A single-qubit position verification protocol that is secure against multi-qubit attacks

Bluhm, A.; Christandl, M.; Speelman, F.

DOI
10.48550/arXiv.2104.06301
10.1038/s41567-022-01577-0

Publication date
2022

Document Version
Submitted manuscript

Published in
Nature Physics

A SINGLE-QUBIT POSITION VERIFICATION PROTOCOL THAT IS SECURE AGAINST MULTI-QUBIT ATTACKS[1]

ANDREAS BLUHM, MATTHIAS CHRISTANDL, AND FLORIAN SPEELMAN

The position of a device or agent is an important security credential in today's society, both online and in the real world. Unless in direct proximity, however, the secure verification of a position is impossible without further assumptions. This is true classically [1], but also in any future quantum-equipped communications infrastructure [2]. We show in this work that minimal quantum resources, in the form of a single qubit, combined with classical communication are sufficient to thwart quantum adversaries that pretend to be at a specific position and have the ability to coordinate their action with entanglement. More precisely, we show that the adversaries using an increasing amount of entanglement can be combatted solely by increasing the number of classical bits used in the protocol. The presented protocols are noise-robust and within reach of current quantum technology.

The difficulty in achieving the verification of a position is best appreciated by considering certain secure-looking protocols and then understanding how they can be broken. For simplicity of the presentation we will consider the verification of the position of an untrusted agent being at midpoint between two verifiers. A protocol for position verification consists of the verifiers each sending messages to the agent who is asked to send responses back. The verification is successful if the responses satisfy certain conditions and if the timing of the signals is right (say in accordance with the speed of light) (see Figure 1).

A first attempt for a secure protocol could consist of a Boolean function f taking the message x from verifier 0 and y (both n-bit strings) from verifier 1 as input and sending the bit $f(x, y)$ back to the verifiers. In order for the agent to return the correct answer, clearly (for most functions f) both x and y are needed, but if the agent was not at midpoint but, say, closer to verifier 0, the agent could never both receive y and send the answer back to verifier 1 in time. Indeed, breaking the protocols entails not one attacking agent, but two, one of which is closer to verifier 0 and one which is closer to verifier 1. Customarily called Alice and Bob, the attacking agents both intercept the input from the verifier they are closest to. Each keeps a copy and forwards another copy of the input to the other partner in crime. When they hold both inputs in hand, they compute the function and return the function value just in time to their respective verifiers (see Figure 1).

This simple attack is indeed the basis of why position verification is not possible in the classical world. Note, however, that the attack directly uses the copying of information. This opens up the possibility of devising protocols based on the exchange of quantum information instead, whose copying is more restricted due to the no-cloning theorem [3, 4]. As Alice and Bob can agree on an attack strategy prior to the start of the protocol, however, they can also distribute entangled particles in order to later coordinate their action. Still, has the balance now tipped and position verification become possible? The plain answer is no [2], as Alice and Bob can immediately upon receipt of the quantum particles engage in an elaborate scheme of back and forth teleportation (with only a single round of crossing classical communication), known as instantaneous non-local

Date: January 24, 2023.

[1]This version of the article has been accepted for publication, after peer review and is subject to Springer Nature's AM terms of use, but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: https://doi.org/10.1038/s41567-022-01577-0

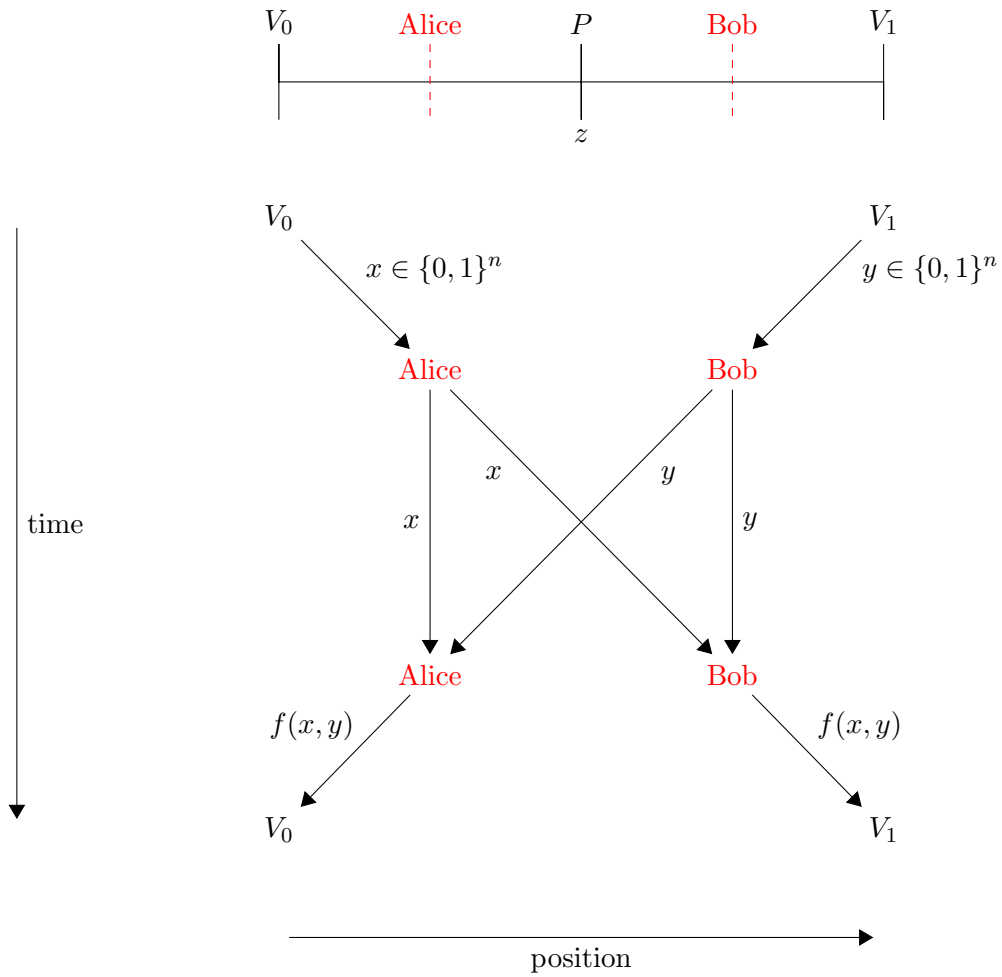

FIGURE 1. **Setup for PV in one spatial dimension (above) and classical attack (below).** The honest agent P is at position z, whereas the verifiers V_0, V_1 are to her left and right. For an attack, P is replaced by the attackers Alice and Bob, which are not at z, but in between z and the verifiers. Upon receiving $x \in \{0,1\}^n$, Alice copies the string and sends a copy on to Bob while Bob does the same with $y \in \{0,1\}^n$. Both attackers compute the function f and send the result back to the closest verifier. From the verifiers' point of view, they are indistinguishable from an honest prover P at z.

computation [5]. In a sense, this means 'game over' for position verification — if only the back and forth teleportation was not so expensive (doubly exponentially many EPR pairs in terms of the size of the protocol). This bound was brought down to singly exponential by use of port-based teleportation instead of standard quantum teleportation [6]. Note that carrying out such attacks is still prohibitively expensive for the attackers. Therefore, such attacks could be seen as unrealistic, forcing us to ask again whether position verification is after all viable in the quantum world.

We give a partial answer to this question by showing that there are protocols that enhance the above classical protocol by a single qubit and that withstand attacks involving roughly n qubits. A specific efficient protocol can withstand attacks of $\log n$ qubits. Thereby, we obtain security of position verification where the ratio of quantum resources required for the attack and of the honest agent is unbounded.

This is a qualitative improvement over prior work, which did show some level of security for an n-qubit protocol inspired by the BB84 protocol [2, 6, 7, 8], but for which an attacker only needs one EPR pair per qubit involved. Other previous works proposed new protocols [9, 10, 11, 12, 13], sometimes with the same scaling as BB84 protocol, and sometimes without security proofs showing how efficient the attacks can be; explicit efficient attacks can be constructed for several of these [14, 15, 13, 16]. Complementary to our results are works that study other security models [17, 18, 19], and that introduce techniques to increase the robustness to photon loss [20, 21, 22]. For the security analysis of a different protocol, see the recent independent work by Junge et al. [23]; we compare our results in Section 7 of the Supplementary Information. The routing protocol was introduced by Kent et al. [3] and studied further by Buhrman et al. [14]. We build on the proof strategy of the latter work.

We consider in this work two closely related protocols, which we will dub the 'routing' and 'measuring' protocols, which are direct enhancements of the classical protocols explained. The protocols are thus specified by a Boolean function f on $2n$ bits. In addition to the verifiers choosing random inputs x and y respectively, in both protocols verifier 0 will prepare a qubit chosen randomly from one of the BB84 states: $|0\rangle, |1\rangle, |+\rangle, |-\rangle$ and send it to the agent along with x. This could for instance be a single polarized photon sent in free space. In the routing protocol, the agent is asked to return the qubit unchanged to the verifier with number $f(x, y)$. Concretely, if $f(x, y) = 1$, the verifier could let the photon pass to verifier 1 and if $f(x, y) = 0$ use a mirror to reflect the photon back to verifier 0. The verifier could then measure the qubit and check whether the measurement result is consistent with the preparation. The protocol is illustrated in Figure 2. In the 'measuring' protocol, instead of routing the qubit, the agent is asked to measure in the $|0\rangle, |1\rangle$ basis in case $f(x, y) = 0$ and in the $|+\rangle, |-\rangle$ basis if $f(x, y) = 1$ and to return the measurement result to both verifiers. The protocol is illustrated in Figure 3.

In a sense, the only difference between the protocols is who carries out the measurement. It turns out that our security arguments therefore only differ in a single place (for more information see the methods section). As is familiar from the security analysis of quantum key distribution protocols, the security analysis of the described prepare and measure protocols is equivalent to their natural entanglement-based versions, which is preferred in formal arguments due to their conceptual simplicity. Here, verifier 0 prepares an EPR pair and sends half of it to the agent and holds on to the other half as a reference qubit. The protocol is otherwise unchanged and in order for the verifier later to compare results, the verifier will measure the reference qubit.

Let us point out that the implementation of the protocols merely requires the honest parties to be able to prepare and measure BB84 states, a task that is routinely carried out in the context of quantum key distribution both in laboratories and commercially. Indeed, the least quantum-technological requirements are demanded from the agent or the agent's device in the routing protocol: namely to either to reflect a photon with a mirror or to measure it.

The routing protocol is even simpler than the measuring protocol in the sense that the honest agent needs to perform no measurements. On the other hand, the reply of the agent in the measuring protocol is completely classical, and here our security proof also applies to the setting where quantum information travels slowly, meaning that only classical messages travel at the speed of light. This requirement fits current technology better, where qubits are transmitted using fiber optics. Thus, both protocols have their pros and cons and it depends on the desired application to determine which one is better suited.

We can show that for an appropriate function f, both the routing and measuring protocol are secure if the attackers do not hold more than $n/2 - 5$ qubits each at the beginning of their attack, when strings x and y of length n are sent by the verifiers. The most general form of the attacks is depicted in Figure 4. Moreover, the protocols can be repeated sequentially to make the probability that the attackers go unnoticed exponentially small. While we cannot give a concrete function f, we show that a uniformly random Boolean function will work with overwhelming probability.

Moreover, we consider the effect of noise on both protocols. The noise in the setup causes the honest agent not to succeed with certainty, but to fail with a 1% chance. In order to deal with the noise, the verifiers will repeat the protocol sequentially a number of times and accept if the protocol succeeds more than a fixed number of times. We can show that such protocols are still secure: An honest prover is rejected with a probability exponentially small in the number of repetitions. On the other hand, attackers controlling at most $n/2-5$ qubits at each round will succeed with probability exponentially small in the number of repetitions. This noise robustness of the single qubit protocols makes them interesting for near-term experimental implementations: for any reasonable bound on the number of qubits, i.e., a standard quantum bounded-storage assumption in cryptography, we have a secure protocol transmitting only a little more *classical* information as well as a single qubit over a *noisy* communication line.

Finally, we give lower bounds for concrete functions f, based on their communication complexity. For example, routing and measuring protocols using the inner product function are secure against attackers with at most $\log(n)/2 - 5$ qubits each. While these bounds for concrete functions are exponentially worse than for random functions, they still exhibit the feature that the ratio of the quantum resources the attackers need compared to the quantum resources an honest prover needs is unbounded in the number of classical bits n involved in the protocol, something not achieved in previous work. Furthermore, this also works in the presence of noise, hence providing us with a

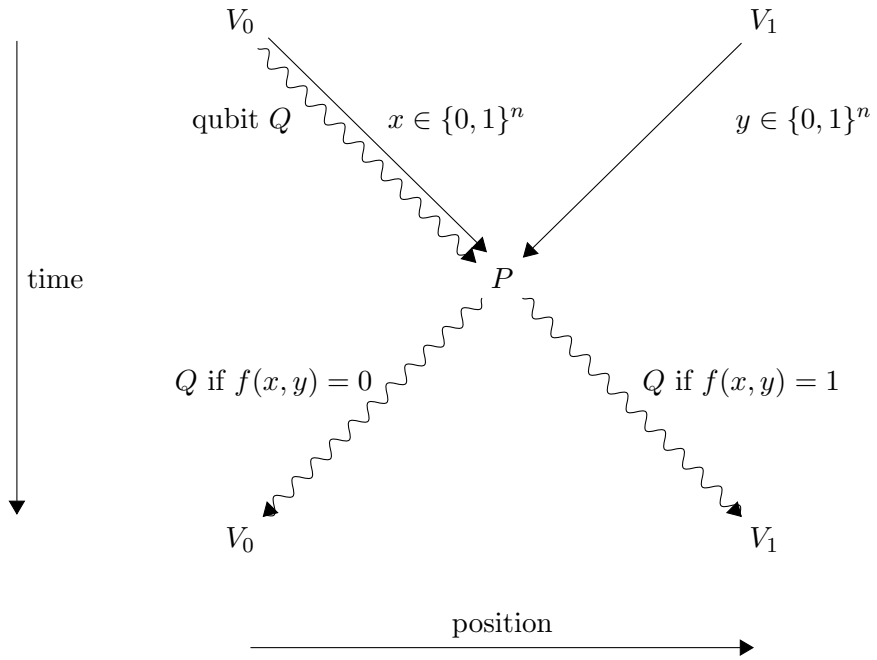

FIGURE 2. **The routing protocol.** In the protocol, the verifier V_0 prepares a qubit Q in one of the four BB84 states uniformly at random. Subsequently, V_0 sends Q together with a random n-bit string x to the agent P at position z and V_1 sends a random n-bit string y. All communication happens at the speed of light and the timing is such that Q, x and y reach position z at the same time. Depending on the outcome $f(x, y)$ of a previously agreed upon Boolean function f on $2n$ bits, the prover has to send the qubit Q received immediately to either verifier V_0 or V_1. The qubit Q has to reach the verifiers on time, i.e. the time of arrival at $V_{f(x,y)}$ has to be consistent with Q being sent from z at the speed of light right after Q has reached z. Straight lines correspond to classical information, while undulated lines correspond to quantum information being sent.

practical protocol that will provide security position verification under a quantum bounded storage assumption. We therefore believe that our work provides a blue print to the near-term realization of a new cryptographic primitive, which has the possibility to enhance our communication infrastructure with verified location as an additional security token.

In order to understand the open questions emerging from this work, note that it is important in our analysis of the random protocol that the Boolean function f we choose in order to run the protocols has to be truly random. This implies that the classical circuit to compute f is of exponential size in n. To decrease the classical resources needed for this protocol, it is therefore highly relevant to know whether it is possible to use pseudo-randomness instead, or whether there is another way to choose f with a circuit of polynomial size.

Finally, the most important open question is the following: When considering the dependence on the number of classical bits n, our lower bound implies that a number of qubits proportional to the number of classical bits sent by the verifiers is needed to attack the scheme. However, the best construction for a general attack takes 2^n EPR pairs [6, 14]. This leaves open the possibility that it could be even harder for attackers to break the security. Can we improve the lower bound to be exponential in n?

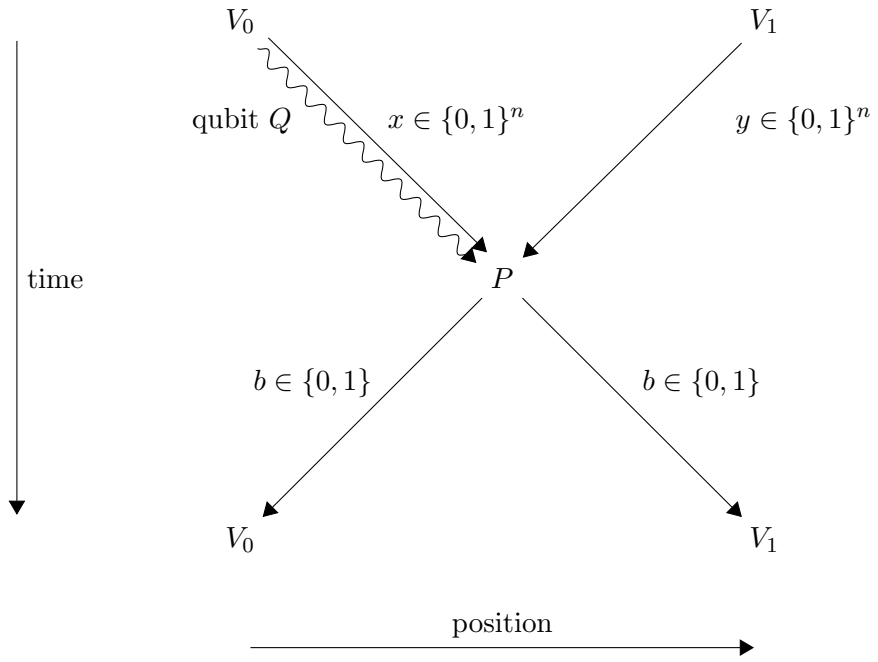

FIGURE 3. **The measuring protocol.** In the protocol, the verifiers V_0 and V_1 choose two random bit strings x, y of length n. If $f(x, y) = 0$, V_0 prepares a qubit Q in one of the computational basis states with equal probability, otherwise, V_0 prepares Q in one of the Hadamard basis states. Then, V_0 sends Q and x to P, V_1 sends y, and the timing is such that Q, x and y reach position z at the same time. If $f(x, y) = 0$, the prover measures Q in the computational basis, otherwise in the Hadamard basis. The outcome bit b of the measurement is subsequently sent back to both verifiers. It has to reach the verifiers on time, i.e. the time of arrival of b has to be consistent with b being sent from z at the speed of light right after Q has reached z. Straight lines correspond to classical information, while undulated lines correspond to quantum information being sent.

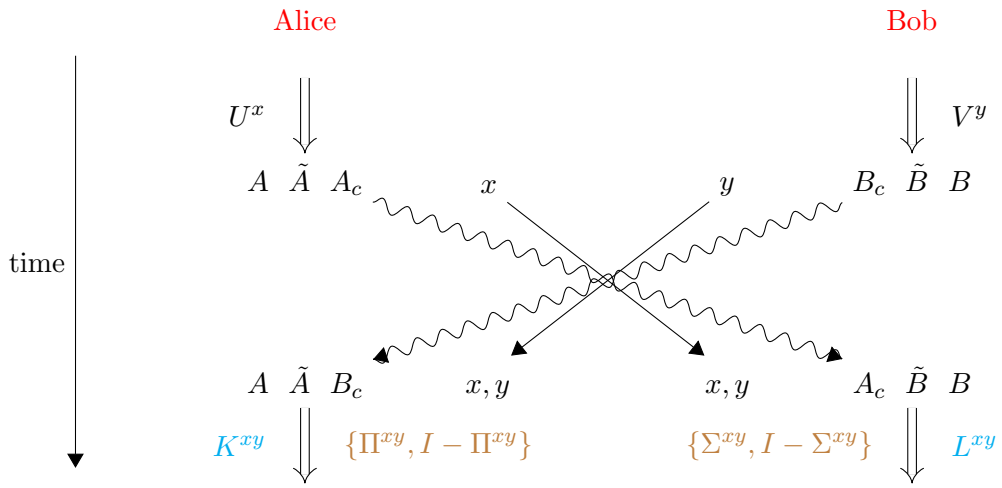

FIGURE 4. **Attack strategies for the routing and the measuring protocol.** The parts in black are the same for both protocols, while cyan belongs to the routing protocol and brown to the measuring protocol. Straight lines correspond to classical information, while undulated lines correspond to quantum information being sent. We assume that Alice and Bob each have a qubit system A and B, respectively. Moreover, Alice and Bob have local quantum registers $\tilde{A}$ and $\tilde{B}$. Due to the constraints imposed by special relativity, Alice and Bob are allowed one round of quantum communication, during which they can exchange systems A_c and B_c. We assume that both Alice and Bob have the same number of qubits. At the beginning of the protocol, Alice intercepts x and stores the qubit Q in A, while Bob intercepts y. The most general attacks are as follows: (1) Alice applies U^x on $A\tilde{A}A_c$, Bob applies V^y to $B\tilde{B}B_c$. (2) Alice sends A_c and x to Bob, Bob sends B_c and y to Alice. **Routing protocol:** (3) Alice applies K^{xy} on $A\tilde{A}B_c$ and Bob applies L^{xy} on $B\tilde{B}A_c$. (4) If $f(x, y) = 0$, Alice returns A to verifier V_0, if $f(x, y) = 1$, Bob returns B to verifier V_1. **Measuring protocol:** (3) Alice measures $\{\Pi^{xy}, I - \Pi^{xy}\}$ on $A\tilde{A}B_c$ and Bob measures $\{\Sigma^{xy}, I - \Sigma^{xy}\}$ on $B\tilde{B}A_c$. (4) Alice sends her measurement outcome to verifier 0, Bob sends his measurement outcome to verifier 1. Here, we take all operators to be unitaries and the superscript indicates which classical strings the unitaries might depend on.

METHODS

To prove our main result, we build on the proof strategy used in [14], overcoming both conceptual and technical difficulties. For simplicity, we will describe the security proof of the routing protocol first and comment on the differences for the measuring protocol at the end of the section. First, we observe that the joint quantum state of the attackers before their mutual communication arrives already suffices to determine where the qubit will be routed to in the given attack. We subsequently discretize the possible quantum strategies of the attackers with the help of ε-nets. Since the number of qubits of the attackers is bounded, the size of the ε-nets is limited. From there, we construct classical rounding functions which capture the essentials of the quantum strategies. In particular, an (ε, q)-classical rounding gives rise to a Boolean function for each attack Alice and Bob could do controlling at most q qubits each. These functions agree with the Boolean function f used in the routing protocol on all pairs of classical bit strings (x, y) on which the attackers succeed with probability at least $1 - \varepsilon^2$. In this sense, the classical rounding captures the information where the qubit is routed to during an attack. The ε-net construction shows that for ε small enough and

$q \in \mathbb{N}$, there exists an (ε, q)-classical rounding of size exponential in q. For the exact definition of an (ε, q)-classical rounding and the details of the proofs, we refer the reader to the Supplementary Information (see Section 3 for the routing protocol and Section 4 for the measuring protocol).

A counting argument that compares the number of (ε, q)-classical roundings to the number of Boolean functions f (on $2n$ bits) used to define the protocol then shows that most Boolean functions are far from any functions produced from classical roundings. More precisely, we show that for $q \leq n/2 - 5$, there exists a function $f : \{0,1\}^{2n} \to \{0,1\}$ that agrees with any function produced from the (ε, q)-classical rounding constructed previously on less than $3/4$ of the possible input pairs (x, y). Moreover, a uniformly random function f has this property with probability at least $1 - 2^{-2^n}$.

Picking f as above, the main result can then be proven by contradiction from the properties of an (ε, q)-classical rounding. Indeed, the counting argument implies that attackers controlling at most $n/2 - 5$ qubits each have to be detected with probability greater than ε^2 on at least $1/4$ of all possible pairs of bit strings (x, y). This shows that cheaters will be detected with constant probability for a random function.

For the measuring protocol, we can show that similarly the attackers have to decide in which basis to measure the qubit Q already before their mutual communication. The argument is based on an entropic uncertainty relation relative to quantum side information [24, 25]. The rest of the proof proceeds as for the routing protocol.

To obtain lower bounds for concrete functions, we consider the distributional communication complexity in the simultaneous message passing model. Here, Alice and Bob receive inputs x and y and send each a message of equal length to a referee. The latter is supposed to compute the value of the function with probability at least $3/4$ if the inputs are drawn from the uniform distribution. The aforementioned communication complexity is the number of bits Alice (or Bob) has to send. We prove that the routing and the measuring protocols are secure for a function with communication complexity at least k against attackers that control at most $\log(k)/2 - 3$ qubits each. The key insight is that any (ε, q)-classical rounding can be converted into a protocol in the communication complexity setting. Since successful attack strategies lead to (ε, q)-classical roundings, the number of qubits q of the attackers cannot be too small, since otherwise very efficient communication protocols would exist, contradicting the lower bound on the communication complexity.

Acknowledgements

The authors would like to thank Adrian Kent for organizing a workshop on relativistic quantum information theory in February 2020 during which part of this work was presented. AB and MC acknowledge financial support from the European Research Council (ERC Grant Agreement No. 81876), VILLUM FONDEN via the QMATH Centre of Excellence (Grant No.10059) and the QuantERA ERA-NET Cofund in Quantum Technologies implemented within the European Union's Horizon 2020 Programme (QuantAlgo project) via the Innovation Fund Denmark.

References

[1] N. Chandran, V. Goyal, R. Moriarty, and R. Ostrovsky, "Position based cryptography," in *Advances in Cryptology - CRYPTO 2009*, pp. 391–407, Springer, 2009. 1

[2] H. Buhrman, N. Chandran, S. Fehr, R. Gelles, V. Goyal, R. Ostrovsky, and C. Schaffner, "Position-based quantum cryptography: Impossibility and constructions," *SIAM Journal on Computing*, vol. 43, no. 1, pp. 150–178, 2014. 1, 3

[3] A. Kent, W. J. Munro, and T. P. Spiller, "Quantum tagging: Authenticating location via quantum information and relativistic signaling constraints," *Physical Review A*, vol. 84, p. 012326, 2011. 1, 3

[4] R. A. Malaney, "Location-dependent communications using quantum entanglement," *Physical Review A*, vol. 81, no. 4, p. 042319, 2010. 1

[5] L. Vaidman, "Instantaneous measurement of nonlocal variables," *Physical Review Letters*, vol. 90, p. 010402, Jan 2003. 2

[6] S. Beigi and R. König, "Simplified instantaneous non-local quantum computation with applications to position-based cryptography," *New Journal of Physics*, vol. 13, no. 9, p. 093036, 2011. 2, 3, 5
[7] M. Tomamichel, S. Fehr, J. Kaniewski, and S. Wehner, "A monogamy-of-entanglement game with applications to device-independent quantum cryptography," *New Journal of Physics*, vol. 15, no. 10, p. 103002, 2013. 3
[8] J. Ribeiro and F. Grosshans, "A tight lower bound for the BB84-states quantum-position-verification protocol," *arXiv-preprint arXiv:1504.07171*, 2015. 3
[9] H.-K. Lau and H.-K. Lo, "Insecurity of position-based quantum-cryptography protocols against entanglement attacks," *Physical Review A*, vol. 83, p. 012322, 2011. 3
[10] K. Chakraborty and A. Leverrier, "Practical position-based quantum cryptography," *Physical Review A*, vol. 92, p. 052304, 2015. 3
[11] R. Malaney, "The quantum car," *IEEE Wireless Communications Letters*, vol. 5, no. 6, pp. 624–627, 2016. 3
[12] S. Das and G. Siopsis, "Practically secure quantum position verification," *New Journal of Physics*, vol. 23, p. 063069, 2021. 3
[13] A. Gonzales and E. Chitambar, "Bounds on instantaneous nonlocal quantum computation," *IEEE Transactions on Information Theory*, vol. 66, no. 5, pp. 2951–2963, 2019. 3
[14] H. Buhrman, S. Fehr, C. Schaffner, and F. Speelman, "The garden-hose model," in *Proceedings of the 4th Conference on Innovations in Theoretical Computer Science*, ITCS '13, pp. 145–158, ACM, 2013. 3, 5, 6
[15] F. Speelman, "Instantaneous non-local computation of low T-depth quantum circuits," in *11th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2016)*, vol. 61 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pp. 9:1–9:24, Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2016. 3
[16] A. Olivo, U. Chabaud, A. Chailloux, and F. Grosshans, "Breaking simple quantum position verification protocols with little entanglement," *arXiv preprint arXiv:2007.15808*, 2020. 3
[17] A. Kent, "Quantum tagging with cryptographically secure tags," *arXiv preprint arXiv:1008.5380*, 2010. 3
[18] F. Gao, B. Liu, and Q.-Y. Wen, "Enhanced no-go theorem for quantum position verification," *arXiv-preprint arXiv:1305.4254*, 2013. 3
[19] D. Unruh, "Quantum position verification in the random oracle model," in *Annual Cryptology Conference*, pp. 1–18, Springer, 2014. 3
[20] B. Qi and G. Siopsis, "Loss-tolerant position-based quantum cryptography," *Physical Review A*, vol. 91, p. 042337, 2015. 3
[21] C. C. W. Lim, F. Xu, G. Siopsis, E. Chitambar, P. G. Evans, and B. Qi, "Loss-tolerant quantum secure positioning with weak laser sources," *Physical Review A*, vol. 94, p. 032315, 2016. 3
[22] R. Allerstorfer, H. Buhrman, F. Speelman, and P. Verduyn Lunel, "New protocols and ideas for practical quantum position verification," *arXiv-preprint arXiv:2106.12911*, 2021. 3
[23] M. Junge, A. M. Kubicki, C. Palazuelos, and D. Pérez-García, "Geometry of Banach spaces: a new route towards position based cryptography," *arXiv-preprint arXiv:2103.16357*, 2021. 3
[24] J. M. Renes and J.-C. Boileau, "Conjectured strong complementary information tradeoff," *Physical Review Letters*, vol. 103, p. 020402, 2009. 7
[25] M. Berta, M. Christandl, R. Colbeck, J. M. Renes, and R. Renner, "The uncertainty principle in the presence of quantum memory," *Nature Physics*, vol. 6, pp. 659–662, 2010. 7

Supplementary material

1. Main results

In this supplementary material, we provide proofs for all results mentioned in the main text. In particular, we prove the following theorems, which are our main results:

Our first result is that the routing protocol $\widetilde{PV}^f_{\text{route}}$ is secure if Alice and Bob control less than $n/2-5$ qubits, where $2n$ classical bits are being sent.

Theorem 1.1. *Let $n \geq 10$. Let us assume that the verifiers choose the bit strings x, y of length n uniformly at random. Then there exists a function $f : \{0,1\}^{2n} \to \{0,1\}$ with the property that, if the number q of qubits each of the attackers controls satisfies*

$$q \leq \frac{1}{2}n - 5,$$

the attackers are caught during $\widetilde{PV}^f_{\text{route}}$ with probability at least $2 \cdot 10^{-2}$. Moreover, a uniformly random function f will have this property (except with exponentially small probability).

Theorem 1.1 follows from Corollary 3.18 in this supplementary material. Already in the original publication of Kent, Munro, and Spiller [1] which proposed the routing protocol, it was shown that it is possible to attack this scheme if attackers share 2^n EPR pairs. Buhrman, Fehr, Schaffner, and Speelman [2] studied this class of protocols further, introducing the garden-hose model of communication complexity, which captures attacks relying on teleportation, and showed that an attack exists on the routing protocol using at most $\text{GH}(f)$ EPR pairs. Here, $\text{GH}(f)$ is the garden-hose complexity of the function f, a measure which is at most polynomial if the function is computable by a log-space Turing machine, but is exponential for a random function.

Our second result is that the measuring protocol PV^f_{meas} is also secure if Alice and Bob control less than $n/2-5$ qubits, where $2n$ classical bits are being sent.

Theorem 1.2. *Let $n \geq 10$. Let us assume that the verifiers choose the bit strings x, y of length n uniformly at random. Then there exists a function $f : \{0,1\}^{2n} \to \{0,1\}$ with the property that, if the number q of qubits each of the attackers controls satisfies*

$$q \leq \frac{1}{2}n - 5,$$

the attackers are caught during PV^f_{meas} with probability at least $2 \cdot 10^{-2}$. Moreover, a uniformly random function f will have this property (except with exponentially small probability).

Theorem 1.2 follows from Theorem 4.10. The results of [2] can be adapted to construct attacks on PV^f_{meas} for which the entanglement required is given by the garden-hose complexity of f (a measure that is polynomial for log-space functions, but can be exponential in general). Thus, a general attack on PV^f_{meas} is possible when attackers share 2^n EPR pairs for any function f. As an aside, we do note that, despite the fact that these specific attacks can be translated, we do not know in general whether an attack on PV^f_{route} can be translated into an attack on PV^f_{meas} or vice-versa.

Finally, we consider concrete instead of random functions f. In particular, for the binary inner product function

$$IP(x,y) = \sum_{i=1}^{n} x_i y_i \pmod 2, \tag{1}$$

we can prove the following:

Theorem 1.3. *Let $n \geq 10$. Let us assume that the verifiers choose the bit strings x, y of length n uniformly at random. If the number q of qubits each of the attackers controls satisfies*

$$q \leq \frac{1}{2}\log n - 5,$$

the attackers are caught during $\widetilde{PV}^{IP}_{\text{route}}$ and PV^{IP}_{meas} with probability at least $2 \cdot 10^{-2}$, respectively.

The statement follows from Theorem 6.2. The supplementary material is organized as follows: Section 2 contains some preliminaries concerning communication matrices and the purified distance between quantum states. Our results concerning the routing protocol appear in Section 3. Subsequently, we consider the measuring protocol in Section 4, before we prove both protocols to be noise robust in Section 5. Lower bounds for concrete instead of random functions for both protocols are proven in Section 6. In Section 7 we discuss the importance of the attack model in results on quantum position verification and compare our results to previous and independent work. Finally, we conclude in Section 8 with some technical results which are needed in the proofs.

2. Preliminaries

2.1. **Communication matrix.** Let $d_H : \{0,1\}^n \times \{0,1\}^n \to \mathbb{N}$ be the Hamming distance. Let us define for a, $n \in \mathbb{N}$,

$$V(n,a) = \sum_{l=0}^{a} \binom{n}{l}.$$

That is the cardinality of the ball of Hamming distance a. Let $\lambda \in (0, 1/2)$ be such that $\lambda n \in \mathbb{N}$. In [3, p.310], we find the useful bound

$$V(n, \lambda n) \leq 2^{nh(\lambda)}, \tag{2}$$

where $h(p) := -p\log p - (1-p)\log(1-p)$ is the binary entropy function. The function log will be the logarithm with respect to base 2 in this paper.

Let $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$. The *communication matrix* of f is defined as

$$(M_f)_{x,y} = f(x,y).$$

It is thus a $2^n \times 2^n$ matrix. The Hamming distance $d_H(M_f, M_g)$ therefore tells you, for how many pairs of bit strings (x,y) the value $g(x,y)$ differs from $f(x,y)$. Note that here we interpret M_f, M_g as strings of length 2^{2n}.

2.2. **Fidelity and purified distance.** Let us define the fidelity between two quantum states as

$$F(\rho,\sigma) := \operatorname{tr}\left[\sqrt{\sqrt{\sigma}\rho\sqrt{\sigma}}\right]. \tag{3}$$

In particular, $F(|\psi\rangle, |\varphi\rangle) = |\langle\psi|\varphi\rangle|$. Here, we write $F(|\psi\rangle, |\varphi\rangle)$ for pure states $|\psi\rangle$, $|\varphi\rangle$ to mean $F(|\psi\rangle\langle\psi|, |\varphi\rangle\langle\varphi|)$ for brevity. Note that sometimes the fidelity is defined as the square of (3). The fidelity can be used to define the purified distance on the set of density matrices [4, Definition 3.8]. For quantum states ρ, σ, it is defined as

$$\mathcal{P}(\rho,\sigma) := \sqrt{1 - F(\rho,\sigma)^2}.$$

Again, we often write $\mathcal{P}(|\psi\rangle, |\varphi\rangle)$ for pure states $|\psi\rangle$, $|\varphi\rangle$ instead of $\mathcal{P}(|\psi\rangle\langle\psi|, |\varphi\rangle\langle\varphi|)$. Unlike the fidelity, the purified distance is a metric on the set of states, which makes it easier to work with (see e.g. [4, Proposition 3.3]). In particular, it satisfies the triangle ineguality.

3. The qubit routing protocol

3.1. **Qubit routing.** Let $f : \{0,1\}^{2n} \to \{0,1\}$. We consider PV in one spatial dimension. We will start by considering a modified version of the routing protocol which uses a maximally entangled pair. Our main result will follow by realizing in Section 3.3 that the entangled and unentangled qubit routing protocols are essentially equivalent. The general setup for the entangled qubit routing protocol PV^f_{route} is the following: The prover P claims to be at position z on a line. To the left and right of z are the verifiers V_0 and V_1. All communication happens at the speed of light. The protocol PV^f_{route} considered in [2] is the following (see Figure 2 of that paper):

(1) V_0 randomly chooses two n-bit strings x, y, computes $f(x,y)$ and sends y on to V_1. Moreover, V_0 prepares a maximally entangled 2-qubit state $|\Omega\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. If $f(x,y) = 0$, V_0 does nothing, if $f(x,y) = 1$, V_0 sends one qubit R of $|\Omega\rangle$ to V_1.
(2) V_0 sends the other qubit Q of $|\Omega\rangle$ together with x to P. V_1 sends y such that it arrives at z at the same time as x and Q sent by V_0.
(3) P sends the qubit Q on to $V_{f(x,y)}$.
(4) V_0 and V_1 accept if the qubit arrives at the correct time at the correct verifier and a Bell measurement of both qubits yields the correct outcome.

The timing of the response from P is deemed correct if it is compatible with the qubit Q originating from z right after it reached that point. An illustration of the protocol can be found in Figure 2 in the main text.

The advantage of this protocol compared to others is that the honest prover only needs to handle one qubit. Note that this qubit could even be presented as a logical qubit in an error-correcting code to combat noise in the communication line. Since the protocol only requires routing this qubit and no further processing of it, any error correcting code (even without fault-tolerant properties) is fine. Errors in creating the qubit states and verifying it on the side of the verifiers, however, need to be carried out in a fault-tolerant manner.

We will now give the form of the most general attack on PV^f_{route} . Note that we can restrict our attention to unitaries by considering the Stinespring dilation of the quantum channels the attackers might wish to perform. There are two attackers Alice and Bob, where Alice is between V_0 and z and Bob is between z and V_1. However, neither of the attackers is actually at z. As explained in Section 1, the verifiers hold a qubit system R, while Alice holds a qubit system A, a local quantum system $\tilde{A}$ and a quantum system used for communication A_c. Bob has similar systems B, $\tilde{B}$ and B_c.

Definition 3.1 (q-qubit strategy for PV^f_{route})**.** *Fix a partition into systems $RA\tilde{A}A_cB\tilde{B}B_c$. Both Alice's and Bob's registers each consist of q qubits. Let d be the combined dimension of this system, therefore $d = 2^{2q+1}$. A* q-qubit strategy for PV^f_{route} *consists of the starting state $|\psi\rangle$ on $RA\tilde{A}A_cB\tilde{B}B_c$ and of unitaries $U^x_{A\tilde{A}A_c}$, $V^y_{B\tilde{B}B_c}$, $K^{xy}_{A\tilde{A}B_c}$ and $L^{xy}_{B\tilde{B}A_c}$ for all x, $y \in \{0,1\}^n$. The superscripts indicate whether the unitaries may depend on the message x that V_0 sends, the message y that V_1 sends or on both messages.*

Note that Alice and Bob only hold equally many qubits at the beginning of the strategy, after the communication phase the numbers can be different. An illustration of the above can be found in Figure 4 in the main text.

Remark 3.2. *From the description of the protocol, it is clear that we only need consider strategies for which $|\psi\rangle = |\Omega\rangle_{AR} \otimes |\psi'\rangle_{\tilde{A}A_cB\tilde{B}B_c}$ to prove the protocol secure. However, it is advantageous to consider more general starting states: In reality, photon signals over fiber travel at around $(2/3)c$, where c is the speed of light, while we want to allow our attackers to signal at speed c. If our proof can quantify over all pre-shared states between $RA\tilde{A}A_cB\tilde{B}B_c$, then conceptually the input message can be 'slow'. We could even imagine the state being available long before the protocol, with Alice*

and Bob distributing the state amongst themselves however they want. The input timing is then only on the classical messages. Alternatively, they could start computing locally before the classical messages x and y are available. All these scenarios lead to a starting state not of the form $|\Omega\rangle\otimes|\psi'\rangle$, which is why considering general states $|\psi\rangle$ within the q-qubit strategies for PV^f_{route} only makes the security notion stronger.

Remark 3.3. *It can easily be seen that shared randomness between Alice and Bob does not help them for a fixed function f. Indeed, if ρ is the reduced state at the end of the protocol on RA if $f(x,y)=0$ or RB if $f(x,y)=1$, the probability that Alice and Bob are not caught by the verifiers is $\langle\Omega|\,\rho\,|\Omega\rangle$. Note that the objective function $\langle\Omega|\,\rho\,|\Omega\rangle$ is linear in ρ and the partial trace is a linear map. Thus, the maximum over convex combinations of strategies is achieved at deterministic strategies $\{U^x, V^y, K^{xy}, L^{xy}\}_{xy}$.*

The main lower-bound result of [2] concerning the entangled qubit routing protocol PV^f_{route} is the following:

Theorem 3.4 ([2, Theorem E.4])**.** *Let $q, n\in\mathbb{N}$. For any q-qubit starting state $|\psi\rangle$ on $RA\tilde{A}A_cB\tilde{B}B_c$, there exists a Boolean function on inputs x, $y\in\{0,1\}^n$ such that any perfect attack on PV^f_{route} requires q to be linear in n.*

On the one hand, this theorem proves that PV^f_{route} is secure in some sense if the number of qubits the attackers control is at most linear in n. On the other hand, it has several features which make it unsuitable to derive any limits for actual attacks on the PV^f_{route} -scheme from it. Firstly, it only discusses perfect attacks, while actual attackers would still be practically successful if they have a small probability of being caught. Secondly, the theorem fixes the state before quantifying over the functions, while actual attackers would be able to choose their entanglement after knowing the function f. This can be interpreted as a violation of Kerkhoffs's principle, since the function f must not be known to the attackers beforehand. Finally, the theorem only shows that there exist an input pair x, y for which the attackers will be detectable, but does not say anything about how many such pairs exist, leaving the possibility that these pairs might only be asked with exponentially small probability. These severe drawbacks make Theorem 3.4 unsuitable for practical applications.

The aim of this work is therefore to improve upon Theorem 3.4 and to provide a version that solves all three problems, thus recovering the statement under a more realistic class of attacks.

3.2. **Lower bounds on the entangled qubit routing protocol.** We start our analysis of the entangled qubit routing protocol by defining an (ε,l)-perfect q-qubit strategy as one which has a high chance of being accepted by the verifiers at the end of the protocol. In [2], only perfect strategies were considered, i.e. $\varepsilon=0$, $l=2^{2n}$. Moreover, we want to allow that the attackers only succeed on l of the 2^{2n} pairs of bit strings (x,y) that the verifiers might send.

Definition 3.5 ((ε,l)-perfect q-qubit strategy for PV^f_{route})**.** *Let $\varepsilon>0$, $l\in\mathbb{N}$. A q-qubit strategy for PV^f_{route} as in Definition 3.1 is* (ε,l)-perfect *if on l pairs of strings (x,y), Alice and Bob are caught by the verifiers with probability at most ε^2.*

Remark 3.6. *Note that Alice and Bob are caught by the verfiers on input (x,y) with probability at most ε^2 if and only if Alice and Bob produce a state $|\tilde{\psi}\rangle$ at the end of the protocol such that $\mathcal{P}(\rho_{RA},|\Omega\rangle\langle\Omega|_{RA})\leq\varepsilon$ if $f(x,y)=0$ and $\mathcal{P}(\rho_{RB},|\Omega\rangle\langle\Omega|_{RB})\leq\varepsilon$ if $f(x,y)=1$, where ρ is the corresponding reduced state of $|\tilde{\psi}\rangle$.*

The following proposition relates the above definition to the purified distance with respect to the state $|\tilde{\psi}\rangle$ as in [2, Appendix E]. It is a direct consequence of Uhlmann's theorem [5, Theorem 3.22].

Proposition 3.7. *For a state $|\tilde{\psi}\rangle_{RA\tilde{A}A_cB\tilde{B}B_c}$, it holds that*

$$\inf_{|\varphi\rangle} \mathcal{P}(|\tilde{\psi}\rangle_{RA\tilde{A}A_cB\tilde{B}B_c}, |\Omega\rangle_{RA} \otimes |\varphi\rangle_{\tilde{A}A_cB\tilde{B}B_c}) = \mathcal{P}(\rho_{RA}, |\Omega\rangle\langle\Omega|_{RA})$$

and

$$\inf_{|\varphi\rangle} \mathcal{P}(|\tilde{\psi}\rangle_{RA\tilde{A}A_cB\tilde{B}B_c}, |\Omega\rangle_{RB} \otimes |\varphi\rangle_{A\tilde{A}A_c\tilde{B}B_c}) = \mathcal{P}(\rho_{RB}, |\Omega\rangle\langle\Omega|_{RB}),$$

where ρ_{RA} and ρ_{RB} are the corresponding reduced density matrices of $|\tilde{\psi}\rangle_{RA\tilde{A}A_cB\tilde{B}B_c}$.

Before we go on, we define the sets of states from which the routed qubit can be recovered by attacker Alice or Bob, to be returned to V_0 or V_1 respectively. Note that we will always write A_X instead $A_X \otimes I_{X^c}$ for ease of notation, where X is a system with complement X^c, A an operator and I the identity operator.

Definition 3.8. *Let $\varepsilon \in [0,1]$. We define $\mathcal{S}_0^{\varepsilon,\text{route}}$ as the set of states $|\varphi\rangle_{RA\tilde{A}A_cB\tilde{B}B_c}$ for which there exists a unitary $K_{A\tilde{A}B_c}$ such that $\mathcal{P}(\rho_{RA}, |\Omega\rangle\langle\Omega|_{RA}) \leq \varepsilon$, where ρ is the reduced state of $K|\varphi\rangle$. Moreover, we define $\mathcal{S}_1^{\varepsilon,\text{route}}$ as the set of states $|\varphi'\rangle_{RA\tilde{A}A_cB\tilde{B}B_c}$ for which there exists a unitary $L_{B\tilde{B}A_c}$ such that $\mathcal{P}(\rho'_{RB}, |\Omega\rangle\langle\Omega|_{RB}) \leq \varepsilon$, where ρ' is the reduced state of $L|\varphi'\rangle$.*

Now, we consider a state that can be used to reveal the qubit at V_0 in the last step of a q-qubit strategy and a state that can be used to reveal the qubit at V_1 in the last step of the strategy. We prove a proposition which formalizes the idea that these two such states have to differ by at least a certain amount. This shows that the sets we just defined are disjoint if we choose ε small enough. This proposition can be seen as a robust version of [2, Lemma E.1].

Proposition 3.9. *Let $0 \leq \varepsilon \leq 0.41$ and let $|\psi_0\rangle_{RA\tilde{A}A_cB\tilde{B}B_c} \in \mathcal{S}_0^{\varepsilon,\text{route}}$, $|\psi_1\rangle_{RA\tilde{A}A_cB\tilde{B}B_c} \in \mathcal{S}_1^{\varepsilon,\text{route}}$. Then,*

$$\mathcal{P}(|\psi_0\rangle, |\psi_1\rangle) > 0.046.$$

Proof. By the definition of the sets in Definition 3.8, there exist unitaries $K_{A\tilde{A}B_c}$ and $L_{B\tilde{B}A_c}$ such that $\mathcal{P}(\rho_0, |\Omega\rangle\langle\Omega|_{RA}) \leq \varepsilon$ and $\mathcal{P}(\rho_1, |\Omega\rangle\langle\Omega|_{RB}) \leq \varepsilon$ for ρ_0 the reduced state on RA of $K_{A\tilde{A}B_c}|\psi_0\rangle$, ρ_1 the reduced state on RB of $L_{B\tilde{B}A_c}|\psi_1\rangle$. By Proposition 3.7 and compactness, we can find states $|\varphi_0\rangle_{\tilde{A}A_cB\tilde{B}B_c}$ and $|\varphi_1\rangle_{A\tilde{A}A_c\tilde{B}B_c}$ such that

$$\mathcal{P}(K_{A\tilde{A}B_c}|\psi_0\rangle_{RA\tilde{A}A_cB\tilde{B}B_c}, |\Omega\rangle_{RA} \otimes |\varphi_0\rangle_{\tilde{A}A_cB\tilde{B}B_c}) = \mathcal{P}(\rho_0, |\Omega\rangle\langle\Omega|_{RA}),$$
$$\mathcal{P}(L_{B\tilde{B}A_c}|\psi_1\rangle_{RA\tilde{A}A_cB\tilde{B}B_c}, |\Omega\rangle_{RB} \otimes |\varphi_1\rangle_{A\tilde{A}A_c\tilde{B}B_c}) = \mathcal{P}(\rho_1, |\Omega\rangle\langle\Omega|_{RB}).$$

Applying the triangle inequality twice and using the fact that $\mathcal{P}(U|\varphi\rangle, U|\psi\rangle) = \mathcal{P}(|\varphi\rangle, |\psi\rangle)$ for any unitary U and any states $|\varphi\rangle$, $|\psi\rangle$, we obtain

$$\mathcal{P}(|\psi_0\rangle, |\psi_1\rangle) \geq \mathcal{P}(K^*|\Omega\rangle \otimes |\varphi_0\rangle, L^*|\Omega\rangle \otimes |\varphi_1\rangle) - \mathcal{P}(K|\psi_0\rangle, |\Omega\rangle \otimes |\varphi_0\rangle) - \mathcal{P}(L|\psi_1\rangle, |\Omega\rangle \otimes |\varphi_1\rangle).$$

We can estimate the last two terms on the right hand side as

$$\mathcal{P}(K|\psi_0\rangle, |\Omega\rangle \otimes |\varphi_0\rangle) \leq \varepsilon,$$
$$\mathcal{P}(L|\psi_1\rangle, |\Omega\rangle \otimes |\varphi_1\rangle) \leq \varepsilon.$$

These inequalities hold by assumption. By the computations of [2, Lemma E.1] (repeated as Lemma 8.1 for completeness), we can estimate the first term as

$$\mathcal{P}(K^*|\Omega\rangle \otimes |\varphi_0\rangle, L^*|\Omega\rangle \otimes |\varphi_1\rangle) \geq \frac{\sqrt{3}}{2}.$$

Thus,

$$\mathcal{P}(|\psi_0\rangle, |\psi_1\rangle) \geq \frac{\sqrt{3}}{2} - 2\varepsilon$$

and the assertion follows using the assumption $\varepsilon \leq 0.41$. □

We will need a final easy lemma to convert between Euclidean distance and purified distance. It is a direct consequence of the fact that $1 - x^2 = (1-x)(1+x) \leq 2(1-x)$ for $x \in [0,1]$.

Lemma 3.10. *Let $|x\rangle$, $|y\rangle \in \mathbb{C}^d$ be two unit vectors. Then,*

$$\mathcal{P}(|x\rangle, |y\rangle) \leq \| |x\rangle - |y\rangle \|_2.$$

We observe that Proposition 3.9 implies that the attackers in some sense already decide before their communication step where the qubit can end up at the end of the protocol. Therefore, if the dimension of the state they share is small enough, a classical description of the first part of their strategy yields a compression of f. The classical compression is captured in the following notion of classical roundings:

Definition 3.11 ((ε, q)-classical rounding)**.** *Let q, k, $n \in \mathbb{N}$, $\varepsilon > 0$. Then,*

$$g : \{0,1\}^{3k} \to \{0,1\}$$

*is an (ε, q)-*classical rounding *of size k if for all $f : \{0,1\}^{2n} \to \{0,1\}$, for all states $|\psi\rangle$ on $2q+1$ qubits, for all $l \in \{1, \ldots, 2^{2n}\}$ and for all (ε, l)-perfect q-qubit strategies for PV^f_{route}, there are functions $f_A : \{0,1\}^n \to \{0,1\}^k$, $f_B : \{0,1\}^n \to \{0,1\}^k$ and $\lambda \in \{0,1\}^k$ such that $g(f_A(x), f_B(y), \lambda) = f(x,y)$ on at least l pairs (x,y).*

The function $\tilde{f}$ defined as

$$\tilde{f}(x,y) := g(f_A(x), f_B(y), \lambda) \qquad \forall x, y \in \{0,1\}^n \tag{4}$$

in a classical rounding hence measures how good the q-qubit strategy Alice and Bob use performs for the qubit routing specified by the function f. For example, if the strategy is an $(\varepsilon, 2^{2n})$-perfect q-qubit strategy, then $f = \tilde{f}$.

Since the following statement holds for both the routing and the measure protocol, which we consider in the next section, we prove it here for both protocols, although the sets $\mathcal{S}_i^{\varepsilon,\text{meas}}$ are only defined in Definition 4.4.

Lemma 3.12. *Let $\# \in \{\text{route}, \text{meas}\}$, $q \in \mathbb{N}$. Furthermore, let $0 \leq \varepsilon \leq \varepsilon_0$, where ε_0 is such that $|\varphi_i\rangle \in \mathcal{S}_i^{\varepsilon,\#}$, $i \in \{1,2\}$ implies $\mathcal{P}(|\varphi_0\rangle, |\varphi_1\rangle) > 0.013$. Then, there is an (ε, q)-classical rounding of size $\log(927) 2^{2q+2}$.*

Proof. We consider $\delta = 0.00216$. Let us choose a δ-net $\mathcal{N}_S$ in Euclidean norm for the set of pure states on $2q+1$ qubits, where the net has cardinality at most 2^k. Likewise, let us choose δ-nets $\mathcal{N}_A$ and $\mathcal{N}_B$ in operator norm for the set of unitaries in dimension 2^q, where the nets have cardinalities at most 2^k each. We will show at the end of the proof that we can choose k as in the assertion.

Let us now construct the (ε, d)-classical rounding g as in Definition 3.11. Let $x' \in \{0,1\}^k$, $y' \in \{0,1\}^k$ and $\lambda \in \{0,1\}^k$ and let $U \in \mathcal{N}_A$ be the element with index x', $V \in \mathcal{N}_B$ be the element with index y' and $|\varphi\rangle \in \mathcal{N}_S$ be the element with index λ. Then, we define $g(x', y', \lambda) = 0$ if $U_{A\tilde{A}A_c} \otimes V_{B\tilde{B}B_c} |\varphi\rangle$ is closer to $\mathcal{S}_0^{\varepsilon,\#}$ than to $\mathcal{S}_1^{\varepsilon,\#}$ in purified distance and $g(x', y', \lambda) = 1$ if $U_{A\tilde{A}A_c} \otimes V_{B\tilde{B}B_c} |\varphi\rangle$ is closer to $\mathcal{S}_1^{\varepsilon,\#}$ than to $\mathcal{S}_0^{\varepsilon,\#}$ in purified distance. If neither is the case, we make the arbitrary choice $g(x', y', \lambda) = 1$. Since the assumption on ε_0 implies that $\mathcal{S}_0^{\varepsilon,\#} \cap \mathcal{S}_1^{\varepsilon,\#} = \emptyset$, this is a well-defined function.

It remains to show that g is indeed an (ε, q)-classical rounding. We consider an arbitrary $f : \{0,1\}^{2n} \to \{0,1\}$ and an arbitrary state on $2q+1$ qubits $|\psi\rangle$. Let $|\psi\rangle$ and $\{U^x, V^y\}_{xy}$ be from a q-qubit strategy for $PV^f_{\#}$. We choose λ as the index of the closest element from $\mathcal{N}_S$ to $|\psi\rangle$ in Euclidean norm. Moreover, we choose $f_A(x)$ to be the closest element from $\mathcal{N}_A$ to U^x in operator norm and $f_B(y)$ to be the closest element from $\mathcal{N}_B$ to V^y in operator norm. If the closest element is not unique, we make an arbitrary choice. We claim that if $U^x \otimes V^y |\psi\rangle \in \mathcal{S}_0^{\varepsilon,\#}$, then $U \otimes V |\varphi\rangle$

is closer to $\mathcal{S}_0^{\varepsilon,\#}$ than to $\mathcal{S}_1^{\varepsilon,\#}$, where $(U, V, |\varphi\rangle)$ are the elements from the nets corresponding to $(f_A(x), f_B(y), \lambda)$. In particular, we will show that for δ as chosen above,

$$\mathcal{P}(U^x \otimes V^y |\psi\rangle, U \otimes V |\varphi\rangle) < 0.0065. \tag{5}$$

Since $\mathcal{P}(|\psi_0\rangle, |\psi_1\rangle) > 0.013$ for $|\psi_0\rangle \in \mathcal{S}_0^{\varepsilon,\#}$ and $|\psi_1\rangle \in \mathcal{S}_1^{\varepsilon,\#}$, the claim then follows by the triangle inequality for the purified distance. Thus, we now prove (5). Let $\Delta_A := U^x - U$, $\Delta_B := V^y - V$ and $|\Delta_S\rangle = |\psi\rangle - |\varphi\rangle$. Note that $\|\Delta_A\|_\infty \leq \delta$, $\|\Delta_B\|_\infty \leq \delta$ and $\| |\Delta_S\rangle \|_2 \leq \delta$. Indeed, using Lemma 3.10,

$$\begin{aligned}\mathcal{P}(U^x \otimes V^y |\psi\rangle, U \otimes V |\varphi\rangle) &\leq \|U^x \otimes V^y |\psi\rangle - U \otimes V |\varphi\rangle\|_2 \\ &\leq \|(U + \Delta_A) \otimes (V + \Delta_B)(|\varphi\rangle + |\Delta_S\rangle) - U \otimes V |\varphi\rangle\|_2 \\ &\leq 3\delta + 3\delta^2 + \delta^3.\end{aligned}$$

In the last line, we have used the triangle inequality together with $\|X \otimes Y |\eta\rangle\|_2 \leq \|X\|_\infty \|Y\|_\infty \| |\eta\rangle \|_2$. For $\delta = 0.00216$, we can compute $3\delta + 3\delta^2 + \delta^3 < 0.0065$ and (5) follows.

Finally, consider an (ε, l)-perfect strategy for $PV_\#^f$ and let (x, y) be such that the attackers are caught with probability at most ε^2. Without loss of generality, let (x, y) be such that $f(x, y) = 0$. Then, it holds in particular that $U^x \otimes V^y |\psi\rangle \in \mathcal{S}_0^{\varepsilon,\#}$. Thus, using (5), it follows that $g(f_A(x), f_B(y), \lambda) = f(x, y)$ on such a pair (x, y). Since there are at least l pairs (x, y) which achieve low detection probability for an (ε, l)-perfect q-qubit strategy, $g(f_A(x), f_B(y), \lambda) = f(x, y)$ on at least l pairs (x, y). Hence, g is an (ε, q)-classical rounding.

In order to conclude the proof, we must estimate k. Lemma 9.5 of [6] implies that $\mathcal{N}_A$, $\mathcal{N}_B$, $\mathcal{N}_S$ can be chosen to have cardinality at most

$$|\mathcal{N}_A| \leq (927)^{2^{2q+1}}, \qquad |\mathcal{N}_B| \leq (927)^{2^{2q+1}} \quad \text{and} \quad |\mathcal{N}_S| \leq (927)^{2^{2q+2}}.$$

Taking the logarithm, the desired bounds on the size of the classical rounding follow. □

The next statement says that if we fix a number of qubits q and an error ε, the attackers will get a large fraction of the inputs (x, y) wrong in any q-qubit strategy if we choose f to be random and if the number of qubits of the state $|\psi\rangle$ in the strategy is not too large.

Lemma 3.13. *Let $\varepsilon \in [0, 1]$ n, k, $q \in \mathbb{N}$, $n \geq 10$. Moreover, fix an (ε, q)-classical rounding g of size k with $k = \log(927) 2^{2q+2}$. Let*

$$q \leq \frac{1}{2} n - 5.$$

Then, a uniformly random $f : \{0, 1\}^{2n} \to \{0, 1\}$ fulfills the following with probability at least $1 - 2^{-2^n}$: For any $f_A : \{0, 1\}^n \to \{0, 1\}^k$, $f_B : \{0, 1\}^n \to \{0, 1\}^k$, $\lambda \in \{0, 1\}^k$, the equality $g(f_A(x), f_B(y), \lambda) = f(x, y)$ holds on less than $3/4$ of all pairs (x, y).

Proof. Let $\tilde{f}$ be as in (4). Definition 3.11 states that given $q \in \mathbb{N}$ and $\varepsilon > 0$, the number of functions $\tilde{f}$ that Alice and Bob can implement only depends on the number of choices for f_A, f_B, λ (since g is fixed given ε and q). Thus, for an (ε, q)-classical rounding of size $k \in \mathbb{N}$, they can implement $2^{(2^{n+1}+1)k}$ possible functions. Therefore, we want to estimate the probability that for a randomly chosen f, we can find f_A and f_B such that the corresponding function $\tilde{f}$ lies within

Hamming distance $1/4 \cdot 2^{2n}$ of f. Hence,

$$\begin{aligned}
&\mathbb{P}(f : \exists f_A, f_B, \lambda \text{ s.t. } d_H(M_f, M_{\tilde{f}}) \leq 2^{2n-2}) \\
&= \frac{|\{f : \exists f_A, f_B, \lambda \text{ s.t. } d_H(M_f, M_{\tilde{f}}) \leq 2^{2n-2}\}|}{|\{f : \{0,1\}^{2n} \to \{0,1\}\}|} \\
&\leq \frac{|\{f : \exists f_A, f_B, \lambda \text{ s.t. } f = \tilde{f}\}| \cdot |V(2^{2n}, 2^{2n-2})|}{|\{f : \{0,1\}^{2n} \to \{0,1\}\}|} \\
&\leq 2^{(2^{n+1}+1)k} 2^{2^{2n} h(1/4)} 2^{-2^{2n}}
\end{aligned}$$

For the first equality, we use the fact that the function f is drawn uniformly at random. For the first inequality, we estimate the numerator by considering a ball in Hamming distance around every function $\tilde{f}$ we can express by suitable f_A, f_B, λ. In the last line, we have used (2). Using $k = \log(927)2^{2q+2}$ and $q \leq n/2 - 5$, we infer that $\mathbb{P}(f : \exists f_A, f_B, \lambda \text{ s.t. } d_H(M_f, M_{\tilde{f}}) \leq 2^{2n-2})$ is strictly bounded from above by 2^{-2^n}. □

In Lemma 3.13, we have shown that for any state $|\psi\rangle$, a random function f has large Hamming distance to any $\tilde{f}$ if the dimension of the state is small enough. In particular, this means that any $(\varepsilon, 3/4 \cdot 2^{2n})$-perfect q-qubit strategy needs a number of qubits which is linear in the number of classical bits. The following proposition makes this precise.

Proposition 3.14. *Let $0 \leq \varepsilon \leq 0.41$ and $n \geq 10$, q, $n \in \mathbb{N}$. Then, a uniformly random function $f : \{0,1\}^{2n} \to \{0,1\}$ has the following property with probability at least $1 - 2^{-2^n}$: Any $(\varepsilon, \frac{3}{4} \cdot 2^{2n})$-perfect q-qubit strategy for PV^f_{route} requires*

$$q > \frac{1}{2}n - 5, \tag{6}$$

where $|\psi\rangle$ is a state on $2q+1$ qubits.

Proof. We prove the statement by contradiction. Let g be the (ε, q)-classical rounding of size k, where $k = \log(927)2^{2q+2}$, which is guaranteed to exist by Proposition 3.9 and Lemma 3.12. Assume that $q \leq \frac{1}{2}n - 5$. Pick a function $f : \{0,1\}^{2n} \to \{0,1\}$ such that for any $f_A : \{0,1\}^n \to \{0,1\}^k$, $f_B : \{0,1\}^n \to \{0,1\}^k$, $\lambda \in \{0,1\}^k$ and $\tilde{f}$ as in (4), the equality $f(x,y) = \tilde{f}(x,y)$ holds on less than $3/4$ of all pairs (x,y). By Lemma 3.13, a uniformly random f will have this property with probability at least $1 - 2^{-2^n}$.

Let $|\psi\rangle$ be a state on $2q+1$ qubits and assume that there is a $(\varepsilon, \frac{3}{4} \cdot 2^{2n})$-perfect q-qubit strategy for PV^f_{route}. Then, the corresponding f_A, f_B, λ satisfy $g(f_A(x), f_B(y), \lambda) = f(x,y)$ on at least $\frac{3}{4} \cdot 2^{2n}$ pairs (x,y). However, this is a contradiction to the choice of f. □

Finally, we can rephrase the previous theorem as a statement about the probability that Alice and Bob are caught by the verifiers.

Theorem 3.15. *Let $f : \{0,1\}^{2n} \to \{0,1\}$, $n \geq 10$ and let*

$$q \leq \frac{1}{2}n - 5.$$

Let us assume that the verifiers choose a function f uniformly at random before the protocol and that they choose x, y uniformly at random during the protocol. Moreover, let us assume that Alice and Bob control at most q qubits each at the beginning of the protocol. Then, the attackers are caught during PV^f_{route} with probability at least $4 \cdot 10^{-2}$.

Proof. Let $0 < \varepsilon \leq 0.41$. By Proposition 3.14, with probability at least $1 - 2^{-2^n}$ the function f is such that there are no $(\varepsilon, 3/4 \cdot 2^{2n})$-perfect q-qubit strategies for PV^f_{route}. That means that for any strategy Alice and Bob can implement with their state, on a fraction at least $1/4$ of the possible bit

strings (x, y), the final reduced state will be at least ε away in purified distance from the maximally entangled state.

That means, that the measurement $\{|\Omega\rangle\langle\Omega|, I_4 - |\Omega\rangle\langle\Omega|\}$ on such an input pair catches them cheating with probability at least

$$1 - F(\rho, |\Omega\rangle\langle\Omega|)^2 > \varepsilon^2.$$

Multiplying all these probabilities and using that $n \geq 10$, we obtain the bound in the assertion. □

In order to increase the probability with which the attackers are caught, it is possible to repeat the protocol sequentially several times, as the following proposition shows. It is important to note that Alice and Bob are not allowed to go to z, the position of the honest prover, during the time the protocol runs. The implicit assumption of PBC is that z is in some secure zone like a bank, for example, that attackers do not have access to.

Proposition 3.16. *Let $f : \{0,1\}^{2n} \to \{0,1\}$, $n \geq 10$, r, q, $n \in \mathbb{N}$ and let*

$$q \leq \frac{1}{2}n - 5.$$

Let us assume that the verifiers choose a function f uniformly at random at the beginning and that they run the protocol PV^f_{route} sequentially r-times, choosing x, y uniformly at random each time. Moreover, let us assume that Alice and Bob control at most q qubits each at the beginning of each iteration of PV^f_{route} . Then, the attackers are caught with probability at least $1 - 0.96^r$.

Proof. Let $X_i \in \{0,1\}$ be random variables where $i \in \{1, \ldots, r\}$. We set $X_i = 0$ if the attackers are detected in round i and $X_i = 1$ if they are not detected. First, we observe that Theorem 3.15 still holds if the state Alice and Bob share is mixed, because it is equivalent to a random mixture of pure states. Thus, the strategy for a mixed state is a random mixture of strategies for pure states. By Remark 3.3, shared randomness does not increase the probability of the attackers to avoid detecion. Between the repetitions, it can happen that the state Alice and Bob share depends on previous iterations. However, the qubit of the maximally entangled pair at the beginning of each round is uncorrelated with that state and the pair (x, y) in each round does not depend on previous rounds. Moreover, the attackers are assumed to control at most q qubits at the beginning of each round. Thus, the probability that the attackers are not detected is at most $\mathbb{P}(X_i = 1 | X_{i-1} = x_{i-1}, \ldots, X_1 = x_1) \leq 0.96$ for any $i \in \{1, \ldots, r\}$ by Theorem 3.15. Since there are r rounds, the probability to escape detection in all rounds is

$$\mathbb{P}(X_r = 1, \ldots, X_1 = 1) = \prod_{i=1}^{r} \mathbb{P}(X_i = 1 | X_{i-1} = 1, \ldots, X_1 = 1) \leq 0.96^r.$$

This proves the assertion. □

3.3. **The qubit routing protocol.** For ease of analysis, we have been considering a protocol where the verifiers hold a reference qubit, but an almost-equivalent protocol exists where the verifiers only need to store classical information. This is the qubit routing protocol considered in the main text. Let $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$. As known from the context of the BB84 protocol [7], we can replace the final measurement $\{|\Omega\rangle\langle\Omega|, I_4 - |\Omega\rangle\langle\Omega|\}$ by the following measurement: With probability $\frac{1}{2}$ each, measure either $\{|{+}{+}\rangle\langle{+}{+}| + |{-}{-}\rangle\langle{-}{-}|, I_4 - (|{+}{+}\rangle\langle{+}{+}| + |{-}{-}\rangle\langle{-}{-}|)\}$ or $\{|00\rangle\langle 00| + |11\rangle\langle 11|, I_4 - (|00\rangle\langle 00| + |11\rangle\langle 11|)\}$. We denote this measurement by M2. Let us therefore consider the slightly altered protocol $\widetilde{PV}^f_{\text{route}}$:

(1) V_0 chooses an n-bit string x uniformly at random, V_1 chooses an n-bit string y uniformly at random. Moreover, V_0 prepares one of the states $|0\rangle$, $|1\rangle$, $|+\rangle$, $|-\rangle$ uniformly at random. Let this state be the qubit Q.

(2) V_0 sends qubit Q together with x to P. V_1 sends y such that it arrives at z at the same time as x and Q sent by V_0.
(3) P sends the qubit Q on to $V_{f(x,y)}$.
(4) If Q was $|0\rangle$ or $|1\rangle$ at step (1), $V_{f(x,y)}$ measures Q in the computational basis. Otherwise, $V_{f(x,y)}$ measures Q in the Hadamard basis. V_0 and V_1 accept if the qubit arrives at the correct time at the correct verifier and if the measurement returns the outcome consistent with the state of Q at step (1).

The following proposition implies that PV^f_{route} and $\widetilde{PV}^f_{\text{route}}$ are essentially equivalent.

Proposition 3.17. *Let $p \in [0,1]$. If the attackers are caught with probability at least p during PV^f_{route}, then the attackers are caught with probability at least $\frac{1}{2}p$ during $\widetilde{PV}^f_{\text{route}}$. Conversely, if the attackers are caught with probability at least p during $\widetilde{PV}^f_{\text{route}}$, they are caught with probability at least p during PV^f_{route}.*

Proof. We begin by noting that preparing $|0\rangle$, $|1\rangle$, $|+\rangle$, $|-\rangle$ with equal probability is equivalent to preparing $|\Omega\rangle$ and measuring one qubit R with probability $1/2$ in the computational basis and with probability $1/2$ in the Hadamard basis. Moreover, the other qubit Q is measured in the same basis in step (4) and any action on R during PV^f_{route} commutes with all the operations the honest prover or the attackers can do. Thus, $\widetilde{PV}^f_{\text{route}}$ is equivalent to PV^f_{route} except for the final measurement, which is M2 instead of $\{|\Omega\rangle\langle\Omega|, I_4 - |\Omega\rangle\langle\Omega|\}$. Hence, the assertions follow from Proposition 8.2. □

In particular, $\widetilde{PV}^f_{\text{route}}$ is secure for an adequate function f since PV^f_{route} is. Indeed, the following corollary is an immediate consequence of Propositions 3.16 and 3.17:

Corollary 3.18. *Let $f : \{0,1\}^{2n} \to \{0,1\}$, $n \geq 10$, r, q, $n \in \mathbb{N}$ and let*

$$q \leq \frac{1}{2}n - 5.$$

Let us assume that the verifiers choose a function f uniformly at random at the beginning and that they run the protocol $\widetilde{PV}^f_{\text{route}}$ sequentially r-times, choosing x, y uniformly at random each time. Moreover, let us assume that Alice and Bob control at most q qubits each at the beginning of each iteration of $\widetilde{PV}^f_{\text{route}}$. Then, the attackers are caught with probability at least $1 - 0.98^r$.

4. The measuring protocol

4.1. **The measuring protocol.** In this section, we consider the measuring protocol. It turns out that we can prove similar security guarantees as for the qubit routing protocol, using essentially the same proof techniques. For the sake of analysis, we consider again a modified protocol in which V_0 sends half of an EPR pair, and measures the other half in the correct basis at the end of the protocol. In this case, the modified protocol is completely equivalent to the original and we will refer to both as PV^f_{meas}. For a Boolean function f on $2n$ classical bits, $n \in \mathbb{N}$, the modified protocol is defined as follows:

(1) V_0 randomly chooses two n-bit strings x, y, computes $f(x,y)$ and sends y on to V_1. Moreover, V_0 prepares a maximally entangled 2-qubit state $|\Omega\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$.
(2) V_0 sends one qubit Q of $|\Omega\rangle$ together with x to P. V_1 sends y such that it arrives at z at the same time as x and Q sent by V_0.
(3) P measures Q in the computational basis if $f(x,y) = 0$ and in the Hadamard basis if $f(x,y) = 1$. Subsequently, P broadcasts the measurement outcome $b \in \{0,1\}$ to both V_0 and V_1.

(4) V_0 and V_1 accept if the classical bit b arrives at the correct time and if a measurement on their qubit (in the computational basis if $f(x,y) = 0$ and in the Hadamard basis if $f(x,y) = 1$) yields the outcome b.

The timing of the response from P is deemed correct if it is compatible with bit b originating from z right after the qubit Q reached that point. An illustration of the protocol can be found in Figure 3 in the main text.

Now, we define attack strategies for the protocol.

Definition 4.1 (q-qubit strategy for PV^f_{meas})**.** *Fix a partition into systems $RA\tilde{A}A_cB\tilde{B}B_c$. Both Alice's and Bob's registers each consist of q qubits. Let d be the combined dimension of this system, therefore $d = 2^{2q+1}$. A* q-qubit strategy for PV^f_{meas} *consists of the starting state $|\psi\rangle$ on $RA\tilde{A}A_cB\tilde{B}B_c$, unitaries $U^x_{A\tilde{A}A_c}$, $V^y_{B\tilde{B}B_c}$, and Alice's and Bob's local two-outcome POVMs $\{\Pi^{xy}_{A\tilde{A}B_c}, I - \Pi^{xy}_{A\tilde{A}B_c}\}$ and $\{\Sigma^{xy}_{B\tilde{B}A_c}, I - \Sigma^{xy}_{B\tilde{B}A_c}\}$, for all x, $y \in \{0,1\}^n$. The superscripts indicate whether the operators may depend on the message x that V_0 sends, the message y that V_1 sends, or on both messages.*

See Figure 4 in the main text for an illustration. We interpret the strategy as follows: First Alice applies U as a function of x and Bob applies V as function of y. Then, Alice and Bob exchange registers A_c and B_c. Finally, Alice (with full knowledge of both x and y) measures her local registers using a POVM given by $\{\Pi, I - \Pi\}$, responding to V_0 with her outcome. Similarly, Bob measures his local registers using $\{\Sigma, I - \Sigma\}$ to determine his response to V_1. Any unitary on Alice's or Bob's side after the communication phase, which may depend on x, y, can be absorbed into the POVMs. The same holds for any classical post-processing. The definition of an (ε, l)-perfect q-qubit strategy is the same as for PV^f_{route} :

Definition 4.2 ((ε,l)-perfect q-qubit strategy for PV^f_{meas})**.** *Let $\varepsilon > 0$, $l \in \mathbb{N}$. A q-qubit strategy for PV^f_{meas} as in Definition 3.1 is* (ε,l)-perfect *if on l pairs of strings (x,y), Alice and Bob are caught by the verifiers with probability at most ε^2.*

4.2. **Lower bounds.** Our main task is to find a proposition which plays the role of Proposition 3.9. We will use entropic uncertainty relations to achieve this task.

Buhrman et al. [8] used an entropic uncertainty principle called the *strong complementary information tradeoff* (CIT) from [9, 10] to bound the attack probability on the basic BB84 quantum PV scheme against unentangled attackers. The following version is also used in [8, Theorem 2.4], where we have relabeled registers and instantiated with $n = 1$.

Theorem 4.3 (CIT)**.** *Let $|\psi_{REF}\rangle \in \mathcal{H}_R \otimes \mathcal{H}_E \otimes \mathcal{H}_F$ be an arbitrary tri-partite state, where $\mathcal{H}_R = \mathbb{C}^2$. Let the hybrid state ρ_{ZEF} be obtained by measuring R in basis $\theta \in \{0,1\}$, and let the hybrid state σ_{ZEF} be obtained by measuring R (of the original state $|\psi_{REF}\rangle$) in the complementary basis $\bar{\theta}$. Then, using conditional quantum entropy,*

$$H(\rho_{ZE}|E) + H(\sigma_{ZF}|F) \geq 1.$$

We start by defining sets of states from which, if they arise after the communication phase, the attackers can successfully attack the protocol.

Definition 4.4. *Let $\varepsilon \in [0,1]$. We define $\mathcal{S}_0^{\varepsilon,\text{meas}}$ as the set of states $|\varphi\rangle_{RA\tilde{A}A_cB\tilde{B}B_c}$ such that there exists a measurement on $A\tilde{A}B_c$ and a measurement on $B\tilde{B}A_c$ which each allow to guess the outcome of a measurement performed on R in the computational basis with probability at least $1 - \varepsilon^2$. Moreover, we define $\mathcal{S}_1^{\varepsilon,\text{meas}}$ as the set of states $|\varphi'\rangle_{RA\tilde{A}A_cB\tilde{B}B_c}$ such that there exists a measurement on $A\tilde{A}B_c$ and a measurement on $B\tilde{B}A_c$ which each allow to guess the outcome of a measurement performed on R in the Hadamard basis with probability at least $1 - \varepsilon^2$.*

Now, note that having a successful attack on PV^f_{meas} for some x, y implies that the corresponding entropy is low:

Lemma 4.5. *Let $\varepsilon \in [0,1]$ and let $\delta = h(\varepsilon^2)$. Let $|\varphi_0\rangle_{RA\tilde{A}A_cB\tilde{B}B_c} \in \mathcal{S}_0^{\varepsilon,\text{meas}}$ and $|\varphi_1\rangle_{RA\tilde{A}A_cB\tilde{B}B_c} \in \mathcal{S}_1^{\varepsilon,\text{meas}}$. Moreover, let $\rho_{ZA\tilde{A}A_cB\tilde{B}B_c}$ be the state that results after measuring register R of $|\varphi_0\rangle$ in the computational basis and $\sigma_{ZA\tilde{A}A_cB\tilde{B}B_c}$ the state that results after measuring register R of $|\varphi_1\rangle$ in the Hadamard basis. Then $H(\rho_{ZA\tilde{A}B_c}|A\tilde{A}B_c) \leq \delta$ and $H(\rho_{ZB\tilde{B}A_c}|B\tilde{B}A_c) \leq \delta$. Likewise, we find that $H(\sigma_{ZA\tilde{A}B_c}|A\tilde{A}B_c) \leq \delta$ and $H(\sigma_{ZB\tilde{B}A_c}|B\tilde{B}A_c) \leq \delta$.*

Proof. First consider $|\varphi_0\rangle$ and Alice's registers $A\tilde{A}B_c$. Abusing notation slightly, we denote by Z the random variable obtained by measuring register R of $|\varphi_0\rangle$ in the computational basis, thus transforming $|\varphi_0\rangle$ into $\rho_{ZA\tilde{A}A_cB\tilde{B}B_c}$. Let W be the random variable denoting Alice's outcome of the POVM measurement on local registers $A\tilde{A}B_c$ which allows to guess Z. This measurement is guaranteed to exist from the definition of $\mathcal{S}_0^{\varepsilon,\text{meas}}$ in Definition 4.4. It transforms $\rho_{ZA\tilde{A}A_cB\tilde{B}B_c}$ into $\rho_{ZWA_cB\tilde{B}}$. For a probability of error $\mathbb{P}(Z \neq W) \leq \varepsilon^2$, by Fano's inequality it holds that $H(Z|W) \leq h(\varepsilon^2)$. Since we have that $H(\rho_{ZA\tilde{A}B_c}|A\tilde{A}B_c) \leq H(Z|W)$ by the data processing inequality for the relative entropy, applied to the mutual information, the statement follows directly. The other three cases can be shown analogously. □

To proceed, we need to recall the continuity of conditional quantum entropy:

Proposition 4.6. *Let R be such that $\dim R = 2$, and let the dimensions of the systems E, F be arbitrary. If $\mathcal{P}(\rho_{REF}, \sigma_{REF}) \leq 0.013$, then $|H(\rho_{RE}|E) - H(\sigma_{RE}|E)| \leq 0.127$.*

Proof. Let $\Delta = 0.013$. The purified distance is an upper bound on the trace distance $\frac{1}{2}\|\cdot\|_1$, see e.g. [4, Lemma 3.17]. Thus, $\frac{1}{2}\|\rho_{RE} - \sigma_{RE}\|_1 \leq \Delta$, where we have used data-processing for the trace distance. The assertion follows then from the Alicki-Fannes-Winter inequality [11, Lemma 2], which yields

$$|H(\rho_{RE}|E) - H(\sigma_{RE}|E)| \leq 2\Delta + (1+\Delta)\, h\left(\frac{1}{1+\Delta}\right).$$

The assertion follows inserting the numerical value for Δ. □

To show security, we can follow a similar strategy as for the entangled routing protocol. A key result is the following proposition:

Lemma 4.7. *Let $\delta \leq h[(0.3)^2]$. Moreover, let $|\varphi_0\rangle_{RA\tilde{A}A_cB\tilde{B}B_c}$ be such that $H(\rho^0_{ZA\tilde{A}B_c}|A\tilde{A}B_c) \leq \delta$, where $\rho^0_{ZA\tilde{A}A_cB\tilde{B}B_c}$ is the state resulting from measuring the R register of $|\varphi_0\rangle$ in the computational basis. Similarly, let $|\varphi_1\rangle_{RA\tilde{A}A_cB\tilde{B}B_c}$ be such that $H(\sigma^1_{ZB\tilde{B}A_c}|B\tilde{B}A_c) \leq \delta$, where $\sigma^1_{ZA\tilde{A}A_cB\tilde{B}A_c}$ is the state resulting from measuring the R register of $|\varphi_1\rangle$ in the Hadamard basis. Then,*

$$\mathcal{P}(|\varphi_0\rangle, |\varphi_1\rangle) > 0.013.$$

Proof. Define $\sigma^0_{ZA\tilde{A}A_cB\tilde{B}B_c}$ analogously to $\rho^0_{ZA\tilde{A}A_cB\tilde{B}B_c}$, except that the R register of $|\varphi_0\rangle$ is measured in the Hadamard basis instead of the computational basis. Then, we can fill in the CIT statement Theorem 4.3 to obtain the inequality we combine with our assumption of $H(\rho^0_{ZA\tilde{A}B_c}|A\tilde{A}B_c) \leq \delta$ to get

$$H(\sigma^0_{ZB\tilde{B}A_c}|B\tilde{B}A_c) \geq 1 - \delta\,.$$

Recall now that we assumed $H(\sigma^1_{ZB\tilde{B}A_c}|B\tilde{B}A_c) \leq \delta$ and that therefore

$$|H(\sigma^0_{ZB\tilde{B}A_c}|B\tilde{B}A_c) - H(\sigma^1_{ZB\tilde{B}A_c}|B\tilde{B}A_c)| \geq 1 - 2\delta > 0.127\,.$$

By the contrapositive of Proposition 4.6, this implies

$$\mathcal{P}(\sigma^0_{ZB\tilde{B}A_c}, \sigma^1_{ZB\tilde{B}A_c}) > 0.013\,.$$

Recall $\sigma^0_{ZB\tilde{B}A_c}$ was obtained from $|\varphi_0\rangle$ by tracing out $A\tilde{A}B_c$ and measuring R in the Hadamard basis. Similarly, $\sigma^1_{ZB\tilde{B}A_c}$ was obtained by applying precisely the same operation to $|\varphi_1\rangle$. Therefore, the lower bound on the distance $\mathcal{P}(\sigma^0_{ZB\tilde{B}A_c}, \sigma^1_{ZB\tilde{B}A_c})$ implies the same lower bound for $\mathcal{P}(|\varphi_0\rangle, |\varphi_1\rangle)$. This follows from data-processing for the fidelity (e.g. [4, Proposition 3.2]). □

Now we are ready to state our replacement for Proposition 3.9.

Proposition 4.8. *Let $0 \leq \varepsilon \leq 0.3$ and let $|\varphi_0\rangle_{RA\tilde{A}A_cB\tilde{B}B_c} \in \mathcal{S}_0^{\varepsilon,\text{meas}}$, $|\varphi_1\rangle_{RA\tilde{A}A_cB\tilde{B}B_c} \in \mathcal{S}_1^{\varepsilon,\text{meas}}$. Then,*

$$\mathcal{P}(|\varphi_0\rangle, |\varphi_1\rangle) > 0.013.$$

Proof. This follows from combining Lemma 4.5 and Lemma 4.7. □

We can now proceed to proving security of the measuring protocol.

Proposition 4.9. *Let $0 \leq \varepsilon \leq 0.3$ and $n \geq 10$, q, $n \in \mathbb{N}$. Then, a uniformly random function $f : \{0,1\}^{2n} \to \{0,1\}$ has the following property with probability at least $1 - 2^{-2^n}$: Any $(\varepsilon, \frac{3}{4} \cdot 2^{2n})$-perfect q-qubit strategy for PV^f_{meas} requires*

$$q > \frac{1}{2}n - 5,$$

where $|\psi\rangle$ is a state on $2q+1$ qubits.

Proof. The statement follows from Proposition 4.8, Lemma 3.12 and Lemma 3.13 in the same way as in the proof of Proposition 3.14. □

Theorem 4.10. *Let $f : \{0,1\}^{2n} \to \{0,1\}$, $n \geq 10$, r, q, $n \in \mathbb{N}$ and let*

$$q \leq \frac{1}{2}n - 5.$$

Let us assume that the verifiers choose a function f uniformly at random at the beginning and that they run the protocol PV^f_{meas} sequentially r-times, choosing x, y uniformly at random each time. Moreover, let us assume that Alice and Bob control at most q qubits each at the beginning of each iteration of PV^f_{meas} . Then, the attackers are caught with probability at least $1 - 0.98^r$.

Proof. The assertion follows from Proposition 4.9 along the lines of the proofs of Theorem 3.15 and Proposition 3.16. □

Remark 4.11. *Note that PV^f_{meas} would still be secure if we replaced the requirement that the honest prover needs to send the bit b to both verifiers at the end of the protocol by requiring that P sends b only to $V_{f(x,y)}$. This can be seen from the proof of Lemma 4.7.*

5. Resistance to noise

Finally, we consider the effect of noise on $\widetilde{PV}^f_{\text{route}}$ and PV^f_{meas} . Let us now assume that the noise in the experiment causes the honest prover to be rejected with probability at most η. In order to deal with the noise, the verifiers will repeat the protocol independently r-times and accept if the individual rounds accept more than $0.996(1-\eta)r$ times. We will call such protocols $PV^f_{\text{noisy},\#}(r)$ *with noise level* η, where $\# \in \{\text{route}, \text{meas}\}$. The next theorem shows that such protocols are still secure.

Theorem 5.1. *Let r, q, $n \in \mathbb{N}$, $n \geq 10$, $0 \leq \eta \leq 10^{-2}$. Assume that a function $f : \{0,1\}^{2n} \to \{0,1\}$ is chosen uniformly at random. Then, an honest prover succeeds in $PV^f_{\text{noisy},\#}(r)$ with noise level η with probability at least*

$$1 - c^r,$$

where $\# \in \{\text{route}, \text{meas}\}$. Attackers controlling at most $q \leq \frac{1}{2}n - 5$ qubits each round will succeed with probability at most

$$c'^r,$$

where $c, c' < 1$ are universal constants. In particular, we can choose $c = c' = \exp\left(-8 \cdot 10^{-6}\right)$.

Proof. Let X_i be random variables which are 1 if the honest prover succeeds at round i and 0 if she fails. Let $X := \sum_{i=1}^r X_i$. Then, the probability that the honest prover succeeds at $PV^f_{\text{noisy},\#}(r)$ with noise level η is $\mathbb{P}[X > 0.996(1-\eta)r]$. Since in each round the honest prover accepts with probability at least 0.99, this implies that $\mathbb{P}(X_i = 1 | X_{i-1} = x_{i-1}, \dots, X_1 = x_1) \geq 0.99$ for any $x_j \in \{0,1\}$, $j \in \{1, \dots, r\}$.

Let X'_i be i.i.d. random variables which are 1 with probability 0.99 and 0 with probability 10^{-2}. Let $X' = \sum_{i=1}^r X'_i$. Then, the probability that the honest prover succeeds can be bounded using the random variable X' as $\mathbb{P}[X' > 0.996(1-\eta)r]$ by Lemma 8.3. We estimate

$$\begin{aligned}\mathbb{P}[X' > 0.996(1-\eta)r] &= 1 - \mathbb{P}[X' \leq 0.996(1-\eta)r] \\ &\geq 1 - e^{-\frac{r(1-\eta)16\cdot 10^{-6}}{2}},\end{aligned}$$

where we have used the Chernoff bound. Inserting the bound on η, the first assertion follows.

Likewise, let Y_i be a random variable which is 1 if the attackers succeed in round i and 0 if they do not, $i \in \{1, \dots r\}$. Let $Y = \sum_{i=1}^r Y_i$. Then, the probability that the attackers succeed is $\mathbb{P}[Y > 0.996(1-\eta)r]$. Using the same argument as in Proposition 3.16, Corollary 3.18 and Theorem 4.10, respectively, yields that $\mathbb{P}(Y_i = 1 | Y_{i-1} = y_{i-1}, \dots, Y_1 = y_1) \leq 0.98$ for any $y_j \in \{0,1\}$, $j \in \{1, \dots, r\}$.

Moreover, let Y'_i be i.i.d. random variables which are 1 with probability 0.98 and 0 with probability $2 \cdot 10^{-2}$. Additionally, let $Y' = \sum_{i=1}^r Y'_i$. Then, by Lemma 8.3, the probability that the attackers succeed is at most $\mathbb{P}[Y' > 0.996(1-\eta)r]$. Let us define $\eta' = 2 \cdot 10^{-2}$. Solving the equation

$$0.996(1-\eta) = (1+\delta')(1-\eta')$$

for δ', we obtain $\delta' \geq 0.99 \cdot \frac{996}{980} - 1 > 0$. We make a similar estimate as before,

$$\begin{aligned}\mathbb{P}[Y' > 0.996(1-\eta)r] &= \mathbb{P}[Y' > (1+\delta')(1-\eta')r] \\ &\leq e^{-\frac{r(1-\eta')(\delta')^2}{3}},\end{aligned}$$

where we have used the Chernoff bound again. Inserting the expressions for η' and bounding $\delta' \geq 5 \cdot 10^{-3}$, the second assertion follows. □

6. Lower bound for concrete functions

In this section, we will finally consider concrete functions f instead of uniformly random ones and prove that $\widetilde{PV}^f_{\text{route}}$ and PV^f_{meas} are still secure against bounded attackers, although the bounds are weaker than for random functions f. The proofs use a connection of classical roundings to the communication complexity of f.

Let us fix some function $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$. We define $D^{1,\mu}_\varepsilon(f)$ as the *one-way distributional communication complexity* of the function f under some distribution μ (see [12, Definition 3.19] for the (two-way) distributional communication complexity $D^\mu_\varepsilon(f)$). This represents the amount of (classical) bits Alice needs to send for Bob in a deterministic protocol, for Bob to be

able to compute $f(x,y)$ correctly with probability $1-\varepsilon$, where the probability is taken over (x,y) pairs drawn from the input distribution μ.

Similarly, let $D_\varepsilon^{\|,\mu}(f)$ be the distributional communication complexity of a function f in the simultaneous message passing (SMP) model. Here Alice and Bob both are allowed to send a single message to a third party, the referee, who has to output the function value given these messages. We take the distributional communication complexity in the SMP model as the length of the longest message, not the sum of the length of both messages. Several lower bounds for this model are given for $D_\varepsilon^{1,\mu}(f)$, but it's easy to see that $D_\varepsilon^{\|,\mu}(f) \geq D_\varepsilon^{1,\mu}(f)$.

In the other lower bounds of this work, we have restricted our analysis to the case of a uniform distribution over the input pairs. The following analysis holds for any input distribution, but for simplicity we will only consider the uniform distribution. Let u denote the uniform distribution over all pairs of n-bit strings (where n will be clear from context). For example, for the inner product function (1), we have that $D_{1/2-\varepsilon}^{1,u}(IP) \geq n/2 - \log(1/\varepsilon) - 1$ [12, Example 3.29], since $D_\varepsilon^{\mu}(f) \leq D_\varepsilon^{1,\mu}(f) + 1$.

By using the classical roundings developed for PV_{route}^f and PV_{meas}^f, we can show that for a wide range of explicit functions, the attackers need to manipulate a number of qubits that is logarithmic in the number of bits n. This bound is exponentially worse than the one we obtain for *random* functions f, but already holds for explicitly-defined easily-computable functions, such as the inner-product function[1]. So, for an explicit easily-computable function, the ratio of entanglement that attackers need also grows unboundedly with the classical information involved (but with a worse bound on the dependence of the number of classical bits n than we obtained for a random function), while the honest parties only need to manipulate a single qubit.

This can be viewed as a robust version of [2, Theorem E.3], which showed that perfect attacks on any injective function (which were effectively functions with maximal deterministic one-way communication complexity) need at least $\Omega(\log(n))$ qubits.

Proposition 6.1. *Let $\varepsilon \leq \varepsilon_0$, where ε_0 is chosen according to the requirements of Lemma 3.12. Moreover, let f be such that $D_{1/4}^{\|,u}(f) \geq k$, where u is the uniform distribution. Then there exists no $(\varepsilon, \frac{3}{4}\cdot 2^{2n})$-perfect q-qubit strategy for either PV_{meas}^f or PV_{route}^f, with*

$$\log(927)2^{2q+2} < k\,,$$

implying no such strategy exists for

$$q \leq \frac{1}{2}\log k - 3\,.$$

Proof. We prove the statement by contradiction: Any assumed strategy on a low number of qubits can be directly converted into a classical communication protocol for solving f. The only required observation is that the classical compression of f that we get as a result of Lemma 3.12 not only encodes a full description of the function f, but its parts can also be evaluated on specific x and y to get a communication protocol for f in the required simple form.

Assume, for a contradiction, that a $(\varepsilon, \frac{3}{4}\cdot 2^{2n})$-perfect q-qubit strategy exists. Then, Lemma 3.12 implies the existence of an (ε, q)-classical rounding of f of size $k = \log(927)2^{2q+2}$. Therefore, by Definition 3.11, there exists a function $g : \{0,1\}^{3k} \to \{0,1\}$, functions $f_A : \{0,1\}^n \to \{0,1\}^k$, $f_B : \{0,1\}^n \to \{0,1\}^k$, and a constant $\lambda \in \{0,1\}^k$ such that for at least $\frac{3}{4}$ of the input pairs (x,y) it holds that

$$f(x,y) = g(f_A(x), f_B(y), \lambda)\,.$$

Given that the function f and the strategy are known beforehand, all parties can precompute these objects in the communication complexity setting.

[1] Recall that for the example of the inner-product function, it is not hard to construct an attack that uses n EPR pairs [2].

The simultaneous-message passing protocol now simply proceeds as follows: Alice sends the k-bit string $s = f_A(x)$ to the referee, and Bob sends the k-bit string $t = f_B(y)$. The referee computes $g(s, t, \lambda)$ and outputs this as function value. □

Theorem 6.2. *Let f be such that $D^{\|,u}_{1/4}(f) \geq k$, where u is the uniform distribution, and let*

$$q \leq \frac{1}{2}\log k - 3.$$

If x, y are chosen uniformly at random during the protocols, then attackers controlling at most q qubits each are detected during $\widetilde{PV}^f_{\text{route}}$ and PV^f_{meas} with probabilities at least $2 \cdot 10^{-2}$, respectively.

Proof. Using Proposition 6.1, this follows from Propositions 3.9 and 4.8 in a similar way as Theorem 3.15. □

For the inner product function, this implies Theorem 1.3.

Remark 6.3. *Replacing the upper bound on q by $q \leq \frac{1}{2}\log k - 3$ for f be such that $D^{\|,u}_{1/4}(f) \geq k$ in Proposition 3.16, Theorem 4.10 and Theorem 5.1, we can derive the corresponding statements on repetition and noise robustness for a concrete function by following the exact same proof strategies. Moreover, we could consider $\widetilde{PV}^f_{\text{route}}$ instead of PV^f_{route} , decreasing the detection probability by a factor $1/2$.*

7. Attack model and comparison to previous work

When analyzing protocols for quantum PV in a resource-bounded setting, care has to be taken with respect to what is counted exactly. A fair comparison will involve weighing the resources required of an honest party compared to those of the attackers, proving hopefully that any attack is much harder to perform than executing the protocol. Which resources are important, and how to weigh them, is not trivial and there are several choices to be made in how to count the resources involved. These choices include:

- Do we only count quantum information manipulated by the attackers and the honest parties, or do we also quantify classical information?
- Do we look at the size of *all* quantum resources required, or do we just want to limit the pre-shared state of the attackers?
- Do we allow quantum communication between the attackers, or do we assume this communication to be classical and subsume these messages in the entanglement by way of teleportation?
- Would it be possible to bound the resources using something else than the number of qubits, such as entanglement entropy?

How the strength of attack resource lower bounds should be interpreted, depends on which choices are made here.

In this work, we only count the quantum information: x and y are distributed amongst the attackers for free[2], and we bound the total number of qubits each of the attackers utilize.

Counting in such a way, Theorems 1.1 and 1.2 imply that the amount of *quantum* resources used by the attackers is *unbounded*[3] as a function of the quantum information manipulated by the honest party. Indeed, our bounds show that the number of qubits manipulated by the attackers grows linearly as the amount of classical information grows while the honest party only manipulates a single qubit.

[2]I.e., the first step of the attack lets Alice see x and Bob see y, and after the messages are exchanged both attackers know (x, y).

[3]This comparison is at most exponential if all information is counted, via the attack of Beigi and König [13], but has no a-priori bound if classical communication is considered free.

Note that there still is a gap between the best known attack (needing 2^n EPR pairs, for an honest protocol with n classical bits and one single qubit) and our lower bounds, when we look at how the requirement itself grows as function of n. This gap is not evident when only looking at how the respective quantum requirements relate as a function of each other.

The choices made in the attack model also influence the comparison to other results. The independent recent work by Junge, Kubicki, Palazuelos, and Pérez-García [14] uses an attack model which is very close to ours. The authors do not count classical communication either, but only compare the quantum resources needed by the honest prover compared to the attackers. In the case that the honest prover has to manipulate quantum systems with $2 \log n$ qubits, the authors can show that the attackers need a quantum system of $\Omega(n^\alpha)$ qubits for some $\alpha > 0$, provided that the attack being used is *smooth*. The classical information that the honest prover has to manipulate during this protocol is n^2 bits. The smoothness requirement covers all known attacks. Furthermore, the authors put forward a conjecture in Banach space theory that, if true, would allow to remove the smoothness assumption, and give evidence for it.

When only comparing quantum resources, our bounds are stronger in the sense that while the ratio of quantum resources in [14] is exponential, the ratio in the qubit routing and measuring protocols is unbounded. The trade-off between the classical information sent and the number of qubits needed by the attackers is similar in all cases. On the other hand, [14] establishes the link to geometric functional analysis, which could allow to tackle the ultimate goal, i.e., showing that the quantum resources the attackers need are exponential in *all* the resources the honest party needs.

When it comes to previous protocols for quantum PV, the best studied is the BB84-type protocol [1, 15, 8, 13, 16, 17], which was the inspiration for our measuring protocol. All the bounds for this protocol are linear in the sense that the honest prover needs to manipulate n qubits while the attackers need $\Omega(n)$ qubits to break the protocol. The improvement of [16] over [13] is that only single qubit measurements are necessary. Our routing protocol is simpler in the sense that the honest party only needs to route one qubit instead of measuring n qubits separately. On the other hand, the BB84-type protocol needs the honest prover to send back merely classical information, whereas the qubit routing protocol requires to send back quantum information. The measuring protocol remedies this fact while the honest prover still needs to manipulate a single qubit. However, the detection probability of attackers can be made arbitrarily large using parallel repetition of the BB84-type protocol as shown in [16], while we can increase the probability in both our protocols only through sequential repetition.

An essential difference with respect to the proof technique is that [13] showed that bounds on the success probability of the attackers without entanglement can be lifted to bound the success probability with pre-shared entanglement. In our case, this technique no longer works and we have to resort to different methods.

Finally, [18] also proves linear lower bounds for protocols based on non-local quantum computation (the BB84-type protocol is of the same type). However, the authors bound the entanglement entropy of the attackers instead of the dimension of their quantum systems. As a downside, their attack model does not allow for quantum communication of the attackers. For the BB84-type protocol, the latter restriction was removed in [16] compared to [13] (but again present in the assumptions of [17]).

8. Technical results

First, we restate and prove [2, Lemma E.1] in order to make the main argument self-contained.

Lemma 8.1. *Let $|\psi_0\rangle$, $|\psi_1\rangle$ be states on $RA\tilde{A}A_cB\tilde{B}B_c$ and such that there are unitaries $K_{A\tilde{A}B_c}$, $L_{B\tilde{B}A_c}$ and states $|\varphi_0\rangle_{\tilde{A}A_cB\tilde{B}B_c}$, $|\varphi_1\rangle_{A\tilde{A}A_c\tilde{B}B_c}$ which satisfy*

$$K_{A\tilde{A}B_c}|\psi_0\rangle_{RA\tilde{A}A_cB\tilde{B}B_c} = |\Omega\rangle_{RA} \otimes |\varphi_0\rangle_{\tilde{A}A_cB\tilde{B}B_c}$$
$$L_{B\tilde{B}A_c}|\psi_1\rangle_{RA\tilde{A}A_cB\tilde{B}B_c} = |\Omega\rangle_{RB} \otimes |\varphi_1\rangle_{A\tilde{A}A_c\tilde{B}B_c}.$$

Then,

$$|\langle\psi_0|\psi_1\rangle| \leq \frac{1}{2}.$$

Proof. Note that K and L commute. We find that

$$\begin{aligned}|\langle\psi_0|\psi_1\rangle| &= |\langle\Omega|_{RA} \otimes \langle\varphi_0| L^*K |\Omega\rangle_{RB} \otimes |\varphi_1\rangle|\\ &= |\langle\Omega|_{RA} \otimes \langle\varphi_0'| |\Omega\rangle_{RB} \otimes |\varphi_1'\rangle|,\end{aligned}$$

where $|\varphi_0'\rangle_{\tilde{A}A_cB\tilde{B}B_c} = L_{B\tilde{B}A_c}|\varphi_0\rangle_{\tilde{A}A_cB\tilde{B}B_c}$ and $|\varphi_1'\rangle_{A\tilde{A}A_c\tilde{B}B_c} = K_{A\tilde{A}B_c}|\varphi_1\rangle_{A\tilde{A}A_c\tilde{B}B_c}$. Now,

$$\langle\Omega|_{RA}|\Omega\rangle_{RB} = \frac{1}{2}(|0\rangle_B\langle 0|_A + |1\rangle_B\langle 1|_A)$$

Note that $|0\rangle_B\langle 0|_A + |1\rangle_B\langle 1|_A$ is a unitary operator from A to B, which transfers a qubit from A to B. Writing

$$|\varphi_1''\rangle_{\tilde{A}A_cB\tilde{B}B_c} = (|0\rangle_B\langle 0|_A + |1\rangle_B\langle 1|_A)|\varphi_1'\rangle_{A\tilde{A}A_c\tilde{B}B_c},$$

we infer

$$\begin{aligned}|\langle\psi_0|\psi_1\rangle| &= \frac{1}{2}|\langle\varphi_0'|\varphi_1''\rangle|\\ &\leq \frac{1}{2},\end{aligned}$$

since both $|\varphi_0'\rangle$ and $|\varphi_1''\rangle$ are states on $\tilde{A}A_cB\tilde{B}B_c$. □

We now show that the measurement of the verifiers at the end of PV^f_{route} can be replaced by a measurement implemented via local measurements and classical post-processing which performs almost as good. This implies in particular that the verifiers need not store their qubit until the end of the protocol but can measure it right away. This fact is well-known in the context of the BB84 protocol [7]. We give a proof here for convenience.

We compare the two measurement procedures for some state ρ on $\mathbb{C}^2 \otimes \mathbb{C}^2$.

- M1: Measure $\{|\Omega\rangle\langle\Omega|, I_4 - |\Omega\rangle\langle\Omega|\}$
- M2: With probability $\frac{1}{2}$ each, either measure each qubit in the computational or Hadamard basis and check whether the measurement outcomes are equal. In other words, measure either $\{|{+}{+}\rangle\langle{+}{+}| + |{-}{-}\rangle\langle{-}{-}|, I_4 - (|{+}{+}\rangle\langle{+}{+}| + |{-}{-}\rangle\langle{-}{-}|)\}$ or $\{|00\rangle\langle 00| + |11\rangle\langle 11|, I_4 - (|00\rangle\langle 00| + |11\rangle\langle 11|)\}$, where the choice is uniformly random.

The two measurements are equivalent in the following sense:

Proposition 8.2. *Let ρ be a quantum state on two qubits and let $\delta > 0$.*

1. *If M1 accepts with probability at least $1-\delta$, then M2 accepts with probability at least $1-\delta$.*
2. *If M2 accepts with probability at least $1-\delta$, then M1 accepts with probability at least $1-2\delta$.*

Proof. Let

$$|\varphi_1\rangle := \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle) \qquad \text{and} \qquad |\varphi_2\rangle := \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle).$$

It can be verified that

$$|\Omega\rangle = \frac{1}{\sqrt{2}}(|{+}{+}\rangle + |{-}{-}\rangle) \qquad \text{and} \qquad |\varphi_1\rangle = \frac{1}{\sqrt{2}}(|{+}{+}\rangle - |{-}{-}\rangle).$$

Thus, $|++\rangle\langle++| + |--\rangle\langle--| = |\Omega\rangle\langle\Omega| + |\varphi_1\rangle\langle\varphi_1|$. Likewise, $|00\rangle\langle 00| + |11\rangle\langle 11| = |\Omega\rangle\langle\Omega| + |\varphi_2\rangle\langle\varphi_2|$ and $|\Omega\rangle$, $|\varphi_1\rangle$, $|\varphi_2\rangle$ are orthogonal.

Let $p_\Omega = \langle\Omega|\,\rho\,|\Omega\rangle$ and $p_i = \langle\varphi_i|\,\rho\,|\varphi_i\rangle$ for $i \in \{1,2\}$. Then, the probability that M1 accepts is p_Ω, whereas the probability that M2 accepts is $p_\Omega + \frac{1}{2}p_1 + \frac{1}{2}p_2$. Thus, the first assertion follows straightforwardly. For the second assertion, note that

$$p_\Omega + p_1 + p_2 \leq 1,$$

since the corresponding states are orthogonal. Thus, rearranging this inequality and combining it with the assumption that M2 accepts with probability at least $1-\delta$,

$$\frac{1}{2} + \frac{1}{2}p_\Omega \geq 1 - \delta.$$

The second assertion thus follows by rearranging the inequality. □

We conclude with a small lemma concerning discrete-time stochastic processes.

Lemma 8.3. *Let $t \in \mathbb{R}$, $r \in \mathbb{N}$, $Y := \sum_{i=1}^r Y_i$, where $Y_1, \ldots, Y_r$ is a discrete-time stochastic process, and $Y_i \in \{0,1\}$. Let $p \in [0,1]$ and $Y' = \sum_{i=1}^r Y_i'$, where $Y_i' := 1$ with probability p and $Y_i' = 0$ with probability $1-p$. It holds that*

1. *If $p(Y_i = 1 | Y_{i-1} = y_{i-1}, \ldots, Y_1 = y_1) \leq p$ for all $y_j \in \{0,1\}$, $j \in \{1,\ldots,i-1\}$ and all $i \in \{1,\ldots,r\}$, then $\mathbb{P}(Y' \geq t) \geq \mathbb{P}(Y \geq t)$*
2. *If $p(Y_i = 1 | Y_{i-1} = y_{i-1}, \ldots, Y_1 = y_1) \geq p$ for all $y_j \in \{0,1\}$, $j \in \{1,\ldots,i-1\}$ and all $i \in \{1,\ldots,r\}$, then $\mathbb{P}(Y' \geq t) \leq \mathbb{P}(Y \geq t)$*

Proof. We will only show the first assertion, since the second follows in a similar manner. Let $X := \sum_{i=1}^r X_i$, where $X_1, \ldots, X_r$ is a discrete-time stochastic process and $X_i \in \{0,1\}$. Set $x_i \in \{0,1\}$, $i \in \{0,\ldots,r\}$ and $\bar{x} := (x_1, \ldots, x_r)$. Moreover, let $|\bar{x}| := x_1 + \ldots + x_r$. We write

$$\mathbb{P}(x_r, \ldots x_1) := \mathbb{P}(X_r = x_r, \ldots, X_1 = x_1)$$

and use a similar notation for conditional expectations. Fix $j \in \mathbb{N}$ and let us assume that $p(x_i | x_{i-1}, \ldots, x_1) = p(x_i)$ for any $i \geq j+1$ and that $p(1|x_{j-1}, \ldots, x_1) \leq p$. We claim that $\mathbb{P}(X \geq t) \leq \mathbb{P}(X' \geq t)$, where $X' := X_j' + \sum_{i \in \{1,\ldots,r\}\setminus\{j\}} X_i$ and X_j' is a random variable independent of $X_1, \ldots, X_r$ such that $\mathbb{P}(X_j' = 1) = p$, $\mathbb{P}(X_j' = 0) = 1-p$. We have thus replaced X_j in X by X_j' to obtain X'. The first assertion then follows from an iterated application of the claim. We now prove the claim. Let $\check{X} := \sum_{i=1, i\neq j}^r X_i$. Then,

$$\begin{aligned} \mathbb{P}(X' \geq t) &= p\mathbb{P}(\check{X} \geq t) + (1-p)\mathbb{P}(\check{X} \geq t) + p\mathbb{P}(\check{X} = t-1) \\ &= \mathbb{P}(\check{X} \geq t) + p\mathbb{P}(\check{X} = t-1), \end{aligned} \tag{7}$$

since the order of the random variables X_i for $i > j$ does not matter and we can put X_j' last. Likewise,

$$\begin{aligned} \mathbb{P}(X \geq t) &= \sum_{\bar{x}: |\bar{x}| \geq t} \mathbb{P}(x_1, \ldots, x_r) \\ &= \sum_{\bar{x}:\ |\bar{x}| - x_j \geq t} \mathbb{P}(x_1, \ldots, x_r) + \sum_{\substack{\bar{x}:\ |\bar{x}| - x_j = t-1, \\ x_j = 1}} \mathbb{P}(x_1, \ldots, x_r) \end{aligned} \tag{8}$$

For the first term, we compute

$$\begin{aligned}
&\sum_{\bar{x}:\ |\bar{x}|-x_j\geq t} \mathbb{P}(x_1,\dots,x_r)\\
&= \sum_{\bar{x}:\ |\bar{x}|-x_j\geq t} \prod_{k=j+1}^{r} \mathbb{P}(x_k) \prod_{i=1}^{j} \mathbb{P}(x_i|x_{i-1},\dots,x_1)\\
&= \sum_{\bar{x}\setminus\{x_j\}:\ |\bar{x}|-x_j\geq t} \prod_{k=j+1}^{r} \mathbb{P}(x_k)[\mathbb{P}(1|x_{j-1},\dots,x_1)+\mathbb{P}(0|x_{j-1},\dots,x_1)] \prod_{i=1}^{j-1} \mathbb{P}(x_i|x_{i-1},\dots,x_1)\\
&= \mathbb{P}(\check{X}\geq t),
\end{aligned}$$

where we have used that $\mathbb{P}(1|x_{j-1},\dots,x_1)+\mathbb{P}(0|x_{j-1},\dots,x_1)=1$. For the second term,

$$\begin{aligned}
\sum_{\substack{\bar{x}:\ |\bar{x}|-x_j=t-1,\\ x_j=1}} \mathbb{P}(x_1,\dots,x_r) &= \sum_{\substack{\bar{x}:\ |\bar{x}|-x_j=t-1,\\ x_j=1}} \prod_{k=j+1}^{r} \mathbb{P}(x_k)\mathbb{P}(1|x_{j-1},\dots,x_1) \prod_{i=1}^{j-1} \mathbb{P}(x_i|x_{i-1},\dots,x_1)\\
&\leq p\mathbb{P}(\check{X}=t-1)
\end{aligned}$$

Thus, inserting the expressions into (8) and using (7),

$$\mathbb{P}(X\geq t)\leq \mathbb{P}(\check{X}\geq t)+p\mathbb{P}(\check{X}=t-1)=\mathbb{P}(X'\geq t).$$

□

References

[1] A. Kent, W. J. Munro, and T. P. Spiller, "Quantum tagging: Authenticating location via quantum information and relativistic signaling constraints," *Physical Review A*, vol. 84, p. 012326, 2011. 9, 25

[2] H. Buhrman, S. Fehr, C. Schaffner, and F. Speelman, "The garden-hose model," in *Proceedings of the 4th Conference on Innovations in Theoretical Computer Science*, ITCS '13, pp. 145–158, ACM, 2013. 9, 11, 12, 13, 23, 25

[3] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*, vol. 16 of *North-Holland Mathematical Library*. North-Holland, 1977. 10

[4] M. Tomamichel, *Quantum Information Processing with Finite Resources*, vol. 5 of *SpringerBriefs in Mathematical Physics*. Springer, 2016. 10, 20, 21

[5] J. Watrous, *The Theory of Quantum Information*. Cambridge University Press, 2018. 12

[6] M. Ledoux and M. Talagrand, *Probability in Banach Spaces: Isoperimetry and Processes*, vol. 23 of *A Series of Modern Surveys in Mathematics Series*. Springer, 1991. 15

[7] C. H. Bennett and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing," in *Proceedings of the International Conference on Computers, Systems and Signal Processing*, vol. 175, pp. 175–179, 1984. 17, 26

[8] H. Buhrman, N. Chandran, S. Fehr, R. Gelles, V. Goyal, R. Ostrovsky, and C. Schaffner, "Position-based quantum cryptography: Impossibility and constructions," *SIAM Journal on Computing*, vol. 43, no. 1, pp. 150–178, 2014. 19, 25

[9] J. M. Renes and J.-C. Boileau, "Conjectured strong complementary information tradeoff," *Physical Review Letters*, vol. 103, p. 020402, 2009. 19

[10] M. Berta, M. Christandl, R. Colbeck, J. M. Renes, and R. Renner, "The uncertainty principle in the presence of quantum memory," *Nature Physics*, vol. 6, pp. 659–662, 2010. 19

[11] A. Winter, "Tight uniform continuity bounds for quantum entropies: Conditional entropy, relative entropy distance and energy constraints," *Communications in Mathematical Physics*, vol. 347, pp. 291–313, 2016. 20

[12] E. Kushilevitz and N. Nisan, *Communication Complexity*. Cambridge University Press, 1996. 22, 23

[13] S. Beigi and R. König, "Simplified instantaneous non-local quantum computation with applications to position-based cryptography," *New Journal of Physics*, vol. 13, no. 9, p. 093036, 2011. 24, 25

[14] M. Junge, A. M. Kubicki, C. Palazuelos, and D. Pérez-García, "Geometry of Banach spaces: a new route towards position based cryptography," *arXiv-preprint arXiv:2103.16357*, 2021. 25

[15] H.-K. Lau and H.-K. Lo, "Insecurity of position-based quantum-cryptography protocols against entanglement attacks," *Physical Review A*, vol. 83, p. 012322, 2011. 25
[16] M. Tomamichel, S. Fehr, J. Kaniewski, and S. Wehner, "A monogamy-of-entanglement game with applications to device-independent quantum cryptography," *New Journal of Physics*, vol. 15, no. 10, p. 103002, 2013. 25
[17] J. Ribeiro and F. Grosshans, "A tight lower bound for the BB84-states quantum-position-verification protocol," *arXiv-preprint arXiv:1504.07171*, 2015. 25
[18] A. Gonzales and E. Chitambar, "Bounds on instantaneous nonlocal quantum computation," *IEEE Transactions on Information Theory*, vol. 66, no. 5, pp. 2951–2963, 2019. 25

Email address: bluhm@math.ku.dk

QMATH, Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen, Denmark

Email address: christandl@math.ku.dk

QMATH, Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen, Denmark

Email address: f.speelman@uva.nl

QuSoft & Informatics Institute, University of Amsterdam, Science Park 904, Amsterdam, the Netherlands