Communication Complexity of Channels in General Probabilistic Theories

A. Montina, M. Pfaffhauser, S. Wolf
Facoltà di Informatica, Università della Svizzera Italiana, Via G. Buffi 13, 6900 Lugano, Switzerland
(Dated: December 24, 2013)

The communication complexity of a quantum channel is the minimal amount of classical communication required for classically simulating the process of preparation, transmission through the channel, and subsequent measurement of a quantum state. At present, only little is known about this quantity. In this paper, we present a procedure for systematically evaluating the communication complexity of channels in any general probabilistic theory, in particular quantum theory. The procedure is constructive and provides the most efficient classical protocols. We illustrate this procedure by evaluating the communication complexity of a quantum depolarizing channel with some finite sets of quantum states and measurements.

Quantum communication has proved to be much more powerful than its classical counterpart. Indeed, quantum channels can provide an exponential saving of communication resources in some distributed computing problems [1], where the task is to evaluate a function of data held by two or more parties. A natural measure of power of quantum communication in a two-party scenario is provided by the communication complexity of a quantum channel, which is defined as the minimal amount of classical communication required for classically simulating the process of preparation, transmission through the channel, and subsequent measurement of a quantum state. Indeed, it is clear that a quantum channel cannot replace an amount of classical communication greater than its communication complexity. Thus, this quantity sets an ultimate limit to the power of quantum communication in a two-party scenario in terms of classical resources.

At present, only little is known about the communication complexity of quantum channels. Toner and Bacon proved that two classical bits are sufficient to simulate the communication of a single qubit [2]. In the case of parallel simulations, the communication can be compressed so that the asymptotic cost per simulation is about 1.28 bits [3]. Simulating the communication of $n$ qubits requires an amount of classical communication greater than or equal to $2^n - 1$ bits [4]. However, no upper bound is known.

In this paper, we present a general procedure for systematically evaluating the communication complexity of channels in any general probabilistic theory, in particular quantum theory. The procedure relies on the reverse Shannon theorem [5] and a strategy discussed in Refs. [3, 6]. There, it was shown that any classical simulation protocol can be turned into a protocol with communication cost equal to the classical mutual information between the quantum state and the communicated variable of the parent protocol. A similar role of the mutual information is played in the context of classical simulations of measurements [7]. We illustrate the procedure by evaluating the communication complexity of a quantum channel with some finite sets of quantum states and measurements.

A protocol simulating a quantum channel actually simulates a process of preparation, transmission through the channel and subsequent measurement of a quantum state. For the sake of simplicity, we will focus on quantum channels, but the following discussion can be easily generalized to any probabilistic theory, as pointed out later. The simulated quantum scenario is as follows. A party, say Alice, prepares $n$ qubits in some quantum state $\ket{\psi}\!\bra{\psi} \equiv \rho$ according to an unknown probability distribution $\rho(\psi)$. Then, she sends the qubits to another party, say Bob, through a quantum channel with associated superoperator $\mathcal{L}$. Finally, Bob generates an outcome by performing a measurement $\mathcal{M} = \{ \hat{E}_1, \hat{E}_2, \ldots \}$, where $\hat{E}_i$ are positive semidefinite self-adjoint operators labeling events of the measurement $\mathcal{M}$. The quantum probability of getting the $w$-th outcome $\hat{E}_w$, given $\ket{\psi}$ and $\mathcal{M}$, is

$$P_C(w|\psi, \mathcal{M}) \equiv \text{Tr} \left[ \hat{E}_w \mathcal{L}(\rho) \right]. \quad (1)$$

In a classical simulation, the quantum channel between Alice and Bob is replaced by classical communication. A classical protocol is as follows. Alice sets a variable, say $k$, according to a probability distribution $\rho(k|y, \psi)$ that depends on the quantum state $\ket{\psi}$ and, possibly, a random variable $y$ shared with Bob. Thus, there is a mapping from the quantum state to a probability distribution of $k$,

$$\ket{\psi} \overset{\mathcal{C}}{\mapsto} \rho(k|y, \psi). \quad (2)$$

Alice sends $k$ to Bob, who simulates a measurement $\mathcal{M}$ by generating an outcome $\hat{E}_w$ with a probability $P(w|k, y, \mathcal{M})$. The protocol exactly simulates the quantum channel if the probability of $\hat{E}_w$ given $\ket{\psi}$ is equal to the quantum probability, that is, if

$$\sum_k \int dy P(w|k, y, \mathcal{M}) \rho(k|y, \psi) \rho(y) = P_C(w|\psi, \mathcal{M}), \quad (3)$$

where $\rho(y)$ is the probability density of the random variable $y$. Let us denote by $\rho(k|y) \equiv \int d\psi \sum \rho(k|y, \psi) \rho(\psi)$ the marginal conditional probability of $k$ given $y$. As defined in Ref. [6], the communication cost, say $C$, of the
classical simulation is the maximum, over the space of distributions $\rho(\psi)$, of the Shannon entropy of the distribution $\rho(k|y)$ averaged over $y$, that is,

$$\mathcal{C} \equiv \max_{\rho(\psi)} H(K|Y),$$

(4) where $H(K|Y) \equiv -\int dy \rho(y) \sum_k \rho(k|y) \log_2 \rho(k|y)$.

Shannon’s source coding theorem [4] establishes an operational meaning of $\mathcal{C}$, as discussed in Ref. [6]. Indeed, suppose that $N$ independent simulations of $N$ quantum channels are performed in parallel. Let $k^i$ be the variable prepared with probability $\rho(k^i|y,\psi)$, where $|\psi\rangle$ is the quantum state prepared for the $i$-th quantum channel. Instead of communicating directly the variables $k^i$, we can encode them into a global $k$, so that the average number of communicated bits per simulation approaches $\mathcal{C}$ with vanishing error as $N$ goes to infinity. The quantity $\mathcal{C}$ is the minimal compression rate for the worst case distribution $\rho(\psi)$. Furthermore, it is possible to show that there is a compression code that is optimal for the worst case and has a compression rate independent of the actual distribution $\rho(\psi)$ and equal to $\mathcal{C}$.

We define the communication complexity [denoted by $C_{\text{min}}(\mathcal{C})$] of a quantum channel $\mathcal{C}$ as the minimal amount of classical communication $\mathcal{C}$ required by an exact classical simulation of the quantum channel, given any measurement $\mathcal{M}$ (in a possible alternative definition, only projective measurements would be allowed). Let $S \equiv \{|\psi_1\rangle,\ldots,|\psi_S\rangle\}$ and $M \equiv \{M_1,\ldots,M_M\}$ be a set of $S$ quantum states and $M$ measurements, respectively. We define the communication complexity, say $C_{\text{min}}(G)$, of the quantum game $(\mathcal{C},S,M) \equiv G$ as the minimal amount of classical communication required to simulate the quantum channel $\mathcal{C}$ with the restriction that the quantum states and the measurements are elements of $S$ and $M$, respectively. The quantities $C_{\text{min}}(\mathcal{C})$ and $C_{\text{min}}(G)$ are functionals of $\mathcal{C}$ and $G$, respectively.

Let us consider the case of $N$ quantum channels. In a general parallel simulation, the communicated variable $k$ is generated according to a probability distribution $\rho(k|y,\psi_1,\psi_2,\ldots,\psi_N)$ depending on the whole set of $N$ prepared quantum states $|\psi_1\rangle,\ldots,|\psi_N\rangle$. Thus, the single-shot map (2) is replaced by

$$\{|\psi_1\rangle,\ldots,|\psi_N\rangle\} \rightarrow \rho(k|y,\psi_1,\psi_2,\ldots,\psi_N).$$

The asymptotic communication cost, say $C_{\text{asy}}$, is equal to $\lim_{N \to \infty} C_{\text{asy}}/N$, $C_{\text{asy}}$ being the cost of the parallelized simulation. The definition of $C_{\text{asy}}$ is similar to that of $\mathcal{C}$, with the difference that the maximization is made over the space of the distributions $\rho(\psi_1,\ldots,\psi_N)$. We define the asymptotic communication complexity, $C_{\text{asy}}(\mathcal{C})$, of a quantum channel $\mathcal{C}$ as the minimal asymptotic communication cost required for simulating the channel. The asymptotic communication complexity $C_{\text{asy}}(G)$ of the game $G$ is similarly defined.

Given a game $G = (\mathcal{L},S,M)$, let $w = \{w_1,\ldots,w_M\}$ be an $M$-dimensional array whose $m$-th element is one of the possible outcomes of the $m$-th measurement $\mathcal{M}_m \in M$. We denote by $s = 1,\ldots,S$ and $m = 1,\ldots,M$ discrete indices labelling the elements of $S$ and $M$, respectively. The summation over every index in $w$ but the $m$-th one, which is set equal to $w$, is concisely written as follows,

$$\sum_{w,\ldots,w_{m-1},\ldots,w_M \equiv w} \rightarrow \sum_{w,\ldots,w_{m-1},\ldots,w_M \equiv w}$$

(6)

**Definition.** Given a game $G = (\mathcal{L},S,M)$, the set $\mathcal{V}(G)$ contains any conditional probability $\rho(w|s)$ over the sequence $w = \{w_1,\ldots,w_M\}$ whose marginal distribution of the $m$-th variable is the quantum distribution of the outcome $w_m$ given the quantum state $s$ and the measurement $m$, for any $s$, $m$. In other words, the set $\mathcal{V}(G)$ contains any $\rho(w|s)$ satisfying the constraints

$$\sum_{w,w_{m-1},\ldots,w_M \equiv w} \rho(w|s) = P_G(w|s,m), \forall s,m,w,$$

(7)

where

$$P_G(w|s,m) \equiv P_L(w|\psi_s,\mathcal{M}_m)$$

is the quantum probability of getting the $w$-th outcome of the measurement $\mathcal{M}_m$ given the quantum state $|\psi_s\rangle$.

The set $\mathcal{V}(G)$ is surely non-empty. A function in $\mathcal{V}(G)$ is $\rho(w|s) = P_G(w_1|s,1) \times \cdots \times P_G(w_M|s,M)$, where the variables $w_1,\ldots,w_M$ are uncorrelated. The definition of $\mathcal{V}(G)$ can be easily extended to any general probabilistic theory, where $P_G(w|s,m)$ is replaced by different conditional probabilities. For the sake of concreteness, we will refer to the quantum case, but the following discussion does not rely on any precise form of $P_G(w|s,m)$ and applies to more general theories.

A pivotal classical protocol for the quantum game $G$ is as follows.

**Master protocol.** Alice generates the array $w$ according to a conditional probability $\rho(w|s) \in \mathcal{V}(G)$. Then, she sends $w$ to Bob. Bob simulates the measurement $\mathcal{M}_m$ by outputting the outcome $w_m$.

The definition of $\mathcal{V}(G)$ implies that this protocol exactly simulates the quantum game $G$. A classical channel from a variable $x_1$ to $x_2$ is defined by the conditional probability of getting $x_2$ given $x_1$. Its capacity is the maximum of the mutual information between $x_1$ and $x_2$ over the space of probability distributions $\rho(x_1)$. Using the strategy discussed in Ref. [3] and the reverse Shannon theorem [3], it is possible to prove that a master protocol can be turned into a child protocol for parallel simulations whose asymptotic communication cost is equal to the capacity of the classical channel $\rho(w|s)$.
Theorem 1. Given a conditional probability \( \rho(w|s) \in \mathcal{V}(G) \), there is a child protocol, simulating in parallel \( N \) quantum games \( G \), whose asymptotic communication cost per game is equal to the capacity of the channel \( \rho(w|s) \) as \( N \) goes to infinity.

Proof. In a parallel simulation of \( N \) games through \( N \) master protocols, Alice sends an array \( w \) to Bob for each game. This array is generated with probability \( \rho(w|s) \). Let \( C(W|S) \) be the capacity of the channel \( T' : s \rightarrow w \). The child protocol is as follows. Instead of sending \( w \), Alice sends an amount of information, say \( C(N) \), that allows Bob to generate \( w \) for every game \( G \) according to the probability \( \rho(w|s) \). The reverse Shannon theorem states that this can be accomplished with a cost \( C(N) \) such that \( \lim_{N \rightarrow \infty} C(N)/N = C(W|S) \), provided that the receiver and sender share some random variable. □

A constructive proof of the reverse Shannon theorem and its one-shot version were provided in Ref. [9]. This gives an explicit procedure for deriving the child protocol associated with \( \rho(w|s) \).

The first main result is the following theorem about the asymptotic communication complexity. Later on, we will consider the single-shot case.

Theorem 1. The asymptotic communication complexity of the game \( G = (L, S, M) \) is the minimum of the capacity of the classical channels \( \rho(w|s) \) in the set \( \mathcal{V}(G) \).

Theorem 1 states that the asymptotic communication complexity of the game \( G \) is equal to the quantity

\[
\mathcal{D}(G) \equiv \min_{\rho(w|s) \in \mathcal{V}(G)} \left( \max_{\rho(s)} I(W;S) \right), \tag{8}
\]

where \( I(W;S) \) is the mutual information between the stochastic variables \( w \) and \( s \). This theorem and Lemma 1 provide a constructive method for deriving the best protocol with communication cost equal to \( \mathcal{D}(G) \). It is sufficient to evaluate the conditional probability \( \rho(w|s) \) that solves the minimax problem stated in Eq. (8) and to use the procedure in Ref. [9] for deriving the associated child protocol. The proof of the theorem is provided in the appendix. It relies on Lemma 1 and the data-processing inequality [1].

Lemma 1 implies that there is a protocol whose communication cost is \( \mathcal{D}(G) \), that is, \( C_{\text{min}}^\text{asym}(G) \leq \mathcal{D}(G) \). Furthermore, the communication cost of any simulation cannot be strictly smaller than \( \mathcal{D}(G) \). This is proved by showing through the data-processing inequality that any simulation protocol with communication cost \( C \) induces a master protocol with associated capacity \( C(W|S) \) smaller or equal to \( C \). Thus, \( C_{\text{min}}^\text{asym}(G) = \mathcal{D}(G) \).

Theorem 1 and Lemma 1 have their one-shot versions.

Lemma 2 (One-shot version of Lemma 1). Given a conditional probability \( \rho(w|s) \in \mathcal{V}(G) \), there is protocol simulating a quantum game \( G \) such that

\[
C_{ch} \leq C \leq C_{ch} + 2 \log_2 (C_{ch} + 1) + 2 \log_2 \epsilon,
\]

where \( C \) and \( C_{ch} \) are the communication cost of the simulation and the capacity of the channel \( \rho(w|s) \).

The proof is similar to that of Lemma 1 and relies on the one-shot version of the reverse Shannon theorem proved in Ref. [9].

Theorem 2 (One-shot version of Theorem 1). The communication complexity \( C_{\text{min}}(G) \) of the game \( G \) satisfies the inequalities

\[
\mathcal{D}(G) \leq C_{\text{min}}(G) \leq \mathcal{D}(G) + 2 \log_2 [\mathcal{D}(G) + 1] + 2 \log_2 \epsilon,
\]

where \( \mathcal{D}(G) \) is given by Eq. (8) and it is equal to the asymptotic communication complexity of the game \( G \) (Theorem 1).

The first inequality is a trivial consequence of Theorem 1, as the asymptotic communication complexity cannot be larger than the communication complexity. The second inequality comes from Lemma 2.

Thus, the communication complexity of a quantum channel is about equal to the asymptotic communication complexity, apart from a possible additional cost that does not grow more than the logarithm of the asymptotic communication complexity. The asymptotic communication complexity of a quantum channel is obtained in the limit \( S, M \rightarrow \infty \) with the sets \( S \) and \( M \) densely covering the space of quantum states and measurements, respectively.

To illustrate these results, we have evaluated the communication complexity of the following game \( G \) for a binary quantum depolarizing channel. The channel is a map from a Bloch vector \( \vec{\gamma} \) to \( \gamma \vec{\gamma} \), where \( 0 \leq \gamma \leq 1 \). The channel is noiseless or completely erasing if \( \gamma = 1 \) or 0, respectively. Let us denote by \( \vec{\gamma}_m \) the tridimensional vectorial function \( (\cos \frac{\pi}{M} m, \sin \frac{\pi}{M} m, \vec{0}) \), where \( x \) is a real number. The measurements are projections in a two-dimensional Hilbert space. The eigenvectors of the \( m \)-th measurement in \( M \) correspond to the Bloch vectors \( \pm \vec{\gamma}_m \) with \( m = 1, \ldots, M \) and outcomes \( w = \pm 1 \). The set \( S \) contains all the \( 2M \) eigenvectors, that is, \( \vec{\gamma}_m \) with \( s = 1, \ldots, 2M \). The quantum probability of getting \( w \) given \( s \) and \( m \) is

\[
P_G(w|s,m) = \frac{1}{2} \left\{ 1 + w \gamma \cos \left[ \frac{\pi}{M} (s - m) \right] \right\}.
\]

Since \( P_G \) is invariant under the transformation \( s \rightarrow s + 1 \) and \( m \rightarrow m + 1 \), the distribution \( \rho(s) \) solving the minimax problem in Eq. (8) is, by symmetry, uniform. Thus the minimax problem is reduced to a minimization problem.

We have evaluated algebraically the asymptotic communication complexity up to \( M = 4 \). The distributions \( \rho(w|s) \in \mathcal{V}(G) \) with minimal capacity for \( M = 2, 3, 4 \) are
summarized by the analytical equation
\[
\rho(w|s) = \sum_{k=1}^{2M} P(w|k) \rho(k|s)
\]
with
\[
P(w|k) = \prod_{m=1}^{M} \theta(w_m \vec{v}_m \cdot \vec{v}_{k+p/2}) ,
\]
\[
\rho(k|s) = f(k,s) + \sqrt{\lambda + f(k,s)}
\]
where \(p = 0\) (1) if \(M\) is odd (even), \(\lambda\) is a constant determined by the normalization \(\sum_{k=1}^{2M} \rho(k|s) = 1\) and
\[
f(k,s) = \frac{\gamma}{2} \sin \left( \frac{\pi}{2M} \vec{v}_{k+p/2} \cdot \vec{v}_s \right).
\]

It is easy to prove that \(\rho(w|s)\) is an element of \(\mathcal{V}(G)\) for any \(M \geq 2\). For \(\gamma = 1\) (noiseless channel), these equations are the discrete version of the Kochen-Specker model [10] with the constraint that the hidden variable \(\vec{v}_{k+p/2}\) is a vector lying on a plane.

Since \(P(w|k)\) is a noiseless channel, the capacity of \(\rho(w|s)\) is equal to the capacity of \(\rho(k|s)\). Thus, the asymptotic communication complexity is
\[
C_{\text{asym}}(G) = \max_{\rho} I(k; s)
\]
In particular, for a noiseless channel
\[
C_{\text{asym}}(G) = N \sum_{n=1}^{M} \cos \left( \frac{\pi n}{M} \right) \log \left[ 2M N \cos \left( \frac{\pi n}{M} \right) \right],
\]
where \(N = \sin \left( \frac{\pi}{2M} \right)\). Note that the sum index \(n\) is not an integer when \(M\) is even. We have numerically verified the validity of the analytical equations for \(M\) up to 20 and some values of \(\gamma\). The simulations are in agreement with Eqs. (12,13,14) within the machine precision. The numerical data for \(\gamma = 1\) and 0.95 are reported in Fig. 1. If we extrapolate Eq. (13) to arbitrary \(M\), we have \(\lim_{M \to \infty} C_{\text{asym}}(G) = 1 + \log_2 \frac{\pi}{2} \approx 1.2088\) (dot line in Fig. 1). This value is the asymptotic communication complexity of a noiseless quantum channel with the constraint that the quantum states and the eigenstates of the measurements correspond to Bloch vectors lying on a plane. In Ref. [3], we found a protocol for any quantum state and projective measurements with communication cost equal to \(\log_2 (4/\sqrt{\pi}) \approx 1.2786\), which is about 6% higher (dashed line in Fig. 1). It is not known if this value is actually the asymptotic communication complexity of the quantum channel for general projective measurements.

In conclusion, we have presented a general procedure for evaluating the communication complexity of channels in any general probabilistic theory, in particular quantum theory. This procedure, which relies on the reverse Shannon theorem and a strategy introduced in Refs. [3,6], is constructive and provides a method to derive the most efficient protocol that classically simulates a channel. More explicitly, given a quantum channel, we have defined a set \(\mathcal{V}\) of classical channels and proved that the minimal classical capacity in \(\mathcal{V}\) is the asymptotic communication complexity of the quantum channel. Thus, the problem of evaluating the communication complexity is reduced to a minimax problem. The channel in \(\mathcal{V}\) with minimal capacity can be turned into the most efficient classical protocol for simulating the quantum channel. We have illustrated this procedure by evaluating the asymptotic communication complexity of a binary quantum depolarizing channel for some finite sets of quantum states and measurements. The procedure is numerically very stable, but the computational time of the minimax routine can grow exponentially with the number of quantum states and measurements. Thus, specific strategies reducing the computational complexity need to be devised in the case of a high number of states and measurements.

At the present, it not known if the communication complexity of noiseless quantum channels is finite, unless the quantum channel capacity is 1 qubit. Our method can help to solve this open problem and, furthermore, to construct explicit simulation protocols. As discussed in Ref. [3], the existence of finite classical communication protocols is also deeply related to the existence of \(\psi\)-epistemic theories, which are being object of recent intense study.

Acknowledgments. This work is supported by the Swiss National Science Foundation, the NCCR QSIT, and the COST action on Fundamental Problems in Quantum Physics. A. M. acknowledges a support in part from Perimeter Institute for Theoretical Physics, where a sketch of the proof of Theorem 1 was conceived. Re-
Finally, from this distribution and \( \rho(k|y, s^1, \ldots, s^N) \), where \( s^i \) is an index labelling the quantum state of the \( i \)-th game (hereafter superscripts label the game). Bob simulates the measurements \( M_{m^1}, \ldots, M_{mN} \) by generating the outcomes \( w^1, \ldots, w^N \) according to a conditional probability \( P(w^1, \ldots, w^N|k, y, m^1, \ldots, m^N) \). Let us denote by \( P^i(w^i|k, y, m^1, \ldots, m^N) \) the marginal probability of the outcome of the \( i \)-th game. We introduce the conditional probabilities

\[
P^i(w_1, w_2, \ldots, w_M|k, y) \equiv \prod_{m=1}^M P^i(w_m|k, y, 1, 1, \ldots, m^i = m, 1),
\]

We will concisely denote \( P^i(w_1, w_2, \ldots, w_M|k, y) \) by \( P^i(w|k, y) \). Note that we have multiplied over the index \( m^i \) and set the other indices equal to 1. For our purposes, any other choice of the values of the \( N - 1 \) indices would be fine. We use \( P^i(w|k, y) \) to build the conditional probability

\[
P(w^1, \ldots, w^N|k, y) = \prod_i P^i(w^i|k, y).
\]

Finally, from this distribution and \( \rho(k|y, s^1, \ldots, s^N) \), we build the conditional probability

\[
\rho(w^1, \ldots, w^N|s^1, \ldots, s^N) = \sum_k \int dy \rho(y) P(w^1, \ldots, w^N|k, y) \rho(k|y, s^1, \ldots, s^N).
\]

From the data-processing inequality, we have that the capacity, say \( C(W^1, \ldots, W^N|S^1, \ldots, S^N) \), of \( \rho(w^1, \ldots, w^N|s^1, \ldots, s^N) \) is smaller than or equal to the communication cost \( NC_0 + o(N) \), that is,

\[
C(W^1, \ldots, W^N|S^1, \ldots, S^N) \leq NC_0 + o(N).
\]

By construction, we have the constraints

\[
\sum_{w^1, \ldots, w^N} \rho(w^1, \ldots, w^N|s^1, \ldots, s^N) = P_G(w|s_i, m),
\]

the left-hand side being the marginal distribution of the variable \( w^i \) (renamed \( w \)) given \( s^1, \ldots, s^N \). Let \( \rho_0(w|s) \) be the probability distribution in \( V(G) \) with minimal capacity \( D(G) \). Then, it is easy to realized that the probability distribution

\[
\rho_{\min}(w^1, \ldots, w^N|s^1, \ldots, s^N) \equiv \prod_i \rho_0(w^i|s^i),
\]

is the channel satisfying constraints with minimal capacity. The minimum is equal to \( ND(G) \). Thus,

\[
ND(G) \leq C(W^1, \ldots, W^N|S^1, \ldots, S^N).
\]

From this inequality and Inequality, we have that

\[
ND(G) \leq NC_0 + o(N).
\]

The theorem is proved. □

References

[1] H. Buhrman, R. Cleve, S. Massar, and R. de Wolf, Rev. Mod. Phys. 82, 665 (2010).
[2] B. F. Toner and D. Bacon, Phys. Rev. Lett. 91, 187904 (2003).
[3] A. Montina, Phys. Rev. Lett. 109, 110501 (2012).
[4] A. Montina, Phys. Rev. A 84, 060303(R) (2011).
[5] C. H. Bennett, P. Shor, J. Smolin, and A. V. Thapliyal, IEEE Trans. Inf. Theory, 48 2637 (2002).
[6] A. Montina, Phys. Rev. A 87, 042331 (2013).
[7] M. Berta, J. M. Renes, and M. M. Wilde, arXiv:1301.1594; M. M. Wilde, P. Hayden, F. Buscemi, M-H. Hsieh, arXiv:1206.4121.
[8] T. M. Cover and J. A. Thomas, Elements of Information Theory (Wiley, New York, 1991).
[9] P. Harsha, R. Jain, D. McAllester, J. Radhakrishnan, IEEE Trans. Inf. Theory 56, 438 (2010).
[10] S. Kochen and E. Specker, J. Math. Mech. 17, 59 (1967).