Security of counterfactual communication

L. Vaidman
Raymond and Beverly Sackler School of Physics and Astronomy
Tel-Aviv University, Tel-Aviv 69978, Israel

At 1991 Avshalom Elitzur asked me a question: "Is it possible to locate a super bomb (an object which explodes when any particle, even a photon, touches it) without destroying it?" This question led us to the interaction-free measurement (IFM) [1]. It provides (at least sometimes) a reliable information about the presence of such an object without exploding it.

A tuned Mach-Zehnder interferometer (MZI) is all that is needed for this task. Placing the object in one of its arms allows a click of a detector in the dark port of the MZI, which could not happen without the object. Penrose [2] coined the term "counterfactual" for such a process since no particle moved towards or from the object. Given the detection at the dark port, any nondemolition measurement of the presence of the particle near the object must show nothing. Thus, a person who places (or does not place) the opaque object (which needs not to be a bomb) in the arm of the interferometer, transfers a bit of information to another person who has access to all the other parts of the interferometer.

The counterfactual communication is *a communication without presence of a particle in the transmission channel*. We do need the transmission channel, but the communication can take place without particles passing through it. The mere possibility of the presence of the particle in this channel is what allows the counterfactual communication.

Josza [3] suggested to replace the bomb by a computer and thus invented a counterfactual computation: we could find the result of a calculation from the mere possibility of running the computer. Noh [4] proposed a protocol for counterfactual key distribution. A modification of Noh's proposal is to use two IFM devices, one representing bit 0 and another representing bit 1, see Fig. 1. Alice, at random, sends a single particle through one of the MZIs, while Bob, randomly, blocks one of the MZIs with an efficient particle detector. If Alice gets a click at one of the dark ports of the MZIs, she makes a public announcement that a bit of their secret key has been established. Indeed, the click of the detector in the dark port might happen only if Alice and Bob make the same choice. Every time Alice makes the announcement, it is a counterfactual communication: the particle could not have been in the transmission channel.

There have been several analyses of the security of Noh's proposal. A particularly interesting issue is the counterfactual attack [5, 6]. Eve performs the IFM detection of the presence of Bob's detector using Bob's mirror of the MZI, but to prevent Bob from revealing her involvement, she modifies the IFM using the Zeno effect [7]. As in the original IFM protocol, in the case of success, Bob is not aware that his detector was discovered. But while in the simple IFM there is a large probability of a failure and in 50% cases the particle is detected, in the method which uses the Zeno effect the probability of success is close to 1. This attack, however, requires multiple consecutive passing of the wave packet through the transmission channel and Bob's site, so making the time that Bob's detector is present in the MZI short provides an efficient defence. For the price of reducing the rate of the key distribution, Alice can send particles at random only in some of the runs, as in the Goldenberg-Vaidman protocol [8], improving the security even more.

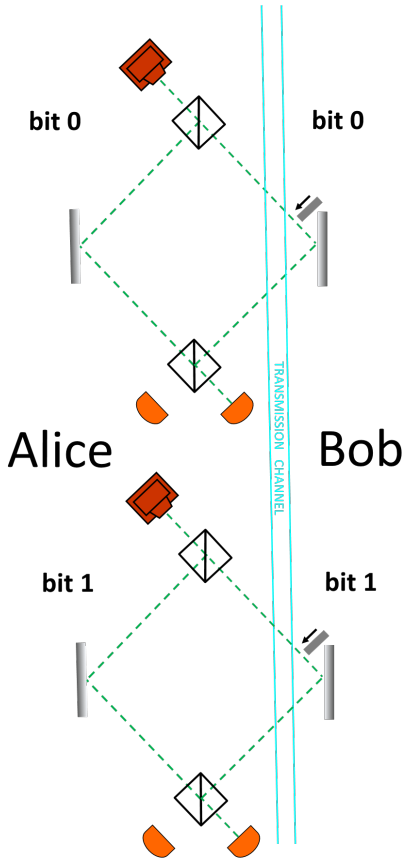

FIG. 1: A simple counterfactual key distribution protocol. Alice randomly sends a particle through one of the tuned MZIs, one of which corresponds to bit 0 and anther to bit 1. Bob randomly places a shutter in one of the interferometers. Only if it happens that they chose the same bit, the particle can be detected at the dark port of the interferometer and when this happens, the particle cannot be in the transmission channel.

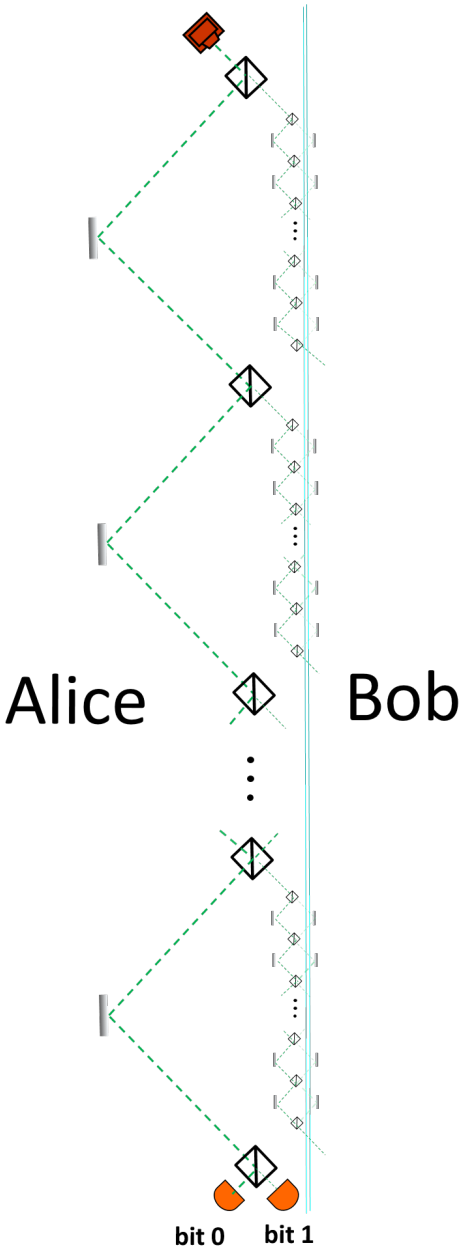

FIG. 2: "Direct counterfactual quantum communication". For bit 0 Bob leaves all inner interferometers undisturbed and then Alice's detector in the right port clicks with probability close to 1. For bit 1 Bob blocks all inner interferometers. In this case, due to destructive interference, the right detector cannot click, and the left detector clicks with probability close to 1.

What makes the counterfactual communication protocol special, is its absolute security against the class of attacks which rely just on eavesdropping to the signals [9], namely, nondemolition measurements performed on particles present in the transmission channel. Indeed, Eve can obtain no information since no particles are present in the channel. Compare this with the standard cryptographic protocol of BB84 [10]. Eve can measure the polarization of the transmitted photon and, if she is lucky with her choice of the basis, she will discover the bit of the key without Alice and Bob revealing her eavesdropping. However, even if the channel is not perfect and causes some decoherence, Eve can learn nothing in the counterfactual protocol. Indeed, if she detects the particle in the channel which Bob chooses, the particle must be detected by Bob, so this "failed" run will not lead to a creation of a bit in the key.

Scenarios in which Eve can learn something involve nonideal devices of Alice and Bob: Alice sends two particles (instead of one) and Bob's detector fails to detect the particle which comes to him. But if we assume that the only imperfect part of the device is the transmission channel and Eve (as it usually assumed) replaces it by a better channel together with an eavesdropping mechanism, then she cannot learn anything.

The original IFM is counterfactual when it successfully operates with the bomb placed in the MZI. Noh's protocol uses only these events. The IFM is clearly not counterfactual when no object is placed in the MZI. Recently, a supposedly counterfactual protocol for both cases, when the bomb is present or not, was proposed [11]. It provides a direct secret communication without prior creation of a secret key. The protocol is based on nested MZIs and employs the quantum Zeno effect [12], see Fig. 2. Alice sends one particle through one port, and it reaches one of her two detectors depending on the choice of Bob: to place shutters in all inner interferometers (bit 1), or to do nothing (bit 0). It is a high efficiency reliable direct communication protocol. The probability of a failure (the particle does not reach one of Alice's detectors) is close to zero and the probability of an error (which might happened only for bit 0) is vanishingly small too.

In this protocol, as in the original IFM, particles passing through the transmission channel cannot reach the final detector, nevertheless, I question its counterfactuality [13–15]. I argue that it is not counterfactual for bit 0 when Bob does not block paths inside the interferometer.

In order to explain how the protocol works and what is my argument against its counterfactuality, it is enough to consider a simplified version which works only sometimes and only for bit 0, see Fig. 3. The inner interferometer is tuned for destructive interference toward the second beam splitter of the external interferometer, see Fig. 3a. The external interferometer is tuned for destructive interference towards the left detector when the lower path of the inner interferometer is blocked, see Fig. 3b. This configuration provides (sometimes) a definite information about value 0 of the bit, namely the absence of the shutter. Indeed, a click in the left detector is possible only if the shutter is not present.

It seems that Alice obtains this information in a counterfactual way, since the particle cannot pass through Bob's site and reach the detector. However, this protocol is not absolutely security against an eavesdropper,

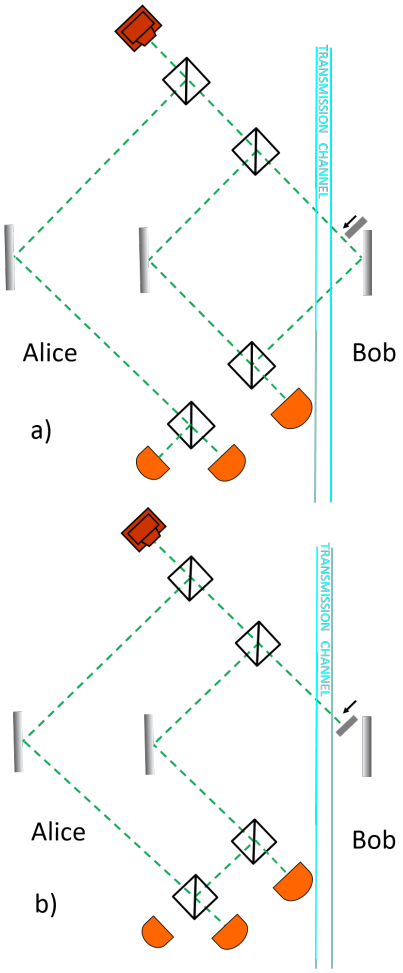

FIG. 3: Counterfactual communication of bit 0. a). The inner interferometer is tuned such that the particle cannot pass through the right arm of the external interferometer. b). There is a destructive interference towards the left detector when the right path of the inner interferometer is blocked. If the left detector clicks we know that the interferometer is not blocked (bit 0) and that the particle could not pass through the transmission channel.

and this provides an argument against its counterfactuality. Indeed, if Eve finds in a nondemolition measurement that the particle is present in the transmission channel, the left detector can click which makes Alice declare a transmission of a bit, believing that it is bit 0. It might happen only if Bob did not place a shutter inside the interferometer, so he also believes that the bit is 0. Thus, Bob transmits the bit 0 to Alice without revealing the presence of Eve.

A possible objection to the above argument is that Eve spoils the counterfactuality of the protocol and it is counterfactual when Eve is not present. If there is no test of the presence of the particle in the transmission channel, the question of counterfactuality, which is defined by this property, might not be settled. But note that this objection does not apply for the IFM and Noh's protocol: eavesdropping does not spoil the counterfactuality of these protocols.

The counterfactual direct communication protocol described in Fig. 2 exhibits the same weakness. When Bob sends bit 1 by putting shutters in all inner interferometers and Eve finds a particle in the transmission channel, the particle cannot reach Alice, so when Eve finds the particle and Alice detects it, Eve learns that the sent bit is 0. Given that Eve does not look at the last inner chain of the interferometers, the probability of an error in the Alice's reading of the bit is close to 0, so Eve discovers the transmitted bit without being revealed. The counterfactual direct communication protocol is not absolutely secure against eavesdropping.

I presented a brief cryptographic security analysis of counterfactual protocols. Their weakness is that they are vulnerable to active counterfactual attacks, but there are efficient defence methods. Their strength is the absolute security against eavesdropping. However, the absolute security is only the property of the counterfactual key distribution protocols and not of the recent proposal of "counterfactual" direct communication protocol.

[1] A. C. Elitzur and L. Vaidman, Found. Phys. **23**, 987 (1993).
[2] R. Penrose, *Shadows of the Mind*. Oxford: Oxford University Press (1994).
[3] R. Jozsa, in *Lecture Notes in Computer Science*, C. P. Williams, ed. (Springer, London, 1998), Vol. 1509, p. 103.
[4] T.-G. Noh, Phys. Rev. Lett. **103**, 230501 (2009).
[5] S. Zhang, J. Wnang, and C.J. Tang, Europhys. Lett. **98**, 30012 (2012).
[6] X. Liu, B. Zhang, J. Wang, C. Tang, J. Zhao, and S. Zhang, Phys. Rev. A **90**, 022318 (2014).
[7] P. Kwiat, H. Weinfurter, T. Herzog, A. Zeilinger, and M.A. Kasevich, Phys. Rev. Lett. **74**, 4763 (1995).
[8] L. Goldenberg and L. Vaidman, Phys. Rev. Lett. **75**, 1239 (1995).
[9] Z.-Q. Yin, H.-W. Li, W. Chen, Z.-F. Han, and G.-C. Guo, Phys. Rev. A **82**, 042335 (2010).
[10] C. H. Bennett and G. Brassard, in Proceedings of IEEE International Conference on Computers, Systems, and Signal Processing, Bangalore, India (IEEE, New York, 1984), p. 175.
[11] H. Salih, Z.H. Li, M. Al-Amri, and M.S. Zubairy, Phys. Rev. Lett. **110**, 170502 (2013).
[12] O. Hosten *et al.*, Nature (London) **439**, 949 (2006).
[13] L. Vaidman, Phys. Rev. Lett. **112**, 208901 (2014).
[14] H. Salih, Z.H. Li, M. Al-Amri, and M.S. Zubairy, Phys. Rev. Lett. **112**, 208902 (2014).
[15] L. Vaidman, arXiv:1410.2723v2 (2015).