \mathbb{1}$ $\Pi^1_{00} = \|1\rangle\langle 1\| \otimes \mathbb{1}$	$\Pi^0_{01} = \mathbb{1} \otimes \|0\rangle\langle 0\|$ $\Pi^1_{01} = \mathbb{1} \otimes \|1\rangle\langle 1\|$	$\Pi^0_{02} = \|0\rangle\langle 0\| \otimes \|0\rangle\langle 0\| + \|1\rangle\langle 1\| \otimes \|1\rangle\langle 1\|$ $\Pi^1_{02} = \|0\rangle\langle 0\| \otimes \|1\rangle\langle 1\| + \|1\rangle\langle 1\| \otimes \|0\rangle\langle 0\|$
1	$\Pi^0_{10} = \mathbb{1} \otimes \|+\rangle\langle +\|$ $\Pi^1_{10} = \mathbb{1} \otimes \|-\rangle\langle -\|$	$\Pi^0_{11} = \|+\rangle\langle +\| \otimes \mathbb{1}$ $\Pi^1_{11} = \|-\rangle\langle -\| \otimes \mathbb{1}$	$\Pi^0_{12} = \|+\rangle\langle +\| \otimes \|+\rangle\langle +\| + \|-\rangle\langle -\| \otimes \|-\rangle\langle -\|$ $\Pi^1_{12} = \|+\rangle\langle +\| \otimes \|-\rangle\langle -\| + \|-\rangle\langle -\| \otimes \|+\rangle\langle +\|$
2	$\Pi^0_{20} = \|1\rangle\langle 1\| \otimes \|+\rangle\langle +\| + \|0\rangle\langle 0\| \otimes \|-\rangle\langle -\|$ $\Pi^1_{20} = \|0\rangle\langle 0\| \otimes \|+\rangle\langle +\| + \|1\rangle\langle 1\| \otimes \|-\rangle\langle -\|$	$\Pi^0_{21} = \|+\rangle\langle +\| \otimes \|1\rangle\langle 1\| + \|-\rangle\langle -\| \otimes \|0\rangle\langle 0\|$ $\Pi^1_{21} = \|+\rangle\langle +\| \otimes \|0\rangle\langle 0\| + \|-\rangle\langle -\| \otimes \|1\rangle\langle 1\|$	$\Pi^0_{22} = \|+i\rangle\langle +i\| \otimes \|+i\rangle\langle +i\| + \|-i\rangle\langle -i\| \otimes \|-i\rangle\langle -i\|$ $\Pi^1_{22} = \|+i\rangle\langle +i\| \otimes \|-i\rangle\langle -i\| + \|-i\rangle\langle -i\| \otimes \|+i\rangle\langle +i\|$

be verified that these measurements always produce outcomes satisfying the parity conditions $x_0 \oplus x_1 \oplus x_2 = 0$ and $y_0 \oplus y_1 \oplus y_2 = 1$ (this holds regardless of the state). Some tedious calculation shows that in fact the output distribution is uniform over all combinations that win the magic square game, i.e. $\Pr(xy|ab) = 1/8$ if $x_b = y_a$ (and x, y satisfy the parity conditions), and $\Pr(xy|ab) = 0$ otherwise.

The above description views Alice and Bob as performing 8-outcome measurements (via a sequence of three 2-outcome measurements). However, since for *any* state the measurements always produce outputs satisfying the parity conditions, we can equivalently suppose Alice and Bob measure to determine only (x_0, x_1) and (y_0, y_1), with the last bit for each determined by the parity conditions. (This is consistent with the way we defined the magic square game earlier.) These are 4-outcome measurements that can be expressed in terms of the Π operators from Table 1 as

$$M^{\mathrm{MS}}_{x_0x_1|a} = \Pi^{x_0}_{a0}\Pi^{x_1}_{a1} \qquad N^{\mathrm{MS}}_{y_0y_1|b} = \Pi^{y_0}_{0b}\Pi^{y_1}_{1b}. \tag{8}$$

It can be checked that each $M^{\mathrm{MS}}_{x_0x_1|a}$ and $N^{\mathrm{MS}}_{y_0y_1|b}$ is a rank-1 projector. Since the measurements for x_0 and x_1 (resp. y_0 and y_1) commute for every a (resp. b), the product of the Π operators in each case can be taken in either order.

Certain nonlocal games exhibit the property that the quantum strategies achieving their optimal values are essentially unique. That is, if a quantum strategy achieves within ε of the optimal value of the game, that strategy must be $\delta(\varepsilon)$-close to the ideal strategy for the game, up to certain local operations. This property of rigidity or self-testing was shown first for the CHSH game [13, 14] and has been shown for other nonlocal games since.

[24] originally gave a proof of the rigidity of a version of the magic square game which is slightly different from ours. [7] showed that the rigidity statement also holds for the version of the magic square game we use. However, both of these results show the self-testing of some operators that are related to Alice and Bob's measurement operators in the magic square game, but not the measurement operators themselves. It is not immediately clear

how to self-test the measurement operators themselves from their results. In the full version we derive the following lemma for self-testing of the measurement operators of the magic square strategy.

▶ **Lemma 9.** *Consider any state $|\rho\rangle$ on registers $\mathcal{X}\mathcal{Y}$ and projective measurements $M_{x|a}, N_{y|b}$ such that $M_{x|a}$ act only on $\mathcal{X}$ and $N_{y|b}$ act only on $\mathcal{Y}$. If this state and measurements win the magic square game with probability $1-\varepsilon$, then there exist local isometries $V_A : \mathcal{X} \to \mathcal{X}_0\mathcal{X}_1\mathcal{J}_A$ and $V_B : \mathcal{Y} \to \mathcal{Y}_0\mathcal{Y}_1\mathcal{J}_B$ and a state $|\text{junk}\rangle$ on $\mathcal{J}_A\mathcal{J}_B$ such that for all a, b, x, y,*

$$\begin{aligned}
&\left\|(V_A \otimes V_B)|\rho\rangle - |\Psi^{\text{MS}}\rangle \otimes |\text{junk}\rangle\right\|_2 \leq O(\varepsilon^{1/4}),\\
&\left\|(V_A \otimes V_B)(M_{x|a} \otimes \mathbb{1})|\rho\rangle - ((M^{\text{MS}}_{x|a} \otimes \mathbb{1})|\Psi^{\text{MS}}\rangle) \otimes |\text{junk}\rangle\right\|_2 \leq O(\varepsilon^{1/4}),\\
&\left\|(V_A \otimes V_B)(\mathbb{1} \otimes N_{y|b})|\rho\rangle - ((\mathbb{1} \otimes N^{\text{MS}}_{y|b})|\Psi^{\text{MS}}\rangle) \otimes |\text{junk}\rangle\right\|_2 \leq O(\varepsilon^{1/4}),
\end{aligned}$$

where $|\Psi^{\text{MS}}\rangle$, $M^{\text{MS}}_{x|a}$, $N^{\text{MS}}_{y|b}$ denote the ideal state and measurements in the magic square game.

2.2 Semidefinite programming

A semidefinite program is an optimization problem of the form

$$p^* = \sup\{\langle C, X\rangle : \Phi(X) = B,\, X \succeq 0\} \tag{9}$$

where Φ is a linear transformation, C and B are Hermitian. When we write $X \succeq Y$, it means that $X - Y$ is (Hermitian) positive semidefinite, noting the special case that $X \succeq 0$ simply means X is positive semidefinite. We use $X \succ Y$ to mean that $X - Y$ is positive definite.

We can define the *dual* of the above SDP as the optimization problem below

$$d^* = \inf\{\langle B, Y\rangle : \Phi^*(Y) = C + S,\, S \succeq 0,\, Y \text{ is Hermitian}\} \tag{10}$$

where use the notation Φ^* to mean the adjoint of the linear operator Φ.

When we deal with an SDP and its dual, we refer to the original SDP as the *primal* SDP. The primal is called *feasible* if the constraints are satisfiable, that is, if

$$\Phi(X) = B \quad \text{and} \quad X \succeq 0 \tag{11}$$

has a solution. Similarly, if

$$\Phi^*(Y) = C + S, \quad S \succeq 0, \quad \text{and} \quad Y \text{ is Hermitian} \tag{12}$$

has a solution, then the dual is said to be feasible.

We can use primal and dual solutions to show *weak duality*, i.e. if X is primal feasible and (Y, S) is dual feasible, then

$$\langle C, X\rangle \leq \langle B, Y\rangle. \tag{13}$$

In particular, $p^* \leq d^*$. Under mild conditions, we can guarantee equality in Eq. (13). For example, if there exists $X \succ 0$ which is primal feasible and (Y, S) which is dual feasible, then one can show that $p^* = d^*$. Alternatively, if there exists (Y, S) with $S \succ 0$ which is dual feasible and X which is primal feasible, then we also have $p^* = d^*$. Either of these conditions is known as *strong duality* and the feasible solution with the positive definite solution is known as a *Slater point*. We refer the reader to the book [4] for a proof of weak and strong duality and for other useful information on the subject.

3 A sample proof

In this section, we prove that when the devices are trusted in Protocol 3, then $\mathrm{P}_A^{\mathrm{XOT}} = 1/2$ and $\mathrm{P}_B^{\mathrm{XOT}} = 3/4$. Before continuing, recall that Protocol 3 is an adaptation of the protocol in [6] where Alice's actions are the exact same and so are the intentions of a dishonest Bob. Therefore, we can also import $\mathrm{P}_B^{\mathrm{XOT}} = 3/4$ directly from the security analysis of that protocol.

The rest of this section is devoted to proving $\mathrm{P}_A^{\mathrm{XOT}} = 1/2$. Since Bob never aborts, all we need to ascertain is the ability for Alice to learn b from her information contained in the first message. To do this, she must infer b from the ensemble

$$\left\{ \left(\frac{1}{3}, \rho_b = \mathrm{Tr}_{\mathcal{Y}}(|\psi_b^+\rangle\langle\psi_b^+|) \right) : b \in \{0,1,2\} \right\}. \tag{14}$$

This is known as the *quantum state discrimination problem*, and the optimal guessing probability can be written as the following SDP:

Primal problem

$$\begin{aligned} \sup \quad & \frac{1}{3}\sum_{b=1}^{3} \mathrm{Tr}(E_b\rho_b) \\ \text{subject to:} \quad & \sum_{b=0,1,2} E_b = \mathbb{1} \\ & \forall b\, E_b \in \mathrm{Pos}(\mathcal{X}) \end{aligned}$$

Dual problem

$$\begin{aligned} \inf \quad & \mathrm{Tr}(\sigma) \\ \text{subject to:} \quad & \forall b\ \frac{1}{3}\rho_b \preceq \sigma \\ & \sigma \in \mathrm{Herm}(\mathcal{X}). \end{aligned}$$

Note that the success probability can easily be seen to be equal to the value of the primal problem as it is a maximization over POVMs and the objective function is the success probability of that POVM measurement.

Now, if Alice uses the POVM

$$\{E_0, E_1, E_2\} = \{|0\rangle\langle 0| + |2\rangle\langle 2|, |1\rangle\langle 1|, 0\}, \tag{15}$$

we can see that

$$\mathrm{P}_A^{\mathrm{XOT}} \geq \frac{1}{3}\sum_{b=1}^{3}\langle E_b, \rho_b\rangle = \frac{1}{3} + \frac{1}{3}\cdot\frac{1}{2} + 0 = \frac{1}{2}. \tag{16}$$

Effectively what this measurement does is measure in the computational basis, then assign the outcomes $|0\rangle$ and $|2\rangle$ to the guess $b = 0$ and the outcome $|1\rangle$ to the guess $b = 1$. Note that $b = 2$ is never guessed in this strategy. All that remains to show is that $\mathrm{P}_A^{\mathrm{XOT}} \leq 1/2$. For this, we use the dual problem. Consider the dual feasible solution

$$\sigma = \frac{1}{6}\mathbb{1}_{\mathcal{X}}. \tag{17}$$

It can be checked that σ satisfies the dual constraints, i.e., $\sigma \succeq \frac{1}{3}\rho_b$ for all $b \in \{0,1,2\}$. Since $\mathrm{Tr}(\sigma) = 1/2$, we have that $\mathrm{P}_A^{\mathrm{XOT}} \leq 1/2$ by weak duality.

Computational platform

Computations were performed using the MATLAB packages QETLAB [9] and YALMIP [12] with solver MOSEK [15]. Some of the calculations reported here were performed using the Euler cluster at ETH Zürich.

References

1 Nati Aharon, André Chailloux, Iordanis Kerenidis, Serge Massar, Stefano Pironio, and Jonathan Silman. Weak coin flipping in a device-independent setting. In *Theory of Quantum Computation, Communication, and Cryptography*, pages 1–12, Berlin, Heidelberg, 2014. doi:10.1007/978-3-642-54429-3_1.

2 Nati Aharon, Serge Massar, Stefano Pironio, and Jonathan Silman. Device-independent bit commitment based on the CHSH inequality. *New Journal of Physics*, 18(2):025014, 2016. doi:10.1088/1367-2630/18/2/025014.

3 Rotem Arnon-Friedman, Frédéric Dupuis, Omar Fawzi, Renato Renner, and Thomas Vidick. Practical device-independent quantum cryptography via entropy accumulation. *Nature Communications*, 9(1):459, 2018. doi:10.1038/s41467-017-02307-4.

4 Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, New York, NY, USA, 2004.

5 André Chailloux, Gus Gutoski, and Jamie Sikora. Optimal bounds for semi-honest quantum oblivious transfer. *Chicago Journal of Theoretical Computer Science*, 2016(13), 2016. URL: http://cjtcs.cs.uchicago.edu/articles/2016/13/contents.html.

6 André Chailloux, Iordanis Kerenidis, and Jamie Sikora. Lower bounds for quantum oblivious transfer. *Quantum Information & Computation*, 13(1-2):158–177, 2013. URL: http://dl.acm.org/citation.cfm?id=2481591.2481600.

7 Matthew Coudron and Anand Natarajan. The parallel-repeated Magic Square game is rigid, 2016. arXiv:1609.06306.

8 Honghao Fu and Carl A. Miller. Local randomness: Examples and application. *Physical Review A*, 97:032324, 2018. doi:10.1103/PhysRevA.97.032324.

9 Nathaniel Johnston. QETLAB: A MATLAB toolbox for quantum entanglement, version 0.9. http://qetlab.com, 2016. doi:10.5281/zenodo.44637.

10 Joe Kilian. Founding cryptography on oblivious transfer. In *Proceedings of the Twentieth Annual ACM Symposium on Theory of Computing*, STOC '88, pages 20—-31, New York, NY, USA, 1988. doi:10.1145/62212.62215.

11 Hoi-Kwong Lo. Insecurity of quantum secure computations. *Physical Review A*, 56:1154–1162, 1997. doi:10.1103/PhysRevA.56.1154.

12 Johan Löfberg. YALMIP : A toolbox for modeling and optimization in MATLAB. In *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004.

13 Dominic Mayers and Andrew Yao. Quantum cryptography with imperfect apparatus. In *Proceedings of the 39th Annual Symposium on Foundations of Computer Science*, FOCS '98, page 503, USA, 1998. URL: https://dl.acm.org/doi/10.5555/795664.796390.

14 Matthew McKague, Tzyh Haur Yang, and Valerio Scarani. Robust self-testing of the singlet. *Journal of Physics A: Mathematical and Theoretical*, 45(45):455304, 2012. doi:10.1088/1751-8113/45/45/455304.

15 MOSEK ApS. *The MOSEK optimization toolbox for MATLAB manual. Version 8.1.*, 2019.

16 Miguel Navascués, Stefano Pironio, and Antonio Acín. A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations. *New Journal of Physics*, 10(7):073013, 2008. URL: http://stacks.iop.org/1367-2630/10/i=7/a=073013.

17 Stefano Pironio, Antonio Acín, Nicolas Brunner, Nicolas Gisin, Serge Massar, and Valerio Scarani. Device-independent quantum key distribution secure against collective attacks. *New Journal of Physics*, 11(4):045021, 2009. doi:10.1088/1367-2630/11/4/045021.

18 Christian Schaffner. Cryptography in the bounded-quantum-storage model, 2007. arXiv:0709.0289.

19 Jamie Sikora, André Chailloux, and Iordanis Kerenidis. Strong connections between quantum encodings, nonlocality, and quantum cryptography. *Physical Review A*, 89:022334, 2014. doi:10.1103/PhysRevA.89.022334.

20 Jonathan Silman, André Chailloux, Nati Aharon, Iordanis Kerenidis, Stefano Pironio, and Serge Massar. Fully distrustful quantum bit commitment and coin flipping. *Physical Review Letters*, 106:220501, 2011. doi:10.1103/PhysRevLett.106.220501.

21 Vilasini Venkatesh, Christopher Portmann, and Lídia del Rio. Composable security in relativistic quantum cryptography. *New Journal of Physics*, 21(4):043057, 2019. doi:10.1088/1367-2630/ab0e3b.

22 Stephanie Wehner, Christian Schaffner, and Barbara M. Terhal. Cryptography from noisy storage. *Phys. Rev. Lett.*, 100:220502, 2008. doi:10.1103/PhysRevLett.100.220502.

23 Stephanie Wehner and Jürg Wullschleger. Composable security in the bounded-quantum-storage model. In *Automata, Languages and Programming*, pages 604–615, 2008. doi:10.1007/978-3-540-70583-3_49.

24 Xingyao Wu, Jean-Daniel Bancal, Matthew McKague, and Valerio Scarani. Device-independent parallel self-testing of two singlets. *Physical Review A*, 93:062121, 2016. doi:10.1103/PhysRevA.93.062121.

Note of the Publisher

Unfortunately, this article was accidentally skipped in the first version of the conference proceedings published on June 8, 2020 and was subsequently published on August 19, 2020.

Dagstuhl Publishing – August 19, 2020.