Simple information-processing tasks with unbounded quantum advantage

Teiko Heinosaari [1,*], Oskari Kerppo [1,†], Leevi Leppäjärvi [2,‡] and Martin Plávala [3,§]
[1] Faculty of Information Technology, University of Jyväskylä, 40100 Jyväskylä, Finland
[2] RCQI, Institute of Physics, Slovak Academy of Sciences, Dúbravská cesta 9, 84511 Bratislava, Slovakia
[3] Naturwissenschaftlich-Technische Fakultät, Universität Siegen, Walter-Flex-Straße 3, 57068 Siegen, Germany

(Received 31 August 2023; accepted 4 March 2024; published 29 March 2024)

Communication scenarios between two parties can be implemented by first encoding messages into some states of a physical system which acts as the physical medium of the communication and then decoding the messages by measuring the state of the system. We show that already in the simplest possible scenarios it is possible to detect a definite, unbounded advantage of quantum systems over classical systems. We do this by constructing a family of operationally meaningful communication tasks, each of which, on the one hand, can be implemented by using just a single qubit but which, on the other hand, require an unboundedly larger classical system for classical implementation. Furthermore, we show that even though, with the additional resource of shared randomness, the proposed communication tasks can be implemented by both quantum and classical systems of the same size, the number of coordinated actions needed for the classical implementation also grows unboundedly. In particular, no finite storage can be used to store all the coordinated actions required to implement all possible quantum communication tasks with classical systems. As a consequence, shared randomness cannot be viewed as a free resource.

DOI: 10.1103/PhysRevA.109.032627

I. INTRODUCTION

One of the most important ongoing investigations in quantum information is to find information-processing tasks that can be used to demonstrate a quantum advantage. This question is equally important for quantum computing, where it is still unclear whether the current noisy intermediate-scale quantum devices [1] can demonstrate a computational advantage over classical computers. Therefore, it is of utmost importance to identify scenarios where quantum devices can offer some benefit over devices operating in the classical domain. Understanding the basic ingredients of quantum advantage forms the foundation for quantum technological applications.

A simple yet important information-processing task is the one-way-communication scenario, where one party, Alice, is trying to send an encoded message to another party, Bob. The message is prepared with a preparation device, and the encoded message, carried by some physical medium, is decoded with a measurement device. While this scenario may seem far too simple at first to reveal anything interesting, large separations between classical and quantum protocols in such scenarios have been shown in entanglement-based setups [2,3], in the presence of shared randomness [4–6], and in other settings [7–17]. Exponential separations are also known in the one-way-communication complexity scenario [18–20], where the aim is to compute the value of a multivariate function; interestingly, it is also known that the classical-quantum separation cannot be arbitrarily large in these scenarios [21,22]. Moreover, one can also see any channel-based quantum computation as a preparation and subsequent measurement of a quantum state and the aforementioned results as proof that quantum computers have an advantage over classical ones when sampling from some conditional probability distributions. These results are especially important in light of boson sampling [23–27], which is a nonuniversal model of quantum computation based on sampling from specific conditional probability distributions.

A customary assumption in the quantum information literature is that the preparation and measurement devices are classically correlated via shared randomness, making the sets of implementable communication scenarios convex and easier to analyze. Importantly, it is known that in the presence of shared randomness quantum communication does not offer any advantage over classical communication as the sets of implementable scenarios are the same [28]. This hints that either quantum physics does not offer a great advantage in one-way communication or shared randomness should not be considered a free resource in these scenarios. In the following we show that the latter is the right conclusion by constructing communication scenarios that require unbounded amounts of shared randomness.

The starting point for our current investigation is the fact that in the absence of shared randomness the sets of implementable communication scenarios are not convex [8,29]; see Fig. 1 for an illustration. It is clear that an information-processing advantage can exist only if the sets of implementable communication scenarios are different

*teiko.heinosaari@jyu.fi
†oskari.e.o.kerppo@jyu.fi
‡leevi.leppajarvi@savba.sk
§martin.plavala@uni-siegen.de

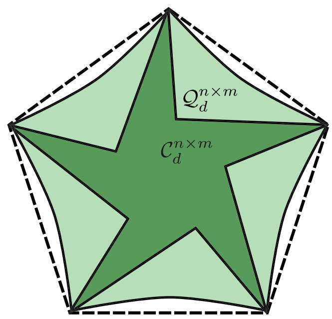

FIG. 1. Illustration of the quantum and classical one-way communication scenarios represented by the sets of communication matrices $\mathcal{Q}_d^{n\times m}$ and $\mathcal{C}_d^{n\times m}$ with fixed size $n \times m$ and the same dimension d, respectively. The convex closures of both sets are the same (dashed area), although the sets are different.

between two communication mediums; it was previously shown that the sets of quantum and classical communication scenarios with systems of fixed dimension are not the same, and they have some rather large discrepancies [29]. This raises the paramount question of whether quantum communication can be simulated with a classical system of larger dimension.

In this work we argue that quantum communication cannot be reliably simulated with larger classical systems. This is shown by introducing a family of simple communication scenarios based on antidistinguishability which can be implemented with a qubit but require arbitrarily large classical systems or an arbitrary amount of classical resources in the form of classical communication or shared randomness in the limit. We then argue that the use of shared randomness cannot be considered a free resource for communication because the difference in the sizes of the quantum and classical systems needed to implement the communication task puts a limit on how much shared randomness is needed to implement quantum communication classically.

II. INFORMATION-PROCESSING TASKS

In a communication scenario we can envision the following setup: Alice will be made aware of the value of a random variable a, while Bob remains unaware of it. Prior to the task, they have the opportunity to meet and establish any strategy they desire. However, once they start, Alice is restricted from freely communicating with Bob. Instead, she is provided with a d-dimensional quantum system or, alternatively, a d-dimensional classical system. Alice is revealed the value of a and has the liberty to prepare this system in any desired state and hand it over to Bob. Subsequently, Bob's task is to specify the value of a. The task can also be more general, in which case, Bob is expected to return different values $b_1, \ldots, b_m$ with some probabilities $p_1, \ldots, p_m$ determined by the input a. For instance, it could be required that for some inputs Bob will erase all information, meaning that he delivers a uniform probability distribution. In a computing scenario the setup is similar, but Alice and Bob can be the same person. In this case, there is a function f, and Alice's task is to produce $f(a)$ for any input a drawn from some set of possible inputs.

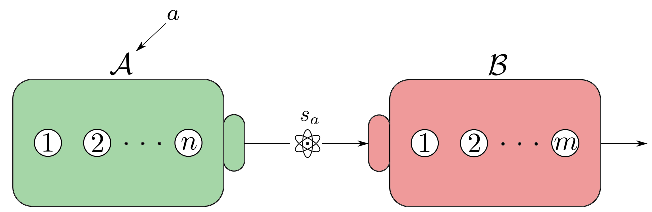

FIG. 2. Basic one-way communication setup between two parties, Alice and Bob. In each round of communication Alice receives a random variable a and prepares a state s_a from her state ensemble. Bob performs a measurement of M and receives an outcome b.

The role of the communication channel is taken by a quantum processor.

To cover all these information-processing scenarios, we use the formalism of communication matrices [29,30], also called channel matrices [28]. We assume that Alice possesses a finite collection of states, referred to as a state ensemble. Alice selects a label a and transmits a quantum (or classical) system in the corresponding state s_a to Bob. Bob then performs a measurement using a fixed measurement device M, which yields b, one of the possible outcomes $\{1, \ldots, m\}$. The entire set of conditional probabilities that describes this preparation-measurement scenario is represented by an $n \times m$ matrix C, given as $C_{ab} = \mathrm{tr}[s_a M_b]$. This matrix is row stochastic, i.e., has non-negative entries, and the sum of each row is 1. This basic communication scenario is illustrated in Fig. 2.

The previously mentioned information-processing tasks can be written as communication matrices. For instance, the task of perfectly distinguishing d states corresponds to the identity matrix $\mathbb{1}_d$. On the other hand, consider a communication task in which there are three possible labels and Alice is able to transmit the first two without error while the third input leads to an ambiguous result. This corresponds to the matrix

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}, \tag{1}$$

where the third row shows that when Alice sends the third label to Bob, Bob randomly gets one of the first labels in his measurement.

To compare the difficulty of the implementation of communication matrices corresponding to different tasks, we recall that there is a physically motivated preorder (i.e., partial order without being antisymmetric) called ultraweak matrix majorization in the set of communication matrices [29]. We write $C \preceq D$ and say a communication matrix C is ultraweakly majorized by another communication matrix D if row-stochastic matrices L and R exist such that $C = LDR$. The matrices L and R have a natural interpretation as pre- and postprocessing matrices of the preparations and measurement outcomes: in the communication task C one uses the mixtures of states used in task D according to the convex weights provided by the matrix L, and the measurement used in task C is a classical postprocessing given by the matrix R of the measurement used in task D. Thus, the preorder gives a precise mathematical definition of the level of difficulty among the tasks: if $C \preceq D$, then performing the task C cannot be any harder than performing D since the implementation of D can be used to implement C as well. In the case when $C \preceq D$ and $D \preceq C$ we say that C and D are ultraweakly equivalent, and then they represent equally hard communication tasks.

In the computational scenario a $0/1$ matrix can be interpreted as a Boolean function, and we will refer to $0/1$ matrices as deterministic matrices. For instance, the logic gates NOT and XOR correspond to the matrices

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}. \tag{2}$$

The NOT setup simply means that Bob reads the opposite label that Alice sends. In the XOR setup Alice can send all ordered pairs from $\{0, 1\}$; hence, the matrix has four rows. It is straightforward to verify that the communication matrices in (2) are ultraweakly equivalent.

There is a relatively widely adopted view that a single qubit on its own is not more powerful than a bit. This view is motivated by Holevo's theorem [31], which states that a single qubit can transmit only one bit of information. Indeed, we make the following observation.

Proposition 1. Let C be a deterministic communication matrix. Then C is implementable with classical and quantum systems of the same size.

Proof. Let C be a $0/1$ matrix, i.e., a deterministic communication matrix. It is straightforward to see that any communication matrix is ultraweakly equivalent with the matrix that is obtained by removing the zero columns (corresponding to measurement outcomes which never occur) and duplicate rows (some of the states that are used in the implementation are the same). Denote by C' the ultraweakly equivalent matrix that is obtained from C in this way. But now C' is clearly a permutation of an identity matrix of some size d, depending on how many duplicate rows were removed. As C' is equivalent to $\mathbb{1}_d$ and the ultraweak relation is clearly transitive, we conclude that C is implementable by d-level classical and quantum systems as both of these systems can transmit d distinct messages without error by the basic decoding theorem [32]. ∎

By Proposition 1 we are forced to conclude that by looking at deterministic communication matrices, quantum systems do not seem to offer any advantage over classical systems. However, as we will see, this is far from the whole truth because there are many interesting and important communication matrices that are not deterministic. In the current framework it becomes natural to look at all possible communication matrices that can be implemented with a d-dimensional quantum or classical system. We denote these sets as $\mathcal{Q}_d$ and $\mathcal{C}_d$, respectively. We note that both of these sets consist of arbitrarily large matrices. Both of them also contain the identity matrix $\mathbb{1}_d$ but not $\mathbb{1}_n$ for any $n > d$.

The sets $\mathcal{Q}_d$ and $\mathcal{C}_d$ do not have convex structure as they have matrices of different sizes. To form convex mixtures, we have to limit ourselves to matrices of a certain size, and we denote by $\mathcal{Q}_d^{n\times m}$ and $\mathcal{C}_d^{n\times m}$ the matrices with size $n \times m$. Forming a convex combination of two communication matrices of the same size corresponds to a coordinated action with Alice and Bob. As mentioned previously, the sets $\mathcal{Q}_d^{n\times m}$ and $\mathcal{C}_d^{n\times m}$ are not convex [8,33], but in the seminal work in [28] it was shown that the convex closures of $\mathcal{Q}_d^{n\times m}$ and $\mathcal{C}_d^{n\times m}$ are the same; the convex structure of these sets is illustrated in Fig. 1. It follows that shared randomness is a genuinely distinguished resource in the setup, but with unlimited use of shared randomness one can simulate any quantum prepare-measure scenario with classical settings. Our main result is to show that this would require infinitely many rounds of coordinated actions.

III. CLASSICAL AND QUANTUM DIMENSIONS

To explain the information-processing tasks that cannot be effectively simulated with a classical system with the same dimension, we need some additional tools. The *quantum dimension* of a communication matrix C, denoted by $\dim_{\mathcal{Q}}(C)$, is the smallest integer d such that $C_{ab} = \mathrm{tr}[s_a M_b]$, where the states $\{s_a\}$ and effects $\{M_b\}$ act on a d-dimensional Hilbert space. As the matrix C has finite size, the quantum dimension of a fixed communication matrix is also always finite. Thus, for a communication matrix C we have $C \in \mathcal{Q}_d$ if and only if $\dim_{\mathcal{Q}}(C) \leqslant d$. It can be shown that the quantum dimension of a communication matrix C is equal to the positive-semidefinite (PSD) rank of C [34,35], denoted by $\mathrm{rank}_{\mathrm{PSD}}(C)$, since a PSD decomposition can always be normalized to give density operators and effects in the decomposition.

The *classical dimension* of a communication matrix C, denoted by $\dim_{\mathcal{C}}(C)$, is the smallest integer d such that C can be obtained using a classical system with d distinct states. Equivalently, it is the smallest integer d such that the implementation of the matrix is obtained by using the standard embedding of classical d-dimensional theory in the quantum formalism by using only diagonal density matrices and effects of dimension d. It then follows that $C \in \mathcal{C}_d$ if and only if $\dim_{\mathcal{C}}(C) \leqslant d$. Furthermore, it can be shown that $\dim_{\mathcal{C}}(C)$ is then the smallest integer d such that row-stochastic matrices L and R exist such that $C = L\mathbb{1}_d R$, i.e., $C \preceq \mathbb{1}_d$ [29]. Thus, the communication tasks with a classical implementation of dimension d are exactly those that can be obtained from the task of perfectly distinguishing d states by means of ultraweak matrix majorization, i.e., in terms of pre- and postprocessing of states and measurement outcomes. From this equivalent formulation it is then clear that mathematically, the classical dimension of a communication matrix C exactly coincides with the concept of a non-negative rank of C [36], denoted by $\mathrm{rank}_+(C)$.

The connection between the quantum and classical communication complexity and the PSD and non-negative ranks is known in the literature (see, e.g., [34]). It is also known that these ranks are very difficult to compute [37,38]. Nontrivial upper bounds on the non-negative rank were derived only relatively recently [39,40]. Our strategy is to define a family of communication matrices with fixed quantum dimension and show that the classical dimension of these matrices cannot be bounded from above. We note that previously, for some matrices (such as the Euclidean distance matrices) an exponential separation between the non-negative and PSD ranks was shown [41,42]. However, our aim is to construct actual communication matrices which have a physically meaningful interpretation as clear communication tasks such that they still experience the same exponential separation. To this end, we make the following definition.

Definition 1. Let $n \geqslant 2$ be an integer. Define Bloch vectors $\{\vec{r}_a\}_{a=1}^n \subset \mathbb{R}^3$ by $\vec{r}_a = (\cos(\frac{2a\pi}{n}), 0, \sin(\frac{2a\pi}{n}))$, so that $s_a = \frac{1}{2}(\mathbb{1} + \vec{r}_a \cdot \vec{\sigma})$, where $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ is composed of the Pauli matrices, defines a pure state for all a. Define the corresponding effects by $M_b = \frac{1}{n}(\mathbb{1} - \vec{r}_b \cdot \vec{\sigma})$, so that $\mathrm{tr}[s_a M_b] = \frac{2}{n}\sin^2(\frac{(a-b)\pi}{n})$. Note that $\mathrm{tr}[s_c M_c] = 0$ for all c, and otherwise, $\mathrm{tr}[s_a M_b] > 0$. We denote by A_n the $n \times n$ matrix with elements $\mathrm{tr}[s_a M_b]$ at the ath row and bth column.

The matrix A_n defines an instance of communication tasks known as antidistinguishability [43–47]. As all diagonal entries of A_n are zero, it corresponds to the task in which upon obtaining outcome b Bob knows that Alice did not send state s_b. Antidistinguishability plays an important role in quantum information and foundations, as evidenced by its role in the influential Pusey–Barrett–Rudolph [48] theorem and its connection to noncontextuality inequalities [49]. Our motivation in defining A_n is that clearly, this matrix has a quantum dimension equal to 2 for any n. We now proceed to show that the classical dimension of A_n cannot be bounded from above.

Lemma 1. The non-negative rank of A_n scales as $\mathrm{rank}_+(A_n) = \Omega(\log_2 n)$.

The result follows from Theorems 5, 6, and 8 of [39]. The details can be found in Appendix A. We now have all the necessary tools to prove one of our main results.

Theorem 1. For all integers $d, m \geqslant 2$ there is a communication matrix C such that $\dim_{\mathcal{Q}}(C) \leqslant d$ while $\dim_{\mathcal{C}}(C) \geqslant m$.

Proof. It is sufficient to consider $d = 2$. Let $m \geqslant 2$ and $n = 2^m$. It is guaranteed by Lemma 1 that $\dim_{\mathcal{C}}(A_n) = \mathrm{rank}_+(A_n) \geqslant m$. ■

Our result shows that there is a sequence of communication tasks that can be implemented with a fixed-size quantum system but which cannot be implemented by any classical system of fixed size. In the limit of these tasks we arrive at an unbounded quantum advantage over any classical implementation. In the literature quantum advantages are typically shown in scenarios involving inputs for both parties of the communication protocol. In the present work we study the scenario where only the party preparing states receives an input, and we are thus able to show the advantage with communication tasks of the simplest type with a very clear and important physical motivation, namely, antidistinguishability of states. This results provides yet another way to see how antidistinguishability can be used to surface some of the fundamental nonclassical aspects of quantum theory as well as what type of information-processing capabilities quantum systems hold. On the other hand, our result can also be used as a witness for the classical dimension of the implementation: if we know that we are using a classical system to implement an antidistiguishability task of a certain size, then we can use the lower bound from Lemma 1 to deduce what the dimension of the system must at least be. Additionally, we consider it important that we managed to use preparations and measurements that belong only to the XZ plane of the Bloch sphere since this immediately shows that our construction is also valid for real quantum theory (and also for PSD rank based on real Hilbert spaces).

From the technical side a natural question is whether it is possible to improve the scaling of the non-negative rank in Lemma 1. A nontrivial upper bound on the non-negative rank of $\lceil \frac{6n}{7} \rceil$ was proven in [40]. Therefore, better scaling could be possible, but we leave that as an open problem.

Example 1. A_7 is a matrix with elements $(A_7)_{ab} = \frac{2}{7}\sin^2(\frac{(a-b)\pi}{7})$, with $a, b \in \{1, 2, \ldots, 7\}$. Define $i = \frac{2}{7}\sin^2(\frac{\pi}{7})$, $j = \frac{2}{7}\sin^2(\frac{2\pi}{7})$, and $k = \frac{2}{7}\sin^2(\frac{3\pi}{7})$. Then

$$A_7 = \begin{bmatrix} 0 & i & j & k & k & j & i \\ i & 0 & i & j & k & k & j \\ j & i & 0 & i & j & k & k \\ k & j & i & 0 & i & j & k \\ k & k & j & i & 0 & i & j \\ j & k & k & j & i & 0 & i \\ i & j & k & k & j & i & 0 \end{bmatrix}. \tag{3}$$

A_7 has the following non-negative factorization because $A_7 = WH$, where

$$W = \begin{bmatrix} 2k & 2j & 0 & 0 & 2i & 0 \\ 0 & 2k & 0 & 0 & \frac{2ik}{k-j} & w \\ 0 & 2j & 2k & 0 & 2i & 0 \\ 0 & 2i & \frac{2k(j-i)}{k-j} & 1-\mathcal{R} & 0 & 0 \\ 0 & 0 & \frac{2ik}{k-j} & 2\left(k - \frac{i^2}{k-j}\right) & 1-\mathcal{R} & w \\ \frac{2ik}{k-j} & 0 & 0 & 2\left(k - \frac{i^2}{k-j}\right) & 1-\mathcal{R} & w \\ \frac{2k(j-i)}{k-j} & 2i & 0 & 1-\mathcal{R} & 0 & 0 \end{bmatrix},$$

$$H = \begin{bmatrix} 0 & \frac{i}{2k} & \frac{j}{2k} & \frac{k-i}{2k} & \frac{k-j}{2k} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{j}{2k} & \frac{i}{2k} & 0 & 0 & 0 & \frac{k-j}{2k} & \frac{k-i}{2k} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ h_1 & 0 & h_1 & h_2 & 0 & 0 & h_2 \end{bmatrix}, \tag{4}$$

where

$$w = 2(i+j) - \frac{2ik}{k-j}, \quad h_1 = \frac{j - \frac{1}{2}\left(k - \frac{i^2}{k-j}\right)}{2(i+j) - \frac{2ik}{k-j}},$$
$$h_2 = \frac{j - \frac{ik}{k-j}}{2(i+j) - \frac{2ik}{k-j}}, \tag{5}$$

and $\mathcal{R}$ stands for the sum of all other elements in the same row. This non-negative factorization was found with the help of a heuristic method [50]. As this decomposition exists, the non-negative rank of A_7 is less than or equal to 6. The bound given in Corollary 4 of [39] can be used to show that equality must hold. The details can be found in Appendix B.

The classical dimension of A_7 provides a counterexample to a conjecture presented in [29], which speculated that the classical dimension of quantum theory equals the quantum dimension squared.

IV. NON-NEGATIVE RANK AND SHARED RANDOMNESS

The first part of our result was to show that the classical dimension of communication matrices with fixed quantum dimension can grow without bound; we now turn to our second result concerning shared randomness. Namely, it is well known that any quantum communication matrix can be obtained using classical communication and shared randomness [28]. Since shared randomness is usually considered a free resource, it would seem that our previous result holds no real consequence and shared randomness could be used as a loophole. Next, we will close this loophole by showing that shared randomness is not, in fact, a free resource.

Suppose Alice and Bob coordinate their actions in the following way: Alice and Bob both have $k \in \mathbb{N}$ different choices for their preparation and measurement devices. That is, in each round of communication, Alice and Bob observe a correlated random variable $k' \in \{1, 2, \dots, k\}$. Alice prepares some state with the preparation device labeled k'. Likewise, Bob uses a measurement device with the label k'. Without loss of generality we can assume that the devices have the same number of inputs and outputs, say, n preparations and m outcomes, so that the coordinated communication matrix can be written as

$$C = \sum_{k'=1}^{k} \alpha_{k'} C_{k'}, \tag{6}$$

where $\alpha_{k'}$ is the probability of sampling k' as the shared variable and $C_{k'}$ is the corresponding communication matrix implemented by those devices. Suppose now that each of the C_i's has a classical dimension equal to d, that is, $C_{k'} = L_{k'} R_{k'}$ for some row-stochastic matrices $L_{k'}$ and $R_{k'}$ of size $n \times d$ and $d \times m$, respectively. Then, clearly,

$$C = [\alpha_1 L_1 \quad \alpha_2 L_2 \quad \dots \quad \alpha_k L_k] \begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_k \end{bmatrix}. \tag{7}$$

This puts an upper bound on the classical dimension of C, namely, $\text{rank}_+(C) \leqslant dk$.

Theorem 2. Suppose Alice and Bob agree beforehand that they will use coordinated actions from a finite set $\{1, 2, \dots, k\}$ and they will communicate with d-dimensional classical systems. Communication matrices with the quantum dimension equal to 2 that they cannot implement classically by coordinating their actions exist.

Proof. The result follows directly from Eq. (7) and Theorem 1, as the maximal classical dimension Alice and Bob can achieve is dk and we have seen in Theorem 1 that the classical dimension of antidistinguishability matrices cannot be bounded. ■

Our second main result thus shows that, ultimately, shared randomness cannot be considered a free resource. In particular, even though the convex closures of $\mathcal{C}_d$ and $\mathcal{Q}_d$ are the same for fixed d [28], one cannot implement all the tasks in $\mathcal{Q}_d$ by using a d-dimensional classical system with any finite amount of shared randomness. In the limit this ultimately leads to the need for an infinite storage for the coordinated actions. It is worth repeating that without this result our first result would lack proper physical and operational interpretation, and hence, it is a key feature that this loophole was closed. Furthermore, there are no further known loopholes remaining in this scenario. On the other hand, our second result can also be used as a witness for shared randomness: there are simple communication tasks which for fixed dimension d can be used to detect how many coordinated actions are minimally needed to implement a given task classically.

V. DISCUSSION

We have shown that it is not possible to simulate quantum communication with larger classical systems reliably. Our result has two consequences. First of all, if the dimension of the communication medium is taken to be proportional to the cost of communication, then quantum communication displays an unbounded advantage. As quantum systems can only give an advantage in the considered simple communication scenarios when the communication matrix is not deterministic, we conclude that the displayed advantage is somehow related to quantum systems generating randomness more effectively than is possible classically. In the example we provided with antidistinguishable matrices the classical dimension scaled at least logarithmically with the size of the matrix. We are unsure whether it is possible to obtain better scaling with other matrices or even what the exact scaling of the antidistinguishable matrices is. We leave this problem for future work.

Second, given that the convex closures of quantum and classical communication matrices coincide, we nonetheless have shown that no finite number of coordinated actions is enough to exactly simulate quantum communication classically when the preparation and measurement devices are correlated. This points us to the conclusion that taking shared randomness as a free resource is an overwhelmingly strong assumption.

ACKNOWLEDGMENTS

T.H. and O.K. acknowledge financial support from Business Finland under the project TORQS, Grant No. 8582/31/2022, and from the Academy of Finland under the mobility funding Grant No. 343228, and under the project DEQSE, Grant No. 349945. L.L. acknowledges support from the European Union's Horizon 2020 research and innovation program under the program SASPRO 2 COFUND, Marie Sklodowska-Curie Grant Agreement No. 945478, as well as from projects APVV-22-0570 (DeQHOST) and VEGA 2/0183/21 (DESCOM). M.P. acknowledges support from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation Projects No. 447948357 and No. 440958198), the Sino-German Center for Research Promotion (Project No. M-0294), the ERC (Consolidator Grant No. 683107/TempoQ), the German Ministry of Education and Research (Project QuKuK, BMBF Grant No. 16KIS1618K), and the Alexander von Humboldt Foundation.

APPENDIX A: PROOF OF LEMMA 1

The restricted non-negative rank of a non-negative matrix A, denoted by $\text{rank}^*_+(A)$, is defined as the smallest

TABLE I. First (relevant) values of the upper bounds for the restricted non-negative rank given in [39].

	r_+				
	3	4	5	6	7
$\phi'(r_+)$	3	4	6	9	14
$\phi_3(r_+)$	3	6	10	18	30

decomposition of A into $A = WH$ with W and H being non-negative and $\mathrm{rank}(A) = \mathrm{rank}(W)$. Clearly, if A is of size $n \times m$, then $\mathrm{rank}(A) \leqslant \mathrm{rank}_+(A) \leqslant \mathrm{rank}_+^*(A) \leqslant m$. The authors of [39] showed that the restricted non-negative rank is related to the nested-polytopes problem, which they used to show bounds on the restricted non-negative rank and the non-negative rank. In particular, they gave the following upper bound on the restricted non-negative rank.

Theorem 3 ([39], Theorem 5). The restricted non-negative rank of a non-negative matrix M with $r = \mathrm{rank}(M)$ and $r_+ = \mathrm{rank}_+(M)$ can be bounded above by

$$\mathrm{rank}_+^*(M) \leqslant \max_{r \leqslant r_u \leqslant r_+} \mathrm{faces}(r_+, r_u - 1, r_u - r) =: \phi_r(r_+), \tag{A1}$$

where $\mathrm{faces}(n, d, k)$ is the maximal number of k faces of a polytope with n vertices in dimension d, which can be calculated as

$$\mathrm{faces}(n, d, k) = \sum_{i=0}^{\frac{d}{2}}{}^{*} \left[\binom{d-i}{k+1-i} + \binom{i}{k+1-d+i} \right] \times \binom{n-d-1+i}{i}, \tag{A2}$$

where $\sum^*$ stands for a sum in which half of the last term is taken for $i = \frac{d}{2}$ if d is even and the whole last term is taken when d is odd, or $i = \frac{d-1}{2}$.

Furthermore, they showed that, in fact, the quantity $\phi_r(r_+)$ can be used as a lower bound for the non-negative rank.

Theorem 4 ([39], Theorem 6). The upper bound $\phi_r(r_+)$ on the restricted non-negative rank of a non-negative matrix M with $r = \mathrm{rank}(M)$ and $r_+ = \mathrm{rank}_+(M)$ satisfies

$$\begin{aligned}\phi_r(r_+) &= \max_{r \leqslant r_u \leqslant r_+} \mathrm{faces}(r_+, r_u - 1, r_u - r) \\ &\leqslant \max_{r \leqslant r_u \leqslant r_+} \binom{r_+}{r_u - r + 1} \\ &\leqslant \binom{r_+}{\lfloor r_+/2 \rfloor} \leqslant 2^{r_+}\sqrt{\frac{2}{\pi r_+}} \leqslant 2^{r_+}.\end{aligned} \tag{A3}$$

By combining Theorems 3 and 4 we see that for a non-negative matrix M it holds that

$$\mathrm{rank}_+^*(M) \leqslant 2^{\mathrm{rank}_+(M)}. \tag{A4}$$

Thus, the restricted non-negative rank gives a lower bound on the actual non-negative rank. In order to use this bound, we must first be able to calculate (or lower bound) the restricted non-negative rank. For a particular class of matrices, which will also suit our purposes, the authors in [39] were able to calculate the restricted non-negative rank exactly.

Let A^i denote the ith row of a matrix A. The sparsity pattern of a row A^i is defined as $S_i = \{k \,|\, A_{ik} = 0\}$. Matrix A is said to have a disjoint sparsity pattern if $S_i \nsubseteq S_j$ for all $i \neq j$.

Theorem 5 ([39], Theorem 8). If M is a rank-3 non-negative square matrix of dimension n whose columns have disjoint sparsity patterns, then $\mathrm{rank}_+^*(M) = n$.

Now we can use all the previously stated results of [39] to prove a lower bound for the antidistinguishability matrices A_n from Example 1.

Lemma 1. The non-negative rank of A_n scales as $\mathrm{rank}_+(A_n) = \Omega(\log_2 n)$.

Proof. We note that, trivially, $\mathrm{rank}_+(A_1) = 1$ and $\mathrm{rank}_+(A_2) = 2$, so that for $n = 1$ and $n = 2$ the result holds. On the other hand, since the communication matrix A_n can be implemented with a restricted qubit (the states are only in the xz plane on the Bloch sphere), it can be shown that the maximum rank of these matrices can be at most 3 (Proposition 3 in [29]). Furthermore, it can be easily checked that for a matrix size of more than 2 the rank cannot be 2. Thus, it follows that $\mathrm{rank}(A_n) = 3$ for all $n > 2$. Thus, for $n > 2$ we have that A_n is a rank-3 matrix which clearly has a disjoint sparsity pattern, so from Theorem 5 we have that $\mathrm{rank}_+^*(A_n) = n$. By using the lower bound for the non-negative rank given by Theorems 4 and 3, namely, Eq. (A4), we have that $\mathrm{rank}_+(A_n) \geqslant \log_2 n$, which completes the proof. ∎

APPENDIX B: PROOF THAT RANK$_+(A_7) = 6$

Since, from Theorem 5, we know that $\mathrm{rank}_+^*(A_7) = 7$, we can use the derived upper bounds from [39] for the restricted non-negative rank based on the actual non-negative rank to lower bound the non-negative rank of A_7. In particular, one could use Theorem 3 for this, but in fact, for rank-3 matrices the authors of [39] provided an even better bound.

Lemma 2 ([39], Corollary 4). For a rank-3 non-negative matrix M with $r_+ = \mathrm{rank}_+(M)$ the restricted non-negative rank is bounded above by

$$\begin{aligned}\mathrm{rank}_+^*(M) &\leqslant \max_{3 \leqslant r_u \leqslant r_+} \min_{i=0,1} \mathrm{faces}(r_+, r_u - 1, r_u - 3 + i) \\ &=: \phi'(r_+) \leqslant \phi_3(r_+).\end{aligned} \tag{B1}$$

It is crucial to note that $\phi'(r_+)$ [and $\phi_r(r_+)$ for any fixed r] is an increasing function of its argument r_+. This is because $\mathrm{faces}(n, d, k)$ increases with increasing n. The first few (relevant) values of ϕ' (and ϕ_3 for comparison) are presented in the Table I. In particular, for A_7 we have that $\mathrm{rank}_+^*(A_7) = 7$, so $\phi'(r_+) \geqslant 7$. From Table I we see that then we must have that $\mathrm{rank}_+(A_7) \geqslant 6$. Our explicit non-negative factorization from Example 1 then shows that, in fact, $\mathrm{rank}_+(A_7) = 6$.

[1] J. Preskill, Quantum computing in the NISQ era and beyond, Quantum 2, 79 (2018).

[2] S. Massar, D. Bacon, N. J. Cerf, and R. Cleve, Classical simulation of quantum entanglement without

local hidden variables, Phys. Rev. A **63**, 052305 (2001).

[3] B. F. Toner and D. Bacon, Communication cost of simulating Bell correlations, Phys. Rev. Lett. **91**, 187904 (2003).

[4] M. J. Renner, A. Tavakoli, and M. T. Quintino, Classical cost of transmitting a qubit, Phys. Rev. Lett. **130**, 120801 (2023).

[5] A. Ambainis, D. Leung, L. Mancinska, and M. Ozols, Quantum random access codes with shared randomness, arXiv:0810.2937.

[6] T. Guha, M. Alimuddin, S. Rout, A. Mukherjee, S. S. Bhattacharya, and M. Banik, Quantum advantage for shared randomness generation, Quantum **5**, 569 (2021).

[7] C. Perry, R. Jain, and J. Oppenheim, Communication tasks with infinite quantum-classical separation, Phys. Rev. Lett. **115**, 030504 (2015).

[8] R. Gallego, N. Brunner, C. Hadley, and A. Acín, Device-independent tests of classical and quantum dimensions, Phys. Rev. Lett. **105**, 230501 (2010).

[9] R. Chaves, J. B. Brask, and N. Brunner, Device-independent tests of entropy, Phys. Rev. Lett. **115**, 110501 (2015).

[10] S. Saha, S. S. Bhattacharya, T. Guha, S. Halder, and M. Banik, Advantage of quantum theory over nonclassical models of communication, Ann. Phys. (Berlin, Ger.) **532**, 2000334 (2020).

[11] P.-E. Emeriau, M. Howard, and S. Mansfield, Quantum advantage in information retrieval, PRX Quantum **3**, 020307 (2022).

[12] H. Buhrman, R. Cleve, J. Watrous, and R. de Wolf, Quantum fingerprinting, Phys. Rev. Lett. **87**, 167902 (2001).

[13] E. F. Galvão and L. Hardy, Substituting a qubit for an arbitrarily large number of classical bits, Phys. Rev. Lett. **90**, 087902 (2003).

[14] A. Ambainis, A. Nayak, A. Ta-Shma, and U. Vazirani, Dense quantum coding and quantum finite automata, J. Assoc. Comput. Mach. **49**, 496 (2002).

[15] S. Rout, N. Sakharwade, S. S. Bhattacharya, R. Ramanathan, and P. Horodecki, Unbounded quantum advantage in one-way strong communication complexity of a distributed clique labelling relation, arXiv:2305.10372.

[16] R. K. Patra, S. G. Naik, E. P. Lobo, S. Sen, T. Guha, S. S. Bhattacharya, M. Alimuddin, and M. Banik, Classical analogue of quantum superdense coding and communication advantage of a single quantum, arXiv:2202.06796.

[17] S. Halder, A. Streltsov, and M. Banik, Quantum vs classical: identifying the value of a random variable unambiguously, arXiv:2211.09194.

[18] D. Gavinsky, J. Kempe, I. Kerenidis, R. Raz, and R. de Wolf, Exponential separation for one-way quantum communication complexity, with applications to cryptography, SIAM J. Comput. **38**, 1695 (2009).

[19] Z. Bar-Yossef, T. S. Jayram, and I. Kerenidis, Exponential separation of quantum and classical one-way communication complexity, SIAM J. Comput. **38**, 366 (2008).

[20] A. Anshu, D. Touchette, P. Yao, and N. Yu, Exponential separation of quantum communication and classical information, in *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing: STOC 2017* (Association for Computing Machinery, New York, 2017), p. 277.

[21] S. Aaronson, Limitations of quantum advice and one-way communication, Theory Comput. **1**, 1 (2005).

[22] H. Klauck, Quantum communication complexity, in *ICALP Workshops 2000: Proceedings of the Satelite Workshops of the 27th International Colloquium on Automata, Languages and Programming, Geneva, Switzerland, July 9–15, 2000* (Carleton Scientific,Waterloo, ON, Canada, 2000), pp. 241–252.

[23] S. Aaronson and A. Arkhipov, The computational complexity of linear optics, in *Proceedings of the Forty-Third Annual ACM Symposium on Theory of Computing: STOC '11* (Association for Computing Machinery, New York, 2011), pp. 333–342.

[24] C. Gogolin, M. Kliesch, L. Aolita, and J. Eisert, Boson-sampling in the light of sample complexity, arXiv:1306.3995.

[25] B. T. Gard, K. R. Motes, J. P. Olson, P. P. Rohde, and J. P. Dowling, An introduction to boson-sampling, in *From Atomic to Mesoscale* (World Scientific Publishing Company, Singapore, 2015), Chap. 8, pp. 167–192.

[26] H. Wang, J. Qin, X. Ding, M.-C. Chen, S. Chen, X. You, Y.-M. He, X. Jiang, L. You, Z. Wang *et al.*, Boson sampling with 20 input photons and a 60-mode interferometer in a 10^{14}-dimensional Hilbert space, Phys. Rev. Lett. **123**, 250503 (2019).

[27] H.-S. Zhong, H. Wang, Y.-H. Deng, M.-C. Chen , L.-C. Peng, Y.-H. Luo, J. Qin, D. Wu, X. Ding, Y. Hu *et al.*, Quantum computational advantage using photons, Science **370**, 1460 (2020).

[28] P. E. Frenkel and M. Weiner, Classical information storage in an n-level quantum system, Commun. Math. Phys. **340**, 563 (2015).

[29] T. Heinosaari, O. Kerppo, and L. Leppäjärvi, Communication tasks in operational theories, J. Phys. A **53**, 435302 (2020).

[30] T. Heinosaari and O. Kerppo, Communication of partial ignorance with qubits, J. Phys. A **52**, 395301 (2019).

[31] A. S. Holevo, Bounds for the quantity of information transmitted by a quantum communication channel, Probl. Peredachi Inf. **9**, 3 (1973).

[32] B. Schumacher and M. Westmoreland, *Quantum Processes, Systems, and Information* (Cambridge University Press, Cambridge 2010).

[33] O. Kerppo, Quantum communication tasks, Ph.D. thesis, University of Turku, 2023.

[34] S. Fiorini, S. Massar, S. Pokutta, H. R. Tiwary, and R. de Wolf, Linear vs. semidefinite extended formulations: Exponential separation and strong lower bounds, in *Proceedings of the Forty-Fourth Annual ACM Symposium on Theory of Computing: STOC '12* (Association for Computing Machinery, New York, 2012), pp. 95–106.

[35] T. Lee, Z. Wei, and R. de Wolf, Some upper and lower bounds on PSD-rank, Math. Program. **162**, 495 (2017).

[36] J. E. Cohen and U. G. Rothblum, Nonnegative ranks, decompositions, and factorizations of nonnegative matrices, Linear Algebra Appl. **190**, 149 (1993).

[37] Y. Shitov, The complexity of positive semidefinite matrix factorization, SIAM J. Optim. **27**, 18981909 (2016).

[38] S. A. Vavasis, On the complexity of nonnegative matrix factorization, SIAM J. Optim. **20**, 1364 (2009).

[39] N. Gillis and F. Glineur, On the geometric interpretation of the nonnegative rank, Linear Algebra Appl. **437**, 2685 (2012).

[40] Y. Shitov, An upper bound for nonnegative rank, J. Comb. Theory, Ser. A **122**, 126 (2014).

[41] P. Hrubeš, On the nonnegative rank of distance matrices, Inf. Process. Lett. **112**, 457 (2012).

[42] H. Fawzi, J. Gouveia, P. A. Parrilo, R. Z. Robinson, and R. R. Thomas, Positive semidefinite rank, Math. Program. **153**, 133 (2015).

[43] T. Heinosaari and O. Kerppo, Antidistinguishability of pure quantum states, J. Phys. A **51**, 365303 (2018).
[44] C. M. Caves, C. A. Fuchs, and R. Schack, Conditions for compatibility of quantum-state assignments, Phys. Rev. A **66**, 062111 (2002).
[45] S. Bandyopadhyay, R. Jain, J. Oppenheim, and C. Perry, Conclusive exclusion of quantum states, Phys. Rev. A **89**, 022336 (2014).
[46] V. Havlíček and J. Barrett, Simple communication complexity separation from quantum state antidistinguishability, Phys. Rev. Res. **2**, 013326 (2020).
[47] V. Russo and J. Sikora, Inner products of pure states and their antidistinguishability, Phys. Rev. A **107**, L030202 (2023).
[48] M. F. Pusey, J. Barrett, and T. Rudolph, On the reality of the quantum state, Nat. Phys. **8**, 475 (2012).
[49] M. S. Leifer and C. Duarte, Noncontextuality inequalities from antidistinguishability, Phys. Rev. A **101**, 062113 (2020).
[50] A. Vandaele, N. Gillis, F. Glineur, and D. Tuyttens, Heuristics for exact nonnegative matrix factorization, J. Global Optim. **65**, 369 (2016).