Zero knowledge convincing protocol on quantum bit is impossible

Paweł Horodecki [1], Michał Horodecki [2], Ryszard Horodecki [1]
[1] *Faculty of Applied Physics and Mathematics, National Quantum Information Centre, Gdańsk University of Technology, 80–233 Gdańsk, Poland*
[2] *Institute of Theoretical Physics and Astrophysics, National Quantum Information Centre, Faculty of Mathematics, Physics and Informatics University of Gdańsk, 80-308 Gdańsk, Poland*

It is one of fundamental features of quantum formalism that on one hand it provides a new information processing resources and on the other hand puts fundamental constraints on the processing of quantum information implying "no-go" theorems for cloning [1–3], bit commitment [4, 5] and deleting [6] in quantum theory. Here we ask about possibility of "zero knowledge" scenario which, for its simplicity, can be considered as a quantum primitive or model scenario for any problems of similar kind. Consider two parties: Alice and Bob and suppose that Bob is given a qubit system in a quantum state ϕ, unknown to him. Alice knows ϕ and she is supposed to convince Bob that she knows ϕ sending some test message. Is it possible for her to convince Bob providing him "zero knowledge" i. e. no information about ϕ he has? We prove that there is no "zero knowledge" protocol of that kind. In fact it turns out that basing on Alice message, Bob (or third party - Eve - who can intercept the message) can synthetize a copy of the unknown qubit state ϕ with nonzero probability. This "no-go" result puts general constrains on information processing where information *about* quantum state is involved.

Consider first the most general test message from Alice. It can involve some *classical information* (some data, function encoded in classical bits) as well as purely *quantum information* represented by quantum register or, in other words, quantum system in some state. In her message she can, for example order to perform quantum computing of some problem and foresee the result or even - in general - she can order Bob to run both quantum and classical Turing machines to check some of her predictions. She must make some predictions however, as the message is supposed to *test* her knowledge. Thus, in general, Bob must perform some *measurement* to check her predictions.

The general form of test message within quantum formalism .- All the above can taken into account in the *test message* (or test in brief) sent form Alice to Bob consisting of three elements (i) *classical* prescription of some quantum operation, (ii) possibly - some ancilla in *quantum* state Alice prepared together with (iii) result of the operation. The quantum operation is supposed to act in general on both ancilla and Bob's state, but this action could be in particular trivial i. e. not affecting some of them at all. Note that all classical information (say classical bits as classical Turing machine etc.) can be included in the description of quantum operation. In fact it can be states of ancillas prepared by Bob. This is because Bob is supposed to perform finally the measurement in which his qubit as well as Alice ancilla are, in general, supposed to subject. All the results of Bob operation (even if there is more than one) predicted by Alice can be included in the output of a general quantum measurement.

The conditions for convincing test messages .- Let us now consider what it would mean to convince Bob in the above scenario or, in other words, which test message from Alice to Bob is *convincing*? It is obvious that if Alice knows the state and she wants to convince Bob about it then *the result of the operation she predicts must occur with probability one i. e. with certainty.* But it is not all: Alice could try to cheat proposing the operation which would give some result with certainty independently on her knowledge about ϕ. For example Alice could order Bob to prepare the spin-$\frac{1}{2}$ in state "up" $|\uparrow_{\hat{\mathbf{z}}}\rangle$ and predict that if he measures the spin component of the particle along $\hat{\mathbf{z}}$ then he will get result "up". Of course, Alice prediction is right with probability one but has nothing to do with her knowledge about ϕ at all. To avoid this the convincing test message should have the property that if Alice does not know ϕ then there is nonzero probability that the result she gives (as prediction) to Bob will not occur i. e. there is nonzero probability of revealing that she cheats. We can summarise the above in the following

Condition 1 .- Any convincing test message from Alice to Bob should have the property that Alice's prediction occurs with certainty if and only if she knows the state ϕ.

There is, however, a problem that Alice can send to Bob the convincing test message, but he may not be able to check whether the condition 1 is satisfied i. e. whether the result of the test is not independent on Alice knowledge about ϕ. On the other hand checking whether the

test satisfies the condition he could be forced to destroy some quantum information (which can not be cloned) about the test and will not be able to carry out the test (or any test equivalent to it). To avoid those two possibilities of that kind we shall postulate the natural condition

Condition 2 .- For any convincing test message Bob must be able to check the condition 1 for the message in such a way that he still can find out whether the test itself works.

Any test message satisfying conditions 1 and 2 we shall call *convincing*. First we shall prove that convincing messages exist. Consider the following protocol. Alice sends Bob only classical message (with no ancilla): "Please measure the spin value along the axis $\hat{\mathbf{n}}$. You will certainly get result "up". This corresponds to the state you have". (where Alice's state $|\phi\rangle$ is "up" eigenvector along $\hat{\mathbf{n}}$ axis). From the above Bob can see himself that the protocol satisfies the condition 1 as if Alice does not know the state exactly (up to some phase factor) then from her point of view it is random with some nontrivial probability distribution. So it is likely that ϕ has both "up" and "down" components nonzero along $\hat{\mathbf{n}}$ i. e. that $\phi = \alpha|\uparrow_{\hat{\mathbf{n}}}\rangle + \beta|\downarrow_{\hat{\mathbf{n}}}\rangle$ with $\beta \neq 0$. Then there is nonzero probability of result "down" contrary to what Alice predicted. The above protocol is based on "convincing" test. Obviously it is not zero knowledge protocol as the full information about $|\phi\rangle$ has been transferred from Alice to Bob.

Fully classical protocols .- Below we shall focus on the protocols, that we call fully classical ones in which there is only classical information transfer from Alice to Bob. We shall briefly prove the following observation

Observation .- To make the convincing test nontrivial i. e. not carrying all information about ϕ the transfer of quantum information (represented by ancilla) form Alice to Bob is necessary. In other words fully classical protocols are completely trivial from "zero knowledge" point of view.

As we discussed before, the whole classical part of message can be included in the classical description of some quantum operation and its result which Alice predicts. The most general quantum operation Bob can perform on the state is the generalised quantum measurement mathematically represented by completely positive map. According to general results of quantum measurement theory it can be written as follows

$$|\phi\rangle\langle\phi| \to \varrho = \sum_i V_i|\phi\rangle\langle\phi|V_i^\dagger \tag{1}$$

where $\sum_i V_i^\dagger V_i$ is equal to 2×2 identity matrix. The indices i correspond to elementary results of the measurement. The most general result predicted by Alice can be that Bob's result i_0 will belong to some subset of indices I. Now we ask about the information which is carried in Alice test message if she does not cheat. Then her prediction must occur with probability one. In quantum measurement theory it means that

$$p = Tr(\sum_{i\in I} V_i^\dagger V_i|\phi\rangle\langle\phi|) = 1 \tag{2}$$

On the other hand, the measurement theory asserts that the hermitian operator $X = \sum_{i\in I} V_i^\dagger V_i$ has eigenvalues λ satisfying $0 \leq \lambda \leq 1$. So (2) means that $|\phi\rangle$ is an eigenvector of X. We have, however, one more condition for test message to be convincing: if Alice does not know the state then there must be nonzero probability of revealing it. This means that in such case the result she predicted *should occur with probability strictly less then one*. This can be very simply expressed as

$$p' = Tr(\sum_{i\in I} V_i^\dagger V_i|\phi'\rangle\langle\phi'|) < 1 \quad \text{for any} \quad \phi' \neq \phi. \tag{3}$$

But, following the spectral property of X it means that the state $|\phi\rangle$ is the only eigenvector of X corresponding to the eigenvalue $\lambda = 1$. Bob (or Eve) does not know the state. But he is given description of V_i-s and the set of indices on the paper (or, say, computer disc). He can now calculate X, diagonalise, find the unique eigenvector corresponding to unit eigenvalue - the vector is nothing but ϕ. He can finally perform the physical measurement of the observable X on his particle to test whether the Alice message protocol works (i. e. whether she does not cheat). Let us summarise. If Alice knows ϕ and wants to convince Bob about it sending only classical information, then any test message form her must contain *full* information about the Bob's state ϕ. Bob is able to get the information and still check whether her test operation works. The crucial observation here is that given the operation and result description in two sets $\{V_i\}$ (of operators) and I (of indices) instead of following the protocol provided by Alice Bob can test Alice knowledge in much simpler way: *calculating, analysing and finally measuring the observable* $X = \sum_{i\in I} V_i V_i^\dagger$. Note, by the way, that if third party "Eve" copies the message then she also gets full information about ϕ.

Now one can ask whether quantum information transfer from Alice to Bob can reduce significantly the information content about ϕ. The answer is positive and is contained in the following protocol.

Symmetric projection convincing protocol .- As an ancilla Alice sends Bob another copy of ϕ (she can prepare it as she knows the state) and says that the joint measurement projecting the states of both original particle and the ancilla onto the symmetric subspace will certainly give positive result. It is easy to see that the above protocol satisfies both conditions 1 and 2 which any convincing test should satisfy. Still Bob can not get more information than the optimal information about unknown ϕ extracted from two copies of it. It is known that such information is far from full one, nevertheless, it is

strictly more than the information one can extract from one copy. Indeed, if one consider ϕ as being chosen randomly by some previous preparer (who was further in contact with Alice and Bob) then one can introduce the *fidelity* of estimation of ϕ [9] :

$$f \equiv \int_{\phi} d\phi |\langle\phi|\phi_{est}(\phi)\rangle|^2 \tag{4}$$

where integral is calculated over uniform distribution of all pure qubit states and $\phi_{est}(\phi)$ is the state estimated under the presence of state ϕ. It is known that optimal extraction of information from one copy (say, before Bob is given an ancilla) is $f_{1copy} = 2/3$, while in the presence of two copies we have $f_{2copies} = 3/4$. So in the above protocol (which can be called *symmetric convincing projection (SCP) protocol* still there is a nontrivial information transfer about the state ϕ from Alice to Bob: after receiving the test message from Alice Bob can learn more about it (on average 3/4 in terms of fidelity) than if he were given the state alone (resp. 2/3). If Eve intercepted the complete message (with the ancilla) she also gets nonzero knowledge about ϕ with fidelity $f_{1copy} = 2/3$ instead of $f_{0copy} = 1/2$. Note that in the case of fully classical protocol the fidelity is $f_{cl} = 1$. So the above SPC protocol is a legitimate convincing one, being much more closer to hypothetical "zero knowledge" than any fully classical protocol. But it is still not a "zero knowledge" one.

Proof of nonexistence of perfect "zero knowledge" protocol .- Here we shall prove that any protocol with test message satisfying condition 1 (and hence any convincing Alice message) has to carry nontrivial information about Bob state ϕ. There is even more than that. As we shall see basing on message satisfying condition 1 Bob (or Eve) can reproduce with some nonzero probability unknown state from the ancilla. Suppose that Alice sends Bob the ancilla in, in general, mixed state ϱ defined on the Hilbert space $\mathcal{H}_{ancilla}$ of arbitrary (may be infinite) dimension, the classical description of quantum operation $\{\tilde{V}_i\}$ (which, in general, is to be carried out on both the Bob qubit and the ancilla), and the set of indices I corresponding to the result $i_0 \in I$. The mixed state ϱ describes the ancilla which can be hydrogen atom, photon with a given state of its polarisation, molecule with the state of nuclear spin prepared etc. The Bob's qubit can be defined as a pure state of spin of spin-half particle, state of effectively two level atom and so on. So the model is completely general from the physical point of view.

After similar considerations as in the case of fully classical protocols it can be seen that condition 1 is satisfied if and only if the mean values of the following observable $A = \sum_{i\in I} \tilde{V}_i^\dagger \tilde{V}_i$ (built on the basis of Alice classical part of message) satisfy:

$$Tr(A\varrho \otimes |\phi\rangle\langle\phi|) = 1, \tag{5}$$

$$Tr(A\varrho \otimes |\phi'\rangle\langle\phi'|) < 1 \text{ for any } \phi' \neq \phi. \tag{6}$$

Note that, as before, the observable A has to have eigenvalues from the interval $[0, 1]$. So the condition (5) says that the joint state $\varrho \otimes |\phi\rangle\langle\phi|$ has eigenvectors belonging to the (may be degenerated) eigensubspace of A corresponding to eigenvalue 1. Let us denote the projector onto that subspace as P_A and the orthogonal projection as $P_A^\perp$. They both correspond to the subspaces in the full Hilbert space $\mathcal{H} = \mathcal{H}_{ancilla} \otimes \mathcal{H}_{qubit}$. First note that P_A can not span the full $\mathcal{H}$ because then A would be identity and clearly the condition (6) could not be satisfied. Hence $P_A^\perp$ is represented by *nonzero* set of its eigenvectors $\{|\Psi_k\rangle\}_{k=1}^N$ from the full space $\mathcal{H}$ where N can be, in general, infinite. From the general theory of Hilbert spaces we have $|\Psi_k\rangle = W_k \otimes I|\Psi_{singlet}\rangle$. Here we have the operators $W_k : \mathcal{H}_{qubit} \to \mathcal{H}_{ancilla}$ (which can be calculated explicitly from Schmidt decomposition of the corresponding vectors Ψ_k) and the familiar two-qubit singlet state $|\Psi_{singlet}\rangle = \frac{1}{\sqrt{2}}(|\uparrow_{\hat{\mathbf{z}}}\downarrow_{\hat{\mathbf{z}}}\rangle - |\downarrow_{\hat{\mathbf{z}}}\uparrow_{\hat{\mathbf{z}}}\rangle)$. Now a little algebra leads to the conclusion that condition (5) is equivalent to

$$Tr(W_k^\dagger \varrho W_k |\phi^\perp\rangle\langle\phi^\perp|) = 0 \tag{7}$$

for all $k = 1, ..., N$. Here $|\phi^\perp\rangle$ represents the (unique) qubit state orthogonal to $|\phi\rangle$. To derive (7) one uses the identity (see, for instance, [7]) $|\phi^\perp\rangle = |\sigma_y\phi^*\rangle$ where $\sigma_y = i|\uparrow_{\hat{z}}\rangle\langle\downarrow_{\hat{z}}| - i|\downarrow_{\hat{z}}\rangle\langle\uparrow_{\hat{z}}|$ is familiar Pauli matrix and star $*$ stands for complex conjugation. Note that we have $W_k^\dagger : \mathcal{H}_{ancilla} \to \mathcal{H}_{qubit}$. So the hermitian operator $W_k^\dagger \varrho W_k$ with positive spectrum is defined on *one qubit* Hilbert space. For (7) is equivalent to (5) at least of 2×2 matrices $\{W_k^\dagger \varrho W_k\}$ must not vanish as otherwise (6) could not be satisfied. So for at least one index k_0 we have

$$p_{k_0} \equiv Tr(W_{k_0}^\dagger \varrho W_{k_0}) > 0 \tag{8}$$

Elementary analysis leads to the conclusion that the matrix $W_{k_0}^\dagger \varrho W_{k_0}$ must represent projection onto the vector *orthogonal* to $\phi^\perp$. But it means that

$$\frac{W_{k_0}^\dagger \varrho W_{k_0}}{p_{k_0}} = |\phi\rangle\langle\phi|. \tag{9}$$

Bob does not know any of k_0, ϱ, ϕ. But he can easily compute all W_k inferring (as it was done above) from P_A calculated from the description of quantum operation which he got from Alice. Suppose now that Alice knows the state and wants to convince Bob about it. Bob first performs the measurement of A. If he got positive result (i. e. corresponding to eigenvalue 1) then he takes the ancilla (in state ϱ) and performs the quantum measurement operation corresponding to operators

$$\tilde{W}_k \equiv W_k^\dagger / \sqrt{Tr(\sum_{k=1}^{N} W_k^\dagger W_k)} \tag{10}$$

With nonzero probability (8) p_{k_0} he will can get the state $\frac{W_{k_0}^\dagger \varrho W_{k_0}}{p_{k_0}} = |\phi\rangle\langle\phi|$. So there is nonzero chance that he will synthetize the second copy of unknown qubit state ϕ. We must emphasise that once Bob gets the second copy of ϕ in his lab, he *knows* about it, as the presence of the second copy is guaranteed whenever he gets positive result of the measurement (10). Note that if Eve gets the full Alice message she have a chance to reproduce the Bob's qubit state in the same manner. So we have proved that any convincing message has to carry highly nontrivial information about ϕ. In the case of SPC protocol this information about ϕ can be made secure against Eve attack just by teleporting the second copy of ϕ from Alice to Bob.

If we assume that ϕ is supposed be completely random we can express the result in terms of fidelities of type (4) [8]. To this end consider the hidden variable η helping Alice to choose the concrete form of convincing message she sends if Bob has state ϕ. The message is indexed by η, ϕ and is chosen, in general with probability $P(\eta, \phi)$. Note that here, in particular, we allow Alice decisions to be completely random. Let $Q(\phi)$ be a uniform distribution on pure qubit states ϕ and let $p(\eta, \phi)$ represent the probability (8) which generally can (indirectly) depend on η and ϕ. Now consider first the information gain Eve can get if she intercept the message. Then she can perform the measurement (10). If she gets the copy (which can happen with nonzero probability) then she performs optimal estimation on basis of one copy which involves choice of random variable $\hat{m}$ - the axis in threedimensional space (see [8] for details). The fidelity of Eve inference can be calculated to be

$$f_{Eve} = f_{0copy} + \int d\eta d\phi d\hat{m} P(\eta,\phi) Q(\phi) p(\eta,\phi) \times$$
$$\times (|\langle\phi|\phi_{est}(\phi,\hat{m})\rangle)|^2 - f_{0copy}) = f_{0copy} + \Delta$$

The integral over random parameter $\hat{m}$ nullifies possible results of deliberated Alice actions to unable eavesdropping inference. After performing the integral we have $\Delta = \int d\eta d\phi P(\eta,\phi) Q(\phi) p(\eta,\phi)(f_{1copy} - f_{0copy}) \equiv \int d\eta d\phi P(\eta,\phi) F(\eta,\phi)$. As there always exists strictly positive p_{k_0} the function $p(\eta,\phi)$ is strictly positive on the whole probability space, hence Δ is an integral on the function F positive everywhere. Thus $\Delta > 0$ and

$$f_{Eve} > f_{0copy}. \tag{11}$$

So we have proved formally in terms of fidelities that Alice her message necessarily carries the information about Bob's state ϕ. The similar analysis can be performed to show that $f_{Bob} > f_{1copy}$.

Finally it is worth to note that Alice is supposed to convince Bob of her *classical knowledge about quantum state*. The tests of that kind can be of practical significance in future. If quantum computers eventually are constructed the question of knowledge about quantum databases content will probably be important from the point of view of data security (for example to test whether someone could have created given data which are under investigation).

On the other hand the analysis of problems of the above kind provides us new features of interrelation of classical and quantum information: as we have seen Alice sends classical text, as well as some quantum state, so the above model problem satisfies the paradigm where classical and quantum levels of information are, in general, supposed to "interact" [10]. Some implications of the present result concerning nature of quantum information will be considered elsewhere. From practical point of view it would be interesting to consider the result in context of quantum computing involving single pure quantum bit [11, 12].

Acknowledgements .- Most of this work was done when the authors visited T. J. Watson Research Centre (IBM) in Yorktown Heights. We thank Barbara Terhal and John Smolin for critical comments. Special thanks are due to Charles Bennett for inspiring discussion and David DiVincenzo for his deep remarks. The work is partially supported by Polish Committee for Scientific Research and by European Community under the grant EQUIP. The work is also supported by John Templeton Foundation.

Note added. In arXiv:1706.06963 E. Adlam and A. Kent have extended our result by providing a thorough quantitative analysis and including to relativistic setup.

[1] W. K. Wootters and W. H. Żurek, Nature **299**, 802 (1982).
[2] D. Dieks, Phys. Lett. A **92**, 271 (1982).
[3] H. Barnum, C. M. Caves, C. A. Fuchs, R. Jozsa and B. Schumacher, Phys. Rev. Lett. **76**, 2818 (1996).
[4] D. Mayers, Phys. Rev. Lett. **78**, 3414 (1997).
[5] H.-K. Lo and H. F. Chau, Phys. Rev. Lett. **78**, 3410 (1997).
[6] A. K. Pati and S. L. Braunstein, Nature **404**, 164 (2000).
[7] N. Gisin and S. Popescu, Phys. Rev. Lett. **83**, 432 (1999).
[8] S. Massar and S. Popescu, Phys. Rev. Lett. **74**, 1259 (1995).
[9] S. Popescu, Phys. Rev. Lett. **72**, 797 (1994).
[10] R. Horodecki, Phys. Lett. A **187**, 145 (1994).
[11] E. Knill and R. Laflamme, Phys. Rev. Lett. **81**, 5672 (1998).
[12] S. Parker and M. B. Plenio, Phys. Rev. Lett. **85**, 3049 (2000)