

ON ENTANGLEMENT ASSISTANCE TO A NOISELESS CLASSICAL CHANNEL

PÉTER E. FRENKEL AND MIHÁLY WEINER

ABSTRACT. For a classical channel, neither the Shannon capacity, nor the sum of conditional probabilities corresponding to the cases of successful transmission can be increased by the use of a non-signaling resource. Yet, perhaps somewhat counterintuitively, entanglement assistance can help and actually elevate the chances of success even in a one-way communicational task that is to be completed by a single-shot use of a noiseless classical channel.

To quantify the help that a non-signaling resource provides to a noiseless classical channel, one might ask how many extra letters should be added to the alphabet of the channel in order to perform equally well *without* the specified non-signaling resource. As was observed by Cubitt, Leung, Matthews, and Winter, there is no upper bound on the number of extra letters required for substituting the assistance of a general non-signaling resource to a noiseless one-bit classical channel. In contrast, here we prove that if this resource is a bipartite quantum system in a maximally entangled state, then an extra classical bit always suffices as a replacement.

1. INTRODUCTION

If a certain two-part resource is *non-signaling*, then — essentially by definition — it cannot be used to exchange messages between its two users. However, as an aid, it might boost the capabilities of an already existing communicational channel between them. For example, in the famous *dense coding protocol* [2], entanglement is used to boost the classical capacity of a quantum channel.

The situation changes somewhat when the channel to be improved is a classical one. This is because it turns out that important quantities such as “information storability” — that is, the sum of conditional

The first author’s research is partially supported by MTA Rényi “Lendület” Groups and Graphs Research Group, by ERC Consolidator Grant 1040085 and by the National Research, Development and Innovation Office of Hungary (NRDI) via the research grant K124152. This latter grant also supports the second author, who is further supported by the Bolyai János Fellowship of the Hungarian Academy of Sciences, the ÚNKP-19-4 New National Excellence Program of the Ministry for Innovation and Technology and by the NRDI grants KH129601 and K132097.

probabilities corresponding to the “output = input” cases; see e.g. [9] — or the Shannon capacity of a classical channel cannot be increased by the additional use of a non-signaling resource [4]. However, entanglement can be used, for example, to increase the zero-error capacity of a noisy classical channel [3]. To put it in another way, it can improve the capability of a noisy classical channel to simulate a noiseless one.

One might also consider the inverse problem of using a noiseless classical channel (together with the possible help of a non-signaling resource) to simulate a noisy one. In [4], an example was given for a classical channel which cannot be simulated by (a single use of) a noiseless one-bit classical channel aided only by shared randomness, but which can be simulated by the same channel if assistance, in the form of using a bipartite quantum system prepared in an entangled state, is allowed. Using the concepts introduced in [5], we may say that the assistance increases the “signaling dimension” of our classical channel.

1.1. Game interpretation. We might view this simulability question from the point of view of one-way communicational tasks. For example, let us consider the following simple game. We have four boxes, two of them empty, two containing (equal) treasures; thus, there are $\binom{4}{2} = 6$ possible configurations regarding the positions of the treasures. Each configuration is equally likely, with the actual (secret) configuration revealed only to Alice, who is allowed to send one classical bit to Bob. After receiving the bit sent by Alice, Bob chooses a box. If it contains a treasure, Alice and Bob win (as a team).

Without a non-signaling resource, relying only on arrangements before the game and a possible use of shared randomness, it is easy to see that the maximum chance of winning can be achieved by a *pure* (deterministic) strategy — in terms of expected reward, shared randomness is of no use.

We may assume that upon receiving the bit sent by Alice, depending on its value, Bob points to either box nr. 1 or box nr. 2. With this agreed, Alice can always send a bit value to Bob that makes him point to a treasure box unless the treasures are in boxes nr. 3 & 4; that is, Alice and Bob will win with a chance of $1 - 1/6 = 5/6$. On the other hand, if during the game Alice and Bob can also make some measurements on a pair of quantum bits (prepared in a maximally entangled state and distributed between them before the start of the game), then there exists a strategy allowing them to win this game with a chance of $(4 + \sqrt{2})/6 > 5/6$ — see the Appendix.

In this protocol involving the use of entanglement, Alice and Bob realize a certain classical channel $\mathcal{N} \in C(X \rightarrow Y)$ with $|X| = 6$ possible inputs and $|Y| = 4$ outputs (with the input being the whereabouts of the treasures revealed to Alice and the output being the box chosen by Bob). The use of this channel allows Alice and Bob to win the game with a chance of $(4 + \sqrt{2})/6$. The fact that $(4 + \sqrt{2})/6 > 5/6$ shows that channel \mathcal{N} cannot be simulated by a single use of a noiseless classical one-bit channel aided only by shared randomness.

Let us consider, in general, for any pair of (finite) sets X, Y , natural number n and non-signaling resource ω the following:

- $C_n(X \rightarrow Y)$, the set of $X \rightarrow Y$ classical channels that can be simulated by a single use of a noiseless classical channel with n different letters (without any other resources),
- $C_n^{SR}(X \rightarrow Y)$, the set of $X \rightarrow Y$ classical channels that can be simulated by a single use of a noiseless classical channel with n different letters together with an unlimited source of shared randomness between the sender and receiver,
- $C_n^\omega(X \rightarrow Y)$, the set of $X \rightarrow Y$ classical channels that can be simulated by a single use of a noiseless classical channel with n different letters and assistance coming from ω ,
- $C_n^{BQ}(X \rightarrow Y)$, the set of $X \rightarrow Y$ classical channels that can be simulated by a single use of a noiseless classical channel with n different letters and assistance from any bipartite quantum system (prepared in any state),
- $C_n^{NS}(X \rightarrow Y)$, the set $X \rightarrow Y$ classical of channels that can be simulated by a single use of a noiseless classical channel with n different letters and assistance from any non-signaling resource.

We postpone to Section 2 the precise definition and detailed description of these sets, and – omitting the $X \rightarrow Y$ indication – note here only that C_n^{SR} is the convex hull of C_n ; C_n^{BQ} and C_n^{NS} are convex sets; and $C_n^{SR} \subseteq C_n^{BQ} \subseteq C_n^{NS}$.

We now generalize the game above and connect the question of (im)possibility of simulations to advantages in one-way communicational games. We begin by describing what we mean by a general one-way communicational game.

Suppose we have a team of players consisting of Alice and Bob. An element x of a (finite) set X is chosen according to a given probability distribution q , and revealed to Alice (but not to Bob). At the end of the game, Bob will need to pick an element y of another (finite) set Y and the team receives a reward, but the actual sum of this reward depends

both on his choice y and on the input x ; it is given by some “reward-function” $R : X \times Y \rightarrow \mathbb{R}$. With both the probability distribution q and reward-function R publicly given, we shall consider how the maximum expected reward (achieved by the best team strategy) depends on the allowed forms of communication and non-signaling resources the team can use.

Once a team strategy (using the given channels and resources) is chosen, the conditional probability of Bob choosing y , given that Alice receives input x , is fixed. Thus, an actual strategy realizes a classical $\mathcal{N} \in C(X \rightarrow Y)$ channel. The expected reward is a linear functional of the realized channel:

$$\mathbb{E}(\text{reward}) = \sum_{x \in X, y \in Y} R(x, y)N(y|x)q(x),$$

with the functional depending on R and q . Since we want to consider all such games, we do not have a restriction on possible reward functions and input probability distributions and thus, for us, the expected reward is just an arbitrary linear functional of the realized channel. It follows that there exists a one-way communicational game in which the single use of a classical noiseless channel with n different letters together with assistance coming from a non-signaling resource ω is more advantageous (in terms of maximal expected rewards) than the single use of a classical noiseless channel with n' different letters together with assistance from a non-signaling resource ω' if and only if $C_n^\omega(X \rightarrow Y)$ is not contained in the convex hull of $C_{n'}^{\omega'}(X \rightarrow Y)$ for some X and Y . In particular, with assistance coming from a non-signaling resource ω , the use of a classical noiseless channel of n letters is never more advantageous than a single use of an unaided classical noiseless channel of m different letters if and only if

$$C_n^\omega(X \rightarrow Y) \subseteq C_m^{SR}(X \rightarrow Y)$$

for all possible sets X and Y of input and output symbols.

1.2. Entanglement vs. generic non-signaling resources. By [4, Proposition 19], for every n there exists a non-signaling resource ω_n such that $C_2^{\omega_n}$ is *not* contained in C_n^{SR} (for some sets of input and output symbols which from now on we shall omit). Thus, in the sense explained, there is no bound on the advantage that a non-signaling resource can give to a one-bit classical noiseless channel. In contrast, in what follows we shall prove that

$$C_2^\omega \subseteq C_4^{SR}$$

whenever ω is realizable by a quantum bipartite system prepared in a maximally entangled state. The result is general in the sense that it holds without any limit on the size of the quantum system (i.e. the dimension of the Hilbert space) used. However, we do exploit that the state is a maximally entangled one. Since there are Bell inequalities whose maximal violation occurs in states which are *not* maximally entangled [1], it remains unclear whether $C_2^{BQ} \subseteq C_4^{SR}$. Nevertheless, we conjecture that this is indeed so; maybe even $C_2^{BQ} \subseteq C_3^{SR}$. Also, it is a natural guess that C_n^{BQ} should always be contained in $C_{n^2}^{SR}$.

If this turns out to be true, then we may say that quantum physics follows the proverb “God helps those who help themselves”. Whereas with generic non-signaling resources, the “help” the resource can give in a one-way communicational game is unlimited, assistance from the use of a bipartite quantum system can give advantage, but — if $C_n^{BQ} \subseteq C_{n^2}^{SR}$ holds — definitely not more than what a simple second shot of the employed classical noiseless channel would offer. Similarly to the so-called “Information Causality” [10], this could be viewed as a fundamental principle limiting the non-signaling resources that can appear in nature.

2. PRELIMINARIES

Given two finite sets X, Y (the “alphabets”), a classical channel from X to Y is a function $\mathcal{N} : Y \times X \rightarrow [0, 1]$ satisfying $\sum_{y \in Y} \mathcal{N}(y|x) = 1$ for all $x \in X$. We interpret the value $\mathcal{N}(y|x)$ as the probability of the channel producing the output y given that the input is x , and we denote by $C(X \rightarrow Y)$ the set of all classical channels from X to Y .

For a natural number n , set $[n] \equiv \{1, \dots, n\}$. We shall say that $\mathcal{N} \in C(X \rightarrow Y)$ can be realized by a single use of a noiseless classical channel with n different letters if there exists a pair of *encoding* and *decoding*, i.e., channels $\mathcal{N}_{enc} \in C(X \rightarrow [n])$ and $\mathcal{N}_{dec} \in C([n] \rightarrow Y)$ such that

$$\mathcal{N}(y|x) = \sum_{r=1}^n \mathcal{N}_{dec}(y|r) \mathcal{N}_{enc}(r|x)$$

for all $x \in X$ and $y \in Y$. We denote the set of all such channels \mathcal{N} by $C_n(X \rightarrow Y)$.

Using a source of randomness shared between the sender and receiver, it is possible to mix different encoding–decoding strategies. Thus, the set of classical channels $C_n^{SR}(X \rightarrow Y)$ realizable by a single use of a noiseless classical channel with n different letters aided by an unlimited source of shared randomness is simply the convex hull of $C_n(X \rightarrow Y)$. We note that $\mathcal{N} \in C_n^{SR}(X \rightarrow Y)$ if and only if the stochastic matrix

A with entries $a_{i,j} := \mathcal{N}(y(i)|x(j))$, where $x : [l] \rightarrow X$ and $y : [k] \rightarrow Y$ are some bijections enumerating the $l = |X|$ and $k = |Y|$ elements of X and Y , is a convex combination of stochastic matrices with at most n nonzero rows. (Throughout this paper, by a *stochastic* matrix we mean a matrix with nonnegative entries whose columns sum to 1; i.e., $A = (a_{i,j})_{i,j}$ is a stochastic matrix if $a_{i,j} \geq 0$ for all i and j , and $\sum_i a_{i,j} = 1$ for all j .)

In our context, a two-part resource ω is just a classical channel with two inputs and two outputs; i.e. an element of $C(X_1 \times X_2 \rightarrow Y_1 \times Y_2)$, where X_1, X_2, Y_1, Y_2 are some finite sets. We say that ω is *non-signaling* if for any $x_1, x'_1 \in X_1$ and $(x_2, y_2) \in X_2 \times Y_2$, we have

$$\sum_{y_1 \in Y_1} \omega(y_1, y_2 | x_1, x_2) = \sum_{y_1 \in Y_1} \omega(y_1, y_2 | x'_1, x_2),$$

i.e., if the choice of the input at access point nr. 1 does not affect the outcome probabilities at access point nr. 2, and, further, the same holds in the other direction as well:

$$\sum_{y_2 \in Y_2} \omega(y_1, y_2 | x_1, x_2) = \sum_{y_2 \in Y_2} \omega(y_1, y_2 | x_1, x'_2)$$

for any $x_2, x'_2 \in X_2$ and $(x_1, y_1) \in X_1 \times Y_1$.

A channel $\mathcal{N} \in C(X \rightarrow Y)$ can be realized by a single use of a noiseless classical channel with n different letters assisted by a non-signaling resource ω if there exist

- a coding for the sender $\mathcal{N}_{in1} \in C(X \rightarrow X_1)$ for selecting an input for the sender's part of the resource,
- an encoding $\mathcal{N}_{enc} \in C(X \times Y_1 \rightarrow [n])$ for the sender for selecting the message (in light of the response of the resource) to be sent,
- a coding for the receiver $\mathcal{N}_{in2} \in C([n] \rightarrow X_2)$ for selecting an input for the receiver's part of the resource,
- a decoding for the receiver $\mathcal{N}_{dec} \in C([n] \times Y_2 \rightarrow Y)$ for selecting the output

such that for all $x \in X$ and $y \in Y$, we have that $\mathcal{N}(y|x) =$

$$\sum_{r, x_1, x_2, y_1, y_2} \mathcal{N}_{dec}(y|r, y_2) \mathcal{N}_{in2}(x_2|r) \mathcal{N}_{enc}(r|x, y_1) \omega(y_1, y_2 | x_1, x_2) \mathcal{N}_{in1}(x_1|x),$$

where the summation is for all $r \in [n]$ and $(x_1, x_2, y_1, y_2) \in X_1 \times X_2 \times Y_1 \times Y_2$. We denote by $C_n^\omega(X \rightarrow Y)$ the set of all such channels and by $C_n^{NS}(X \rightarrow Y)$ the union of these sets taken over all non-signaling resources ω . Note that this latter set is automatically convex; this is because shared randomness is also a particular case of a non-signaling resource.

Recall that a *partition of unity* (also known as a *positive operator valued measure*; POVM) on a Hilbert space is a collection of positive semidefinite operators summing to the identity operator $\mathbf{1}$.

Let \mathcal{H}_A and \mathcal{H}_B be two (complex) Hilbert spaces and ρ a *density operator* on $\mathcal{H}_A \otimes \mathcal{H}_B$; i.e. a positive semidefinite operator with $\text{tr } \rho = 1$. A two-part resource $\omega \in C(X_A \times X_B \rightarrow Y_A \times Y_B)$ is realizable by the use of a bipartite quantum system (with parts corresponding to the spaces \mathcal{H}_A and \mathcal{H}_B) in state ρ if, for each $x_a \in X_A$ and $x_b \in X_B$, there exist a partition of unity $\left(F_{y_a}^{(x_a)}\right)_{y_a \in Y_A}$ on \mathcal{H}_A and a partition of unity $\left(E_{y_b}^{(x_b)}\right)_{y_b \in Y_B}$ on \mathcal{H}_B such that

$$\omega(y_a, y_b | x_a, x_b) = \text{tr } \rho \left(F_{y_a}^{(x_a)} \otimes E_{y_b}^{(x_b)} \right)$$

for all $(x_a, x_b, y_a, y_b) \in X_A \times X_B \times Y_A \times Y_B$. We note that such a resource is automatically non-signaling, and introduce $C_n^{BQ}(X \rightarrow Y)$ as the union of the sets $C_n^\omega(X \rightarrow Y)$ with ω ranging over all non-signaling resources realizable by the use of some bipartite quantum system prepared in some state.

Let us consider the linear map Φ_ρ with domain $\mathcal{B}(\mathcal{H}_A)$ defined by the formula

$$(2.1) \quad \Phi_\rho(Z) \equiv \text{tr}_A \rho(Z \otimes \mathbf{1}),$$

where tr_A denotes the *partial trace* corresponding to \mathcal{H}_A . It is easy to check that this map is well defined, takes values in $\mathcal{B}(\mathcal{H}_B)$, and is a *positive map*: if $Z \geq \mathbf{0}$, then $\Phi_\rho(Z) \geq \mathbf{0}$. Let us now introduce, in the previous construction of the non-signaling resource ω , the operator

$$\beta_{y_a}^{(x_a)} \equiv \Phi_\rho \left(F_{y_a}^{(x_a)} \right).$$

Then, for each $x_a \in X_A$, the operators $\left(\beta_{y_a}^{(x_a)}\right)_{y_a \in Y_A}$ form a *positive decomposition* of $\rho_B \equiv \text{tr}_A \rho$; i.e., $\beta_{y_a}^{(x_a)} \geq \mathbf{0}$ for all x_a and y_a , and

$$\sum_{y_a \in Y_A} \beta_{y_a}^{(x_a)} = \rho_B \equiv \text{tr}_A \rho$$

for all x_a . With these newly introduced operators, we can express ω as follows:

$$(2.2) \quad \omega(y_a, y_b | x_a, x_b) = \text{tr } E_{y_b}^{(x_b)} \beta_{y_a}^{(x_a)}$$

for all $(x_a, x_b, y_a, y_b) \in X_A \times X_B \times Y_A \times Y_B$. This shows that for a non-signaling resource ω to be realizable by the use of a bipartite quantum system, with parts corresponding to the Hilbert spaces \mathcal{H}_A and \mathcal{H}_B , prepared in the state given by the density operator ρ , there must exist,

for each $x_a \in X_A$ and $x_b \in X_B$, a positive decomposition $\left(\beta_{y_a}^{(x_a)}\right)_{y_a \in Y_a}$ of $\rho_B = \text{tr}_A \rho$ and a partition of unity $\left(E_{y_b}^{(x_b)}\right)_{y_b \in Y_b}$ such that (2.2) holds. In some cases, we can turn this construction the other way around.

Lemma 1. *Let \mathcal{H}_A and \mathcal{H}_B be separable Hilbert spaces, ρ a density operator on $\mathcal{H}_A \otimes \mathcal{H}_B$, Φ_ρ the map defined by (2.1), $\rho_B = \text{tr}_A \rho$ and finally $\mathcal{K} = \rho_B^{1/2} \mathcal{B}(\mathcal{H}_B) \rho_B^{1/2}$. If ρ is pure (i.e., it is an orthogonal projection of rank 1), then there exists a linear map $\Gamma_\rho : \mathcal{K} \rightarrow \mathcal{B}(\mathcal{H}_A)$ such that*

- $\Phi_\rho \circ \Gamma_\rho = \text{id}_{\mathcal{K}}$; i.e., Γ_ρ is a right-inverse of Φ_ρ ,
- $\Gamma_\rho(K) \geq \mathbf{0}$ whenever $K \geq \mathbf{0}$; i.e., Γ_ρ is a positive map,
- $\Gamma_\rho(\rho_B) = \mathbf{1}$.

Hence for every positive decomposition $(\beta_y)_{y \in Y}$ of ρ_B , the formula $F_y := \Gamma_\rho(\beta_y)$ defines a POVM for which $\Phi_\rho(F_y) = \beta_y$ holds for all $y \in Y$.

Proof. Suppose $\rho = |\Psi\rangle\langle\Psi|$, where $\Psi \in \mathcal{H}_A \otimes \mathcal{H}_B$ is a unit vector. By the existence of a Schmidt decomposition, we have a countable set S , an orthonormal system $(e_n^A)_{n \in S}$ in \mathcal{H}_A , another one $(e_n^B)_{n \in S}$ in \mathcal{H}_B , and some positive numbers $(\lambda_n)_{n \in S}$ such that

$$\Psi = \sum_{n \in S} \lambda_n e_n^A \otimes e_n^B.$$

Moreover, we have that $\rho_A \equiv \text{tr}_B \rho = \sum_{n \in S} \lambda_n^2 |e_n^A\rangle\langle e_n^A|$ and similarly, $\rho_B = \sum_{n \in S} \lambda_n^2 |e_n^B\rangle\langle e_n^B|$. Let us further consider the partial isometry

$$V = \sum_{n \in S} |e_n^A\rangle\langle e_n^B|$$

and the orthogonal projections $Q^A = VV^*$ and $Q^B = V^*V$ onto the closed subspaces spanned by $\{e_n^A | n \in S\}$ and $\{e_n^B | n \in S\}$, respectively. Finally, we choose an anti-unitary map $J : \mathcal{H}_A \rightarrow \mathcal{H}_A$ satisfying $Je_n^A = e_n^A$ for every $n \in S$, and define Γ_ρ by setting

$$\Gamma_\rho(K) = \Gamma_\rho\left(\rho_B^{1/2} Z \rho_B^{1/2}\right) := J V Z^* V^* J + (\text{tr } K) (\mathbf{1} - Q^A)$$

for any

$$K = \rho_B^{1/2} Z \rho_B^{1/2} \in \rho_B^{1/2} \mathcal{B}(\mathcal{H}_B) \rho_B^{1/2} = \mathcal{K}.$$

By the above formula, it is evident that Γ_ρ is well defined (note that $\rho_B^{1/2} Z \rho_B^{1/2} = \rho_B^{1/2} \tilde{Z} \rho_B^{1/2}$ implies $V Z V^* = V \tilde{Z} V^*$), that it is linear (because both the adjoint map and J are anti-linear), that it is a positive map from \mathcal{K} to \mathcal{H}_A , and that $\Gamma_\rho(\rho_B) = \mathbf{1}$.

It is easy to see that

$$\rho((\mathbf{1} - Q^A) \otimes \mathbf{1}) = 0,$$

and hence that the part $(\text{tr } K)(\mathbf{1} - Q^A)$ appearing in the definition of $\Gamma_\rho(K)$, can be ignored when considering the composition $\Phi_\rho \circ \Gamma_\rho$. Thus, for any $T \in \mathcal{B}(\mathcal{H}_B)$, and for Z and K as before, we have

$$\begin{aligned} \text{tr } \Phi_\rho(\Gamma_\rho(K))T &= \text{tr } \rho(\Gamma_\rho(K) \otimes T) = \langle \Psi, (JVZ^*V^*J \otimes T)\Psi \rangle \\ &= \sum_{n,m \in S} \lambda_n \lambda_m \langle e_n^A \otimes e_n^B, (JVZ^*V^*J \otimes T)(e_m^A \otimes e_m^B) \rangle \\ &= \sum_{n,m \in S} \lambda_n \lambda_m \langle Z^* e_m^B, e_n^B \rangle \langle e_n^B, T e_m^B \rangle \\ &= \sum_{n,m \in S} \lambda_m \langle e_m^B, Z \rho_B^{1/2} T e_m^B \rangle = \text{tr } \rho_B^{1/2} Z \rho_B^{1/2} B = \text{tr } KT, \end{aligned}$$

showing that $\Phi_\rho(\Gamma_\rho(K)) = K$ as claimed. \square

Suppose now that ω is realizable by the use of a bipartite quantum system — with parts corresponding to the Hilbert spaces \mathcal{H}_A and \mathcal{H}_B — prepared in the state given by the density operator ρ . When defining $C_n^\omega(X \rightarrow Y)$, we needed to consider all protocols involving four different kinds of codings (two on the sender side and two on the receiver side). It is not difficult to see that all these codings can be incorporated into the choice of partitions / positive operator valued measures, and hence that $\mathcal{N} \in C_n^\omega(X \rightarrow Y)$ if and only if, for each $x \in X$ and $r \in [n]$, there exist a partition of unity $(F_s^{(x)})_{s \in [n]}$ on \mathcal{H}_A and a partition of unity $(E_y^{(r)})_{y \in Y}$ on \mathcal{H}_B such that

$$\mathcal{N}(y|x) = \sum_{r=1}^n \text{tr } \rho(F_r^{(x)} \otimes E_y^{(r)})$$

for all $x \in X$ and $y \in Y$. In particular, if ρ is a density operator corresponding to a *maximally entangled state*; i.e., if $d := \dim \mathcal{H}_A = \dim \mathcal{H}_B < \infty$, ρ is pure and $\text{tr}_A \rho = (1/d)\mathbf{1}$, then, by Lemma 1, the channel \mathcal{N} is in $C_n^\omega(X \rightarrow Y)$ if and only if, for each $x \in X$ and $r \in [n]$, there exists a positive decomposition $(\beta_s^{(x)})_{s \in [n]}$ of $\mathbf{1}/d$ and a partition of unity $(E_y^r)_{y \in Y}$ such that

$$\mathcal{N}(y|x) = \sum_{r=1}^n \text{tr } E_y^r \beta_r^{(x)}$$

for all $x \in X$ and $y \in Y$. In what follows, we will apply the above formula specifically with $n = 2$, and use the notation E_y^\pm and $\beta_\pm^{(x)}$ rather than E_y^r ($r = 1, 2$) and $\beta_r^{(x)}$ ($r = 1, 2$).

3. MAIN RESULT

Our goal is to show that a classical bit assisted by a maximally entangled quantum state can be simulated by two classical bits assisted only by shared randomness. The proof relies on the method that was used in [6, 7] to obtain simulation results. We shall need the following trace inequality.

Lemma 2. *For any operators $\mathbf{0} \leq E^\pm \leq \mathbf{1}$ and $\beta_\pm \geq \mathbf{0}$ such that $\beta_+ + \beta_- =: \rho_B$ is a density operator, we have*

$$|\operatorname{tr} E^+ E^- \rho_B|^2 \leq \operatorname{tr} E^+ \beta_+ + \operatorname{tr} E^- \beta_-.$$

Proof. Set $c_\pm = \operatorname{tr} E^\pm \beta_\pm$ and $t_\pm = \operatorname{tr} \beta_\pm$; then c_\pm and t_\pm are all non-negative, and $t_+ + t_- = 1$. Using the Cauchy–Schwarz inequality $|\operatorname{tr} AB|^2 \leq (\operatorname{tr} A^* A) \cdot (\operatorname{tr} B^* B)$, we have

$$\begin{aligned} |\operatorname{tr} E^+ E^- \beta_+|^2 &= \left| \operatorname{tr} \beta_+^{1/2} E^+ E^- \beta_+^{1/2} \right|^2 \leq \operatorname{tr} ((E^+)^2 \beta_+) \cdot \operatorname{tr} ((E^-)^2 \beta_+) \\ &\leq (\operatorname{tr} E^+ \beta_+) \cdot \operatorname{tr} \beta_+ = c_+ t_+, \end{aligned}$$

and, similarly, $|\operatorname{tr} \beta_- E^+ E^-|^2 \leq c_- t_-$ by interchanging $+$ and $-$ throughout. Therefore,

$$\begin{aligned} |\operatorname{tr} E^+ E^- \rho_B|^2 &= |\operatorname{tr} E^+ E^- \beta_+ + \operatorname{tr} E^+ E^- \beta_-|^2 \\ &\leq (|\operatorname{tr} E^+ E^- \beta_+| + |\operatorname{tr} E^+ E^- \beta_-|)^2 \\ &\leq \left(\sqrt{c_+ t_+} + \sqrt{c_- t_-} \right)^2. \end{aligned}$$

Computing this last square we find that

$$\begin{aligned} \left(\sqrt{c_+ t_+} + \sqrt{c_- t_-} \right)^2 &= c_+ t_+ + c_- t_- + 2 \sqrt{(c_+ t_+)(c_- t_-)} \\ &\leq c_+ t_+ + c_- t_- + 2 \frac{c_+ t_- + c_- t_+}{2} = c_+ + c_- \end{aligned}$$

by the inequality between the geometric and arithmetic means and the fact that $t_+ + t_- = 1$. Putting together the last two inequalities, we have $|\operatorname{tr} E^+ E^- \rho_B|^2 \leq c_+ + c_-$, as claimed. \square

Theorem 3. *Let ω be a non-signaling resource realizable by the use of a bipartite quantum system prepared in a maximally entangled state. Then $C_2^\omega(X \rightarrow Y) \subseteq C_4^{SR}(X \rightarrow Y)$ for any finite alphabets X and Y ,*

i.e., a classical bit assisted by ω can be simulated by two classical bits assisted only by shared randomness.

Proof. Let $l = |X|$ and $k = |Y|$. The Theorem is equivalent to the statement that any $k \times l$ matrix $A = (a_{ij})_{i,j}$ with entries $a_{ij} = \text{tr } E_i^+ \beta_+^{(j)} + \text{tr } E_i^- \beta_-^{(j)}$, where E_i^\pm and $\beta_\pm^{(j)}$ are $d \times d$ positive semidefinite matrices with $E_1^+ + \dots + E_k^+ = E_1^- + \dots + E_k^- = \mathbf{1}$ and $\beta_+^{(j)} + \beta_-^{(j)} = \mathbf{1}/d$ for all $j \in [l]$, is a convex combination of stochastic matrices with at most four non-zero rows.

For $I = (i_1, i_2, i_3, i_4) \in [k]^4$, put

$$(3.1) \quad p_I = \frac{1}{d^2} (\text{tr } E_{i_1}^+ E_{i_2}^-) (\text{tr } E_{i_3}^+ E_{i_4}^-).$$

We have $p_I \geq 0$ for all I . Thus, we get a measure P on $[k]^4$ defined by $P(T) = \sum_{I \in T} p_I$. Due to the multilinear nature of (3.1) and the assumption that E_1^\pm, \dots, E_k^\pm is a partition of unity (POVM), we see that

$$P([k]^4) = \frac{1}{d^2} (\text{tr}(\mathbf{1}^2)) (\text{tr}(\mathbf{1}^2)) = 1,$$

so P is a probability measure. Now set $E_S^\pm := \sum_{i \in S} E_i^\pm$ for any $S \subseteq [k]$. Since $\mathbf{0} \leq E_S^\pm \leq \mathbf{1}$, we may apply Lemma 2 with $\rho_B = \mathbf{1}/d$ to get

$$P(S^4) = \frac{1}{d^2} (\text{tr } E_S^+ E_S^-)^2 \leq \text{tr } E_S^+ \beta_+^{(j)} + \text{tr } E_S^- \beta_-^{(j)}$$

for all j . The right hand side here is $A_j(S)$, where A_j is the probability measure on $[k]$ given by the numbers a_{ij} ($i \in [k]$); i.e. the j^{th} column of the matrix A . So we have

$$A_j(S) \geq P(S^4) \quad \text{for all } S \subseteq [k].$$

Let us connect $I \in [k]^4$ to $i \in [k]$ by an edge if i occurs in I . This gives us a bipartite graph. The neighborhood of any set $T \subseteq [k]^4$ is the set $S \subseteq [k]$ of indices occurring in some element of T . We always have $T \subseteq S^4$, whence $A_j(S) \geq P(S^4) \geq P(T)$. Thus, by the Supply–Demand Theorem [8, 2.1.5. Corollary], and using the fact that both A_j and P are probability measures, there exists a probability measure P_j on $[k]^4 \times [k]$ which is supported on the edges of the graph and has marginals P and A_j . Whenever $p_I \neq 0$, let $B(I)$ be the $k \times l$ stochastic matrix whose j -th column is given by the conditional distribution $P_j|I$ on $[k]$. Now $B(I)$ has at most four nonzero rows, and $A = \sum p_I B(I)$, as desired. \square

Remark 4. Suppose that our bipartite quantum system is *not* in a maximally entangled state, and hence ρ_B is not (necessarily) a multiple

of the identity. Still, the above proof could be virtually copied if we had a bilinear, scalar-valued map D satisfying

- (i) $D(Z_1, Z_2) \geq 0$ whenever $Z_1, Z_2 \geq \mathbf{0}$,
- (ii) $D(\mathbf{1}, \mathbf{1}) = 1$,
- (iii) $|D(E^+, E^-)|^2 \leq \text{tr } E^+ \beta_+ + \text{tr } E^- \beta_-$ whenever $\mathbf{0} \leq E^\pm \leq \mathbf{1}$ and $\rho_B = \beta_+ + \beta_-$ is a positive decomposition of ρ_B .

Indeed, having such a bilinear map, we could replace (3.1) by setting

$$p_I = D(E_{i_1}^+, E_{i_2}^+) D(E_{i_3}^+, E_{i_4}^+)$$

and continue the rest of the argument unchanged. Actually, in the proof we *did* set p_I to be of the mentioned form; specifically, with D being the bilinear map given by the formula $D(Z_1, Z_2) = (1/d) \text{tr } Z_1 Z_2$.

When ρ_B is not necessarily $\mathbf{1}/d$, one could try to replace the previous formula by $D(Z_1, Z_2) = (\text{tr } Z_1 Z_2 \rho_B + \text{tr } Z_2 Z_1 \rho_B)/2$. This reduces to the previous one when $\rho_B = \mathbf{1}/d$, and it satisfies requirements (ii) and (iii); this latter one follows from Lemma 2 and the fact that for self-adjoint operators E^\pm , we have

$$|D(E^+, E^-)|^2 = (\text{Re}(\text{tr } E^+ E^- \rho_B))^2 \leq |\text{tr } E^+ E^- \rho_B|^2.$$

However, this D does not satisfy the positivity condition (i) — unless of course ρ_B is a multiple of the identity.

Another idea is to try setting $D(Z_1, Z_2) = \text{tr } Z_1 \rho_B^{1/2} Z_2 \rho_B^{1/2}$, which again reduces to the formula used in our proof in case ρ_B is a multiple of the identity. The thus defined D is evidently bilinear and satisfies both the positivity (i) and the normalization (ii) requirements. However, examples show that in general it fails to satisfy requirement (iii) — unless, for example, if ρ_B is a multiple of a projection.

Having experimented with various candidate formulas, we grew skeptical about the possibility of simultaneously satisfying all listed requirements. Thus, while we still believe that the theorem remains true even if arbitrary entangled states are allowed, we expect the general proof to follow a somewhat different direction.

APPENDIX A. THE “TWO WINNING, TWO LOSING BOXES” GAME

Let $\rho = |\Psi\rangle\langle\Psi|$, where $\Psi = \frac{1}{\sqrt{2}}(e_1 \otimes e_2 - e_2 \otimes e_1)$ and (e_1, e_2) is the standard basis of \mathbb{C}^2 . Before the game begins, Alice and Bob prepare a pair of quantum bits in the state given by ρ ; Alice then takes the first, Bob the second quantum bit with herself / himself. Upon learning the positions $a, b \in \{1, 2, 3, 4\}$ of treasures, Alice performs the measurement corresponding to the 2×2 partition of unity $F_+^{\{a,b\}}$, $F_-^{\{a,b\}}$ and sends the result, a + or a - sign, to Bob via the noiseless one-bit channel.

For the specific protocol we want to describe, we will have $F_+^{\{1,2\}} = \mathbf{1}$, $F_-^{\{1,2\}} = \mathbf{0}$ (i.e., in case the treasures are in the first two boxes, Alice will surely send a “+” to Bob), $F_\pm^{\{1,3\}} = (1/2)(\mathbf{1} \pm \sigma_z)$, $F_\pm^{\{2,4\}} = (1/2)(\mathbf{1} \mp \sigma_z)$, $F_\pm^{\{1,4\}} = (1/2)(\mathbf{1} \pm \sigma_x)$, $F_\pm^{\{2,3\}} = (1/2)(\mathbf{1} \mp \sigma_x)$, where σ_z and σ_x are two Pauli matrices, and, finally, $F_+^{\{3,4\}} = 0$, $F_-^{\{3,4\}} = I$ (so that in case the treasures are in the last two boxes, Alice will surely send a “−” to Bob).

After receiving the + or − sign from Alice, Bob performs the measurement corresponding to the partition of unity $E_1^\pm, E_2^\pm, E_3^\pm, E_4^\pm$ and chooses the box according to the result. We will specifically have $E_1^+ = (1/2)(\mathbf{1} - (\sigma_z + \sigma_x)/\sqrt{2})$, $E_2^+ = \mathbf{1} - E_1^+$, $E_3^+ = 0$, $E_4^+ = 0$ and $E_1^- = 0$, $E_2^- = 0$, $E_3^- = (1/2)(\mathbf{1} + (\sigma_z - \sigma_x)/\sqrt{2})$, $E_4^- = \mathbf{1} - E_3^-$.

As $E_3^+ = E_4^+ = 0$ and likewise, $E_1^- = E_2^- = 0$, Bob will always choose one of the first two boxes if he receives a +, and one of the last two boxes if he receives a − sign. Hence if the two treasure boxes are either the first two or the last two, they will win with certainty. On the other hand, if the treasures are e.g. in boxes 1 and 3, then they win with probability

$$\text{tr } \rho \left(F_+^{\{1,3\}} \otimes (E_1^+ + E_3^+) \right) + \text{tr } \rho \left(F_-^{\{1,3\}} \otimes (E_1^- + E_3^-) \right),$$

which, after substitution, turns out to be $\frac{1}{2} + \frac{1}{4}\sqrt{2}$. It turns out that all other cases result in the same probability of success, yielding the claimed overall winning probability of $(4 + \sqrt{2})/6$. We finish the discussion of this example by pointing out that all listed measurements are either trivial or projective; the entire protocol can be easily realized experimentally using e.g. a pair of spin-half particles prepared in the zero-total-spin state and spin measurements performed on individual particles.

REFERENCES

- [1] A. Acín, T. Durt, N. Gisin and J.I. Latorre: Quantum nonlocality in two three-level systems. *Phys. Rev. A* **65** (2002), 052325.
- [2] C. H. Bennett and S. J. Wiesner: Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states. *Phys. Rev. Lett.* **69** (1992), pg. 2881–2884.
- [3] T. S. Cubitt, D. Leung, W. Matthews and A. Winter: Improving Zero-Error Classical Communication with Entanglement. *Phys. Rev. Lett.* **104** (2010), 230503.
- [4] T. S. Cubitt, D. Leung, W. Matthews and A. Winter: Zero-Error Channel Capacity and Simulation Assisted by Non-Local Correlations. *IEEE Trans. Inf. Theory* **57** (2011), pg. 5509–5523.

- [5] M. Dall’Arno, S. Brandsen, A. Tosini, F. Buscemi and V. Vedral: No-hypersignaling principle. *Phys. Rev. Lett.* **119** (2017), 020401.
- [6] P. E. Frenkel, Classical simulations of communication channels, arXiv:2101.10985
- [7] P.E.Frenkel and M.Weiner: Classical information storage in an n -level quantum system. *Commun. Math. Phys.* **340** (2015), pg. 563–574.
- [8] L. Lovász and M. D. Plummer: Matching Theory. North-Holland, 1986.
- [9] K. Matsumoto and G. Kimura: Information storing yields a point-asymmetry of state space in general probabilistic theories, arXiv:1802.01162
- [10] M. Pawłowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter and M. Zukowski: Information causality as a physical principle. *Nature* **461** (2009), pg.1101–1104.

EÖTVÖS LORÁND UNIVERSITY, PÁZMÁNY PÉTER SÉTÁNY 1/C, BUDAPEST,
1117 HUNGARY, AND RÉNYI INSTITUTE, BUDAPEST, REÁLTANODA U. 13-15,
1053 HUNGARY

Email address: `frenkelp265@gmail.com`

BUDAPEST UNIVERSITY OF TECHNOLOGY AND ECONOMICS (BME), DEPARTMENT OF ANALYSIS, H-1111 BUDAPEST MŰEGYETEM RKP. 3–9 HUNGARY, AND MTA-BME LENDÜLET QUANTUM INFORMATION THEORY RESEARCH GROUP

Email address: `mweiner@math.bme.hu`