Classical Simulation of Entanglement Swapping with Bounded Communication

Cyril Branciard, Nicolas Brunner, Harry Buhrman, Richard Cleve, Nicolas Gisin, Samuel Portmann, Denis Rosset, and Mario Szegedy

Classical simulation of entanglement swapping with bounded communication

Cyril Branciard,[1] Nicolas Brunner,[2] Harry Buhrman,[3] Richard Cleve,[4, 5]
Nicolas Gisin,[6] Samuel Portmann,[6] Denis Rosset,[6] and Mario Szegedy[7]

[1]School of Mathematics and Physics, The University of Queensland, St Lucia, QLD 4072, Australia
[2]H.H. Wills Physics Laboratory, University of Bristol, Bristol, BS8 1TL, United Kingdom
[3]Centrum Wiskunde & Informatica, and University of Amsterdam,
Science Park 123, 1098 XG Amsterdam, The Netherlands
[4]Institute for Quantum Computing and School of Computer Science,
University of Waterloo, 200 University Avenue West, Waterloo, Ontario, Canada N2L 3G1
[5]Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, Ontario, Canada N2L 2Y5
[6]Group of Applied Physics, University of Geneva,
Chemin de Pinchat 22, CH-1211 Geneva 4, Switzerland
[7]Department of Computer Science, Rutgers, the State University of NJ,
110 Frelinghuysen Road, Piscataway, NJ 08854-8019 USA

Entanglement appears under two different forms in quantum theory, namely as a property of states of joint systems and as a property of measurement eigenstates in joint measurements. By combining these two aspects of entanglement, it is possible to generate nonlocality between particles that never interacted, using the protocol of entanglement swapping. We show that even in the more constraining bilocal scenario where distant sources of particles are assumed to be independent, i.e. to share no prior randomness, entanglement swapping can be simulated classically with bounded communication, using only 9 bits in total. Our result thus provides an upper bound on the nonlocality of the entanglement swapping process.

By performing suitably chosen local measurements on an entangled quantum state, distant observers can establish nonlocal correlations, as witnessed by the violation of a Bell inequality [1]. This means that quantum statistics cannot be simulated by classically correlated systems, unless some classical communication is added to the model. Although experiments give strong evidence that nature does not use classical communication to establish correlations [2], it is nevertheless interesting from a fundamental perspective to ask how much communication is required to reproduce quantum correlations. Generally referred to as classical simulation of entanglement, this provides a natural approach to the problem of quantifying quantum nonlocality.

Nonlocality is a fundamental aspect of quantum mechanics, hence quantifying it is much desirable. Besides being one of the most striking and counterintuitive feature of the theory, it is also a powerful resource, allowing for instance for the reduction of communication complexity [3], as well as for information processing in the 'device-independent' setting [4–8], where one wants to achieve an information task and prove its security without any assumption on the devices used in the protocol.

Several works [9–12] underwent the task of estimating how much communication is needed to simulate the correlations of a maximally entangled state of two qubits under all possible projective measurements. This research culminated in 2003, when Toner and Bacon [13] showed that one bit of communication is enough. Importantly this single bit of communication is not an average value, but represents the exact amount that is to be used at each round. Thus the corresponding simulation protocol is said to have bounded communication. The communication costs of other states have been explored as well [14, 15]. Notably, Regev and Toner [16] have shown that the correlations obtainable from dichotomic measurements on any bipartite entangled state can be simulated with only two bits of communication, which are proven to be necessary [17]. Note however that their protocol does not reproduce the correct marginal distributions, and that simulating more general measurements is much more costly in terms of communication [3, 10]. The simulation of multipartite entanglement also attracted some attention [18–21], and two of the authors [22] recently showed that the correlations of equatorial measurements on a tripartite GHZ state can be simulated with 3 bits of communication, thus reproducing in particular the Mermin-GHZ paradox, which is arguably the strongest demonstration of the nonlocality of this state [23, 24].

Quantum mechanics allows not only for entangled states of distant systems, but also for entangled measurements. In such a measurement the initial state is arbitrary—it could be entangled or not—but the final state is entangled, that is the eigenstates of the operator that represents such a measurement are entangled. This second aspect of entanglement is in itself independent of nonlocality—although it leads to nonlocality when combined with entangled states [25]. It demonstrates another nonclassical feature of entanglement, which is, loosely speaking, the possibility to ask two (or more) quantum systems questions about their relations without gaining

any information about the individual properties of each subsystem [26].

It is therefore a natural question to ask whether or not the classical simulation of protocols involving both entangled states and entangled measurements is possible with bounded communication, and how much communication is required. Here, we investigate this question for the scenario of entanglement swapping [27], where quantum particles that never interacted become nonlocally correlated after their twins underwent a joint measurement. We work in the scenario of *bilocality* [28, 29], in which the shared randomness we will use in the simulation protocol is assumed to originate from two independent sources (see figure 1). Since entanglement swapping can be achieved with fully uncorrelated quantum sources, even experimentally [30, 31], it is indeed natural to impose an equivalent constraint on the simulation model.

Importantly the constraint of bilocality makes the simulation of entanglement swapping a challenging problem. The goal is basically to generate singlet nonlocal correlations between two parties which are initially fully uncorrelated. At first sight this may even appear to be impossible with finite communication, in the light of a result by Massar et al. [32], showing that simulating singlet correlations requires either infinite shared randomness (between the two parties simulating the state) or infinite communication. Here however, we will show that the process of entanglement swapping can be simulated with 9 bits, by presenting an explicit protocol. Since the communication cost of the cheapest protocol can be considered as a measure of its nonlocality, our result provides an upper bound on the nonlocality of entanglement swapping.

The entanglement swapping process. We consider three distant parties, Alice (A), Bob (B) and a Referee (R). R shares two maximally entangled qubit pairs in the state $|\Psi_-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$ with A and B, respectively. The first pair is produced by a source located between A and R, the second one is produced by an independent source located between B and R, see Fig. 1 (i). Accordingly, Alice and Bob are initially uncorrelated. By performing a Bell state measurement, i.e. a two-qubit joint measurement which features four maximally entangled Bell states as eigenstates, R projects A and B's particles onto one of the Bell states. R can thus 'swap' entanglement to A and B. This protocol is essentially identical to the celebrated quantum teleportation protocol [33]; entanglement swapping is basically a *teleportation of entanglement.*

To complete the protocol, the Referee needs to communicate the result of his measurement r—by sending two bits of classical communication—to (say) Bob, who can then apply a suitable unitary local transformation to his qubit to finally share a definite Bell state with Alice, say $|\Psi_-\rangle$. Upon receiving measurement settings $\boldsymbol{x}$ and $\boldsymbol{y}$ (represented by unit vectors on the Bloch sphere) and performing the corresponding projective measurements on their respective qubit, Alice and Bob obtain binary
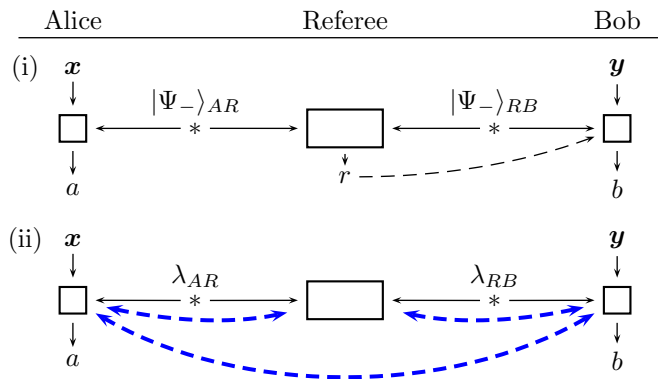

FIG. 1: (i) The scenario of entanglement swapping with two fully independent sources of $|\Psi_-\rangle$ states. The Referee sends to Bob the result r of his Bell state measurement. Upon receiving these two bits of communication, Bob can apply the adequate unitary operation to his qubit such that Alice and Bob finally share a singlet state. (ii) The classical simulation of entanglement swapping in the bilocal scenario, i.e. with two uncorrelated sources of shared random variables λ_{AR} and λ_{RB}. The three parties exchange messages symbolized by the thick dashed arrows. Here we present a simulation protocol using 9 bits of communication in total.

measurement outcomes $a = \pm 1$ and $b = \pm 1$ respectively. These outcomes exhibit nonlocal correlations of the form

$$E(\boldsymbol{x}, \boldsymbol{y}) \ = \ P(a = b|\boldsymbol{x}, \boldsymbol{y}) - P(a \neq b|\boldsymbol{x}, \boldsymbol{y}) = -\boldsymbol{x} \cdot \boldsymbol{y}, \ (1)$$

with random marginals.

A classical simulation of this protocol in the bilocal scenario would then amount to the following (see Fig. 1 (ii)): Alice and Bob receive as inputs the measurement directions $\boldsymbol{x}$ and $\boldsymbol{y}$, independently of the random variables λ_{AR} and λ_{RB} they may each share (independently) with R. After a finite amount of information exchange between the three parties, Alice and Bob produce their outputs which, after averaging over the shared random variables, must be correlated as in (1) [34].

At this point, it is instructive to recall the result of Massar *et al.* [32], which implies that correlations of the form (1) are impossible to simulate with finite communication when A and B only have *finite shared* randomness (in the sense that their shared randomness could be established with bounded communication)—even if they may each have an additional source of infinite randomness, independent for A and B. The present scenario may at first glance appear similar: initially, A and B each only have access to independent sources of *infinite* randomness, λ_{AR} and λ_{RB}, respectively, *shared with the Referee.* After any finite exchange of communication between the parties (including R), *Alice and Bob* could only share some *finite* amount of randomness. Intuition suggests that the result of [32] implies that, from such randomness, correlations of the form (1) cannot be simulated with finite communication, and entanglement swapping

is impossible to simulate in a bilocal manner [35]. However, the present scenario differs from that of [32] in that the third party, R, can share an infinite amount of randomness with A and B independently—by having access to both λ_{AR} and λ_{RB}—and can exchange some (finite number of) bits with A and B. This seemingly subtle difference dramatically changes the situation: despite the above heuristic reasoning, entanglement swapping can be simulated with finite communication in a bilocal manner, as we now show.

Simulation protocol. We start by deriving a protocol for the simulation of equatorial measurements $\boldsymbol{x} = (\cos\phi_A, \sin\phi_A, 0)$ and $\boldsymbol{y} = (\cos\phi_B, \sin\phi_B, 0)$. In this case, the correlation (1) is equivalent to

$$P(a=b|\phi_A, \phi_B) = \frac{1-\cos(\phi_A - \phi_B)}{2} = \sin\left(\frac{\phi_A - \phi_B}{2}\right)^2. \quad (2)$$

Note in particular that if $\phi_A = \phi_B$, then $P(a=b|\phi_A = \phi_B) = 0$, as expected from the anticorrelations of the singlet state.

Let us divide the equator of the Bloch sphere into $2m$ sectors of the same size, and assume that Alice and the Referee (resp. the Referee and Bob) share a random variable λ_{AR} (resp. λ_{RB}), with λ_{AR} and λ_{RB} uniformly distributed in $[0, \frac{\pi}{m}]$. Let Alice tell Bob in which sector her setting (modulo π) lies, and send him the additional bit $c_A = [(\phi_A \mod \frac{\pi}{m}) < \lambda_{AR}]$ (we use the Iverson bracket, with $[\mathcal{E}] = 1$ if the expression $\mathcal{E}$ is true, $[\mathcal{E}] = 0$ otherwise); let the Referee send to Bob the bit $c_R = [\lambda_{AR} < \lambda_{RB}]$; and let Bob calculate the bit $c_B = [\lambda_{RB} < (\phi_B \mod \frac{\pi}{m})]$.

To warm up and gain some intuition, assume that Alice and Bob's settings both lie in $[0, \frac{\pi}{m}]$, and let Alice simply output $a = +1$. If $c_A = c_R = c_B$, then Bob knows that (with probability 1) $\phi_A \neq \phi_B$; let him then also (for now) output $b = a = +1$. On the other hand, if $c_A \neq c_R$ or $c_R \neq c_B$, it could be the case that $\phi_A = \phi_B$; to ensure that $P(a = b|\phi_A = \phi_B) = 0$, Bob must output $b = -a = -1$. The probability that Alice and Bob give the same output is then (assuming $\phi_A < \phi_B$; the case $\phi_A \geq \phi_B$ being similar)

$$\begin{aligned} P(a=b|\phi_A < \phi_B) &= P(c_A = c_B = c_R|\phi_A < \phi_B) \\ &= P(\phi_A < \lambda_{AR} < \lambda_{RB} < \phi_B|\phi_A < \phi_B) \\ &= \int_{\phi_A}^{\phi_B} \frac{m}{\pi} d\lambda_{AR} \int_{\lambda_{AR}}^{\phi_B} \frac{m}{\pi} d\lambda_{RB} = \frac{m^2}{2\pi^2}(\phi_A - \phi_B)^2. (3) \end{aligned}$$

If $m \geq \frac{\pi}{\sqrt{2}}$, this probability is too large (for any $\phi_A, \phi_B \in [0, \frac{\pi}{m}]$), compared to the desired value of $\sin\left(\frac{\phi_A - \phi_B}{2}\right)^2$ (2). Now, one can actually make use of some freedom in the case $c_A = c_R = c_B$, and ask Bob to output $b = a = +1$ only with some probability $\wp$ (and to output $b = -a = -1$ with probability $1 - \wp$). By letting $\wp$ depend on ϕ_B and λ_{RB}, and take the particular form

$\wp(\phi_B - \lambda_{RB}) = \frac{\pi^2}{2m^2}\cos(\phi_B - \lambda_{RB})$, we obtain, as desired,

$$\begin{aligned} P(a=b|\phi_A, \phi_B) &= \frac{m^2}{\pi^2} \int_{\phi_A}^{\phi_B} d\lambda_{AR} \int_{\lambda_{AR}}^{\phi_B} d\lambda_{RB} \, \wp(\phi_B - \lambda_{RB}) \\ &= \sin\left(\frac{\phi_A - \phi_B}{2}\right)^2. \quad (4) \end{aligned}$$

In order now to extend this protocol to all possible equatorial measurements $\phi_A, \phi_B \in [0, 2\pi]$, one needs to find adequate strategies for Bob (i.e. adequate functions $\wp(\phi_B - \lambda_{RB}) \in [0, 1]$) for all possible sectors where Alice and Bob's settings may lie, and for all possible values of the bits c_A, c_R and c_B. We were able to find such strategies for the case $m = 4$, see Table I. This leads us to define the following protocol:

Protocol 1: *Let Alice and the Referee (resp. the Referee and Bob) share a random variable λ_{AR} (resp. λ_{RB}), with λ_{AR} and λ_{RB} uniformly distributed in $[0, \frac{\pi}{4}]$. After reception of Alice and Bob's measurement settings ϕ_A, ϕ_B, the three parties proceed as follows:*

- *Alice calculates $a = \text{sign}(\sin\phi_A) = \pm 1$, $j_A = \lfloor \frac{4}{\pi}(\phi_A \mod \pi)\rfloor \in \{0, 1, 2, 3\}$, $\phi'_A = (\phi_A - j_A\frac{\pi}{4} \mod \pi) \in [0, \frac{\pi}{4}[$ and $c_A = [\phi'_A < \lambda_{AR}] \in \{0, 1\}$. She sends j_A (2 bits) and c_A (1 bit) to Bob, and outputs a.*

- *The Referee calculates $c_R = [\lambda_{AR} < \lambda_{RB}] \in \{0, 1\}$ and sends it to Bob (1 bit).*

- *After reception of j_A, Bob calculates $\beta = \text{sign}[\sin(\phi_B - j_A\frac{\pi}{4})]$ and $\phi'_B = (\phi_B - j_A\frac{\pi}{4} \mod \pi) \in [0, \pi[$. He determines the index $j_B = \lfloor \frac{4}{\pi}\phi'_B\rfloor \in \{0, 1, 2, 3\}$ of the sector where his angle ϕ'_B lies, and the bit $c_B = [\lambda_{RB} < \phi'_B - j_B\frac{\pi}{4}] \in \{0, 1\}$. Depending on $j_B, c_A, c_R, c_B, \phi'_B$ and λ_{RB}, he outputs $b = \beta$ with probability $\wp^{j_B}_{c_A c_R c_B}(\phi'_B - \lambda_{RB})$, and $b = -\beta$ with probability $1 - \wp^{j_B}_{c_A c_R c_B}(\phi'_B - \lambda_{RB})$, for the functions $\wp^{j_B}_{c_A c_R c_B}$ defined in Table I.*

As explicitly shown in the Supplementary Material, the above protocol gives the desired probability $P(a = b|\phi_A, \phi_B) = \sin\left(\frac{\phi_A - \phi_B}{2}\right)^2$ for all possible equatorial measurements, using 4 bits of communication. It can then be extended in the following way to all measurement directions on the Bloch sphere, using a similar technique as in [10]:

Protocol 2: *For measurement directions $\boldsymbol{x} = (\sin\theta_A \cos\phi_A, \sin\theta_A \sin\phi_A, \cos\theta_A)$ and $\boldsymbol{y} = (\sin\theta_B \cos\phi_B, \sin\theta_B \sin\phi_B, \cos\theta_B)$, the three parties*

- *run Protocol 1 with input angles ϕ_A and ϕ_B; Alice and Bob obtain intermediate outputs a_0 and b_0.*

- *run Protocol 1 a second time (using a new set of independent variables $\lambda_{AR}, \lambda_{RB}$), now with input angles $a_0\theta_A$ and $-b_0\theta_B$; Alice and Bob output the outcomes a and b of this second run of Protocol 1.*

$c_A\,c_R\,c_B$	$j_B=0,\ \gamma\in[-\frac{\pi}{4},\frac{\pi}{4}]$	$j_B=1,\ \gamma\in[0,\frac{\pi}{2}]$	$j_B=2,\ \gamma\in[\frac{\pi}{4},\frac{3\pi}{4}]$	$j_B=3,\ \gamma\in[\frac{\pi}{2},\pi]$
0 0 0	$\frac{\pi^2}{32}\cos\gamma$	0	$\frac{1}{2}+\frac{\pi^2}{64}[1-(2+\sqrt{2})\sin\gamma]$	$1+\frac{\pi^2}{32}[\cos\gamma+2\cos(\gamma+\frac{\pi}{4})]$
0 0 1	0	0	$\frac{1}{2}+\frac{\pi^2}{64}[1-\sqrt{2}\sin\gamma]$	$1+\frac{\pi^2}{32}\cos\gamma$
0 1 0	0	0	$\frac{1}{2}-\frac{\pi^2}{64}[1+\sqrt{2}\sin(\gamma+\frac{\pi}{4})]$	$1+\frac{\pi^2}{32}\cos(\gamma+\frac{\pi}{4})$
0 1 1	0	$\frac{\pi^2}{32}\cos(\gamma-\frac{\pi}{4})$	$\frac{1}{2}-\frac{\pi^2}{64}[1-\sqrt{2}\sin(\gamma-\frac{\pi}{4})]$	1
1 0 0	0	0	$\frac{1}{2}+\frac{\pi^2}{64}[1-\sqrt{2}\sin(\gamma+\frac{\pi}{4})]$	$1+\frac{\pi^2}{32}\cos(\gamma+\frac{\pi}{4})$
1 0 1	0	$\frac{\pi^2}{32}\cos(\gamma-\frac{\pi}{4})$	$\frac{1}{2}+\frac{\pi^2}{64}[1+\sqrt{2}\sin(\gamma-\frac{\pi}{4})]$	1
1 1 0	0	$\frac{\pi^2}{32}\cos\gamma$	$\frac{1}{2}-\frac{\pi^2}{64}[1-\sqrt{2}\sin\gamma]$	1
1 1 1	$\frac{\pi^2}{32}\cos\gamma$	$\frac{\pi^2}{32}[\cos\gamma+2\cos(\gamma-\frac{\pi}{4})]$	$\frac{1}{2}-\frac{\pi^2}{64}[1-(2+\sqrt{2})\sin\gamma]$	1

TABLE I: Functions $\wp_{c_A c_R c_B}^{j_B}(\gamma)$, for all values of j_B, c_A, c_R and c_B, defining the probability $\wp_{c_A c_R c_B}^{j_B}(\phi'_B-\lambda_{RB})$ that Bob outputs $b=\beta$ in Protocol 1. Note that in all cases, $\wp_{c_A c_R c_B}^{j_B}(\gamma)\in[0,1]$ for all possible values of $\gamma=\phi'_B-\lambda_{RB}$ in the given interval.

This second protocol now simulates the desired correlation $E(\boldsymbol{x},\boldsymbol{y})=-\boldsymbol{x}\cdot\boldsymbol{y}$ for all possible projective measurements by Alice and Bob, with 8 bits of communication; for more details on the calculations, see Supplementary Material. Note that Protocols 1 and 2 do not simulate the correct marginals. In order to randomize the marginals, Alice can—at the very end of the protocol—generate a random bit and send it to Bob; depending on the value of this bit, they will both flip their outcomes or not. All in all, the entanglement swapping correlations can thus be simulated with 9 bits of communication.

Discussion. We thus have proved that remarkably, the entanglement swapping process can be simulated with bounded communication, even in a bilocal scenario where Alice and Bob are (as in the quantum case) completely uncorrelated before the protocol is run, and therefore do not have any prior shared randomness. Our protocol provides an upper bound on the nonlocality of entanglement swapping in terms of its communication cost. It is an open question whether fewer bits of communication are actually sufficient: it might indeed be possible to simulate equatorial measurements more efficiently than with Protocol 1, or to find a more direct protocol that does not treat separately the azimuth and zenith angles of the measurement settings, more in the spirit of the Toner-Bacon simulation protocol for singlet correlations [13].

Next, it is natural to consider the simulation of multistage entanglement swapping, which is essential for long distance quantum communication. Now, N referees (R_1, R_2, ..., R_N) are placed on a line between Alice and Bob. Two neighboring referees share a singlet state, while R_1 and R_N share singlet states with A and B, resp.; each referee performs a joint measurement, leaving at the end the particles of Alice and Bob entangled. Whereas the quantum protocol has a straightforward and nice iterative character, we were not able to find a simulation protocol with a finite amount of communication in a $(N+1)$-locality scenario [29]. Consider for

instance the case with one additional referee R_2. Analogously to our Protocol 1, assume that Alice and R_1 share the random variable λ_{AR_1}, R_1 and R_2 share $\lambda_{R_1 R_2}$, and R_2 and Bob share $\lambda_{R_2 B}$, all uniformly and independently distributed on some interval $[0,\frac{\pi}{m}]$. After some finite communication, Bob could for instance (as in our first attempt, before Eq. (3)) output $b=a=1$ if and only if $[\phi_A<\lambda_{AR_1}]=[\lambda_{AR_1}<\lambda_{R_1 R_2}]=[\lambda_{R_1 R_2}<\lambda_{R_2 B}]=[\lambda_{R_2 B}<\phi_B]$. This would result in the probability $P(a=b|\phi_A,\phi_B\in[0,\frac{\pi}{m}])=\frac{m^3}{6\pi^3}|\phi_A-\phi_B|^3$, which scales cubically with $\phi_A-\phi_B$, and is therefore too small when ϕ_A is close to ϕ_B. It is unclear how to change the cubic scaling with finite communication. The following questions remain open: can multistage entanglement swapping be simulated with finite communication? Or can one prove, that above a certain value of N, an infinite amount of communication is necessary?

---

[1] J. S. Bell, Physics **1**, 195 (1964).
[2] D. Salart, A. Baas, C. Branciard, N. Gisin, and H. Zbinden, Nature **454**, 861 (2008).
[3] H. Buhrman, R. Cleve, S. Massar, and R. de Wolf, Rev. Mod. Phys. **82**, 665 (2010).
[4] A. Ekert, Phys. Rev. Lett. **67**, 661 (1991).
[5] J. Barrett, L. Hardy, and A. Kent, Phys. Rev. Lett. **95**, 10503 (2005).
[6] A. Acín et al., Phys. Rev. Lett. **98**, 230501 (2007).
[7] S. Pironio et al., Nature **464**, 1021 (2010).
[8] R. Colbeck and A. Kent, Journal of Physics A: Mathematical and Theoretical **44**, 095305 (2011).
[9] T. Maudlin, PSA: Proceedings of the Biennial Meeting

of the Philosophy of Science Association **1992**, pp. 404 (1992).

[10] G. Brassard, R. Cleve, and A. Tapp, Phys. Rev. Lett. **83**, 1874 (1999).

[11] B. Gisin and N. Gisin, Phys. Lett. A **260**, 323 (1999).

[12] M. Steiner, Phys. Lett. A **270**, 239 (2000).

[13] B. Toner and D. Bacon, Phys. Rev. Lett. **91**, 187904 (2003).

[14] S. Pironio, Phys. Rev. A **68**, 062102 (2003).

[15] J. Degorre, S. Laplante, and J. Roland, Phys. Rev. A **75**, 012309 (2007).

[16] O. Regev and B. Toner, in *FOCS '07: Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science* (IEEE Computer Society, Washington, DC, USA, 2007), pp. 384–394.

[17] T. Vértesi and E. Bene, Phys. Rev. A **80**, 062316 (2009).

[18] T. Tessier, C. Caves, I. Deutsch, B. Eastin, and D. Bacon, Phys. Rev. A **72**, 032305 (2005).

[19] J. Barrett, C. Caves, B. Eastin, M. Elliott, and S. Pironio, Phys. Rev. A **75**, 012103 (2007).

[20] A. Broadbent, P. Chouha, and A. Tapp, in *Quantum, Nano and Micro Technologies, 2009. ICQNM'09. Third International Conference on* (IEEE, 2009), pp. 59–62.

[21] J. Bancal, C. Branciard, and N. Gisin, Adv. Math. Phys. **2010**, 293245 (2010).

[22] C. Branciard and N. Gisin, Phys. Rev. Lett. **107**, 020401 (2011).

[23] D. Greenberger, M. Horne, and A. Zeilinger, in *Bell's Theorem, Quantum Theory, and Conceptions of the Universe*, edited by M. Kafatos (Dordrecht: Kluwer, 1989), pp. 73–76.

[24] N. Mermin, Am. J. Phys. **58**, 731 (1990).

[25] A. Grudka, M. Horodecki, P. Horodecki, R. Horodecki, and M. Piani, Phys. Rev. A **77**, 060307 (2008).

[26] N. Gisin and S. Iblisdir, Eur. Phys. J. D **39**, 321 (2006).

[27] M. Zukowski, A. Zeilinger, M. Horne, and A. Ekert, Phys. Rev. Lett. **71**, 4287 (1993).

[28] C. Branciard, N. Gisin, and S. Pironio, Phys. Rev. Lett. **104**, 170401 (2010).

[29] C. Branciard, D. Rosset, N. Gisin, and S. Pironio, Phys. Rev. A **85**, 032119 (2012).

[30] M. Halder, A. Beveratos, N. Gisin, V. Scarani, C. Simon, and H. Zbinden, Nature Physics **3**, 692 (2007).

[31] R. Kaltenbaek, R. Prevedel, M. Aspelmeyer, and A. Zeilinger, Phys. Rev. A **79**, 040302 (2009).

[32] S. Massar, D. Bacon, N. Cerf, and R. Cleve, Phys. Rev. A **63**, 052305 (2001).

[33] C. H. Bennett et al., Phys. Rev. Lett. **70**, 1895 (1993).

[34] Note that we do not explicitly consider the simulation of the Referee's Bell state measurement (BSM). One could easily include this in the simulation protocol, by letting for instance the Referee and Bob share a random result for the BSM, and by letting Bob apply an approriate transformation to his measurement setting directly.

[35] This intuition led some of us to post an erroneous preprint: N. Brunner, C. Branciard, N. Gisin, and D. Rosset, arXiv:1103.5058.