

## Partial-ignorance communication tasks in quantum theory

Oskari Kerppo <sup>\*</sup>*Turku Centre for Quantum Physics, Department of Physics and Astronomy, University of Turku, FI-20014 Turku, Finland*

(Received 25 February 2022; accepted 7 June 2022; published 28 June 2022)

We introduce a generalization of the communication of partial ignorance where both parties of a prepare-and-measure setup receive inputs from a third party before a success metric is maximized over the measurements and preparations. We call these communication scenarios partial ignorance communication tasks of type  $T_{n,m}$ , where  $n$  indicates the number of possible correct answers in the communication task and  $m$  loosely represents the information about the correct answer that is revealed to the party preparing states. Various methods are used to obtain bounds on the success metrics, including semidefinite programs, ultraweak monotones for communication matrices, and frame theory for quantum states. The simplest scenarios in this alternative generalized prepare-and-measure setting, partial-ignorance communication tasks of types  $T_{3,1}$ ,  $T_{4,1}$ , and  $T_{4,2}$ , are analyzed exhaustively for bits and qudits. Finally, this generalized setting allows the introduction of operational equivalences to the preparations and measurements, allowing us to analyze and observe a contextual advantage for quantum theory in the partial ignorance communication task of type  $T_{4,1}$ .

DOI: [10.1103/PhysRevA.105.062607](https://doi.org/10.1103/PhysRevA.105.062607)

### I. INTRODUCTION

In the usual communication setting one party is interested in sending an encoded message to another party. The receiving party will try to decode the message so that information can be transmitted between the parties. In quantum mechanics these parties are typically called Alice and Bob. The limiting factor of this communication is usually the communication medium, which could be a classical bit or the quantum analog, a qubit, for instance. If Bob successfully decodes the message sent by Alice, the communication between them is deemed successful.

Perhaps the most basic scenario is the following: Alice has a preparation device that can prepare  $n$  distinct states of the communication medium. Hence each of the states encodes a unique message. Bob, on the other hand, has a measurement device with  $n$  distinct outcomes. If each outcome of Bob's device identifies with certainty the state prepared by Alice, the states are called *distinguishable*. The Basic Decoding Theorem [1] states that, in the quantum case, whenever there are more possible messages than distinguishable states the error in this kind of communication is at least  $1 - \frac{d}{n}$ , where  $d$  is the respective Hilbert space dimension of the quantum state and  $n$  is the number of states in Alice's preparation device.

In a recent work [2], a variation of the basic communication scenario was studied. In this variant there is a third party, Charlie, who acts as a game master to the following communication game. Charlie has  $n$  empty boxes and he hides a prize inside one of them. He reveals at least one empty box to Alice who must then communicate this information to Bob. Alice and Bob win the game if Bob chooses the box with the prize in it. The communication between Alice and Bob is successful if

Bob has the same chance of finding the prize as Alice would. Hence Alice's encoded message should contain information on which choices Bob should avoid, as even Alice does not know where the prize is. This communication of choices to be avoided was called *communication of partial ignorance* by the authors. The qubit case was perfectly characterized and some general results concerning qudits were presented in [2]. The main takeaway was that, perhaps remarkably, perfect communication of partial ignorance requires an entirely different setup than the basic communication scenario, where Bob tries to identify the state sent by Alice directly. However, the work in [2] was limited to studying the case where Bob's measurement is fixed. Indeed, a much richer set of communication scenarios is possible when Bob is free to choose his measurement, but this was left out of the scope of the previous work.

In the present article, we continue the work done in [2] by modifying the communication game in a major way. In this new game it's not only Alice who is revealed information about the location of the prize, but also Bob. Charlie could, for instance, reveal one empty box to Alice and one empty box to Bob. This seemingly simple modification has dramatic effects on the strategies that Alice and Bob must employ. Moreover, the difficulty of finding viable communication strategies and proving their optimality is significantly increased. Typically Bob will have multiple choices for his measurement based on the information that is revealed to him. Thus it will become possible to analyze the effects of contextuality as a resource to this game. In general, entirely different mathematical tools are required to analyze this alternative scenario compared to the communication of partial ignorance.

We call our setup of modified communication of partial ignorance simply partial-ignorance communications tasks. To explicitly state how many boxes Charlie has and how many of the boxes are shown to be empty to Alice we use the notation

<sup>\*</sup>oekoer@utu.fi

$T_{n,m}$  for the game where there are  $n$  boxes,  $m$  of which are shown to be empty to Alice.

The rest of the article is structured as follows. In Sec. II the alternative communication game is presented in more detail, mostly focusing on the game  $T_{3,1}$ . We then go on to present the required mathematical methods in Sec. III. A brief overview to semidefinite programming, contextuality, and frame theory is given. Using the introduced mathematical methods we analyze the communication games  $T_{4,1}$  and  $T_{4,2}$  in Sec. IV. Numerical and analytical bounds on successful communication will be presented in the classical, contextual, and general quantum cases. Finally, in Sec. V, we end the article with some final thoughts and discussion on possible future directions.

## II. PARTIAL IGNORANCE COMMUNICATION TASKS

Let's consider a general communication setting first. Suppose Charlie has an  $n$ -bit string,  $s$ . The objective of the communication task is that Bob must guess the value of at least one bit in  $s$ . Alice will receive a bit string, or input, of her own,  $a$ . The bit string  $a$  can be of any fixed but finite length, and in general it will contain some information on the bit string  $s$ . Alice then sends an encoded message,  $r(a)$ , to Bob. In the classical case the message will be a bit string of fixed length. A qudit  $\rho_a$  of fixed dimension  $d$  will be sent in the quantum case. Upon receiving the message  $r(a)$  sent by Alice, Bob will also receive a bit string  $b$  of fixed length from Charlie. Once communication is done Bob must produce a guess on the values of at least some bits in  $s$ .

An example of the above communication setting would be a random access code (RAC) [3–5]. In a RAC Alice's input  $a$  will coincide with Charlie's bit string  $s$ . Typically, Alice is only allowed to send one bit or a qubit to Bob. The input  $b$  given to Bob will contain the index of the bit whose value Bob must guess. If we denote by  $g[b, r(a)]$  the guess Bob produces based on his input  $b$  and the message  $r(a)$ , Alice and Bob win the game if  $g[b, r(a)] = s_b$ , where  $s_b$  denotes the value of the bit string  $s$  at index  $b$ .

Communication of partial ignorance, with the notation we are using, would be presented in the following way. Charlie has an  $n$ -bit string  $s$  containing exactly one 1, the index of which will indicate the location of the prize. Alice's input  $a$  will reveal the index of at least one of the 0's. Bob does not receive any input except for the message sent by Alice. Alice and Bob win if Bob's guess of  $g[r(a)]$  equals the index of the 1 in  $s$ .

It is very important to note that, in communication of partial ignorance, the game master Charlie is not obliged to sample the bit-string  $s$  according to any fixed distribution. In fact, he is allowed to freely choose it even after Alice has already received her input. However, he must not give conflicting information to Alice. Thus perfect communication of partial ignorance requires Alice and Bob to minimize the worst-case error probability with respect to all possible inputs. The strategy that Alice and Bob must implement leads to the concept of communication matrices, which were extensively studied in [6] along with their operational hierarchy.

We can now define the generalization of communication of partial ignorance we are interested in.

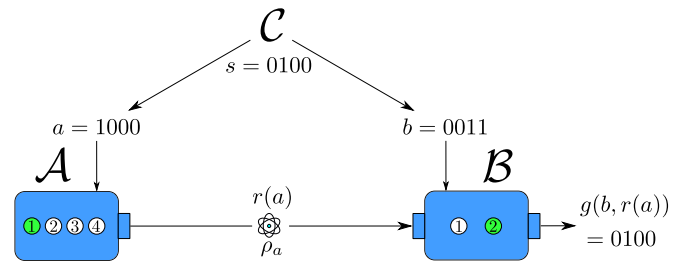


FIG. 1. An exemplary setting of a partial ignorance communication task of type  $T_{4,1}$ . The 1 in  $a$  reveals the first 0 in  $s$ . The two 1's in  $b$  reveal the last two 0's in  $s$ . Bob should guess the correct index of the 1 in  $s$ .

*Definition 1.* A partial-ignorance communication task of type  $T_{n,m}$  is a communication game where a game master Charlie chooses an  $n$ -bit string  $s$  with exactly one 1. Charlie then sends an input  $a$  revealing the indices of  $m$  0's in  $s$  to Alice. Bob will be revealed the remaining indices of 0's in input  $b$ . Both inputs  $a$  and  $b$  are bit strings of the same length as  $s$ , and a value of 1 in both inputs will reveal a 0 in  $s$  at the corresponding index. Alice and Bob are allowed to communicate according to predefined rules. After communication Bob must produce a guess on the index of the 1 in  $s$ . Alice and Bob win if Bob's guess is correct.

An illustration of the partial ignorance communication task of type  $T_{4,1}$  is given in Fig. 1.

The limiting factor in partial ignorance communication tasks is the communication medium. In the classical version Alice is allowed to send a classical message  $r(a)$  containing a predefined number of bits to Bob. In the cases studied in this article the number of bits will be just 1. In the quantum version a qudit  $\rho_a$  of fixed dimension can be sent.

It should be noted that any type of partial ignorance communication task will be a game of complete information in the sense that Alice and Bob's inputs together will reveal the index of the 1. We could also study the case where this is not the case, but for now we will only study the informationally complete version. For the remainder of this article, whenever we are talking about a task  $T_{n,m}$ , we will always be referring to the partial ignorance communication task of type  $T_{n,m}$  as defined in Definition 1.

Let's consider the simplest task  $T_{3,1}$  as an example before moving on to the mathematical methods.

### Classical and quantum strategies for $T_{3,1}$

Let's begin by introducing the best possible classical strategy for the task  $T_{3,1}$ .

*Example 1. (Classical strategy for  $T_{3,1}$ ).* There are three different choices for the string  $s$ , namely 100, 010, and 001. These strings also coincide with the possible inputs  $a$  and  $b$  for Alice and Bob. Table I presents a general form of the strategy that Alice and Bob will implement. A specific strategy is obtained by replacing the variables  $i$ ,  $j$  and  $k$  in Table I by values 0 or 1, and by listing the values of the guesses  $g[b, r(a)]$  in the middle table.

By looking at the rows of Table I where the input  $b$  is identical, it becomes apparent that to never make a mistake the

TABLE I. General classical strategy for the task  $T_{3,1}$ .

$a$	$r(a)$	$b$	$g(b, 0)$	$g(b, 1)$	$s$	$a$	$b$	$r(a)$	$g[b, r(a)]$
100	$i$	100	$l$	$x$	100	010	001	$j$	$g(001, j)$
010	$j$	010	$m$	$y$	100	001	010	$k$	$g(010, k)$
001	$k$	001	$n$	$z$	010	100	001	$i$	$g(001, i)$
					010	001	100	$k$	$g(100, k)$
					001	100	010	$i$	$g(010, i)$
					001	010	100	$j$	$g(100, j)$

variables  $i, j$  and  $k$  should fulfill the following restrictions:

$$j \neq i, \quad k \neq i, \quad k \neq j.$$

Otherwise Bob is forced to produce the same guess with different strings  $s$  because his inputs are identical. However, the variables  $i, j$ , and  $k$  can only take on the binary values of 0 or 1, and hence at least one of the restrictions must be violated. This means that any classical strategy, formed by giving concrete values to Table I, must contain at least one mistake. The average classical success probability in task  $T_{3,1}$  is therefore at most  $5/6$ . Table II shows an example of an optimal classical strategy that saturates the upper bound for success probability. Therefore, the classical average success probability for  $T_{3,1}$  is exactly  $5/6$ .

Note that the worst-case success probability for all classical strategies for the task  $T_{3,1}$  is always zero. Interestingly, the strategy listed in Table II has an unused capacity for communication because the guess  $g(001, 0)$  never occurs. Perfect communication is nevertheless not possible.

Before moving on to the next section, let us introduce the optimal quantum strategy for  $T_{3,1}$ . The states and measurements used in the following example are illustrated in Fig. 2.

*Example 2. (Quantum strategy for  $T_{3,1}$ ).* No matter which string  $s$  Charlie chooses, Bob's measurement will try to distinguish between a pair of states. For instance, upon receiving input 001 from Charlie, Bob's measurement should try to distinguish Alice's states  $\varrho_1$  and  $\varrho_2$  because Bob knows Alice is going to prepare one of those states. Following this logic, Alice's preparation device should contain states that can be pairwise distinguished as well as possible. With qubits these states can be chosen from a Bloch sphere plane with equal angles between the state vectors.

We can choose Alice's first state to correspond to the Bloch vector  $[0 \ 0 \ 1]$ . The other two states are obtained by 120

TABLE II. An optimal classical strategy for the task  $T_{3,1}$ . The erroneous value  $g(001, 1) = 1$  is shown in bold. The guess  $g(001, 0)$  is not listed in the middle table because it is not needed in this implementation of the strategy.

$a$	$r(a)$	$b$	$g(b, 0)$	$g(b, 1)$	$s$	$a$	$b$	$r(a)$	$g[b, r(a)]$
100	1	100	2	3	100	010	001	1	1
010	1	010	1	3	100	001	010	0	1
001	0	001	—	1	010	100	001	1	<b>1</b>
					010	001	100	0	2
					001	100	010	1	3
					001	010	100	1	3

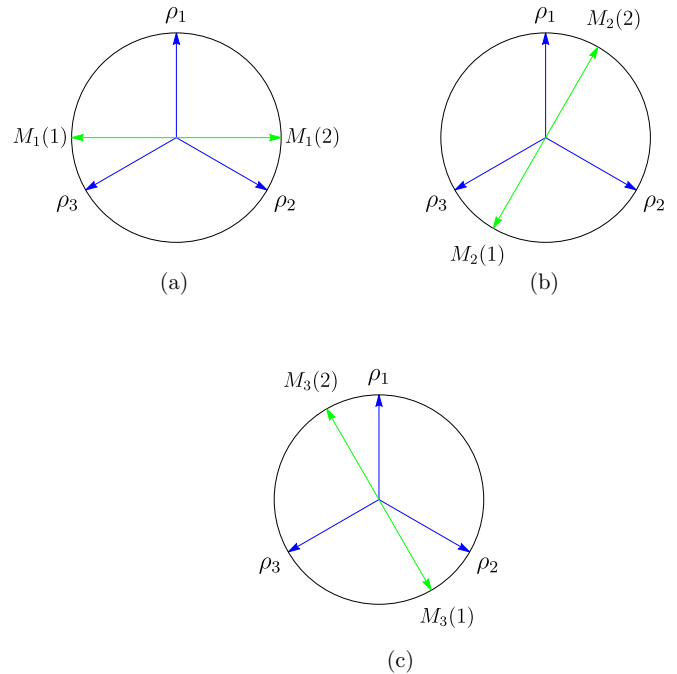


FIG. 2. Alice's states and Bob's measurements for the task  $T_{3,1}$ , presented in the  $xz$  plane of the Bloch sphere for different values of input  $b$ . The value of input  $b$  equals (a) 100, (b) 010, and (c) 001.

degree clockwise rotations in the  $xz$  plane. As density matrices these states can be written as

$$\begin{aligned} \varrho_1 &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \varrho_2 = \frac{1}{4} \begin{pmatrix} 1 & \sqrt{3} \\ \sqrt{3} & 3 \end{pmatrix}, \\ \varrho_3 &= \frac{1}{4} \begin{pmatrix} 1 & -\sqrt{3} \\ -\sqrt{3} & 3 \end{pmatrix}. \end{aligned} \quad (1)$$

Bob's measurements that best distinguish Alice's states in a pairwise manner, written in terms of positive operator-valued measures (POVMs), are the following:

$$\begin{aligned} M_1(1) &= \frac{1}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \\ M_1(2) &= \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \\ M_2(1) &= \frac{1}{2} \begin{pmatrix} 1 - \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 + \frac{\sqrt{3}}{2} \end{pmatrix}, \\ M_2(2) &= \frac{1}{2} \begin{pmatrix} 1 + \frac{\sqrt{3}}{2} & \frac{1}{2} \\ \frac{1}{2} & 1 - \frac{\sqrt{3}}{2} \end{pmatrix}, \\ M_3(1) &= \frac{1}{2} \begin{pmatrix} 1 - \frac{\sqrt{3}}{2} & \frac{1}{2} \\ \frac{1}{2} & 1 + \frac{\sqrt{3}}{2} \end{pmatrix}, \\ M_3(2) &= \frac{1}{2} \begin{pmatrix} 1 + \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 - \frac{\sqrt{3}}{2} \end{pmatrix}. \end{aligned} \quad (2)$$

Each POVM is obtained by a rotation of  $\pm 90$  degrees in the  $xz$  plane from the corresponding state vector, e.g., the effects of POVM  $M_2$  are obtained by rotating the Bloch vector of  $\varrho_2$ . For the success probabilities we can calculate

that, for instance,  $\text{tr}[\varrho_3 \mathbf{M}_1(1)] = \frac{1}{2}(1 + \frac{\sqrt{3}}{2}) \approx 0.933013$  and  $\text{tr}[\varrho_3 \mathbf{M}_1(2)] = \frac{1}{2}(1 - \frac{\sqrt{3}}{2}) \approx 0.0669873$ . Because the states and POVMs were chosen in a symmetrical manner we can conclude that the average success probability is equal to  $\frac{1}{2}(1 + \frac{\sqrt{3}}{2}) \approx 0.933013$  while the worst-case error probability is equal to  $\frac{1}{2}(1 - \frac{\sqrt{3}}{2}) \approx 0.0669873$ .

As we can see, there is a drastic difference in average and worst-case success probabilities between the bit and the qubit in the task  $T_{3,1}$ . This case was, however, quite easy to analyze. To analyze any task  $T_{n,m}$  with  $n > 3$  we are going to need some mathematical tools.

### III. MATHEMATICAL METHODS

In this section we will introduce three tools that can be used to analyze the communication tasks presented in the previous section. Semidefinite programming is a powerful numerical tool that can be used to obtain numerical bounds on success probabilities in various communication and computation tasks involving quantum resources. Contextuality is one such resource and we will use it to provide an alternative comparison between classical and quantum strategies. While semidefinite programs (SDPs) provide valuable insight to many problems, it turns out we can also explain some bounds obtained by SDPs analytically with the help of frame theory. To this end we will use Grassmannian frames.

#### A. Semidefinite programming

There are many excellent sources on the theory of SDPs [7–15]. We will be using the unitary SDP hierarchy introduced in [16] largely due to the fact that it gives very good results already on the first level of the hierarchy, i.e., we only have to consider moment matrices with monomials of length 1.

Consider the following optimization problem:

$$\begin{aligned}
 & \max \sum_{i,j,k} p_{ijk} \text{tr}[\varrho_i \mathbf{M}_j(k)], \\
 & \text{s.t. } \varrho_i \geq 0 \forall i, \\
 & \text{tr}[\varrho_i] = 1 \forall i, \\
 & \sum_i (\alpha_i^r - \beta_i^r) \varrho_i = 0 \forall r \in \mathcal{O}_P, \\
 & 0 \leq \mathbf{M}_j(k) \leq 1 \forall j, k \\
 & \sum_k \mathbf{M}_j(k) = \mathbb{1} \forall j \\
 & \sum_{i,j} (\alpha_{i,j}^s - \beta_{i,j}^s) \mathbf{M}_i(j) = 0 \forall s \in \mathcal{O}_M, \quad (3)
 \end{aligned}$$

where the weights  $p_{ijk}$  define a success metric on the behavior  $p(k|\varrho_i, \mathbf{M}_j) \equiv \text{tr}[\varrho_i \mathbf{M}_j(k)]$  and  $\mathcal{O}_P$  (indexed by  $r$ ) contains restrictions of the form  $\sum_i \alpha_i^r \varrho_i = \sum_j \beta_j^r \varrho_j$  on the preparations for some convex weights  $\alpha_i^r$  and  $\beta_j^r$ . Similarly  $\mathcal{O}_M$  (indexed by  $s$ ) contains restrictions on the effects for some convex weights  $\alpha_{i,j}^s$  and  $\beta_{i,j}^s$ . All optimal quantum strategies for communication tasks  $T_{n,m}$  can be obtained by solving an optimization problem of the form (3). Notice that the sets  $\mathcal{O}_P$  and  $\mathcal{O}_M$  may be empty. Unfortunately, this problem is not an

SDP. We can, however, use the following algorithm to give the success metric  $\sum_{i,j,k} p_{ijk} \text{tr}[\varrho_i \mathbf{M}_j(k)]$  a dimension based lower bound.

*Definition 2.* See-saw algorithm.

(1) Fix a dimension  $d$  for the Hilbert space of the quantum states and POVMs.

(2) Generate random rank-1 states and fix the states  $\varrho_i$  as constants.

(3) Optimize Eq. (3) as an SDP over POVM elements  $\mathbf{M}_j(k)$ . Calculate  $x_1 = \sum_{i,j,k} p_{ijk} \text{tr}[\varrho_i \mathbf{M}_j(k)]$ .

(4) Fix POVM elements  $\mathbf{M}_j(k)$  as constants with values obtained in the previous step. Optimize Eq. (3) as an SDP over the states  $\varrho_i$ . Calculate  $x_2 = \sum_{i,j,k} p_{ijk} \text{tr}[\varrho_i \mathbf{M}_j(k)]$ .

(5) Repeat steps 3 and 4 until a predefined threshold value  $x_2 - x_1 < \epsilon$  is achieved. The value  $x_2$  is now a dimension-based lower bound on the success metric of Eq. (3).

To the best of our knowledge the see-saw algorithm lacks any kind of serious theory behind it. Therefore, there are no performance guarantees. The see-saw algorithm offers a dimension-based lower bound, or the inner bound from here on, and an explicit construction for the states and POVMs that achieve this bound. Some mentions of this method can be found in the literature [17–21]. In practice the see-saw algorithm gives very good results. As we will see later, the see-saw algorithm is able to give optimal implementations within numerical accuracy to all scenarios studied in the present article. However, one still has to prove that the implementations given by the see-saw algorithm are truly optimal.

To obtain a result of optimality we need to produce an upper bound on the success metric of Eq. (3), or an outer bound. This can be done in the following way.

Suppose we have a quantum behavior  $p(k|\varrho_i, \mathbf{M}_j) = \text{tr}[\varrho_i \mathbf{M}_j(k)]$ . As was proven in [16], we can always find unitary matrices such that

$$\mathbf{M}_j(k) = \frac{1}{2} \mathbb{1} + \frac{U_k^j + U_k^{j\dagger}}{4}. \quad (4)$$

The moment matrices  $(\Gamma_i^\mathcal{O})_{j,k} = (\Gamma_i^\mathcal{O})_{\varrho_j, \varrho_k} \equiv \text{tr}[\varrho_i \mathcal{O}_j^\dagger \mathcal{O}_k]$  (with monomials of  $\mathbb{1}$  and the unitary operators  $U_k^j$  and  $U_k^{j\dagger}$ ) have the following properties:

$$\begin{aligned}
 & \forall i : \Gamma_i^\mathcal{O} \geq 0, \\
 & \forall i, r \in \mathcal{O}_P : \sum_j (\alpha_j^r - \beta_j^r) \Gamma_i^\mathcal{O} = 0, \\
 & \forall i, j, k : (\Gamma_i^\mathcal{O})_{\mathbb{1}, U_k^j} + (\Gamma_i^\mathcal{O})_{\mathbb{1}, U_k^{j\dagger}} = 4 \left( p(k|i, j) - \frac{1}{2} \right), \\
 & \forall i, j : (\Gamma_i^\mathcal{O})_{j,j} = 1. \quad (5)
 \end{aligned}$$

Additionally, the possible constraints from  $\mathcal{O}_M$  must be encoded into the moment matrices via conditions on the unitary matrices.

The existence of such moment matrices is a necessary condition for the behavior  $p(k|\varrho_i, \mathbf{M}_j) = \text{tr}[\varrho_i \mathbf{M}_j(k)]$  to be realizable in a quantum experiment. Moreover, the existence of such moment matrices is a semidefinite feasibility problem, which can be solved with efficient methods [43]. The optimization of any success metric over feasible behaviors is again an SDP.

The semidefinite feasibility problem defined by Eq. (5) defines a first level in a hierarchy of SDPs. That is, a behavior  $p(k|i, j)$  obtained from maximizing a success metric over the feasible set is not necessarily realizable in any quantum experiment. What we do get is an upper bound on the success metric (no quantum behavior can exceed this bound). Moreover, by considering monomials of length greater than 1 we obtain a converging hierarchy that converges on the set of quantum behaviors [16]. In practice we get very good results already on the first level. The notation  $\mathcal{U}_1$  is used to specify that a solution is obtained by using the first level of the unitary hierarchy.

### B. Noncontextual polytope of correlations

Contextuality is understood today as an important resource for quantum computation [16,22–31]. We will now give a brief introduction to the topic, focusing on recent developments. The reader is encouraged to check [32,33] for a more complete introduction to the framework of ontological models.

Suppose there is a state space  $\Lambda$ , called the ontic state space. Every time a quantum state is prepared, a state  $\lambda \in \Lambda$  is produced according to some probability distribution. The ontic states are considered complete descriptions of nature, but they are, generally speaking, inaccessible to all observers. Therefore we associate each preparation  $P$  with a corresponding probability measure  $\mu_P$  over the state space  $\Lambda$ .

When a measurement of some POVM  $M$  is performed, it is supposed that the value  $\lambda$  completely determines the outcome. Note that this does not imply determinism. Instead we associate each effect  $M(i)$  with the response function

$$\int_{\Lambda} \xi_M(i|\lambda) \mu_{\varrho}(\lambda) d\lambda = \text{tr}[\varrho M(i)]. \quad (6)$$

The response function  $\xi_M(i|\lambda)$  determines the probability of obtaining outcome  $i$  when  $\varrho$  was prepared and a measurement of  $M$  was performed.

*Definition 3.* An ontological model consist of a measure space  $(\Lambda, d\lambda)$ <sup>1</sup> and two functions  $\Delta$  and  $\Xi$ . The function  $\Delta$  maps every quantum state  $\varrho$  to a set of probability measures  $\Delta_{\varrho}$ . Likewise the function  $\Xi$  maps every POVM  $M$  to a set of indicator functions  $\Xi_M$ . Moreover, for each  $\lambda \in \Lambda$  and for every  $M$  and  $\xi \in \Xi_M$ :

$$\sum_i \xi_M(i|\lambda) = 1. \quad (7)$$

The ontological model is said to reproduce quantum predictions if it respects the Born rule defined in Eq. (6).

Quantum theory has the property that mixed states do not have unique decompositions into pure states. Suppose we have a preparation device that can prepare a state  $\varrho = \frac{1}{2}(\varrho_1 + \varrho_2) = \frac{1}{2}(\varrho_3 + \varrho_4)$  through two distinct decompositions. It is

a basic fact of quantum information theory that these decompositions are indistinguishable from each other. The principle of noncontextuality states that whenever two states are operationally indistinguishable, they should also be ontologically indistinguishable. On the ontological level this means that the probability measure associated with  $\varrho$  should be unique. For measurements the principle of noncontextuality states that each effect should be associated with a unique response function, no matter which POVM the effect is a part of.

It is widely believed, in light of quantum theory, that nature does not allow a deterministic description. Therefore, the nonclassical features of quantum mechanics, such as contextuality and nonlocality, only manifest themselves in a statistical manner. The question is then how can contextuality be detected from a set of experimental data. Let  $p(x|P, M)$  denote the probability distribution of outcomes  $x$  given that a measurement of  $M$  followed the preparation procedure  $P$ .

*Definition 4.* Two preparation procedures  $P_1$  and  $P_2$  are operationally equivalent if  $p(x|P_1, M) = p(x|P_2, M)$  for all  $M$ . Likewise, two measurements  $M_1$  and  $M_2$  are operationally equivalent if  $p(x|P, M_1) = p(x|P, M_2)$  for all  $P$ .

The notation " $\simeq$ " is used to denote operationally equivalent procedures for both preparations and measurements.

Whenever two procedures are operationally equivalent, the principle of noncontextuality implies that the procedures should also be ontologically equivalent. For preparations this means that  $\mu_{P_1}(\lambda) = \mu_{P_2}(\lambda)$  for all  $\lambda$  whenever  $P_1 \simeq P_2$ . For measurements, the principle on noncontextuality implies that operationally equivalent effects should be represented by a unique response function

$$\xi_{M_1}(i|\lambda) = \xi_{M_2}(i|\lambda),$$

for all  $\lambda$  whenever  $M_1(i) \simeq M_2(i)$ .

It is convenient to collect all operational equivalences in two distinct sets:  $\mathcal{O}_P$  will contain all operational equivalences of the form

$$\sum_i \alpha_i^r P_i = \sum_j \beta_j^r P_j,$$

for some distinct sets of convex weights  $\{\alpha_i^r\}_i$  and  $\{\beta_j^r\}_j$ . The variable  $r$  indexes different operational equivalences. For every operational equivalence in  $\mathcal{O}_P$  a corresponding ontological restriction must be satisfied by all noncontextual models

$$\forall \lambda : \sum_i \alpha_i^r \mu_{P_i}(\lambda) = \sum_j \beta_j^r \mu_{P_j}(\lambda).$$

Likewise for measurements, a set  $\mathcal{O}_M$  will contain all operational equivalences of the form

$$\sum_{i,j} \alpha_{i,j}^s M_i(j) = \sum_{i',j'} \beta_{i',j'}^s M_{i'}(j'),$$

for some set of convex weights  $\{\alpha_{i,j}^s\}_{i,j}$  and  $\{\beta_{i',j'}^s\}_{i',j'}$ , where the variable  $s$  indexes the set of operational equivalences.

Once the sets of operational equivalences  $\mathcal{O}_P$  and  $\mathcal{O}_M$  are determined for a behavior  $p(x|P_i, M_j)$ , the question if a noncontextual model exists for the behavior can be presented in a compact form.

<sup>1</sup>It would be more rigorous to say that the measure space is  $(\Lambda, \Sigma)$  where  $\Sigma$  is a  $\sigma$  algebra. Here we assume the existence of a canonical measure  $d\lambda$  that dominates each probability measure defined on  $(\Lambda, \Sigma)$ . There are some ontological models that do not allow this, but this assumption makes the notation somewhat simpler.

*Definition 5.* A noncontextual ontological model exists for a behavior  $p(x|P_i, \mathbf{M}_j)$  if there exists an ontic state space  $\Lambda$ ,

epistemic states  $\{\mu_{P_i}(\lambda)\}_i$  and response functions  $\{\xi_{M_j}(k|\lambda)\}_{j,k}$  such that

$$\begin{aligned} \forall \lambda, j, k : \xi_{M_j}(k|\lambda) &\geq 0 \text{ (positivity of response functions),} \\ \forall \lambda, j : \sum_k \xi_{M_j}(k|\lambda) &= 1 \text{ (normalization of response functions),} \\ \forall \lambda, s : \sum_{j,k} (\alpha_{j,k}^s - \beta_{j,k}^s) \xi_{M_j}(k|\lambda) &= 0 \text{ (operational equivalences } \mathcal{O}_M), \\ \forall \lambda, i : \mu_{P_i}(\lambda) &\geq 0 \text{ (positivity of epistemic states),} \\ \forall i : \int_{\Lambda} \mu_{P_i}(\lambda) d\lambda &= 1 \text{ (normalization of epistemic states),} \\ \forall \lambda, r : \sum_i (\alpha_i^r - \beta_i^r) \mu_{P_i}(\lambda) &= 0 \text{ (operational equivalences } \mathcal{O}_P), \\ \forall i, j, k : \int_{\Lambda} \xi_{M_j}(k|\lambda) \mu_{P_i}(\lambda) d\lambda &= p(k|P_i, \mathbf{M}_j) \text{ (model reproduces data).} \end{aligned}$$

A detailed description of methods that can be used to demonstrate a contextual advantage from any behavior is presented in [27]. We will use those methods to examine whether the communication tasks can be used to demonstrate a contextual advantage for quantum theory.

The first step is to characterize the so-called measurement assignment polytope. Suppose all measurements in a prepare-and-measure setup have  $d$  outcomes and that there are  $l$  distinct measurement procedures. Then the  $ld$ -dimensional vector

$$\vec{\xi} = [\xi_{M_1}(1|\lambda^*) \dots \xi_{M_1}(d|\lambda^*) \xi_{M_2}(1|\lambda^*) \dots \xi_{M_l}(d|\lambda^*)],$$

defined for a specific ontic state  $\lambda^*$ , defines a noncontextual measurement assignment if it satisfies the first three conditions of Definition 5. The set of all such assignments defines the measurement-assignment polytope, which we must characterize by its vertices. The use of mathematical optimization software is encouraged to perform vertex enumeration.<sup>2</sup> The vertices are the extremal points of the convex measurement-assignment polytope.

Once vertex enumeration is performed, a key observation is that any noncontextual model, no matter the size of the ontic state space  $\Lambda$ , can be reconstructed into a model defined by probability assignments to the finite set of extremal points of the measurement-assignment polytope. That is, each preparation defines one epistemic state for each vertex in the measurement-assignment polytope. Let  $\kappa$  enumerate the vertices and let  $\nu_P(\kappa)$  denote the epistemic states defined on the vertices.

*Definition 6.* A noncontextual ontological model exists for a behavior  $p(x|P_i, \mathbf{M}_j)$  if there exists epistemic states  $\{\nu_{P_i}(\kappa)\}_{i,\kappa}$  such that

$$\forall i, \kappa : \nu_{P_i}(\kappa) \geq 0 \text{ (positivity of epistemic states),}$$

$$\forall i : \sum_{\kappa} \nu_{P_i}(\kappa) = 1 \text{ (normalization of epistemic states),}$$

$$\begin{aligned} \forall r, \kappa : \sum_i (\alpha_i^r - \beta_i^r) \nu_{P_i}(\kappa) \\ = 0 \text{ (operational equivalences } \mathcal{O}_P), \end{aligned}$$

$$\begin{aligned} \forall i, j, k : \sum_{\kappa} \tilde{\xi}_{M_j}(k|\kappa) \nu_{P_i}(\kappa) \\ = p(k|P_i, \mathbf{M}_j), \text{ (model reproduces data),} \end{aligned}$$

where  $\tilde{\xi}_{M_j}(j|\kappa)$  are the extremal response functions defined by the noncontextual measurement-assignment polytope.

Another key observation from Definition 6 is that each operational probability  $p(k|P_i, \mathbf{M}_j)$  is a linear combination of a finite amount of variables. We could now proceed as in [27] and perform linear quantifier elimination to find all the noncontextual inequalities that define the entire noncontextual polytope of correlations. It is more convenient, however, to treat the equations of Definition 6 as a linear program.

The last three equations of Definition 6 can be encoded in a single matrix equation  $\mathbf{M}\mathbf{x} = \mathbf{b}^*$ , where  $\mathbf{M}$  contains the operational equivalence parameters  $\alpha_i^r - \beta_i^r$  and the extremal measurement assignments  $\tilde{\xi}_{M_j}(k|\kappa)$ ,  $\mathbf{x}$  contains the epistemic states  $\nu_{P_i}(\kappa)$  as a vector, and  $\mathbf{b}^*$  contains the operational probabilities  $p(k|P_i, \mathbf{M}_j)$  as well as 0s and 1s corresponding to the normalization of epistemic states and operational equivalences  $\mathcal{O}_P$ . The positivity of the epistemic states corresponds to the inequality  $\mathbf{x} \geq 0$ . Hence, the equations of Definition 6 define a linear program (LP) whose primal feasibility is determined by the existence of  $\mathbf{x}$  such that

$$\begin{aligned} \mathbf{M}\mathbf{x} &= \mathbf{b}^*, \\ \mathbf{x} &\geq 0. \end{aligned} \tag{8}$$

Whenever such an  $\mathbf{x}$  exists that Eq. (8) is fulfilled, we can be sure that a noncontextual ontological model exists for the given operational probabilities.

<sup>2</sup>We used the function FMEL from a free software called PORTA.

By Farkas' lemma [34], it must be true that either the primal LP is feasible or a certificate of primal feasibility is negative. We can define the certificate of primal infeasibility as the solution to the Farkas dual of the primal LP:

$$\begin{aligned} \min_{\mathbf{y}} \mathbf{b}^{*\top} \mathbf{y}, \\ \mathbf{M}^\top \mathbf{y} \geq 0. \end{aligned} \quad (9)$$

Whenever the primal LP is infeasible it is guaranteed that we can find a  $\mathbf{y}$  such that  $\mathbf{b}^{*\top} \mathbf{y} < 0$ . By finding the minimum value among these we will simultaneously find the noncontextuality inequality that is most violated by the operational probabilities.

### C. Grassmannian frames

Let  $V$  be an inner product space. A sequence  $\{f_i\}_{i=1}^n$  is called a frame if there exists frame bounds  $A$  and  $B$  such that

$$A\|v\|^2 \leq \sum_{i=1}^n |\langle v, f_i \rangle|^2 \leq B\|v\|^2, \quad (10)$$

for all  $v \in V$ . Some special cases of frames are listed below.

A frame is called:

- (1) tight whenever the choice  $A = B$  is possible;
- (2) uniform whenever  $\|f_i\| = 1$  for all  $i$ ;
- (3) equiangular whenever  $|\langle f_i, f_j \rangle| = c$  for all  $i \neq j$  and  $c \geq 0$ .

Additionally, whenever the number of frame elements  $n$  coincides with the dimension  $d$  of  $V$ , the frame is also a basis for  $V$ . An example of an equiangular uniform frame would be an orthonormal basis.

We are not particularly interested in frames that are also bases. There are many reasons to study overcomplete frames [35–37]. An important concept is the correlation between frame elements.

*Definition 7.* Let  $\{f_i\}_{i=1}^n$  be a uniform frame in  $V$ . Then

$$\mathcal{M}(\{f_i\}_{i=1}^n) = \max_{j,k,j \neq k} |\langle f_j, f_k \rangle| \quad (11)$$

is the maximal frame correlation of  $\{f_i\}_{i=1}^n$ .

The maximal frame correlation measures the maximal overlap between frame elements. A Grassmannian frame is simply a frame that minimizes the maximum overlap [38,39].

*Definition 8.* Let  $\{f_i\}_{i=1}^n$  be a uniform frame in  $V$ . Then  $\{f_i\}_{i=1}^n$  is called a Grassmannian frame if it is a solution to the problem

$$\min \{ \mathcal{M}(\{f_i\}_{i=1}^n) \}. \quad (12)$$

At first glance it does not seem probable that there would exist analytical bounds on maximal frame correlations that only depend on the dimension of the inner product space and the number of frame elements. Yet such a result exists which we will make use of later on.

*Proposition 1.* Let  $\{f_i\}_{i=1}^n$  be a uniform frame for  $\mathbb{C}^d$  or  $\mathbb{R}^d$ . Then

$$\mathcal{M}(\{f_i\}_{i=1}^n) \geq \sqrt{\frac{n-d}{d(n-1)}}. \quad (13)$$

TABLE III. General classical strategy for the task  $T_{4,1}$ .

$a$	$r(a)$	$b$	$g(b, 0)$	$g(b, 1)$
1000	$i$	1100	$x_1$	$y_1$
0100	$j$	1010	$x_2$	$y_2$
0010	$k$	1001	$x_3$	$y_3$
0001	$l$	0110	$x_4$	$y_4$
		0101	$x_5$	$y_5$
		0011	$x_6$	$y_6$
$s$	$a$	$b$	$r(a)$	$g[b, r(a)]$
1000	0100	0011	$j$	$g(0011, j)$
1000	0010	0101	$k$	$g(0101, k)$
1000	0001	0110	$l$	$g(0110, l)$
0100	1000	0011	$i$	$g(0011, i)$
0100	0010	1001	$k$	$g(1001, k)$
0100	0001	1010	$l$	$g(1010, l)$
0010	1000	0101	$i$	$g(0101, i)$
0010	0100	1001	$j$	$g(1001, j)$
0010	0001	1100	$l$	$g(1100, l)$
0001	1000	0110	$i$	$g(0110, i)$
0001	0100	1010	$j$	$g(1010, j)$
0001	0010	1100	$k$	$g(1100, k)$

Moreover, the equality in Eq. (13) is achieved if and only if  $\{f_i\}_{i=1}^n$  is a tight equiangular frame. The equality in Eq. (13) requires  $n \leq d^2$  for  $\mathbb{C}^d$  and  $n \leq \frac{d(d+1)}{2}$  for  $\mathbb{R}^d$ .

A proof for Proposition 1 can be found from [40].

Proposition 1 concludes our brief introduction to frames. Additional information on frame theory can be found from various excellent sources, such as [41].

## IV. BOUNDS ON COMMUNICATION SUCCESS

We are now ready to analyze the tasks  $T_{4,1}$  and  $T_{4,2}$ . Let us first consider the classical version with a single bit as the communication medium.

### A. Classical bounds

We will follow a similar technique for the tasks  $T_{4,1}$  and  $T_{4,2}$  as was used for  $T_{3,1}$ . The task  $T_{4,1}$  involves four possibilities for the strings  $s$  and input  $a$  while there are six possibilities for the input  $b$ . The entire strategy can be conveniently presented in table form. By looking at the rows of Table III where the input  $b$  is identical, we get the following constraints on the messages  $r(a)$ :

$$i \neq j, \quad i \neq k, \quad i \neq l, \quad j \neq k, \quad j \neq l, \quad l \neq k.$$

That means that, to never make a mistake, each message  $r(a)$  needs to be distinct. However, the message  $r(a)$  can only take on the binary values of 0 and 1, and hence at least two of the inequalities must be violated. Each violation leads to one mistake because each row in Table III with identical  $b$  has a different string  $s$ . Therefore, any classical strategy formed by giving concrete values to Table III must contain at least two mistakes. An optimal classical strategy for the task  $T_{4,1}$  is presented in Table IV. The average classical success probability for the task  $T_{4,1}$  is  $5/6$ .

TABLE IV. An optimal classical strategy for the task  $T_{4,1}$ . Bold values in the last column indicate erroneous guesses by Bob. The guesses  $g(1100, 0)$  and  $g(0011, 1)$  are not listed because they are not needed in this specific implementation of the strategy.

$a$	$r(a)$	$b$	$g(b, 0)$	$g(b, 1)$
1000	0	1100	—	3
0100	0	1010	4	2
0010	1	1001	3	2
0001	1	0110	4	1
		0101	3	1
		0011	1	—
$s$	$a$	$b$	$r(a)$	$g[b, r(a)]$
1000	0100	0011	0	1
1000	0010	0101	1	1
1000	0001	0110	1	1
0100	1000	0011	0	<b>1</b>
0100	0010	1001	1	2
0100	0001	1010	1	2
0010	1000	0101	0	3
0010	0100	1001	0	3
0010	0001	1100	1	3
0001	1000	0110	0	4
0001	0100	1010	0	4
0001	0010	1100	1	<b>3</b>

As the final classical example, let us introduce the task  $T_{4,2}$ . Now there are six possibilities for input  $a$  while there are four possibilities for the string  $s$  and input  $b$ . Again by looking at the rows of Table V where the input  $b$  is identical, we obtain the following constraints:

$$i \neq j \neq l, \quad i \neq k \neq m, \quad j \neq k \neq n, \quad l \neq m \neq n. \quad (14)$$

Any violation of the above inequalities instantly leads to at least one mistake in the strategy, as each inequality was

TABLE V. General classical strategy for the task  $T_{4,2}$ .

$a$	$r(a)$	$b$	$g(b, 0)$	$g(b, 1)$
1100	$i$	1000	$x_1$	$y_1$
1010	$j$	0100	$x_2$	$y_2$
1001	$k$	0010	$x_3$	$y_3$
0110	$l$	0001	$x_4$	$y_4$
0101	$m$			
0011	$n$			
$s$	$a$	$b$	$r(a)$	$g[b, r(a)]$
1000	0110	0001	$l$	$g(0001, l)$
1000	0101	0010	$m$	$g(0010, m)$
1000	0011	0100	$n$	$g(0100, n)$
0100	1010	0001	$j$	$g(0001, j)$
0100	1001	0010	$k$	$g(0010, k)$
0100	0011	1000	$n$	$g(1000, n)$
0010	1100	0001	$i$	$g(0001, i)$
0010	1001	0100	$k$	$g(0100, k)$
0010	0101	1000	$m$	$g(1000, m)$
0001	1100	0010	$i$	$g(0010, i)$
0001	1010	0100	$j$	$g(0100, j)$
0001	0110	1000	$l$	$g(1000, l)$

TABLE VI. An optimal classical strategy for the task  $T_{4,2}$ . The bold values indicate erroneous guesses by Bob.

$a$	$r(a)$	$b$	$g(b, 0)$	$g(b, 1)$
1100	0	1000	2	3
1010	1	0100	1	3
1001	1	0010	4	1
0110	1	0001	3	1
0101	1			
0011	0			
$s$	$a$	$b$	$r(a)$	$g[b, r(a)]$
1000	0110	0001	1	1
1000	0101	0010	1	1
1000	0011	0100	0	1
0100	1010	0001	1	<b>1</b>
0100	1001	0010	1	<b>1</b>
0100	0011	1000	0	2
0010	1100	0001	0	3
0010	1001	0100	1	3
0010	0101	1000	1	3
0001	1100	0010	0	4
0001	1010	0100	1	<b>3</b>
0001	0110	1000	1	<b>3</b>

derived from rows with different strings  $s$ . By checking all possible choices for the messages  $r(a)$  we find that the minimum number of violations is four in any classical strategy. An example of such a strategy is presented in Table VI. The average classical success probability for the task  $T_{4,2}$  is therefore  $2/3$  (4 mistakes out of 12 guesses).

## B. Contextual bounds

By introducing operational equivalences to Alice's preparations<sup>3</sup> we can produce noncontextuality inequalities that bound the success chance of noncontextual models on communication tasks of any type with  $n \geq 4$ . In this way we can compare noncontextual models to the simplest case of one bit as the communication medium, while simultaneously observing a quantum advantage using SDP methods.

In the partial-ignorance communication task of type  $T_{4,1}$  Charlie sends one of four strings to Alice: 1000, 0100, 0010, or 0001. Alice prepares a state from the set  $\{\varrho_i\}_{i=1}^4$ , where the subscript  $i$  corresponds with the index of the 1 in Alice's input string. Bob will receive one of the following input strings from Charlie: 1100, 1010, 1001, 0110, 0101, or 0011. For each input string Bob will have a binary POVM  $M_{ij}$ , where the subscript  $ij$  will indicate the indexes of the 0's in Bob's input. When Bob's measurements are labeled in this way all Bob's POVMs will contain the index of Alice's preparation in the label. In other words, the success metric for the task  $T_{4,1}$  will be a combination of terms  $\pm \text{tr}[\varrho_i M_{jk}(1)]$  where either  $i = j$  or  $i = k$ . The term will be positive if  $i = k$  and negative if  $i = j$ . By maximizing this success metric Bob's output will coincide with the correct answer as often as possible on average.

<sup>3</sup>We could also consider operational equivalences for Bob's measurements. However, it is sufficient to only consider operational equivalences for the preparations to observe a quantum advantage.



*Example 3. (Contextual  $T_{4,1}$ ).* Let us consider the task  $T_{4,1}$  with a nontrivial operational constraint between Alice's preparations. As there are four possible states that Alice can prepare, there are three natural choices for operational equivalences. We choose the operational equivalence  $\varrho_1 + \varrho_2 \simeq \varrho_3 + \varrho_4$ . With this choice the problem we wish to optimize is the following:

$$\begin{aligned} \max \quad & -\text{tr}[\varrho_1 \mathbf{M}_{12}(1)] - \text{tr}[\varrho_1 \mathbf{M}_{13}(1)] - \text{tr}[\varrho_1 \mathbf{M}_{14}(1)] \\ & + \text{tr}[\varrho_2 \mathbf{M}_{12}(1)] - \text{tr}[\varrho_2 \mathbf{M}_{23}(1)] - \text{tr}[\varrho_2 \mathbf{M}_{24}(1)] \\ & + \text{tr}[\varrho_3 \mathbf{M}_{13}(1)] + \text{tr}[\varrho_3 \mathbf{M}_{23}(1)] - \text{tr}[\varrho_3 \mathbf{M}_{34}(1)] \\ & + \text{tr}[\varrho_4 \mathbf{M}_{14}(1)] + \text{tr}[\varrho_4 \mathbf{M}_{24}(1)] + \text{tr}[\varrho_4 \mathbf{M}_{34}(1)], \end{aligned} \quad (15a)$$

$$\text{subject to } \varrho_i \geq 0, \quad i = 1, 2, 3, 4, \quad (15b)$$

$$\text{tr}[\varrho_i] = 1, \quad i = 1, 2, 3, 4, \quad (15c)$$

$$\varrho_1 + \varrho_2 = \varrho_3 + \varrho_4, \quad (15d)$$

$$\mathbf{M}_j(1), \mathbf{M}_j(2) \geq 0, \quad j = 12, 13, 14, 23, 24, 34, \quad (15e)$$

$$\mathbf{M}_j(1) + \mathbf{M}_j(2) = \mathbb{1}, \quad j = 12, 13, 14, 23, 24, 34, \quad (15f)$$

where Eq. (15d) represents the operational equivalence that we chose, Eqs. (15b) and (15c) correspond to positivity and unit trace conditions for the states and Eqs. (15e) and (15f) correspond to positivity and normalization of the effects. Note that the success metric in Eq. (15a) is written in terms of the first effect of each POVM.

The unitary SDP hierarchy  $\mathcal{U}_1$  converges on 4.828427123, while the see-saw algorithm converges on 4.828427104 for qubits. Hence the outer bound on the first level of the unitary SDP hierarchy matches the inner bound for qubits by seven decimals. Note that the outer bound does not depend on the dimension. Instead it is valid for all quantum systems that conform to the operational equivalence between preparations. The average success probability, as given by the implementation found by the see-saw algorithm, is 0.9023689, much higher than the average classical success probability.

Performing vertex enumeration we find that the measurement assignment polytope has 64 vertices<sup>4</sup>. Hence there are 256 variables  $v_p(\kappa)$ . We then from the matrix  $\mathbf{M}$  and query a linear program for the optimal solution to Eq. (15a) within noncontextual ontological models. We find that the answer is 4 within numerical accuracy. By forming the Farkas' dual we find that the noncontextual inequality that is most violated by the optimal quantum implementation is

$$\begin{aligned} & + 0.000764p_{1,12} + 0.000764p_{2,12} \\ & - 0.000764p_{3,12} - 0.000764p_{4,12} \\ & + 0.250121p_{1,13} - 0.249879p_{2,13} \end{aligned}$$

<sup>4</sup>We did not specify any operational equivalences between the effects. As is explained in [27], in this case the extremal points of the measurement assignment polytope are the deterministic assignments consisting of zeros and ones. As we have six binary measurements the number of vertices is 2<sup>6</sup>.

$$\begin{aligned} & - 0.250121p_{3,13} + 0.249879p_{4,13} \\ & + 0.249258p_{1,14} - 0.249258p_{2,14} \\ & + 0.249258p_{3,14} - 0.249258p_{4,14} \\ & - 0.250742p_{1,23} + 0.250742p_{2,23} \\ & - 0.250742p_{3,23} + 0.250742p_{4,23} \\ & - 0.250121p_{1,24} + 0.249879p_{2,24} \\ & + 0.250121p_{3,24} - 0.249879p_{4,24} \\ & - 0.000764p_{1,34} - 0.000764p_{2,34} \\ & + 0.000764p_{3,34} + 0.000764p_{4,34} \leq 1, \end{aligned}$$

where  $p_{i,jk}$  is shorthand for  $p(1|P_i, M_{jk})$ . Our quantum implementation achieves a value of 1.414213561 for this inequality, or  $\sqrt{2}$  within numerical accuracy, a violation of  $\sqrt{2} - 1$  for the noncontextual bound.

As a final observation, we can directly compare the noncontextual bound on the success metric of Eq. (15a) to the optimal bit implementation as follows. The terms in the success metric in Eq. (15a) are written in the same order as the rows in Table IV. A correct value in the bit implementation corresponds to a 1 or 0 depending on the sign of the term in the success metric. An incorrect value in the bit implementation leads to a degradation of the success metric by 1. As it happens, the bit implementation also obtains a value of 4 for the success metric.

In the partial-ignorance communication task of type  $T_{4,2}$  Charlie still has four input strings  $s$ : 1000, 0100, 0010, and 0001. The inputs of Alice and Bob are reversed in a sense. Alice will receive one of the following input strings: 1100, 1010, 1001, 0110, 0101, and 0011. Bob, on the other hand, will receive one of the following input strings: 1000, 0100, 0010, or 0001. Hence Alice will need to prepare states with six distinct labels although some of these states may be identical and Bob will need to perform four ternary measurements. Alice will label her states as  $\varrho_{ij}$ , where the subscript  $ij$  will indicate the indices of the 1's in her input string. Bob will label his POVMs as  $M_k$ , where the  $k$  will correspond to the index of the 1 in Bob's input. Each of Bob's POVM  $M_k$  will have the set  $\{1, 2, 3, 4\} \setminus k$  as the outcome set. Thus, upon receiving measurement outcome  $l$ , Bob will guess that the index of the 1 in  $s$  was  $l$ . The success metric that Alice and Bob try to maximize is

$$\sum_{\substack{i,j,k,l=1 \\ i \neq j \neq k \neq l}}^4 \text{tr}[\varrho_{ij} M_k(l)].$$

We will, however, relabel the effects with the following rule:

$$M_k(l) = \begin{cases} M_k(l), & \text{if } l < k, \\ M_k(l-1), & \text{if } l > k. \end{cases}$$

With this relabeling all Bob's measurements will have the outcome set  $\{1, 2, 3\}$ . This makes the SDPs somewhat easier to program, but as a drawback Bob's outcome will not directly correspond with his guess, but has to be interpreted according to the relabeling.

*Example 4. (Contextual  $T_{4,2}$ ).* The success metric for the task  $T_{4,2}$ , written in terms of the first two effects, is written as

$$\begin{aligned} \max & \operatorname{tr}[\varrho_{23}\mathbf{M}_4(1)] + \operatorname{tr}[\varrho_{24}\mathbf{M}_3(1)] + \operatorname{tr}[\varrho_{34}\mathbf{M}_2(1)] \\ & + \operatorname{tr}[\varrho_{13}\mathbf{M}_4(2)] + \operatorname{tr}[\varrho_{14}\mathbf{M}_3(2)] + \operatorname{tr}[\varrho_{34}\mathbf{M}_1(1)] \\ & + 1 - \operatorname{tr}[\varrho_{12}(\mathbf{M}_4(1) + \mathbf{M}_4(2))] \\ & + \operatorname{tr}[\varrho_{14}\mathbf{M}_2(2)] + \operatorname{tr}[\varrho_{24}\mathbf{M}_1(2)] \\ & + 3 - \operatorname{tr}[\varrho_{12}(\mathbf{M}_3(1) + \mathbf{M}_3(2))] \\ & - \operatorname{tr}[\varrho_{13}(\mathbf{M}_2(1) + \mathbf{M}_2(2))] \\ & - \operatorname{tr}[\varrho_{23}(\mathbf{M}_1(1) + \mathbf{M}_1(2))], \end{aligned} \quad (16a)$$

$$\text{subject to } \varrho_{ij} \geq 0, \quad ij = 12, 13, 14, 23, 24, 34, \quad (16b)$$

$$\operatorname{tr}[\varrho_{ij}] = 1, \quad ij = 12, 13, 14, 23, 24, 34, \quad (16c)$$

$$\varrho_{12} + \varrho_{13} = \varrho_{14} + \varrho_{23} = \varrho_{24} + \varrho_{34}, \quad (16d)$$

$$\mathbf{M}_k(1), \mathbf{M}_k(2), \mathbf{M}_k(3) \geq 0, \quad k = 1, 2, 3, 4, \quad (16e)$$

$$\mathbf{M}_k(1) + \mathbf{M}_k(2) + \mathbf{M}_k(3) = \mathbb{1}, \quad k = 1, 2, 3, 4. \quad (16f)$$

The order of terms in the success metric of Eq. (16a) is the same as the rows in Table VI. A nontrivial operational equivalence between preparations is defined in Eq. (16d)

Querying a linear program for the maximal value to Eq. (16a) we obtain a maximal value of 8 within numerical accuracy for all noncontextual models. Interestingly, this again coincides with the optimal bit implementation of Table VI. The see-saw method also converges on exactly 8 for both qubits and qutrits. This leads us to suspect that no contextual advantage is possible in this task. The unitary hierarchy  $\mathcal{U}_1$  seems to break down for this task, so we will need to prove the outer bound with analytical methods.

*Proposition 2.* The outer bound for the success metric of Eq. (16a) equals exactly 8.

*Proof.* To prove the outer bound we will write the outcome distributions of the first POVM as a communication matrix:

$$A_1 = \begin{array}{c|ccc} & \mathbf{M}_1(1) & \mathbf{M}_1(2) & \mathbf{M}_1(3) \\ \hline \varrho_{12} & \operatorname{tr}[\varrho_{12}\mathbf{M}_1(1)] & \operatorname{tr}[\varrho_{12}\mathbf{M}_1(2)] & \operatorname{tr}[\varrho_{12}\mathbf{M}_1(3)] \\ \varrho_{13} & \operatorname{tr}[\varrho_{13}\mathbf{M}_1(1)] & \operatorname{tr}[\varrho_{13}\mathbf{M}_1(2)] & \operatorname{tr}[\varrho_{13}\mathbf{M}_1(3)] \\ \varrho_{14} & \operatorname{tr}[\varrho_{14}\mathbf{M}_1(1)] & \operatorname{tr}[\varrho_{14}\mathbf{M}_1(2)] & \operatorname{tr}[\varrho_{14}\mathbf{M}_1(3)] \\ \varrho_{23} & \operatorname{tr}[\varrho_{23}\mathbf{M}_1(1)] & \operatorname{tr}[\varrho_{23}\mathbf{M}_1(2)] & \mathbf{tr}[\varrho_{23}\mathbf{M}_1(3)] \\ \varrho_{24} & \operatorname{tr}[\varrho_{24}\mathbf{M}_1(1)] & \mathbf{tr}[\varrho_{24}\mathbf{M}_1(2)] & \operatorname{tr}[\varrho_{24}\mathbf{M}_1(3)] \\ \varrho_{34} & \mathbf{tr}[\varrho_{34}\mathbf{M}_1(1)] & \operatorname{tr}[\varrho_{34}\mathbf{M}_1(2)] & \operatorname{tr}[\varrho_{34}\mathbf{M}_1(3)] \end{array}.$$

The matrix  $A_1$  is row-stochastic, each row sums up to 1. The terms of  $A_1$  that show up in the success metric Eq. (16a) are highlighted in bold. Those three terms are the ones that we need to maximize in this communication matrix. From the operational equivalence  $\varrho_{14} + \varrho_{23} \simeq \varrho_{24} + \varrho_{34}$  it follows that

$$\begin{aligned} 2 & = \operatorname{tr}[\varrho_{24}(\mathbf{M}_1(1) + \mathbf{M}_1(2) + \mathbf{M}_1(3))] \\ & \quad + \operatorname{tr}[\varrho_{34}(\mathbf{M}_1(1) + \mathbf{M}_1(2) + \mathbf{M}_1(3))] \\ & = \operatorname{tr}[\varrho_{24}(\mathbf{M}_1(1) + \mathbf{M}_1(2))] + \operatorname{tr}[\varrho_{34}(\mathbf{M}_1(1) + \mathbf{M}_1(2))] \\ & \quad + \operatorname{tr}[(\varrho_{14} + \varrho_{23})\mathbf{M}_1(3)] \\ & \geq \operatorname{tr}[\varrho_{34}\mathbf{M}_1(1)] + \operatorname{tr}[\varrho_{24}\mathbf{M}_1(2)] + \operatorname{tr}[\varrho_{23}\mathbf{M}_1(3)]. \end{aligned}$$

TABLE VII. Inner bounds on the tasks  $T_{4,1}$  and  $T_{4,2}$  as produced by the see-saw algorithm.

$d$	2	3	4
$T_{4,1}$	10.89897946	11.65685425	12.0
$T_{4,2}$	8.0	12.0	12.0

Hence in any operational theory that satisfies the operational equivalence  $\varrho_{14} + \varrho_{23} \simeq \varrho_{24} + \varrho_{34}$  the inequality  $\operatorname{tr}[\varrho_{34}\mathbf{M}_1(1)] + \operatorname{tr}[\varrho_{24}\mathbf{M}_1(2)] + \operatorname{tr}[\varrho_{23}\mathbf{M}_1(3)] \leq 2$  holds. It turns out that a similar argument holds for the other POVMs as well so that the maximum value for Eq. (16a) is 8 for all operational theories that that satisfies the operational equivalence (16d).

The previous proposition concludes our contextual examples. In the next section we will study the tasks  $T_{4,1}$  and  $T_{4,2}$  without the operational equivalences.

### C. General bounds

We can produce inner bounds on the tasks  $T_{4,1}$  and  $T_{4,2}$  without the operational equivalences (15d) and (16d) by using the see-saw algorithm. Table VII collects these results. The bounds on  $T_{4,1}$  in Table VII are presented in terms of both effects, i.e., the negative terms in Eq. (15a) are replaced by the corresponding positive term.

We can see from Table VII that there is a major discrepancy between the tasks  $T_{4,1}$  and  $T_{4,2}$ . The inner bounds on task  $T_{4,1}$  gradually increase from the minimum value obtained by qubits to the maximum value of 12 obtained by four-dimensional quantum states. The value obtained by qubits is slightly greater than the contextual limit for the task.

For the task  $T_{4,2}$  the inner bound is 8 for qubits within numerical accuracy, the same as the outer bound found in Proposition 2. For qutrits the inner bound already reaches the maximum value of 12. We will now try to make sense of these numbers.

*Definition 9.* Let  $A$  be a communication matrix, i.e.,  $A_{ij} = \operatorname{tr}[\varrho_i\mathbf{M}(j)]$  for some finite set of quantum states  $\{\varrho_i\}_i$  and a POVM  $\mathbf{M}$  with a finite outcome set. The function  $\lambda_{\max}$  is defined as  $\lambda_{\max}(A) := \sum_j \max_i(A_{ij})$ .

The function  $\lambda_{\max}$  is an ultraweak monotone on the set of communication matrices. Ultraweak monotones were extensively studied in [6]. However, for our present investigation a more important result is the following.

*Proposition 3.* Let  $A$  be a communication matrix. Then

$$\operatorname{rank}_{\text{psd}}(A) \geq \lambda_{\max}(A),$$

where  $\operatorname{rank}_{\text{psd}}$  is the positive semidefinite rank of a matrix.

The above proposition was proved in [44]. For communication matrices it holds that a necessary and sufficient condition for a communication matrix to have a  $d$ -dimensional quantum implementation is that the positive-semidefinite rank of the matrix is not greater than  $d$  [6]. With this knowledge we can prove the outer bound on task  $T_{4,2}$ .

*Proposition 4.* The outer bound on task  $T_{4,2}$  is 8 for qubits.

*Proof.* The task  $T_{4,2}$  consists of four communication matrices. The communication matrix of the first POVM is the

following:

	$M_1(1)$	$M_1(2)$	$M_1(3)$
$A_1 =$	$\text{tr}[\varrho_{12}M_1(1)]$	$\text{tr}[\varrho_{12}M_1(2)]$	$\text{tr}[\varrho_{12}M_1(3)]$
	$\text{tr}[\varrho_{13}M_1(1)]$	$\text{tr}[\varrho_{13}M_1(2)]$	$\text{tr}[\varrho_{13}M_1(3)]$
	$\text{tr}[\varrho_{14}M_1(1)]$	$\text{tr}[\varrho_{14}M_1(2)]$	$\text{tr}[\varrho_{14}M_1(3)]$
	$\text{tr}[\varrho_{23}M_1(1)]$	$\text{tr}[\varrho_{23}M_1(2)]$	$\text{tr}[\varrho_{23}M_1(3)]$
	$\text{tr}[\varrho_{24}M_1(1)]$	<b><math>\text{tr}[\varrho_{24}M_1(2)]</math></b>	$\text{tr}[\varrho_{24}M_1(3)]$
	<b><math>\text{tr}[\varrho_{34}M_1(1)]</math></b>	$\text{tr}[\varrho_{34}M_1(2)]$	$\text{tr}[\varrho_{34}M_1(3)]$

The bold highlights those elements of  $A_1$  that should be maximized. If  $A_1$  is implementable with qubits, it follows that  $\text{rank}_{\text{psd}}(A_1) \leq 2$ . But then by Proposition 3,  $\lambda_{\max}(A_1) \leq 2$ . This means that the sum of the elements highlighted in bold cannot be greater than 2. A similar argument holds for the other POVMs so the total outer bound on task  $T_{4,2}$  for qubits is exactly 8.

We could use Proposition 4 to produce the outer bound 12 for qutrits. However, 12 is already the maximum value that the success metric can obtain, so there is no need to prove the outer bound for qutrits.

Let us now return to the task  $T_{4,1}$ . In this task Bob is using dichotomic measurements, that is, Bob is always trying to distinguish between two possible input states from Alice. The best strategy, therefore, tries to minimize the maximal overlap between Alice's states, so that they can be distinguished pair-wise as well as possible. This ensures there will not be any weakness in Alice and Bob's strategy that Charlie could exploit. We need just one more definition before we can prove that the bounds in Table VII are tight for  $T_{4,1}$ .

*Definition 10.* The ambiguous distinguishability of two pure states  $|\varphi_1\rangle\langle\varphi_1|$  and  $|\varphi_2\rangle\langle\varphi_2|$  is defined as

$$P_{\text{success}} = \frac{1}{2}(1 + \sqrt{1 - |\langle\varphi_1|\varphi_2\rangle|^2}).$$

*Proposition 5.* The inner bounds presented in Table VII are tight for the task  $T_{4,1}$ .

*Proof.* Notice that the success metric (15a) consists of six terms of the form  $\text{tr}[\varrho_j M_{ij}(1)] - \text{tr}[\varrho_i M_{ij}(1)]$ . We can interpret this term as  $P_{\text{success}} - P_{\text{fail}} = 2P_{\text{success}} - 1$  for the POVM  $M_{ij}$ . Let  $\{\varphi_i\}_{i=1}^4$  be a Grassmannian frame for their span in  $\mathbb{C}^d$ .

Then  $P_{\text{success}} \leq \frac{1}{2}(1 + \sqrt{1 - \frac{4-d}{3d}})$  by Proposition 1. Equality can be achieved if and only if  $\{\varphi_i\}_{i=1}^4$  is equiangular and that is the best possible success chance for Bob while minimizing the maximum overlap between Alice's states. For qubits we get  $P_{\text{success}} \leq \frac{1}{2}(1 + \sqrt{\frac{2}{3}})$ . Omitting the  $-1$  terms we get an outer

bound of  $12P_{\text{success}} = 6(1 + \sqrt{\frac{2}{3}}) \approx 10.89897949$ , matching the value in Table VII by six decimals. For qutrits the corresponding values are  $P_{\text{success}} \leq \frac{1}{2}(1 + \frac{2\sqrt{2}}{3})$  and outer bound  $12P_{\text{success}} = 6(1 + \frac{2\sqrt{2}}{3}) \approx 11.65685425$ , again matching the value in Table VII within numerical accuracy. For  $d = 4$  the maximal value of 12 is achieved as there are four distinguishable states.

## V. FINAL THOUGHTS

In this article we continued the work done in [2] by introducing an input for Bob in the setting of communication of partial ignorance. We called these new types of

communication tasks simply partial-ignorance communication tasks of type  $T_{n,m}$ , where  $n$  is an integer representing the length of Charlie's string  $s$  and  $m$  represents loosely the amount of information on  $s$  that Charlie reveals to Alice. Whatever information Charlie does not reveal to Alice he will instead reveal to Bob, so that the partial ignorance communication task of type  $T_{n,m}$  is always informationally complete, that is, Alice and Bob's combined knowledge is enough to solve the communication task perfectly.

We began by analyzing the simplest communication task of type  $T_{3,1}$  for bits and qubits. This proved to be a relatively straightforward task. The tasks  $T_{4,1}$  and  $T_{4,2}$  were much more complicated. We used various methods to establish inner and outer bounds on success metrics. These methods included SDPs, ultraweak monotones on communication matrices, and frame theory for quantum states. We found out that the bit was as good as a communication medium as any noncontextual ontological model for both tasks  $T_{4,1}$  and  $T_{4,2}$ . For  $T_{4,1}$  we observed a contextual advantage for qubits and proved tight general bounds on the success metric for qubits and qutrits. For the task  $T_{4,2}$  we proved that a contextual advantage was impossible for the operational equivalence we chose and that the qubit did not perform any better than the bit.

The complexity involved in solving the inner and outer bounds for communication tasks of type  $T_{n,m}$  grows rapidly with  $n$ . Actually, already for  $T_{4,2}$  we could have chosen another set of operational equivalences for preparations and possibly included operational equivalences for measurements as well. In general as  $n$  increases the number of preparations and measurements increases as well, and this makes the analysis more difficult. Additionally, with large  $n$  the optimal strategies might require systems with higher dimensions and this increases the computational demands of numerical methods. However, we believe that the methods presented in this article serve as a good starting point when solving communication tasks similar to those studied in this work.

There were several questions that ultimately could not be included in the scope of this article. The first question is what happens if we abandon the requirement of informational completeness so that Alice and Bob's combined knowledge does not determine the correct answer completely. This would certainly increase the complexity involved in solving for the optimal strategies, but we believe the current methods would suffice to analyze those cases as well. The second question involves shared randomness. It is known that shared randomness is a powerful resource for communication [42]. Alice and Bob's access to shared randomness would most certainly increase the effectiveness of their best strategies, as this would allow them to mix strategies in a way that would at the very least increase the worst case success probabilities for the bit. However, the effect of shared randomness is already tedious to analyze at the level of communication matrices. We leave this case for future research.

As we discussed earlier, the bit turned out to be as effective as a communication medium as any noncontextual ontological model in the tasks  $T_{4,1}$  and  $T_{4,2}$ . It is not entirely clear to us why this is the case. We did not impose the bit to respect any operational equivalences and the measurements simply consist of receiving the bit that was sent. In this sense the bit makes up a poor comparison for noncontext-

tual models. Nevertheless, the average success probabilities are the same. However, it must be kept in mind that the worst case success probability is always zero for the bit. The worst case success probability is greater than zero for qubits in the task  $T_{4,2}$ , but it is not clear what the actual worst-case probability is as the see-saw method is only capable of optimizing the average success chance (linear functions of states and effects). Frame theory does not help either because the equality in Eq. (13) cannot be achieved. We leave the worst-case success chance for qubits in  $T_{4,2}$  as an open problem.

As a final question we could consider possible connections of the communication tasks studied in this work to

other similar communication tasks. We believe that some partial-ignorance communication tasks could be mapped to a corresponding random access code. It might even be possible to map all RACs to some subset of partial-ignorance communication tasks of particular types. However, this connection is not clear to us at present time and requires further thought.

#### ACKNOWLEDGMENTS

O.K. personally thanks Teiko Heinosaari for his support during the research for this article and his insightful comments. O.K. acknowledges financial support from the Turku University Foundation during the research of this article.

- 
- [1] B. Schumacher and M. Westmoreland, *Quantum Processes, Systems, and Information* (Cambridge University Press, Cambridge, England, 2010).
  - [2] T. Heinosaari and O. Kerppo, Communication of partial ignorance with qubits, *J. Phys. A: Math. Theor.* **52**, 395301 (2019).
  - [3] S. Wiesner, Conjugate coding, *SIGACT News* **15**, 78 (1983).
  - [4] A. Ambainis, A. Nayak, A. Ta-Shma, and U. Vazirani, Dense quantum coding and quantum finite automata, *J. ACM* **49**, 496 (2002).
  - [5] A. Ambainis, D. Leung, L. Mancinska, and M. Ozols, Quantum random access codes with shared randomness, [arXiv:0810.2937](https://arxiv.org/abs/0810.2937).
  - [6] T. Heinosaari, O. Kerppo, and L. Leppäjärvi, Communication tasks in operational theories, *J. Phys. A: Math. Theor.* **53**, 435302 (2020).
  - [7] M. Navascués, S. Pironio, and A. Acín, Bounding the Set of Quantum Correlations, *Phys. Rev. Lett.* **98**, 010401 (2007).
  - [8] M. Navascués, S. Pironio, and A. Acín, A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations, *New J. Phys.* **10**, 073013 (2008).
  - [9] M. Navascués and T. Vértesi, Bounding the Set of Finite Dimensional Quantum Correlations, *Phys. Rev. Lett.* **115**, 020501 (2015).
  - [10] M. Navascués, A. Feix, M. Araújo, and T. Vértesi, Characterizing finite-dimensional quantum behavior, *Phys. Rev. A* **92**, 042117 (2015).
  - [11] P. Wittek, Algorithm 950: Ncpol2sdpa-Sparse Semidefinite Programming Relaxations for Polynomial Optimization Problems of Noncommuting Variables, *ACM Trans. Math. Softw.* **41**, 1 (2015).
  - [12] P. Mironowicz, Applications of semi-definite optimization in quantum information protocols, Ph.D. thesis, Gdansk University of Technology, 2015.
  - [13] Y. Wang, I. W. Primaatmaja, E. Lavie, A. Varvitsiotis, and C. C. W. Lim, Characterising the correlations of prepare-and-measure quantum networks, *npj Quantum Inf.* **5**, 17 (2019).
  - [14] A. Tavakoli, E. Cruzeiro, R. Uola, and A. Abbott, Bounding and simulating contextual correlations in quantum theory, *PRX Quantum* **2**, 020334 (2021).
  - [15] A. Tavakoli, E. Zambrini Cruzeiro, E. Woodhead, and S. Pironio, Informationally restricted correlations: a general framework for classical and quantum systems, *Quantum* **6**, 620 (2022).
  - [16] A. Chaturvedi, M. Farkas, and V. Wright, Characterising and bounding the set of quantum behaviours in contextuality scenarios, *Quantum* **5**, 484 (2021).
  - [17] R. F. Werner and M. M. Wolf, Bell Inequalities and Entanglement, *Quantum Inf. Comput.* **1**, 1 (2001).
  - [18] Y. C. Liang and A. C. Doherty, Better Bell-inequality violation by collective measurements, *Phys. Rev. A* **73**, 052116 (2006).
  - [19] Y. C. Liang and A. C. Doherty, Bounds on quantum correlations in Bell-inequality experiments, *Phys. Rev. A* **75**, 042103 (2007).
  - [20] Y. C. Liang, C. W. Lim, and D. L. Deng, Reexamination of a multisetting Bell inequality for qudits, *Phys. Rev. A* **80**, 052116 (2009).
  - [21] A. Ambainis, M. Banik, A. Chaturvedi, D. Kravchenko, and A. Rai, Parity oblivious d-level random access codes and class of noncontextuality inequalities, *Quantum Inf. Process.* **18**, 111 (2019).
  - [22] J. S. Bell, On the problem of hidden variables in quantum mechanics, *Rev. Mod. Phys.* **38**, 447 (1966).
  - [23] S. Kochen and E. P. Specker, The Problem of hidden variables in quantum mechanics, *J. Math. Mech.* **17**, 59 (1967).
  - [24] R. W. Spekkens, Contextuality for preparations, transformations, and unsharp measurements, *Phys. Rev. A* **71**, 052108 (2005).
  - [25] R. W. Spekkens, D. H. Buzacott, A. J. Keehn, B. Toner, and G. J. Pryde, Preparation Contextuality Powers Parity-Oblivious Multiplexing, *Phys. Rev. Lett.* **102**, 010401 (2009).
  - [26] M. D. Mazurek, M. F. Pusey, R. Kunjwal, K. J. Resch, and R. W. Spekkens, An experimental test of noncontextuality without unphysical idealizations, *Nat. Commun.* **7**, 11780 (2016).
  - [27] D. Schmid, R. Spekkens, and E. Wolfe, All the noncontextuality inequalities for arbitrary prepare-and-measure experiments with respect to any fixed set of operational equivalences, *Phys. Rev. A* **97**, 062103 (2018).
  - [28] D. Schmid and R. Spekkens, Contextual Advantage for State Discrimination, *Phys. Rev. X* **8**, 011015 (2018).
  - [29] D. Saha and A. Chaturvedi, Preparation contextuality as an essential feature underlying quantum communication advantage, *Phys. Rev. A* **100**, 022108 (2019).
  - [30] R. Kunjwal, M. Lostaglio, and M. F. Pusey, Anomalous weak values and contextuality: Robustness, tightness, and imaginary parts, *Phys. Rev. A* **100**, 042116 (2019).
  - [31] A. Tavakoli and R. Uola, Measurement incompatibility and steering are necessary and sufficient for

- operational contextuality, *Phys. Rev. Research* **2**, 013011 (2020).
- [32] M. F. Pusey, J. Barrett, and T. Rudolph, On the reality of the quantum state, *Nat. Phys.* **8**, 475 (2012).
- [33] M. Leifer, Is the quantum state real? An extended review of  $\psi$ -ontology theorems, *Quanta* **3**, 67 (2014).
- [34] E. D. Andersen, Certificates of Primal or Dual Infeasibility in Linear Programming, *Comp. Opt. Applic.* **20**, 171 (2001).
- [35] R. J. Duffin and A. C. Schaeffer, A Class of Nonharmonic Fourier Series, *Trans. Am. Math. Soc.* **72**, 341 (1952).
- [36] I. Daubechies, Ten Lectures on Wavelets, *CBMS-NSF Regional Conference Series in Applied Mathematics* (SIAM, Philadelphia, 1992).
- [37] R. Balan, P. Casazza, C. Heil, and Z. Landau, Density, overcompleteness, and localization of frames, *Electron. Res. Announc. Am. Math. Soc.* **12**, 71 (2006).
- [38] T. Strohmer and R. W. Heath, Jr., Grassmannian frames with applications to coding and communication, *Appl. Comput. Harmon. Anal.* **14**, 257 (2003).
- [39] N. Leonhard, Correlation minimizing frames, Ph.D. thesis, University of Houston, 2016.
- [40] L. R. Welch, Lower bounds on the maximum cross-correlation of signals, *IEEE Trans. Inform. Theory* **20**, 397 (1974).
- [41] O. Christensen, *Frames and Bases: An Introductory Course* (Birkhäuser, Basel, Switzerland, 2008).
- [42] P. Frenkel and M. Weiner, Classical Information Storage in an n-Level Quantum System, *Commun. Math. Phys.* **340**, 563 (2015).
- [43] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, Cambridge, England, 2004).
- [44] T. Lee, Z. Wei, and R. de Wolf, Some upper and lower bounds on PSD-rank, *Math. Program.* **162**, 495 (2017).