UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Complete insecurity of quantum protocols for classical two-party computation

Buhrman, H.; Christandl, M.; Schaffner, C.

DOI
10.1103/PhysRevLett.109.160501

Publication date
2012

Document Version
Final published version

Published in
Physical Review Letters

Link to publication

Citation for published version (APA):
Buhrman, H., Christandl, M., & Schaffner, C. (2012). Complete insecurity of quantum protocols for classical two-party computation. *Physical Review Letters*, *109*(16), 160501. https://doi.org/10.1103/PhysRevLett.109.160501

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:26 Jan 2022

Complete Insecurity of Quantum Protocols for Classical Two-Party Computation

Harry Buhrman,[1] Matthias Christandl,[2] and Christian Schaffner[1]

[1]*University of Amsterdam and CWI Amsterdam, Amsterdam, The Netherlands*

[2]*Institute for Theoretical Physics, ETH Zurich, Wolfgang-Pauli-Strasse 27, CH-8093 Zurich, Switzerland*

(Received 4 January 2012; published 17 October 2012)

A fundamental task in modern cryptography is the joint computation of a function which has two inputs, one from Alice and one from Bob, such that neither of the two can learn more about the other's input than what is implied by the value of the function. In this Letter, we show that any quantum protocol for the computation of a classical deterministic function that outputs the result to both parties (two-sided computation) and that is secure against a cheating Bob can be completely broken by a cheating Alice. Whereas it is known that quantum protocols for this task cannot be completely secure, our result implies that security for one party implies complete insecurity for the other. Our findings stand in stark contrast to recent protocols for weak coin tossing and highlight the limits of cryptography within quantum mechanics. We remark that our conclusions remain valid, even if security is only required to be approximate and if the function that is computed for Bob is different from that of Alice.

DOI: 10.1103/PhysRevLett.109.160501 PACS numbers: 03.67.Dd, 03.67.Hk

Traditionally, cryptography has been understood as the art of "secret writing," i.e., of sending messages securely from one party to another. Today, the research field of cryptography comprises much more than encryption and studies all aspects of secure communication and computation among players that do not trust each other, including tasks such as electronic voting and auctioning. Following the excitement that the exchange of quantum particles may allow for the distribution of a key that is unconditionally secure [1,2], a level of security unattainable by classical means, the question arose whether other fundamental cryptographic tasks could be implemented with the same level of security using quantum mechanical effects. For oblivious transfer and bit commitment, it was shown that the answer is negative [3,4]. Interestingly, however, a weak version of a coin toss can be implemented by quantum mechanical means [5].

In this Letter, we study the task of secure two-party computation. Here, two mistrustful players, Alice and Bob, wish to compute the value of a classical deterministic function f, which takes an input u from Alice and v from Bob, in such a way that both learn the result of the computation and that none of the parties can learn more about the other's input, even by deviating from the protocol. As our main result, we show that any quantum protocol which is secure against a cheating Bob can be completely broken by a cheating Alice. Formally, we design an attack by Alice which allows her to compute the value of the function f for all of her inputs (rather than only a single one, which would be required from a secure protocol).

Our result strengthens the impossibility result for two-sided secure two-party computation by Colbeck, where he showed that Alice can always obtain more information about Bob's input than what is implied by the value of the function [6]. In a similar way, we complement a result by Salvail *et al.* [7] showing that any quantum protocol for a nontrivial primitive necessarily leaks information to a dishonest player. Our result is motivated by Lo's impossibility result for the case where only Alice obtains the result of the function (one-sided computation) [8]. Lo's approach is based on the idea that Bob does not have any output; hence, his quantum state cannot depend on Alice's input. Then, Bob has learned nothing about Alice's input, and a cheating Alice can therefore still change her input value (by purifying the protocol) and thus cheat.

In the two-sided case, this approach to proving the insecurity of two-party computation fails as Bob knows the value of the function and has thus some information about Alice's input. In order to overcome this problem, we develop a new approach. We start with a formal definition of security based on the standard real–ideal-world paradigm from modern cryptography. In our case of a classical functionality, this definition guarantees the existence of a classical input for Bob in the ideal world, even if he is, in the real world, dishonestly purifying his steps of the protocol. Since real and ideal are indistinguishable for a secure protocol and since a purification of the classical input cannot be part of Bob's systems, Alice can now obtain a copy of this input by applying a unitary—constructed with help of Uhlmann's theorem—to her output registers and, henceforth, break the protocol.

We wish to emphasize that the above conclusion remains valid if the protocol is not required to be perfectly secure

Published by the American Physical Society under the terms of the Creative Commons Attribution 3.0 License. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

(nor perfectly correct). More precisely, if the protocol is secure up to a small error against cheating Bob, then Alice is able to compute the value of the function for all of her inputs with only a small error. Since the error is independent of the number of inputs that both Alice and Bob have, our analysis improves over Lo's result in the one-sided case. In fact, our results apply to this case since, more generally, they remain true should Bob receive the output of a function g, different from Alice's f, as a careful look at our argument reveals.

Security definition.— Alice and Bob, at distant locations and only connected with a quantum channel, wish to execute a protocol that takes an input u from Alice and an input v from Bob and that outputs the value $f(u, v)$ of a classical deterministic function f to both of them. Since Alice does not trust Bob, she wants to be sure that the protocol does not allow him to extract more information about her input than what is implied by the output value of the function. The same should be true if Alice is cheating and Bob is honest. Whereas for simple functions this intuitive notion of security can be made precise by stating a list of security requirements for certain quantum states of Alice and Bob, such an approach seems very complicated and prone to pitfalls for general functions f, in particular, if we want to consider protocols that are only approximately secure. We therefore follow the modern literature on cryptography where such situations have been in the center of attention for many years (cf., zero knowledge, composability) and where a suitable notion of security, known as the real–ideal-world paradigm, has been firmly established.

In this paradigm, we first define an ideal situation in which everything is computed perfectly and securely and call this the ideal functionality. Informally, a two-party protocol is secure if it looks to the outside world just like the ideal functionality it is supposed to implement. More concretely, a protocol is deemed secure if for every adversarial strategy, or real adversary, there exists an ideal adversary interacting only with the ideal functionality such that the execution of the protocol in the real world is indistinguishable from this ideal world. If such a security guarantee holds, it is clear that a secure protocol can be treated as a call to the ideal functionality, and hence, it is possible to construct and prove secure more complicated protocols in a modular fashion. See Refs. [9–15] for further information about this concept of security in the context of classical and quantum protocols, respectively.

There exist different meaningful ways to make the above informal notion of the real–ideal-world paradigm precise. All these notions have in common that the execution of the protocol by the honest and dishonest players is modeled by a completely positive trace-preserving (CPTP) map. Likewise, every ideal adversary interacting with the ideal functionality is composed out of CPTP maps modeling the pre- and postprocessing of the in- and outputs to the ideal functionality (which is a CPTP map itself). A desirable notion of security is the following: for every real adversary there exists an ideal adversary such that the corresponding CPTP maps are (approximately) indistinguishable. The natural measure of distinguishability of CPTP maps in this context is the diamond norm, since it can be viewed as the maximal bias of distinguishing real and ideal world by supplying inputs to the CPTP maps and attempting to distinguish the outputs by measurements (i.e., by interacting with an environment). This rather strong notion of security naturally embeds into a composable framework for security in which also quantum key distribution can be proven secure (see, e.g., Ref. [16]).

Since our goal is the establishment of a no-go theorem, we consider a notion of security which is weaker than the above in two respects. First, we do not allow the environment to supply an arbitrary input state but only the purification of a classical input (see definition of ρ_{UVR} below), and second, we consider a different order of quantifiers: instead of "$\forall$ real adversary $\exists$ ideal adversary $\forall$ input, the output states are indistinguishable" as a security requirement we only require "$\forall$ real adversary $\forall$ input $\exists$ ideal adversary, the outputs states are indistinguishable." This notion of security is closely related to notions of security considered in Ref. [13,15] and is further discussed in the Supplemental Material [17].

We will now give a formal definition of security. Following the notation of Ref. [15], we denote by A and B the real honest Alice and Bob and add a prime to denote dishonest players A', B' and a hat for the ideal versions $\hat{\mathsf{A}}$, $\hat{\mathsf{B}}$. The CPTP map corresponding to the protocol for honest Alice and dishonest Bob is denoted by $\pi_{\mathsf{A},\mathsf{B}'}$. Both honest and dishonest players obtain an input, in Alice's case u (in register U) and in Bob's case v (in register V) drawn from the joint distribution $p(u, v)$. The output state of the protocol, augmented by the reference R, takes the form $\mathrm{id}_R \otimes \pi_{\mathsf{A},\mathsf{B}'}(\rho_{UVR})$, where ρ_{UVR} is a purification of $\sum_{u,v} p(u, v)|u\rangle\langle u|_U|v\rangle\langle v|_V$.

Since we are faced with the task of the secure evaluation of a classical deterministic function, we consider an ideal functionality $\mathcal{F}$ which measures the inputs in registers $\tilde{U}$ and $\tilde{V}$ and outputs orthogonal states in registers $\tilde{X}$ and $\tilde{Y}$ that correspond to the function values. Formally, $\mathcal{F}(|u\rangle\langle u'|_{\tilde{U}}|v\rangle\langle v'|_{\tilde{V}}) := \delta_{u,u'}\delta_{v,v'}|f(u, v)\rangle\langle f(u, v)|_{\tilde{X}}|f(u, v)\rangle\langle f(u, v)|_{\tilde{Y}}$, where δ denotes the Kronecker delta function. When an ideal honest $\hat{\mathsf{A}}$ and an ideal adversary $\hat{\mathsf{B}}'$ interact with the ideal functionality, we denote the joint map by $\mathcal{F}_{\hat{\mathsf{A}},\hat{\mathsf{B}}'}: UV \to XY'$ (see Fig. 1). $\hat{\mathsf{A}}$ just forwards the in- and outputs to and from the functionality, whereas $\hat{\mathsf{B}}'$ pre- and postprocesses them with CPTP maps $\Lambda^1_{V\to\tilde{V}K}$ and $\Lambda^2_{K\tilde{Y}\to Y'}$ resulting in a joint map $\mathcal{F}_{\hat{A},\hat{\mathsf{B}}'} = [\mathrm{id}_{\tilde{X}\to X} \otimes \Lambda^2_{K\tilde{Y}\to Y'}] \circ [\mathcal{F}_{\tilde{U}\tilde{V}\to\tilde{X}\tilde{Y}} \otimes \mathrm{id}_K] \circ [\mathrm{id}_{U\to\tilde{U}} \otimes \Lambda^1_{V\to\tilde{V}K}]$, where $\circ$ denotes sequential application of CPTP maps.

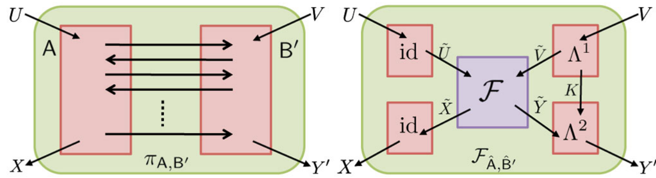

FIG. 1 (color online). Illustration of the security definition. A protocol is secure against Bob if the real protocol (left) can be simulated as an interaction with the ideal functionality $\mathcal{F}$ (right).

In the following, we let $\varepsilon \geq 0$ and write $\rho \approx_\varepsilon \sigma$ if $C(\rho, \sigma) \leq \varepsilon$. $C(\rho, \sigma)$ is the purified distance, defined as $\sqrt{1 - F(\rho, \sigma)^2}$ for $F(\rho, \sigma) := \mathrm{tr}\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}$ the fidelity. We say that a (two-party quantum) protocol π for f is ε-correct if for any distribution $p(u, v)$ of the inputs $[\mathrm{id}_R \otimes \pi_{\mathsf{A},\mathsf{B}}](\rho_{UVR}) \approx_\varepsilon [\mathrm{id}_R \otimes \mathcal{F}_{\hat{\mathsf{A}},\hat{\mathsf{B}}}](\rho_{UVR})$ and ε-secure against dishonest Bob if for any $p(u, v)$ and for any real adversary B' there exists an ideal adversary $\hat{\mathsf{B}}'$ such that $[\mathrm{id}_R \otimes \pi_{\mathsf{A},\mathsf{B}'}](\rho_{UVR}) \approx_\varepsilon [\mathrm{id}_\mathrm{R} \otimes \mathcal{F}_{\hat{\mathsf{A}},\hat{\mathsf{B}}'}](\rho_{UVR})$. ε-security against dishonest Alice is defined analogously.

Since $\mathcal{F}$ is classical, we can augment it so that it outputs $\tilde{v}$ in addition. More precisely, we define $\mathcal{F}_{\mathrm{aug}}:\tilde{U}\tilde{V}\to\tilde{X}\tilde{Y}\tilde{V}$ by $\mathcal{F}_{\mathrm{aug}}(|u\rangle\langle u'|_{\tilde{U}}|v\rangle\langle v'|_{\tilde{V}}) := \delta_{u,u'}\delta_{v,v'}|f(u,v)\rangle\langle f(u,v)|_{\tilde{X}}|f(u,v)\rangle\langle f(u,v)|_{\tilde{Y}}|v\rangle\langle v|_{\tilde{V}}$ which has the property that $\mathcal{F} = \mathrm{tr}_{\tilde{V}} \circ \mathcal{F}_{\mathrm{aug}}$. For a concrete input distribution we define $\sigma_{RX\tilde{V}Y'} := [\mathrm{id}_R \otimes \mathcal{F}_{\hat{\mathsf{A}},\hat{\mathsf{B}}',\mathrm{aug}}](\rho_{UVR})$ which satisfies $\sigma_{RXY'} \approx_\varepsilon \rho_{RXY'}$ for $\rho_{RXY'} := [\mathrm{id}_R \otimes \pi_{\mathsf{A},\mathsf{B}'}](\rho_{UVR})$ if the protocol is secure against cheating Bob. We call $\sigma_{RX\tilde{V}Y'}$ a secure state for $p(u, v)$.

Main results.—The proofs of our main results build upon the following lemma which constructs a cheating strategy for Alice that works on average over the input distribution $p(u, v)$.

Lemma.—If a protocol π for the evaluation of f is ε-correct and ε-secure against Bob, then for all input distributions $p(u, v)$ there is a cheating strategy for Alice such that she obtains $\tilde{v}$ with some probability distribution $q(\tilde{v}|u, v)$ satisfying $\sum_{u,v,\tilde{v}} p(u, v)q(\tilde{v}|u, v)\delta_{f(u,v),f(u,\tilde{v})} \geq 1 - 6\varepsilon$. Furthermore, $q(\tilde{v}|u, v)$ is almost independent of u; i.e., there exists a distribution $\tilde{q}(\tilde{v}|v)$ such that $\sum_{u,v,\tilde{v}} p(u, v)|q(\tilde{v}|u, v) - \tilde{q}(\tilde{v}|v)| \leq 6\varepsilon$.

Proof.—We first construct a "cheating unitary" T for Alice and then show how Alice can use it to cheat successfully.

Let Alice and Bob play honestly, but let them purify their protocol with purifying registers X_1' and Y_1', respectively. We assume without loss of generality that honest parties measure their classical input, and hence, X_1' and Y_1' contain copies of u and v, respectively. We denote by $|\Phi\rangle_{RXX_1'Y_1'Y}$ the state of all registers at the end of the protocol. Notice that tracing out X_1' from $|\Phi\rangle_{RXX_1'Y_1'Y}$ results in a state $\mathrm{tr}_{X_1'}|\Phi\rangle\langle\Phi|_{RXX_1'Y_1'Y} = \rho_{RXY_1'Y}$ which is exactly the final state when Alice played honestly and Bob played dishonestly with the following strategy: he plays the honest but purified strategy and outputs the purification of the protocol (register Y_1') and the output values $f(u, v)$ (register Y). His combined dishonest register is $Y' = Y_1'Y$. Since the protocol is ε-secure against Bob by assumption, there exists a secure state $\sigma_{RX\tilde{V}Y'}$ satisfying $\sigma_{RXY'} \approx_\varepsilon \rho_{RXY'}$. Let $|\Psi\rangle_{RXP\tilde{V}Y'}$ be a purification of $\sigma_{RX\tilde{V}Y'}$ with purifying register P. Note that $|\Psi\rangle_{RXP\tilde{V}Y'}$ is also a purification of $\sigma_{RXY'}$, this time with purifying registers $P\tilde{V}$. Recall that $|\Phi\rangle_{RXX_1'Y'}$ purifies $\rho_{RXY'}$ with purifying register X_1'. Since $\sigma_{RXY'} \approx_\varepsilon \rho_{RXY'}$ we can use Uhlmann's theorem [18] to conclude that there exists an isometry $T \equiv T_{X_1'\to P\tilde{V}}$ (with induced CPTP map $\mathcal{T} \equiv \mathcal{T}_{X_1'\to P\tilde{V}}$) such that

$$[\mathcal{T}_{X_1'\to P\tilde{V}} \otimes \mathrm{id}_{RXY'}](|\Phi\rangle\langle\Phi|_{RXX_1'Y'}) \approx_\varepsilon |\Psi\rangle\langle\Psi|_{RXP\tilde{V}Y'}. \quad (1)$$

We will now show how Alice can use T to cheat. Notice that tracing out Y_1' from $|\Phi\rangle_{RXX_1'YY_1'}$ results exactly in the final state when Bob played honestly and Alice played dishonestly with the following strategy: she plays the honest but purified strategy and outputs the purification of the protocol (register X_1') and the output values $f(u, v)$ (register X). She then applies $T_{X_1'\to P\tilde{V}}$, measures register $\tilde{V}$ in the computational basis, and obtains a value $\tilde{v}$. It remains to argue that Alice can compute $f(u, v)$ with good probability based on the value $\tilde{v}$ that she has obtained from measuring register $\tilde{V}$.

Let $\mathcal{M}_{R\tilde{V}X}$ be the CPTP map that measures registers R, X, and $\tilde{V}$ in the computational basis. Tracing over PY' and applying $\mathcal{M}_{R\tilde{V}X}$ on both sides of Eq. (1), we find

$$\begin{aligned}&[\mathcal{M}_{RX\tilde{V}} \otimes \mathrm{tr}_{PY'}]([\mathcal{T}_{X_1'\to P\tilde{V}} \otimes \mathrm{id}_{RXY'}](|\Phi\rangle\langle\Phi|_{RXX_1'Y'}))\\ &\quad\approx_\varepsilon [\mathcal{M}_{RX\tilde{V}} \otimes \mathrm{tr}_{PY'}](|\Psi\rangle\langle\Psi|_{RXP\tilde{V}Y'}) \quad (2)\end{aligned}$$

by the monotonicity of the purified distance under CPTP maps. The right-hand side of Eq. (2) equals $\sum_{u,v,\tilde{v}} p(u, v)\tilde{q}(\tilde{v}|v)|uv\rangle\langle uv|_R|\tilde{v}\rangle\langle\tilde{v}|_{\tilde{V}}|f(u, \tilde{v})\rangle\langle f(u, \tilde{v})|_X$ for some probability distribution $\tilde{q}(\tilde{v}|v)$ that is conditioned only on Bob's input v, since $|\Psi\rangle_{RXP\tilde{V}Y'}$ is a purification of the secure state $\sigma_{RX\tilde{V}Y'}$. The left-hand side of Eq. (2) equals $\sum_{u,v,\tilde{v},x} p(u, v)q(\tilde{v}|u, v)|uv\rangle\langle uv|_R|\tilde{v}\rangle\langle\tilde{v}|_{\tilde{V}}r(x|u, v, \tilde{v})|x\rangle\langle x|_X$ for some conditional probability distributions $q(\tilde{v}|u, v)$ and $r(x|u, v, \tilde{v})$. Because of the correctness of the protocol, this state is ε-close to

$$\sum_{u,v,\tilde{v}} p(u, v)\bar{q}(\tilde{v}|u, v)|uv\rangle\langle uv|_R|\tilde{v}\rangle\langle\tilde{v}|_{\tilde{V}}|f(u, v)\rangle\langle f(u, v)|_X, \quad (3)$$

for some conditional probability distribution $\bar{q}(\tilde{v}|u, v)$. Noting that therefore also $p(\cdot, \cdot)q(\cdot|\cdot, \cdot)$ and $p(\cdot, \cdot)\bar{q}(\cdot|\cdot, \cdot)$ (when interpreted as quantum states) are ε-close in purified distance, we can replace $p(\cdot, \cdot)\bar{q}(\cdot|\cdot, \cdot)$ in Eq. (3) by $p(\cdot, \cdot)q(\cdot|\cdot, \cdot)$ increasing the purified distance to the left-hand side of Eq. (2) only to 2ε. Putting things together, Eq. (2) implies

$$\sum_{u,v,\tilde{v}} p(u,v)q(\tilde{v}|u,v)|uv\rangle\langle uv|_R|\tilde{v}\rangle\langle\tilde{v}|_{\tilde{V}}|f(u,v)\rangle\langle f(u,v)|_X$$
$$\approx_{3\varepsilon} \sum_{u,v,\tilde{v}} p(u,v)\tilde{q}(\tilde{v}|v)|uv\rangle\langle uv|_R|\tilde{v}\rangle\langle\tilde{v}|_{\tilde{V}}|f(u,\tilde{v})\rangle\langle f(u,\tilde{v})|_X.$$

Sandwiching both sides with $\mathrm{tr}[Z\cdot]$, where $Z = \sum_{u,v,\tilde{v}}|uv\rangle\langle uv|_R|\tilde{v}\rangle\langle\tilde{v}|_{\tilde{V}}|f(u,\tilde{v})\rangle\langle f(u,\tilde{v})|_X$, we find the first claim since the purified distance of two distributions upper bounds their total variation distance and since the latter does not increase under $\mathrm{tr}[Z\cdot]$. The second claim follows similarly by tracing out register X from the last displayed equation. ∎

Applying the lemma to the uniform distribution we immediately obtain our impossibility result for perfectly secure protocols.

Theorem 1.—If a protocol π for the evaluation of f is perfectly correct and perfectly secure ($\varepsilon = 0$) against Bob, then, if Bob holds input v, Alice can compute $f(u, v)$ for all u.

We note that this notion of insecurity implies that Alice can completely break the security for nontrivial functions f.

Proof. Letting $p(u, v) = \frac{1}{|U||V|}$ and $\varepsilon = 0$ in the lemma results in the statement that if Alice has input u_0, then she will obtain $\tilde{v}$ from the distribution $q(\tilde{v}|u_0, v)$ which equals $\tilde{q}(\tilde{v}|v)$. But since also $q(\tilde{v}|u, v) = \tilde{q}(\tilde{v}|v)$ for all u, we have $\frac{1}{|U||V|}\sum_{u,v,\tilde{v}} q(\tilde{v}|u_0, v)\delta_{f(u,v),f(u,\tilde{v})} = 1$. In other words, all $\tilde{v}$ that occur (i.e., that have $\tilde{q}(\tilde{v}|v) > 0$) satisfy for all u, $f(u, v) = f(u, \tilde{v})$. Alice can therefore compute the function for all u. ∎

The impossibility result for the case of imperfect protocols is also based on the lemma but requires a subtle swap in the order of quantifiers (from "$\forall$ input $\exists$ ideal adversary" to "$\exists$ ideal adversary $\forall$ input") which we achieve by use of von Neumann's minimax theorem.

Theorem 2.—If a protocol π for the evaluation of f is ε-correct and ε-secure against Bob, then there is a cheating strategy for Alice (where she uses input u_0 while Bob has input v) which gives her $\tilde{v}$ distributed according to some distribution $Q(\tilde{v}|u_0, v)$ such that for all u: $Pr_{\tilde{v}\sim Q}[f(u, v) = f(u, \tilde{v})] \geq 1 - 28\varepsilon$.

Proof.—The argument is inspired by Ref. [19]. For a finite set $\mathcal{S}$, we denote by $\Delta(\mathcal{S})$ the simplex of probability distributions over $\mathcal{S}$. Denote by $\mathcal{W}$ the set of pairs (u, v). Consider a finite ε-net $\mathcal{D}$ of $\Delta(\mathcal{W})$ in total variation distance and to each distribution in $\mathcal{D}$ the corresponding cheating unitary T constructed in the proof of the lemma. We collect all these unitaries in the (finite) set $\mathcal{E}$ and assume that T determines p uniquely, as we could include the value p into T. For each such T, let $q(\tilde{v}|u, v, T)$ and $\tilde{q}(\tilde{v}|v, T)$ be the distributions from the lemma. Define the payoff function $g(u, v, T) := \sum_{\tilde{v}} q(\tilde{v}|u, v, T)\delta_{f(u,v),f(u,\tilde{v})} - \sum_{\tilde{v}}|q(\tilde{v}|u, v, T) - \tilde{q}(\tilde{v}|v, T)|$. The lemma then yields $1 - 12\varepsilon \leq \min_{p\in\mathcal{D}}\max_{T\in\mathcal{E}}\sum_{u,v} p(u, v)g(u, v, T)$ which is at most $2\varepsilon + \min_{p'\in\Delta(\mathcal{W})}\max_{T\in\mathcal{E}}\sum_{u,v} p'(u,v)g(u,v,T)$, since replacing p by p' incurs only an overall change in the value by 2ε [as $-1 \leq g(u, v, T) \leq 1$]. By von Neumann's minimax theorem, this last term equals $2\varepsilon + \max_{p''\in\Delta(\mathcal{E})}\min_{(u,v)\in\mathcal{W}}\sum_T g(u, v, T)p''(T)$ [20].

Hence, we have shown that there is a strategy for Alice, where she chooses her cheating unitary T with probability $p''(T)$, such that (for some $\varepsilon_1 + \varepsilon_2 \leq 14\varepsilon$) for all u, v,

$$\sum_{\tilde{v}} Q(\tilde{v}|u, v)\delta_{f(u,v),f(u,\tilde{v})} \geq 1 - \varepsilon_1 \tag{4}$$

and $\sum_{\tilde{v}}|Q(\tilde{v}|u, v) - \tilde{Q}(\tilde{v}|v)| \leq \sum_{\tilde{v},T} p(T)|q(\tilde{v}|u, v, T) - \tilde{q}(\tilde{v}|v, T)| \leq \varepsilon_2$, where $Q(\tilde{v}|u, v) := \sum_T p(T)q(\tilde{v}|u, v, T)$ and $\tilde{Q}(\tilde{v}|v) := \sum_T p(T)\tilde{q}(\tilde{v}|v, T)$. This implies that for all u, v, $\sum_{\tilde{v}}|Q(\tilde{v}|u_0, v) - Q(\tilde{v}|u, v)| \leq 2\varepsilon_2$. Combining this inequality with Eq. (4), we find for all u, v, $\sum_{\tilde{v}} Q(\tilde{v}|u_0, v)\delta_{f(u,v),f(u,\tilde{v})} \geq 1 - \varepsilon_1 - 2\varepsilon_2 \geq 1 - 28\varepsilon$. ∎

One might wonder whether Theorem 2 can be strengthened to obtain, with probability $1 - O(\varepsilon)$, a $\tilde{v}$ such that for all u: $f(u,v) = f(u,\tilde{v})$. It turns out that this depends on the function f: when f is equality $\mathrm{EQ}(u, v) = 1$ iff $u = v$] and inner product [$\mathrm{IP}(u,v) = \sum_i u_i \cdot v_i \bmod 2$], the stronger conclusion is possible. However, for disjointness [$\mathrm{DISJ}(u, v) = 0$ iff $\exists\, i : u_i = v_i = 1$] such a strengthening is not possible showing that our result is tight in general.

For EQ, we reason as follows. Set $u = v$ in Theorem 2. Alice is able to sample a $\tilde{v}$ such that $\sum_{\tilde{v}} Q(\tilde{v}|u_0, v)\delta_{EQ(v,v),EQ(v,\tilde{v})} \geq 1 - 28\varepsilon$. Since $\delta_{\mathrm{EQ}(v,v),\mathrm{EQ}(v,\tilde{v})} = 1$ iff $v = \tilde{v}$, $Q(v|u_0, v) \geq 1 - 28\varepsilon$. When f is IP, we pick u uniform at random and obtain $\sum_{\tilde{v}} Q(\tilde{v}|u_0,v)(2^{-n}\sum_u \delta_{\mathrm{IP}(u,v),\mathrm{IP}(u,\tilde{v})}) \geq 1-28\varepsilon$. Using $2^{-n}\sum_u \delta_{\mathrm{IP}(u,v),\mathrm{IP}(u,\tilde{v})} = 1$ if $\tilde{v} = v$, and $\frac{1}{2}$ if $\tilde{v} \neq v$, we find $Q(v|u_0, v) + \frac{1}{2}[1 - Q(v|u_0, v)] \geq 1 - 28\varepsilon$, which implies $Q(v|u_0, v) \geq 1 - 56\varepsilon$. Interestingly, for DISJ such an argument is not possible. Assume that we have a protocol that is ε-secure against Bob. Bob could now run the protocol normally on strings v with Hamming weight $|v| \leq n/2$, but on inputs v with $|v| > n/2$ he could flip, at random, $\sqrt{n}$ of v's bits that are 1. It is not hard to see that this new protocol is still ε-secure and $\varepsilon + O(\frac{1}{\sqrt{n}})$-correct. The loss in the correctness is due to the fact that, on high Hamming–weight strings, the protocol may, with a small probability, not be correct. On the other hand, on high Hamming–weight inputs, the protocol can not transmit or leak the complete input v to Alice, simply because Bob does not use it.

We thank Ivan Damgård, Frédéric Dupuis, Louis Salvail, Christopher Portmann, and Renato Renner for valuable discussions and an anonymous referee for suggesting an example presented in the Supplemental Material. M. C. is supported by the Swiss National Science Foundation (Grants No. PP00P2_128455 and No. 20CH21_138799), the NCCR "Quantum Science and Technology," and the German Science Foundation (Grant No. CH 843/2-1). C. S. is supported by a NWO Veni grant. H. B. was supported by an NWO vici grant and by EU project QCS.

[1] C. H. Bennett and G. Brassard, in *Proceedings of IEEE International Conference on Computers, Systems, and Signal Processing* (IEEE, New York, 1984), p. 175.
[2] A. K. Ekert, Phys. Rev. Lett. **67**, 661 (1991).
[3] H.-K. Lo and H. F. Chau, Phys. Rev. Lett. **78**, 3410 (1997).
[4] D. Mayers, Phys. Rev. Lett. **78**, 3414 (1997).
[5] C. Mochon, arXiv:0711.4114.
[6] R. Colbeck, Phys. Rev. A **76**, 062308 (2007).
[7] L. Salvail, M. Sotáková, and C. Schaffner, in *Advances in Cryptology—ASIACRYPT*, Lecture Notes in Computer Science Vol. 5912 (Springer-Verlag, Berlin, 2009), pp. 70–87.
[8] H.-K. Lo, Phys. Rev. A **56**, 1154 (1997).
[9] R. Canetti, J. Cryptol. **13**, 143 (2000).
[10] R. Canetti, Ph.D. thesis, The Weizmann Institute of Science, 1996.
[11] O. Goldreich, *Foundations of Cryptography* (Cambridge University Press, Cambridge, England, 2004), Vol. 2.
[12] D. Unruh, arXiv:quant-ph/0409125.
[13] D. Unruh, in *Advances in Cryptology EUROCRYPT*, Lecture Notes in Computer Science Vol. 6110 (Springer, New York, 2010), pp. 486–505.
[14] M. Ben-Or and D. Mayers, arXiv:quant-ph/0409062.
[15] S. Fehr and C. Schaffner, in *Theory of Cryptography Conference (TCC)* (Springer, New York, 2009), Vol. 5444, p. 350.
[16] M. Christandl, R. König, and R. Renner, Phys. Rev. Lett. **102**, 020504 (2009).
[17] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.109.160501 for a discussion about the details of our security definition.
[18] A. Uhlmann, Rep. Math. Phys. **9**, 273 (1976).
[19] G. M. D'Ariano, D. Kretschmann, D. Schlingemann, and R. F. Werner, Phys. Rev. A **76**, 032328 (2007).
[20] In order to apply von Neumann's theorem, note that the initial term equals $\min_{p'\in\Delta(\mathcal{W})}\max_{p''\in\Delta(\mathcal{E})}\sum_{u,v} p'(u,v)\times g(u,v,T)p''(T)$ since the maximal value of the latter is attained at an extreme point. Von Neumann's minimax theorem [21] allows us to swap minimization and maximization leading to $\max_{p''\in\Delta(\mathcal{E})}\min_{p\in\Delta(\mathcal{W})}\sum_{u,v,T}\times p(u,v)g(u,v,T)p''(T)$ without changing the value. This expression corresponds to the final term since the minimization can without loss of generality be restricted to its extreme points.
[21] J. v. Neumann, Math. Ann. **100**, 295 (1928).