

# Optimal amount of entanglement to distinguish quantum states instantaneously

Berry Groisman and Sergii Strelchuk

*Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 0WA, U.K.*

We introduce a new aspect of nonlocality which arises when the task of quantum states distinguishability is considered under local operations and shared entanglement in the absence of classical communication. We find the optimal amount of entanglement required to accomplish the task perfectly for sets of orthogonal states and argue that it quantifies information nonlocality.

## I. INTRODUCTION

According to the postulates of quantum mechanics any set of orthogonal states of a *single* quantum system in principle can be perfectly (i.e. with probability 1) distinguished by an appropriate projective measurement. A composite system can be treated effectively as a single system if one can implement a global projection measurement on it.

It is interesting to look at nontrivial scenarios, where the class of possible global projection measurements is restricted. Such scenarios arise naturally in quantum information theory when two or more quantum systems are separated in space, in which case implementation of a global operation on these systems will be restricted by the available resources of communication (quantum and classical). The implications of these restrictions on distinguishability properties of sets of orthogonal joint [1] states have been used extensively as a tool to study fundamental properties of quantum systems [2–5].

Communication resources in quantum information theory can be divided into classical and quantum. The abundance of both types of resources trivialises the task of quantum state distinguishability, since all the systems can be teleported to one location where a required global operation can be implemented, after which the systems are teleported back to their original locations. Restrictions on available resources affect distinguishability and particularly interesting operational settings arise when one type of resource is completely prohibited. It is under these restrictions that the nature of quantumness of the states is revealed. The first case corresponds to local operations, accompanied by classical communication, without access to shared entanglement (LOCC). The second case is a dual to LOCC — where we are restricted to local operations and shared entanglement (LOSE).

It is natural to expect that the nonlocal nature of composite systems with parts separated in space will manifest itself in their distinguishability properties. In the past it was widely believed that entangled states are the only states that exhibit nonlocal behaviour. However, it was shown that there are aspects of quantum states which, albeit not being associated with entanglement, are difficult to reconcile with locality. It might not come as a surprise that some sets of orthogonal entangled states cannot be distinguished by LOCC. What does come as a surprise is that certain sets of mutually orthogonal *product* (i.e. nonentangled) multipartite states cannot be perfectly distinguished by LOCC. This is contrary to what one would have expected from a set of classical states. Thus, despite the absence of entanglement in the states themselves there is a certain aspect of nonclassicality (quantumness) to such states. The fact that their distinguishability cannot be facilitated without utilizing entangled resources prompted researchers to conclude that there is a new type of nonlocality involved, *nonlocality without entanglement* [6].

The change of paradigm from LOCC to LOSE has profound implications on distinguishability. There exist families of sets which are perfectly distinguishable by LOCC with communicating as little as one classical bit, yet, cannot be perfectly distinguished by LOSE with arbitrary large but finite resources. As LOSE does not involve classical communication, corresponding protocols can be considered as instantaneous. The parties can implement their local interactions — unitary operations and measurements — in a very short time: that is, these operations occur in spacelike separated regions. During this process all the (classical) results are recorded locally. The parties still need to combine their local classical records in order to identify the initial state. However, the latter step is not considered to be a part of the measurement process [7].

Distinguishability under LOCC is well understood [2–5]. However, little is known about distinguishability under LOSE with limited resources. It was shown that for finite-dimensional systems any multipartite mutually orthogonal set can be discriminated with unlimited resources of entanglement [8, 9]. The latter works did not address the question of finding optimal amounts of entanglement required to distinguish particular sets. The characteristic feature of the above protocols is that they consume all pre-shared entanglement. Recently, an alternative scheme had been proposed [10] where a finite number of ebits is consumed *on average*. The protocol is designed to halt as soon as all required local operations are completed successfully leaving the rest of the pre-shared entanglement resources intact. It should be noted, however, that with nonzero probability such protocols consume entanglement well beyond their expectation value. As such, their worst-case performance is still very far from the optimal.

In this paper we find the minimum amount of resources which must be *available* to the parties beforehand in order to accomplish the task perfectly. This leads us to introducing the new aspect of nonlocality. We show that the algorithm of remote instantaneous rotations introduced in [8] requires an optimal amount of entanglement with respect to this aspect. The method of instantaneous teleportations suggested in [9] will consume the same amount of entanglement.

Several recent works endeavour to construct efficient protocols and to find better bounds on the amount of entanglement required [10, 11]. One important application of the improved bounds for entanglement consumption is attacks for position-based cryptography. They rely on the ability to perform state transformations efficiently. It is currently known that an adversary having access to an exponential amount of entanglement can successfully attack a number of position-based verification protocols. It is currently not known whether one can do better. Our work provides additional evidence that only a polynomial amount of entanglement may be required to perform successful attacks on position-based cryptography protocols [12].

This article is organised as follows. In Section II we formulate the problem in the form of a steering game and define information nonlocality. Section III presents a general steering algorithm. Section IV provides a proof of optimality of that algorithm in terms of entanglement resources. In Section V we discuss generalizations of the steering sets.

## II. STEERING GAME AND NEW MEASURE OF INFORMATION NONLOCALITY

Consider a bipartite quantum state drawn from the known ensemble  $\{|\psi_i\rangle, p_i\}_{i=1}^m$  of mutually orthogonal bi-partite states shared by two parties. Their uncertainty about the state is quantified by its entropy.

In the absence of classical communication, the manner in which the information about the state can be accessed locally dictates the nonlocality property of the set. We say that information encoded in the ensemble (set) is *bi-localised* if it can be accessed by the parties using local operations (here accessed does not imply interpreted cf. [7]) and *de-localised* if it is not accessible by local operations alone. In the latter case we say that the ensemble is associated with *information nonlocality*. The aim of the parties in our setup is to bi-localise the nonlocal information encoded in the set of quantum states by means of LOSE. Thus, we will refer to it as *information nonlocality of instantaneous bi-localisation*. It is ultimately linked with perfect distinguishability, which operationally depends on parameter(s) characterising the states  $\{|\psi_i\rangle\}$ , but not on  $\{p_i\}$ . Hence we expect that a quantitative measure of information nonlocality should not depend on the latter. Thus, we consider the task of state discrimination in the presence of de-localised information. The optimal amount of resources required to complete the task is quantified by the number of ebits needed to access, i.e., bi-localise, the de-localised information. Clearly, LOSE is an appropriate framework to analyse the bi-localization of classical information because it prohibits the classical communication between two parties which would otherwise trivialise the task.

We now introduce a *steering game* and a new quantity which will serve as a figure of merit for this game. The latter quantifies the information nonlocality in sets of orthogonal quantum states. To illustrate this task, consider two parties, Alice and Bob, who share a state from the known set

$$\mathcal{A}[\alpha] = \{|0\rangle_A|0\rangle_B, |0\rangle_A|1\rangle_B, |1\rangle_A|+\alpha\rangle_B, |1\rangle_A|-\alpha\rangle_B\}, \quad (1)$$

where  $|+\alpha\rangle = \cos\alpha|0\rangle + \sin\alpha|1\rangle$ ,  $|-\alpha\rangle = \sin\alpha|0\rangle - \cos\alpha|1\rangle$ . When it does not cause confusion, we will omit the subscripts which denote the label of the subsystem. The state is chosen and prepared by an external party and is unknown to Alice and Bob. By means of LOSE parties try to steer set  $\mathcal{A}$  to set  $\mathcal{B} = \{|0\rangle|0\rangle, |0\rangle|1\rangle, |1\rangle|0\rangle, |1\rangle|1\rangle\}$  (with no particular ordering of states) minimizing the amount of entanglement used in the process. We denote the task of converting the states from set  $\mathcal{A}$  to set  $\mathcal{B}$  with probability 1 as  $\mathcal{A} \rightarrow \mathcal{B}$ . For both sets, the overall phase in front of the states will play no role and will further be ignored.

**Definition 1** We say that the set of two-qubit states  $\mathcal{A}[\alpha]$  contains  $\mathcal{I}(\mathcal{A}[\alpha]) = k$  bits of nonlocality of instantaneous bi-localisation, where  $\mathcal{I}(\mathcal{A}[\alpha])$  is the length of the binary expansion of  $\alpha \bmod \frac{\pi}{2}$ :

$$\mathcal{I}(\mathcal{A}[\alpha]) := |(\bar{\alpha})_2|. \quad (2)$$

It was demonstrated in [8] how this game is won using  $\mathcal{I}(\mathcal{A}[\alpha])$  pre-shared ebits. In Section IV we prove that this number of ebits is indeed optimal. Thus, we treat  $\mathcal{I}$  as a measure, quantifying the optimal amount of resources to steer the set  $\mathcal{A}$  defined in (1) to the set  $\mathcal{B}$ .

Hence, for finite length of the binary expansion of  $\alpha \bmod \frac{\pi}{2}$  one can steer perfectly with a finite amount of entanglement [8]. However, when the expansion of  $\alpha \bmod \frac{\pi}{2}$  is infinite, then all known algorithms use an infinite amount of entanglement required to achieve  $\mathcal{A} \rightarrow \mathcal{B}$  [13].

The use of entanglement to accomplish the task is essential, although not immediately clear. One cannot perform  $\mathcal{A} \rightarrow \mathcal{B}$  using LO with a weaker resource such as common randomness because in the absence of classical communication it can be seen to represent a pre-agreed strategy. As such, common randomness is equivalent to Alice and Bob committing to a sequence of LO in advance.

Our definition is not sensitive to the probability distribution of the four states (1) because we require the probability of success  $p = 1$ . If they are equiprobable, i.e.  $p_i = 1/4 \forall i$ , then they encode two classical bits of which  $1 - H_2(\cos^2 \alpha)$  are de-localised. This quantity, known as zero-way quantum information deficit [14], is unsuitable to quantify information nonlocality of the set for two main reasons. First, the above deficit depends on the probability distribution  $\{p_i\}$  of the corresponding states, which is irrelevant as far as distinguishability is concerned. Clearly, physical resources needed to manipulate this set of states cannot depend on that probability distribution. Second, it is natural to expect that the measure is related to uncertainty about the basis on Bob's side in (1). Thus, our operational understanding of information nonlocality of the set  $\{|\psi_i\rangle, p_i\}$  is quite different from the previously suggested ways of quantifying nonlocality of the mixtures  $\rho_{AB} = \sum p_i |\psi_i\rangle\langle\psi_i|$ , where  $|\psi_i\rangle \in \mathcal{A}$ . Our measure also differs from accessible information [15], which quantifies mutual information between an ensemble and the outcomes of an optimal measurement, and as such depends on the probability distribution. In our distinguishability scenario accessible information is always maximal, i.e. two bits, while the figure of merit is the amount of entanglement required to carry out a measurement, which achieves this value. A similar comparison can be made with relative entropy of quantumness [16] and quantum discord [17].

The parametrization of  $\mathcal{A}$  can be easily extended to multiple angles:

$$\begin{aligned} \mathcal{A}[\alpha_1, \dots, \alpha_n] = \{ & |0\rangle|0\rangle, |0\rangle|1\rangle, \\ & |1\rangle|+\alpha_1\rangle, |1\rangle|-\alpha_1\rangle, \\ & |2\rangle|+\alpha_2\rangle, |2\rangle|-\alpha_2\rangle, \\ & \vdots \\ & |n\rangle|+\alpha_n\rangle, |n\rangle|-\alpha_n\rangle\} \end{aligned} \quad (3)$$

where in general  $\alpha_i \neq \alpha_j$  when  $i \neq j$ . In this case we define  $\mathcal{I}(\mathcal{A}[\alpha_1, \dots, \alpha_n]) = \max_{1 \leq i \leq n} |(\bar{\alpha}_i)_2|$ . Then, the set  $\mathcal{B}$  becomes  $\mathcal{B} = \{|i\rangle|0\rangle, |i\rangle|1\rangle\}_{i=0}^n$ . In Sec. III we will show how the task can be accomplished using  $\mathcal{I}(\mathcal{A}[\alpha_1, \dots, \alpha_n])$  pre-shared ebits.

The above extension yields the following two properties of  $\mathcal{I}$  (here we assume  $|(\bar{\alpha}_i)_2| < \infty$ ):

- $\mathcal{I}$  is *faithful*:  $\mathcal{I}(\mathcal{A}[\alpha_1, \dots, \alpha_n]) = 0$  iff  $\forall i \alpha_i = 0$ .
- $\mathcal{I}$  is *subadditive* with respect to  $\alpha_i$ :

$$\mathcal{I}(\mathcal{A}[\alpha_1, \alpha_2]) \leq \mathcal{I}(\mathcal{A}[\alpha_1]) + \mathcal{I}(\mathcal{A}[\alpha_2]). \quad (4)$$

It trivially holds for  $\alpha_1 = \alpha_2$ . Considering sets  $\mathcal{A}$  with more nonlocal states corresponding to different values of  $\alpha_i$  increases the complexity of the steering task. However, as we show in Sec. III the required number of ebits is  $\max_i |(\bar{\alpha}_i)_2|$ . This is due to the fact that Alice and Bob may take advantage of the collective local transformation in order to reset the corresponding bit in the binary expansion regardless of the phase. The collective local operation of each of the parties will result in resetting the lowest bit in the binary expansion of each of the phases without the knowledge about the precise identity of the shared state.

### III. GENERAL STEERING ALGORITHM

In this section we present the protocol  $\mathcal{Q}^{(k)}[\alpha_1, \dots, \alpha_n]$ , which steers the set  $\mathcal{A}[\alpha_1, \dots, \alpha_n]$  to  $\mathcal{B}$  using  $k$  ebits, and show that it is optimal with respect to  $\mathcal{I}$ . Recall that Alice and Bob are given one of the elements from  $\mathcal{A}[\alpha_1, \dots, \alpha_n]$ , and their task is by performing local operations to turn it into one of the states of the set  $\mathcal{B}$  using the minimal amount of the shared entanglement. We present a protocol for the steering game and in the subsequent section prove its optimality.

#### A. A strategy for $\mathcal{A}[\alpha_1, \dots, \alpha_n] \rightarrow \mathcal{B}$

The basic building block of the protocol is a controlled instantaneous probabilistic rotation introduced in [8]. The parties share  $k$  ebits,  $|\Phi^+\rangle_{a_t b_t}$  ( $t = 1, \dots, k$ ), where  $k = \max_i \mathcal{I}(\mathcal{A}[\alpha_i])$  (we assume that  $k$  is finite). The algorithm

they execute constitutes of the following pre-agreed sequence of actions, which they perform  $k$  times. Each iteration  $t$  utilises one entangled pair  $|\Phi^+\rangle_{a_t b_t}$  and results in changing the set from  $\mathcal{A}_t [\alpha_1^{(t)}, \dots, \alpha_n^{(t)}]$  to  $\mathcal{A}_{t+1} [\alpha_1^{(t+1)}, \dots, \alpha_n^{(t+1)}]$ .

1. Bob implements a C-NOT interaction between the qubit  $b_t$  and the target qubit  $B$ :

$$U_{b_t B}^{CNOT} = |0\rangle\langle 0|_{b_t} \otimes \mathbb{I}_B + |1\rangle\langle 1|_{b_t} \otimes \sigma_y^B, \quad (5)$$

where index  $t$  denotes the iteration round of the protocol.

2. Bob then measures  $\sigma_x$  on particle  $b_t$  and records the result,  $r_{b_t} = \pm 1$ .
3. Alice performs a controlled rotation

$$U_{A a_t} = \sum_{j=0}^n |j\rangle\langle j|_A \otimes R_{a_t}^x(\alpha_j^{(t)}), \quad (6)$$

where  $R_{a_t}^x(\alpha_j^{(t)}) \equiv \exp\{i\alpha_j^{(t)}\sigma_x^{a_t}\}$ , and  $\alpha_0 = 0$ , which results in rotating qubit  $B$  by  $r_{b_t}\alpha_j^{(t)}$ , depending on the state  $|j\rangle$  of Alice's qubit  $A$ .

4. If  $r_{b_t} = +1$ , then the set is transformed to  $\mathcal{A}_{t+1}^+[0, \dots, 0] = \mathcal{B}$ , while the case of  $r_{b_t} = -1$  corresponds to  $\mathcal{A}_{t+1}^- [2\alpha_1^{(t)}, \dots, 2\alpha_n^{(t)}]$ . In the former case, Bob ceases his actions (Alice continues with her actions, but they do not affect the state of qubit  $A$ ). In the latter case the protocol proceeds to the next iteration step.

Thus, on each step there is a 50% chance of performing the map  $\mathcal{A}_t \rightarrow \mathcal{B}$ . The algorithm terminates with certainty at  $t = k$ , because  $\mathcal{A}_{k+1}^\pm = \mathcal{B}$ . (It is easy to see that when  $n = 1$  the protocol above reduces to the protocol described in [8].)

The actions of Alice and Bob take place in two spacelike separated regions, therefore the protocol can be presented as a tensor product of two unitary transformations (see Sec. IV A). In addition, the Appendix provides a very elegant presentation of the protocol in the language of state-operators — *stators* — introduced in [18]. The latter is a universal resource for remote LOCC and LOSE operations.

## B. Binary representation

It is convenient to represent the set  $\mathcal{A}[\alpha_1, \dots, \alpha_n]$  as a table in which row  $i$  contains the binary expansion of  $\alpha_i$  (modulo  $\pi/2$ ). Alice and Bob compute  $k$ , and construct a table with binary expansions of  $\alpha_i$  aligned by the binary order. When the expansion for  $\alpha_i$  is less than  $k$  they pad it with zeros until its length is  $k$ . The example for the set  $\mathcal{A}[\frac{7\pi}{16}, \frac{5\pi}{16}, \frac{3\pi}{16}, \frac{1\pi}{4}]$  is shown in Table I.

$i$	$\frac{\alpha_i}{\pi}$	$(\frac{\alpha_i}{\pi} \bmod \frac{1}{2})_2$
1	7/16	0.111
2	5/16	0.101
3	3/16	0.011
4	1/4	0.100

TABLE I. Binary representation of  $\alpha_i$  for  $\mathcal{A}[\frac{7\pi}{16}, \frac{5\pi}{16}, \frac{3\pi}{16}, \frac{1\pi}{4}]$ . The values are calculated modulo  $1/2$ .

The ultimate aim of the protocol is to obtain a table with all entries being zero, which corresponds to  $\mathcal{B}$ . There is a simple global algebraic operation which nullifies all the nonzero entries. Its effect is equivalent to multiplication by  $2^k$ .

The advantage of the binary representation is to clarify and simplify the action of the protocol. Each iteration step results in the entries of the table being multiplied either by  $2^{k-t}$  (if  $r_{b_t} = +1$ ) or by 2 (if  $r_{b_t} = -1$ ). Alice and Bob update their tables accordingly. Bob updates the table depending on the value of  $r_{b_t}$ . If  $r_{b_t} = +1$  then no further action is required on his side. Alice, who does not have access to  $r_{b_t}$ , updates her table as if  $r_{b_t} = -1$  and carries on to the next iteration, where she implements (6) with new values of  $\alpha_i^{(t+1)}$  according to her new table. The algorithm

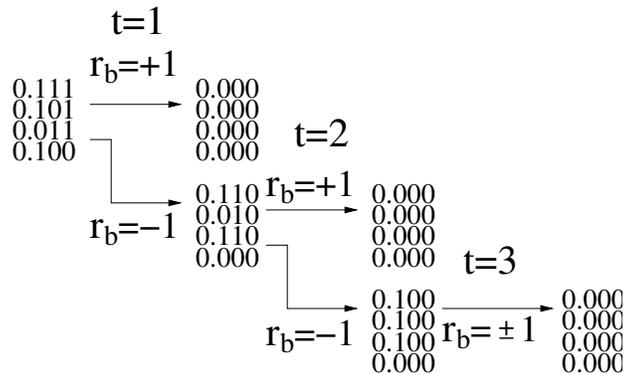


FIG. 1. Schematic representation of the steering protocol for  $\mathcal{A} \left[ \frac{7\pi}{16}, \frac{5\pi}{16}, \frac{3\pi}{16}, \frac{\pi}{4} \right]$ . On each  $2^{k-t}$  or 2 depending on the value of  $r_{b_t}$ . The resulting values are then taken modulo  $1/2$ .

is guaranteed to terminate at  $t = k$ . The example of the protocol for the set given in Table I is schematically shown in Figure 1.

The result of the application of the above protocol resembles that of [19] where authors learned each digit of the binary expansion sequentially.

Our protocol can be viewed as one which bi-localises information bit-by-bit and this may be interpreted as the resetting of delocalised information with the cost of resetting quantified by the number of ebits expended, much in the spirit of Landauer's principle.

### C. Approximate steering

In practical applications, we would not seek perfect discrimination of states, i.e. we would not need the perfect steering, namely turning the set  $\mathcal{A}[\alpha]$  to  $\mathcal{B}$ . Instead, we may want to obtain a guarantee that we are close to  $\mathcal{B}$  with respect to some notion of proximity defined below. This is especially relevant if there exists a state with a very large length of binary expansion or in the case when  $\alpha/\pi$  is an irrational number — infinite expansion.

**Definition 2** *The set  $\mathcal{A}[\alpha_1, \dots, \alpha_n]$  is  $\epsilon$ -close to  $\mathcal{B}$  if the error probability of discriminating any pair of the states from  $\mathcal{A}$  under local operations (with no entanglement and classical communication) is  $p_e \leq \epsilon$ .*

Consider  $n = 1$  and each of the states in  $\mathcal{A}[\beta]$  is equiprobable. In this case, the optimal discrimination of states from  $\mathcal{A}[\beta]$  is achieved when Bob simply measures qubit  $B$  in the basis  $\{|+\beta/2\rangle, |-\beta/2\rangle\}$  [20], with probability of error  $p_e = \sin^2(\beta/2) = \epsilon$ . Thus,  $\forall \alpha \in (0, \pi/4)$  and  $\epsilon \in (0, 1)$  there exists  $\beta$  such that  $\mathcal{A}[\beta]$  is  $\epsilon$ -close to  $\mathcal{B}$  and we can steer  $\mathcal{A}[\alpha]$  to  $\mathcal{A}[\beta]$ . It is easy to show that for a nonuniform distribution of the states in  $\mathcal{A}[\alpha]$  the error probability  $p_e$  is upper bounded by  $\epsilon$ .

Alice and Bob perform iterations of the algorithm of Sec. III A which resets the digits starting from the position which corresponds to the position of the most significant digit in the expansion of  $2\epsilon$ . Hence, the number of ebits which the parties consume is equal to the length of the binary expansion of the portion of  $\alpha$  until the position of the most significant digit in the expansion of  $2\epsilon$ .

## IV. PROOF OF OPTIMALITY

We now prove our main result — the optimality of the steering algorithm for set  $\mathcal{A}[\alpha]$ . It generalises to multiple sets of angles in a straightforward manner. As a preparatory step we introduce several useful definitions and describe generic properties of protocols.

**Definition 3** *Consider a set  $\mathcal{A}[\alpha]$  and  $k$  ebits  $|\Phi^+\rangle_{a_i b_i}$ ,  $i = 1, \dots, k$ , shared by Alice and Bob. Assume that there exists a protocol,  $\mathcal{P}^{(k)}[\alpha]$ , that is a sequence of pre-agreed local operations performed by Alice and Bob, that achieves  $\mathcal{A}[\alpha] \rightarrow \mathcal{B}$ . Without loss of generality, any such protocol can be represented as a two-step process.*

(1) *A tensor product of local unitary operations,  $V_{A\bar{a}_k} \otimes U_{B\bar{b}_k}$ , acting on qubits  $A, B$  and all ancillas  $a_i, b_i$ , which results in transformation*

$$V_{A\tilde{a}_k} \otimes U_{B\tilde{b}_k} \left( \mathcal{A}[\alpha] \bigotimes_{i=1}^k |\Phi^+\rangle_{a_i b_i} \right) = \mathcal{C}^{(k)}[\alpha], \quad (7)$$

where  $\mathcal{C}^{(k)}[\alpha]$  is the resulting set of entangled states of all particles and  $\tilde{a}_k = a_1 \dots a_k$ ,  $\tilde{b}_k = b_1 \dots b_k$ .

(2) Projection measurements on all ancillas  $a_i$ ,  $b_i$  and the original qubits  $A$ ,  $B$  in local computational bases.

Henceforth we will only consider the first step of the protocol.

**Definition 4** Two sets of bipartite states  $\mathcal{C}^{(k)}[\alpha]$  and  $\mathcal{C}^{(k)}[\beta]$  are said to be locally equivalent (LE) if they can be interconverted via local operations only. Without loss of generality any such local operation can be assumed to be unitary.

It is useful to introduce the following notation for the original set supplemented by  $k$  resource ebits:

$$\mathcal{A}^{(k)}[\alpha] \equiv \mathcal{A}[\alpha] \bigotimes_{i=1}^k |\Phi^+\rangle_{a_i b_i}. \quad (8)$$

Consider two sets  $\mathcal{A}^{(k)}[\alpha]$  and  $\mathcal{A}^{(k)}[\beta]$  and assume that there exist protocols,  $\mathcal{P}^{(k)}[\alpha]$  and  $\mathcal{P}^{(k)}[\beta]$ , which steer  $\mathcal{A}[\alpha]$  and  $\mathcal{A}[\beta]$  to  $\mathcal{C}^{(k)}[\alpha]$  and  $\mathcal{C}^{(k)}[\beta]$  respectively. Sets  $\mathcal{A}^{(k)}[\alpha]$  and  $\mathcal{C}^{(k)}[\alpha]$  (and similarly,  $\mathcal{A}^{(k)}[\beta]$  and  $\mathcal{C}^{(k)}[\beta]$ ) are locally equivalent by virtue of existence of  $\mathcal{P}^{(k)}[\alpha]$  and  $\mathcal{P}^{(k)}[\beta]$ .

**Proposition 5** Assume there exist two distinct protocols  $\mathcal{P}^{(k)}[\alpha]$  and  $\tilde{\mathcal{P}}^{(k)}[\alpha]$ , which steer  $\mathcal{A}^{(k)}[\alpha]$  to  $\mathcal{C}^{(k)}[\alpha]$  and  $\tilde{\mathcal{C}}^{(k)}[\alpha]$ , respectively. Then,  $\mathcal{C}^{(k)}[\alpha]$  and  $\tilde{\mathcal{C}}^{(k)}[\alpha]$  are locally equivalent.

**Proof** Observe that

$$\mathcal{C}^{(k)}[\alpha] \xrightarrow{(\mathcal{P}^{(k)}[\alpha])^{-1}} \mathcal{A}^{(k)}[\alpha] \xrightarrow{\tilde{\mathcal{P}}^{(k)}[\alpha]} \tilde{\mathcal{C}}^{(k)}[\alpha], \quad (9)$$

i.e.  $\mathcal{C}^{(k)}[\alpha]$  and  $\tilde{\mathcal{C}}^{(k)}[\alpha]$  are related by the unitary  $\tilde{\mathcal{P}}^{(k)}[\alpha](\mathcal{P}^{(k)}[\alpha])^{-1}$ . □

Observation (5) will be instrumental for our further analysis as it will allow us to restrict the discussion to a single set.

**Observation 6** Consider two orthogonal subspaces spanned by the first two states in  $\mathcal{A}[\alpha]$ ,  $\mathcal{A}_0[\alpha] = \{|00\rangle, |01\rangle\}$  and the last two states,  $\mathcal{A}_1[\alpha] = \{|1+\alpha\rangle, |1-\alpha\rangle\}$ . An optimal protocol  $\mathcal{P}^{(k)}[\alpha] = \mathcal{P}_1^{(k)}[\alpha] \oplus \mathcal{P}_2^{(k)}[\alpha]$  can be defined by its action on the respective subspaces, i.e.

$$\begin{cases} \mathcal{A}_0^{(k)}[\alpha] \rightarrow_{\mathcal{P}_0^{(k)}[\alpha]} \mathcal{C}_0^{(k)}[\alpha], \\ \mathcal{A}_1^{(k)}[\alpha] \rightarrow_{\mathcal{P}_1^{(k)}[\alpha]} \mathcal{C}_1^{(k)}[\alpha]. \end{cases} \quad (10)$$

$\mathcal{A}_0[\alpha]$  and  $\mathcal{A}_1[\alpha]$  are locally equivalent. [Indeed, they are related by a transformation  $\mathbb{I}_A \otimes R_B^y(\alpha)$ .]

### A. The instantaneous rotations protocol

Earlier, in Section III, we described the step-by-step protocol  $\mathcal{Q}^{(k)}[\alpha]$  based on instantaneous rotations, which steers  $\mathcal{A}^{(k)}[\alpha]$  to  $\mathcal{C}^{(k)}[\alpha]$  for  $\alpha = \pi/2^{k+1}$ . This protocol can be written as the product of local unitaries of the form

$$\mathcal{Q}^{(k)}[\alpha] = \left( \prod_{i=1}^k \Lambda_{Aa_i}(2^{i-1}\alpha) \right) \otimes \left( \left[ \prod_{i=2}^k \left( (1 - P_1^{(i-1)}) \otimes \mathbb{I}_{Bb_i} + P_1^{(i-1)} \otimes \Omega_{Bb_i} \right) \right] \Omega_{Bb_1} \right), \quad (11)$$

where

$$\begin{aligned} \Lambda_{Aa}(\theta) &= |0\rangle\langle 0|_A \otimes \mathbb{I}_a + |1\rangle\langle 1|_A \otimes R_a^x(\theta), \\ \Omega_{Bb} &= \mathbb{I}_B \otimes |+\rangle\langle 0|_b + \sigma_y^B \otimes |-\rangle\langle 1|_b, \\ P_1^{(j)} &= |1_{b_1} 1_{b_2} \dots 1_{b_j}\rangle\langle 1_{b_1} 1_{b_2} \dots 1_{b_j}|, \end{aligned} \quad (12)$$

with  $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$ . The task can be accomplished with  $m > k$  ebits by acting on the remaining  $m - k$  ebits by  $Z = \mathbb{I}_{a_j} \otimes \Omega_{Bb_j}$ . We can compare any protocol that requires  $k$  ebits with the one that requires  $m > k$  ebits by padding the former with the  $m - k$  ebits and acting on them by  $Z$ . This will result in the same number of terms in respective sets  $\mathcal{C}$ .

As a simple example, consider the action of  $\mathcal{Q}^{(1)}[\pi/4] = \Lambda_{Aa}(\pi/4) \otimes \Omega_{Bb}$ , which results in  $\mathcal{C}^{(1)}[\pi/4]$ :

$$\begin{aligned}
|00\rangle_{AB}|\Phi^+\rangle_{ab} &\longrightarrow_Q \frac{1}{2} [|00\rangle_{Aa}|00\rangle_{Bb} + |00\rangle_{Aa}|01\rangle_{Bb} + i|01\rangle_{Aa}|10\rangle_{Bb} - i|01\rangle_{Aa}|11\rangle_{Bb}], \\
|01\rangle_{AB}|\Phi^+\rangle_{ab} &\longrightarrow_Q \frac{1}{2} [|00\rangle_{Aa}|10\rangle_{Bb} + |00\rangle_{Aa}|11\rangle_{Bb} - i|01\rangle_{Aa}|00\rangle_{Bb} + i|01\rangle_{Aa}|01\rangle_{Bb}], \\
|1+\frac{\pi}{4}\rangle_{AB}|\Phi^+\rangle_{ab} &\longrightarrow_Q \frac{1}{2} [|10\rangle_{Aa}|00\rangle_{Bb} + |10\rangle_{Aa}|11\rangle_{Bb} + i|11\rangle_{Aa}|10\rangle_{Bb} + i|11\rangle_{Aa}|01\rangle_{Bb}], \\
|1-\frac{\pi}{4}\rangle_{AB}|\Phi^+\rangle_{ab} &\longrightarrow_Q \frac{1}{2} [|10\rangle_{Aa}|10\rangle_{Bb} + |10\rangle_{Aa}|01\rangle_{Bb} + i|11\rangle_{Aa}|00\rangle_{Bb} - i|11\rangle_{Aa}|11\rangle_{Bb}].
\end{aligned} \tag{13}$$

### B. Proof of optimality for $k = 1$ and $\alpha = \pi/4$

Assume that  $\mathcal{P}^{(1)}(\beta)$  with  $\beta \neq 0, \frac{\pi}{4}$  exists. It steers  $\mathcal{A}^{(1)}(\beta)$  to the set with elements

$$\tilde{\mathcal{C}}^{(1)}(\beta) = \sum_{j=1}^N a_{ij} |\omega_{ij}\rangle_{Aa} |\xi_{ij}\rangle_{Bb}, \tag{14}$$

where  $i = 1, \dots, 4$ ,  $a_{ij}$  are complex amplitudes, and  $\omega_{ij}, \xi_{ij}$  are binary strings which correspond to particular sets of results obtained by Alice and Bob, respectively. These strings obey the following relations:

(1)

$$\langle \omega_{1j} | \omega_{2j} \rangle = \langle \omega_{3j} | \omega_{4j} \rangle = 1, \tag{15}$$

which follow from the fact that initial states of Alice are indistinguishable within the two pairs. An instantaneous operation cannot increase their distinguishability.

(2) On the other hand, the orthogonality of Alice's states across the pairs 1 – 2 and 3 – 4 must be preserved; hence

$$\langle \omega_{1j} | \omega_{3j} \rangle = \langle \omega_{1j} | \omega_{4j} \rangle = \langle \omega_{2j} | \omega_{3j} \rangle = \langle \omega_{2j} | \omega_{4j} \rangle = 0. \tag{16}$$

(3) Condition (1) implies that on Bob's side the corresponding pairs need to remain orthogonal, i.e.,

$$\langle \xi_{1j} | \xi_{2j} \rangle = \langle \xi_{3j} | \xi_{4j} \rangle = 1, \tag{17}$$

(4) We assume that the final result of the nonlocal measurement is a 4-valued function of all output bits; that is, none of the bits  $A, B, a, b$  is redundant. This will be taken into account when analysing the general structure of possible unitaries.

In the following derivation, we will make use of the corollary of the nonsignaling principle.

**Corollary 7** *The reduced density matrix on Bob's side,  $\rho_{Bb}$ , is invariant under local actions on systems  $A$  and  $a$ , and vice versa.*

As a trivial illustration of the corollary 7 let us show that the required map cannot be accomplished for  $\alpha = \pi/4$  without ancillary systems. In this case there will be only four output binary strings (up to permutations), i.e.,

$$\begin{aligned}
|00\rangle_{AB} &\longrightarrow a_{11}|0\rangle_A|0\rangle_B, \\
|01\rangle_{AB} &\longrightarrow a_{21}|0\rangle_A|1\rangle_B, \\
|1+\rangle_{AB} &\longrightarrow a_{31}|1\rangle_A|0\rangle_B, \\
|1-\rangle_{AB} &\longrightarrow a_{41}|1\rangle_A|1\rangle_B,
\end{aligned} \tag{18}$$

where  $|a_{ij}| = 1$  due to normalization. Now, assume that the initial state is  $|00\rangle_{AB}$ . After the protocol is executed the output state remains unchanged. The matrix element on Bob's side  $\rho_B^{(22)} = 0$ . If, however, Alice flips her spin shortly before the protocol is executed, the state becomes  $|10\rangle = 1/\sqrt{2}(|1+\rangle + |1-\rangle)$  and the output state is  $|1\rangle_A|+\rangle_B$ , with  $\rho_B^{(22)} = 1/4$ . Hence, such an operation is forbidden [21].

### C. One ebit is necessary and sufficient for $\alpha = \pi/4$

Let us show that an ebit is necessary and sufficient. In this case, there are 16 output binary strings  $|\omega_{ij}\rangle_{Aa}|\xi_{ij}\rangle_{Bb}$ . Taking into account the conditions (1)-(4) the action of  $V_{Aa} \otimes U_{Bb}$  can be represented (up to trivial permutations) as

$$\begin{aligned} |00\rangle_{AB}|\Phi\rangle_{ab} &\longrightarrow a_{11}|00\rangle_{Aa}|00\rangle_{Bb} + a_{12}|00\rangle_{Aa}|01\rangle_{Bb} + a_{13}|01\rangle_{Aa}|10\rangle_{Bb} + a_{14}|01\rangle_{Aa}|11\rangle_{Bb} \\ |01\rangle_{AB}|\Phi\rangle_{ab} &\longrightarrow a_{21}|00\rangle_{Aa}|10\rangle_{Bb} + a_{22}|00\rangle_{Aa}|11\rangle_{Bb} + a_{23}|01\rangle_{Aa}|00\rangle_{Bb} + a_{24}|01\rangle_{Aa}|01\rangle_{Bb} \\ |1+\rangle_{AB}|\Phi\rangle_{ab} &\longrightarrow a_{31}|10\rangle_{Aa}|00\rangle_{Bb} + a_{32}|10\rangle_{Aa}|11\rangle_{Bb} + a_{33}|11\rangle_{Aa}|10\rangle_{Bb} + a_{34}|11\rangle_{Aa}|01\rangle_{Bb} \\ |1-\rangle_{AB}|\Phi\rangle_{ab} &\longrightarrow a_{41}|10\rangle_{Aa}|10\rangle_{Bb} + a_{42}|10\rangle_{Aa}|01\rangle_{Bb} + a_{43}|11\rangle_{Aa}|00\rangle_{Bb} + a_{44}|11\rangle_{Aa}|11\rangle_{Bb}. \end{aligned} \quad (19)$$

**Proposition 8**  $|a_{ij}| = 1/4 \forall i, j$ .

**Proof** The proof rests on a straightforward application of Corollary 7. Let  $a_{ij} = |a_{ij}|e^{i\theta_{ij}}$ . Assume that Alice and Bob are given  $|00\rangle_{AB}$ . After they implement the protocol the reduced density matrix on Bob's side is

$$\rho_{Bb} = \begin{pmatrix} |a_{11}|^2 & |a_{11}||a_{12}|e^{i(\theta_{11}-\theta_{12})} & 0 & 0 \\ |a_{11}||a_{12}|e^{-i(\theta_{11}-\theta_{12})} & |a_{12}|^2 & 0 & 0 \\ 0 & 0 & |a_{13}|^2 & |a_{14}||a_{13}|e^{i(\theta_{13}-\theta_{14})} \\ 0 & 0 & |a_{14}||a_{13}|e^{-i(\theta_{13}-\theta_{14})} & |a_{14}|^2 \end{pmatrix}. \quad (20)$$

However, if Alice flips her qubit shortly before the protocol is implemented, then the initial state is

$$|10\rangle_{AB} = |1\rangle_A(\cos\alpha|+\rangle_B + \sin\alpha|-\rangle_B). \quad (21)$$

For simplicity let us consider the element in the first row and the first column of the resulting density matrix  $\rho'_{Bb}$  on Bob's side:

$$\rho_{Bb}'^{(11)} = \cos^2\alpha|a_{31}|^2 + \sin^2\alpha|a_{43}|^2. \quad (22)$$

By Corollary (7) it has to match the corresponding element in (20), i.e.  $\rho_{Bb}^{(11)} = |a_{11}|^2$ . The same argument can be applied to three other starting states of the set, which gives rise to the following set of simultaneous equations:

$$\begin{cases} |a_{11}|^2 = \cos^2\alpha|a_{31}|^2 + \sin^2\alpha|a_{43}|^2, \\ |a_{23}|^2 = \sin^2\alpha|a_{31}|^2 + \cos^2\alpha|a_{43}|^2, \\ |a_{31}|^2 = \cos^2\alpha|a_{31}|^2 + \sin^2\alpha|a_{43}|^2, \\ |a_{43}|^2 = \sin^2\alpha|a_{31}|^2 + \cos^2\alpha|a_{43}|^2, \end{cases} \quad (23)$$

which yields

$$|a_{11}| = |a_{23}| = |a_{31}| = |a_{43}|. \quad (24)$$

Similarly, one can show equality of the remaining four quadruples. Also, notice that  $|a_{11}|^2 = |a_{12}|^2 = |a_{13}|^2 = |a_{14}|^2$  and the final result follows.  $\square$

Now, we are ready to prove our main result of this section:

**Theorem 9** For  $\alpha = \pi/4$  the necessary and sufficient resource is a maximally entangled state of  $a$  and  $b$ , where  $a, b$  are both qubits.

**Proof** Sufficient condition follows from existence of explicit protocols, which achieve that goal. For the necessary condition observe that each of the output state contains one ebit. Indeed, by Proposition 8 for  $i = 1$  the state can be written as

$$|0\rangle_A \otimes \frac{1}{\sqrt{2}} \left( |0\rangle_a \frac{e^{i\theta_{11}}|00\rangle_{Bb} + e^{i\theta_{12}}|01\rangle_{Bb}}{\sqrt{2}} + |1\rangle_a \frac{e^{i\theta_{13}}|10\rangle_{Bb} + e^{i\theta_{14}}|11\rangle_{Bb}}{\sqrt{2}} \right), \quad (25)$$

which clearly contains one ebit. As local unitary operations  $V_{Aa}, U_{Bb}$  cannot change entanglement of the state we conclude that  $|\Phi\rangle_{ab}$  must contain an ebit.  $\square$

#### D. Two qubits are necessary and sufficient for $\alpha = \pi/8$

Similarly to the proof of Theorem 9 two qubits are sufficient because there exists an explicit protocol  $\mathcal{Q}^{(2)}[\pi/8]$ . It remains to prove the necessary condition.

**Proposition 10** For  $\alpha \neq \pi/4$  one ebit is not sufficient.

**Proof** Assume again that Alice and Bob are given  $|00\rangle_{AB}$  and compare the reduced matrix on Bob's side with the case when Alice flips her spin. By Proposition 8

$$\rho_{Bb} = \frac{1}{4} \begin{pmatrix} 1 & e^{i(\theta_{11}-\theta_{12})} & 0 & 0 \\ e^{-i(\theta_{11}-\theta_{12})} & 1 & 0 & 0 \\ 0 & 0 & 1 & e^{i(\theta_{13}-\theta_{14})} \\ 0 & 0 & e^{-i(\theta_{13}-\theta_{14})} & 1 \end{pmatrix}. \quad (26)$$

Now, in case Alice flips her qubit consider an off-diagonal element

$$\rho'_{Bb}{}^{(12)} = \frac{1}{8} \sin 2\alpha (e^{i(\theta_{31}-\theta_{42})} + e^{i(\theta_{43}-\theta_{34})}) \quad (27)$$

of the new density matrix  $\rho'_{Bb}$ . By Corollary 7 it must be equal to  $\rho_{Bb}^{(12)}$ , i.e.

$$2e^{i(\theta_{11}-\theta_{12})} = \sin 2\alpha (e^{i(\theta_{31}-\theta_{42})} + e^{i(\theta_{43}-\theta_{34})}). \quad (28)$$

Equating real and imaginary components yields

$$\begin{cases} 2 \cos(\theta_{11} - \theta_{12}) = \sin 2\alpha (\cos(\theta_{31} - \theta_{42}) + \cos(\theta_{43} - \theta_{34})), \\ 2 \sin(\theta_{11} - \theta_{12}) = \sin 2\alpha (\sin(\theta_{31} - \theta_{42}) + \sin(\theta_{43} - \theta_{34})). \end{cases} \quad (29)$$

After adding the squared equations we obtain

$$2 = \sin^2 2\alpha (1 + \cos(\theta_{31} - \theta_{42} - \theta_{43} + \theta_{34})). \quad (30)$$

The maximum of the RHS is  $2\sin^2 2\alpha$ , therefore the equality can be only achieved for  $\alpha = \pi/4$ .  $\square$

Although we believe that the above proof technique could be used to prove optimality of  $\mathcal{Q}^{(k)}[\pi/2^{k+1}]$  for any value of  $k > 2$ , it becomes computationally intractable.

#### E. Proof for arbitrary $k$

**Theorem 11** For  $\alpha$  such that  $|\bar{\alpha}|_2 = k$  consider the protocol  $\mathcal{Q}^{(k)}[\alpha]$  (described in the earlier sections), which steers  $\mathcal{A}^{(k)}[\alpha]$  to  $\mathcal{C}^{(k)}[\alpha]$ . Then  $\mathcal{Q}^{(k)}[\alpha]$  is an optimal protocol.

**Proof** We prove by induction on  $k$ . We have shown earlier that the statement is true for  $k = 1, 2$ . Now, let us assume that  $\mathcal{Q}^{(k-1)}[2\alpha]$  is optimal and prove that  $\mathcal{Q}^{(k)}[\alpha]$  is optimal too. We prove by contradiction. Assume that there exists a more economical protocol,  $\mathcal{P}^{(k-1)}[\alpha]$ , which steers  $\mathcal{A}^{(k-1)}[\alpha]$  to  $\tilde{\mathcal{C}}^{(k-1)}[\alpha]$ . The most general form of  $\mathcal{P}$  is

$$\mathcal{P}^{(m)}[\alpha] = V_{A\tilde{a}_m} \otimes U_{B\tilde{b}_m} = (|0\rangle\langle 0|_A \otimes V_{0,\tilde{a}_m} + |1\rangle\langle 1|_A \otimes V_{1,\tilde{a}_m}) \otimes U_{B\tilde{b}_m}. \quad (31)$$

where  $\tilde{a}_m \equiv a_1 \dots a_m$ ,  $\tilde{b}_m \equiv b_1 \dots b_m$  [22]. The specific form of  $V_{A\tilde{a}_m}$  follows from the fact that it is determined by its action on the two orthogonal subspaces of the  $2^k$ -dimensional Hilbert space. In general,  $V_{0,\tilde{a}_m}$ ,  $V_{1,\tilde{a}_m}$  and  $U_{B\tilde{b}_m}$  depend on  $\alpha$ , while the known optimal protocol for  $\mathcal{A}^{(k-1)}[2\alpha]$  which utilises  $k-1$  ebits,  $\mathcal{Q}^{(k-1)}[2\alpha]$ , has  $V_{0,\tilde{a}_{k-1}} = \mathbb{I}_{\tilde{a}_{k-1}}$ ,  $V_{1,\tilde{a}_{k-1}} = \bigotimes_{i=1}^{k-1} R_{\alpha_i}^x(2^i\alpha)$ ,  $U_{B\tilde{b}_{k-1}} = \left[ \prod_{i=2}^{k-1} \left( (1 - P_1^{(i-1)}) \otimes \mathbb{I}_{Bb_i} + P_1^{(i-1)} \otimes \Omega_{Bb_i} \right) \right] \Omega_{Bb_1} \equiv \Omega_{B\tilde{b}_{k-1}}$ .

*Step 1: we reduce (31) to a special form (36) below.*

Let us analyse how  $\mathcal{P}^{(k-1)}[\alpha]$  acts on  $\mathcal{A}_0^{(k-1)}[\alpha]$  (recall that it is independent of  $\alpha$ ). It is helpful to rewrite (31) as

$$\mathcal{P}^{(k-1)}[\alpha] = \left( \mathbb{I}_A \otimes V_{0,\tilde{a}_{k-1}} \otimes U_{B\tilde{b}_{k-1}} \Omega_{B\tilde{b}_{k-1}}^\dagger \right) \left( |0\rangle\langle 0|_A \otimes \mathbb{I}_{\tilde{a}_{k-1}} \otimes \Omega_{B\tilde{b}_{k-1}} + |1\rangle\langle 1|_A \otimes V_{0,\tilde{a}_{k-1}}^{-1} V_{1,\tilde{a}_{k-1}} \otimes \Omega_{B\tilde{b}_{k-1}} \right). \quad (32)$$

Now

$$\left( |0\rangle\langle 0|_A \otimes \mathbb{I}_{\tilde{a}_{k-1}} \otimes \Omega_{B\tilde{b}_{k-1}} \right) \mathcal{A}_0^{(k-1)}[\alpha] = \mathcal{C}_0^{(k-1)}[2\alpha], \quad (33)$$

where  $\mathcal{C}_0^{(k-1)}[2\alpha]$  is the set achieved by the standard protocol  $\mathcal{Q}^{(k-1)}[2\alpha]$  implemented on  $\mathcal{A}_0^{(k-1)}[2\alpha]$ . This is true because  $\mathcal{A}_0^{(k-1)}[\alpha] = \mathcal{A}_0^{(k-1)}[2\alpha]$ . But we also know from (32) that

$$\mathcal{P}^{(k-1)}[\alpha] \mathcal{A}_0^{(k-1)}[\alpha] = \tilde{\mathcal{C}}_0^{(k-1)}[\alpha] \neq \mathcal{C}_0^{(k-1)}[2\alpha], \quad (34)$$

due to action of the first term in the brackets on the RHS, where

$$\tilde{\mathcal{C}}_0^{(k-1)}[\alpha] = \left( \mathbb{I}_A \otimes V_{0,\tilde{a}_{k-1}} \otimes U_{B\tilde{b}_{k-1}} \Omega_{B\tilde{b}_{k-1}}^{-1} \right) \mathcal{C}_0^{(k-1)}[2\alpha]. \quad (35)$$

Also,  $\tilde{\mathcal{C}}^{(k-1)}[\alpha]$  is still a set which satisfies standard conditions. Hence we observe that  $\mathbb{I}_A \otimes V_{0,\tilde{a}_{k-1}} \otimes U_{B\tilde{b}_{k-1}} \Omega_{B\tilde{b}_{k-1}}^{-1}$  only reshuffles the product basis states and, perhaps, introduces phases. We conclude that, there exists a protocol which acts on  $\mathcal{A}_0^{(k-1)}[\alpha]$  as the standard optimal protocol  $\mathcal{Q}^{(k-1)}[2\alpha]$ . In other words, the hypothetical protocol  $\mathcal{P}^{(k-1)}[\alpha]$  can be assumed to be acting on  $\mathcal{A}_0^{(k-1)}[\alpha]$  in the same way as  $\mathcal{Q}^{(k-1)}[2\alpha]$  does on  $\mathcal{A}_0^{(k-1)}[2\alpha]$ . In other words, we can assume that

$$\mathcal{P}^{(k-1)}[\alpha] = (|0\rangle\langle 0|_A \otimes \mathbb{I}_{\tilde{a}_{k-1}} + |1\rangle\langle 1|_A \otimes V_{\tilde{a}_{k-1}}) \otimes \Omega_{B\tilde{b}_{k-1}}, \quad (36)$$

where  $V_{\tilde{a}_{k-1}} \equiv V_{0,\tilde{a}_{k-1}}^{-1} V_{1,\tilde{a}_{k-1}}$ . It would have been identical to  $\mathcal{Q}^{(k-1)}[2\alpha]$  if  $V_{0,\tilde{a}_{k-1}}^{-1} V_{1,\tilde{a}_{k-1}} = \bigotimes_{i=1}^{k-1} R_{\alpha_i}^x(2^i\alpha)$ . Clearly, this cannot be the case.

*Step 2: We extend the protocol to make its action comparable to that of  $\mathcal{Q}^{(k)}[\alpha]$ .*

Let us return to the original protocol  $\mathcal{Q}^{(k)}[\alpha]$ , which gives  $\mathcal{C}^{(k)}[\alpha] = \mathcal{C}_0^{(k)} \oplus \mathcal{C}_1^{(k)}[\alpha]$ . Now consider extension of  $\mathcal{P}^{(k-1)}[\alpha]$ , with an additional ebit as follows:

$$\mathbb{I}_A \otimes \mathbb{I}_{\alpha_k} \otimes \left( (1 - P_1^{(k-1)}) \otimes \mathbb{I}_{Bb_k} + P_1^{(k-1)} \otimes \Omega_{Bb_k} \right) \mathcal{P}^{(k-1)}[\alpha] \equiv \tilde{\mathcal{P}}^{(k)}[\alpha]. \quad (37)$$

This extended protocol, albeit consuming a redundant ebit, still accomplishes the task, i.e.

$$\tilde{\mathcal{P}}^{(k)}[\alpha] \mathcal{A}^{(k)}[\alpha] = \tilde{\mathcal{C}}^{(k)}[\alpha] = \mathcal{C}_0^{(k)} \oplus \tilde{\mathcal{C}}_1^{(k)}[\alpha], \quad (38)$$

and can now be compared with the action of  $\mathcal{Q}^{(k)}[\alpha]$ ,

$$\mathcal{Q}^{(k)}[\alpha] \mathcal{A}^{(k)}[\alpha] = \mathcal{C}^{(k)}[\alpha] = \mathcal{C}_0^{(k)} \oplus \mathcal{C}_1^{(k)}[\alpha], \quad (39)$$

where  $\mathcal{C}^{(k)}[\alpha]$  and  $\tilde{\mathcal{C}}^{(k)}[\alpha]$  are locally equivalent by Proposition 5 hence there exists a permutation of the basis vectors which reduces  $\mathcal{Q}$  to  $\tilde{\mathcal{P}}$ .

*Step 3: We show that  $\mathcal{Q}$  is not reducible to  $\tilde{\mathcal{P}}$ .*  
We can represent  $\mathcal{Q}$  and  $\tilde{\mathcal{P}}$  as follows:

$$\mathcal{Q}^{(k)}[\alpha] = (|0\rangle\langle 0|_A \otimes \mathbb{I}_{\tilde{a}_k} + |1\rangle\langle 1|_A \otimes_{i=1}^k R_{a_i}^x [2^i \alpha]) \otimes \Omega_{Bb} \quad (40)$$

$$\tilde{\mathcal{P}}^{(k)}[\alpha] = (|0\rangle\langle 0|_A \otimes \mathbb{I}_{\tilde{a}_k} + |1\rangle\langle 1|_A \otimes V_{\tilde{a}_{k-1}} \otimes \mathbb{I}_{a_k}) \otimes \Omega_{Bb}. \quad (41)$$

Consider the ancillary subspace  $\mathcal{H}_{|1\rangle\langle 1|} \otimes \mathcal{H}_{\tilde{a}_k}$ . On this subspace  $\mathcal{Q}$  acts as a tensor product of  $k$  single-qubit rotations  $\otimes_{i=1}^k R(2^i \alpha)$ , and  $\tilde{\mathcal{P}}$  acts as  $V_{\tilde{a}_{k-1}} \otimes \mathbb{I}_{a_k}$ . There exists  $\alpha$  for which the spectrum of the total rotation consists of  $2^k$  different eigenvalues. The spectrum of the unitary  $V_{\tilde{a}_{k-1}} \otimes \mathbb{I}_{a_k}$  consists of at most  $2^{k-1}$  different eigenvalues. Thus, the latter protocol cannot be reduced to the former which contradicts our initial assumption.  $\square$

## V. A (POSSIBLE) GENERALIZATION OF THE STEERING SETS

Here we give two examples of sets of entangled states with distinguishability properties equivalent to the sets of product states studied earlier.

**Example 12** *The following set of two orthogonal entangled states is equivalent to  $\mathcal{A}[\alpha_1, \dots, \alpha_n]$  [23].*

$$\begin{aligned} |\psi\rangle &= c_0|0\rangle_A|0\rangle_B + c_1|1\rangle_A|+\alpha_1\rangle_B + \dots + c_n|n\rangle_A|+\alpha_n\rangle_B \\ |\phi\rangle &= d_0|0\rangle_A|1\rangle_B + d_1|1\rangle_A|-\alpha_1\rangle_B + \dots + d_n|n\rangle_A|-\alpha_n\rangle_B \end{aligned} \quad (42)$$

Clearly, a local projection measurement in the basis  $\{|i\rangle_A\}_{i=1}^n$  by Alice maps this set to  $\mathcal{A}[\alpha_1, \dots, \alpha_n]$ . Hence, its distinguishability is necessary and sufficient for distinguishability of (3).

**Example 13** *The set of four entangled states*

$$\mathcal{D}[\alpha] := \{|\Psi_i\rangle\}_{i=1}^4, \quad (43)$$

where

$$\begin{aligned} |\Psi_1\rangle &= \cos \alpha|00\rangle + \sin \alpha|11\rangle, \\ |\Psi_2\rangle &= \sin \alpha|00\rangle - \cos \alpha|11\rangle, \\ |\Psi_3\rangle &= \cos \alpha|01\rangle + \sin \alpha|10\rangle, \\ |\Psi_4\rangle &= \sin \alpha|01\rangle - \cos \alpha|10\rangle. \end{aligned}$$

Let us show that  $\mathcal{D} \rightarrow \mathcal{B}$  requires the same amount of entanglement resources as  $\mathcal{A} \rightarrow \mathcal{B}$ . First stage is to steer  $\mathcal{D}$  to  $\mathcal{A}[2\alpha]$ . The protocol which achieves such a map is a modification of the procedure invented in [8]. Alice and Bob start with implementing instantaneous nonlocal C-NOT with  $B$  and  $A$  being a control and a target respectively (as a part of this step they obtain corresponding results of local measurements –  $x_a = \pm 1$  and  $z_b = \pm 1$ ). This is followed by Bob rotating his qubit  $B$  by  $\alpha$  about the  $y$  axis. The resulting mapping is summarised in Table II.

Here the states of  $A$  and  $b$  are in direct product with the set

$$\begin{aligned} &|-\rangle_a|0\rangle_B \\ &|-\rangle_a|1\rangle_B \\ &|+\rangle_a|+2\alpha\rangle_B \\ &|+\rangle_a|-2\alpha\rangle_B, \end{aligned} \quad (44)$$

which is equivalent to  $\mathcal{A}[2\alpha]$  and requires  $\mathcal{I}(\alpha) - 1$  ebits to steer it to  $\mathcal{B}$  on the second stage, bringing the total cost to  $\mathcal{I}(\alpha)$ .

$x_a \setminus z_b$	$z_b = +1$	$z_b = -1$
$x_a = +1$	$ 0\rangle_A   +2\alpha\rangle_B$	$ 1\rangle_A   +2\alpha\rangle_B$
	$ 0\rangle_A   -2\alpha\rangle_B$	$ 1\rangle_A   -2\alpha\rangle_B$
	$ 1\rangle_A   +2\alpha\rangle_B$	$ 0\rangle_A   +2\alpha\rangle_B$
	$ 1\rangle_A   -2\alpha\rangle_B$	$ 0\rangle_A   -2\alpha\rangle_B$
$x_a = -1$	$ 0\rangle_A   0\rangle_B$	$ 1\rangle_A   0\rangle_B$
	$ 0\rangle_A   1\rangle_B$	$ 1\rangle_A   1\rangle_B$
	$ 1\rangle_A   0\rangle_B$	$ 0\rangle_A   0\rangle_B$
	$ 1\rangle_A   1\rangle_B$	$ 0\rangle_A   1\rangle_B$

TABLE II. The resulting map after first stage.

## VI. DISCUSSION

We have considered the question of distinguishability of sets of mutually orthogonal states under LOSE. Our main objects of interest were specific families of sets of bipartite product states (3) characterised by  $n$  real parameters  $\{\alpha_i\}_{i=1}^n$ , with (1) being a simplest case of a single parameter  $\alpha$ . The task had been represented as a steering game and we have provided an algorithm which achieves the final state of the game, i.e. a set composed of tensor-product local bases states, using optimal amount of entanglement. For the steering game we defined a figure of merit  $\mathcal{I}$  which quantifies the exact amount of entanglement required to steer the initial set to the given one and hence can be regarded as a measure of information nonlocality in the set. We have shown that it is a maximal length of binary expansions of  $\alpha_i$ .

We have provided examples of sets of entangled states, which can be steered using the same number of ebits as the sets of product states with the same values of corresponding parameters. In order to show that the type of nonlocality that we study is not associated with entanglement of a set, but depends solely on the values of parameters characterising them, let us consider particular mixtures of the states.

**Example 14** *Alice and Bob are given a state drawn from  $\mathcal{D}[\alpha]$  (43) with a probability distribution such that the resulting mixed state is*

$$\rho_W = \frac{1-F}{3} (\Psi_1 + \Psi_2 + \Psi_3) + F\Psi_4 = \frac{1-F}{3} \mathbb{I} + \frac{4F-1}{3} \Psi_4, \quad (45)$$

where  $0 \leq F \leq 1$ . It is known that for  $\alpha = \pi/4$  the mixed state  $\rho_W$  is not entangled when  $0 \leq F \leq \frac{1}{2}$  and entangled when  $\frac{1}{2} < F \leq 1$  [24] with  $\Psi = |\Psi\rangle\langle\Psi|$  and  $\Psi_1 = \Phi^+$ ,  $\Psi_2 = \Phi^-$ ,  $\Psi_3 = \Psi^+$ ,  $\Psi_4 = \Psi^-$  – the four Bell states. However, as was shown earlier, the steering properties of  $\rho_W$  depends only on the value of  $\alpha$ , but not on  $F$ .

Let us discuss our results in relation to past works. The question of designing more efficient protocols had been addressed in [10]. The distinguishing feature of their proposal is that the protocol halts with increasing probability as the number of rounds grows larger. Thereby, it does not consume the remaining entanglement resources. However, there is non-zero probability that the protocol will consume all available entanglement. Thus, the protocol had been designed to minimise the expected number of ebits consumed. Our goal is different: we aim at minimizing the amount of entanglement that has to be available beforehand.

Two correlation measures related to our scenario — *quantum deficit* and *classical zero-way deficit* — have been introduced in [14].

We note that they are unsuitable measures of the nonlocality of  $\mathcal{A}[\alpha]$ : quantum deficit may be made large or small depending on  $\alpha$ , and independent of the binary expansion of the latter. Classical zero-way deficit depends on the probability distribution of the states drawn from  $\mathcal{A}[\alpha]$  in (1), whereas the amount of nonlocality of the given set as quantified by a steering task is not — it depends only on  $\alpha$ .

Our results are closely related to the definition of quantumness of sets of quantum states proposed in [16]. The sets considered in our work fall into the category of nonclassical (quantum) sets, and the presence of de-localised information can be interpreted as a signature of quantumness. However, the measure of information nonlocality proposed by us is profoundly different from the *relative entropy of quantumness* proposed in [16]. The latter, being a distance-measure, depends on the probability distribution of the members of the set. It is important to note that information nonlocality captures the information-theoretic aspect of sets in terms of the length of binary representation of corresponding parameter(s) and as such cannot serve as an alternative measure of quantumness in a distance sense — two very close sets can have very different values of information nonlocality.

Our work raises a number of questions. First, what are the most general orthogonal sets which one can steer to  $\mathcal{B}$  and how does our measure  $\mathcal{I}$  generalise to these sets? Second, what are the applications of the steering algorithm to other tasks of quantum information processing? Third, whether one can relate the number of ebits needed to bi-localise information to the cost of resetting delocalised information in the sense of Landauer’s principle. Finally, what is the meaningful generalization of the steering procedure to the multipartite states for three parties and more?

### ACKNOWLEDGMENTS

The authors would like to thank Sidney Sussex College, Cambridge for financial support.

### Appendix: the instantaneous rotations protocol – the stator formalism

The instantaneous steering protocol of Sec. III can be presented in the state-operator (stator) formalism which was first introduced in [18]. Steps 1 and 2 result in the preparation of the following stator:

$$|0\rangle_{b_t} \otimes S_+^{(t)} + |1\rangle_{b_t} \otimes S_-^{(t)}, \quad (\text{A.1})$$

where  $S_{\pm}^{(t)} = |0\rangle_{a_t} \otimes \mathbb{I}_B \pm |1\rangle_{a_t} \otimes \sigma_y^B$  and we omit normalization. Both  $S_{\pm}^{(t)}$  satisfy eigen-operator equation  $\sigma_x^a S_{\pm}^{(t)} = \pm \sigma_y^B S_{\pm}^{(t)}$ ; thus

$$\sigma_x^a \left( |0\rangle_{b_t} \otimes S_+^{(t)} + |1\rangle_{b_t} \otimes S_-^{(t)} \right) = \sigma_y^B \left( |0\rangle_{b_t} \otimes S_+^{(t)} - |1\rangle_{b_t} \otimes S_-^{(t)} \right), \quad (\text{A.2})$$

which in turn implies that

$$R_a(\theta) \left( |0\rangle_{b_t} \otimes S_+^{(t)} + |1\rangle_{b_t} \otimes S_-^{(t)} \right) = |0\rangle_{b_t} \otimes R_B(\theta) S_+^{(t)} + |1\rangle_{b_t} \otimes R_B(-\theta) S_-^{(t)}. \quad (\text{A.3})$$

This ability to “propagate” transformations from  $a$  to  $B$  lies in the heart of stator formalism.

The sequence of Bob’s actions on  $B$  and  $k$  ebits  $|\Phi^+\rangle^{\otimes k}$  [the unitary on the RHS of Eq. (11)] results in what we call the “superstator”:

$$S_{sup} = \sum_{t=1}^k \frac{1}{2^{t/2}} |\Phi^+\rangle_{a_k b_k} \dots |\Phi^+\rangle_{a_{t+1} b_{t+1}} |0\rangle_{b_t} |1\rangle_{b_{t-1}} \dots |1\rangle_{b_1} S_+^{(t)} S_-^{(t-1)} \dots S_-^{(1)} + \Theta, \quad (\text{A.4})$$

where  $\Theta = 2^{-k/2} |1\rangle_{b_k} |1\rangle_{b_{k-1}} \dots |1\rangle_{b_1} S_-^{(k)} S_-^{(k-1)} \dots S_-^{(1)}$ . For each term  $S_{sup}$  the actions on Alice’s side generate a desired rotation of  $B$ :

$$\left[ \bigotimes_{i=1}^t R_{a_i}^x(2^{i-1}\alpha) \right] S_+^{(t)} S_-^{(t-1)} \dots S_-^{(1)} = R_B^y(\alpha) S_+^{(t)} S_-^{(t-1)} \dots S_-^{(1)}, \quad (\text{A.5})$$

$$\left[ \bigotimes_{i=1}^k R_{a_i}^x(2^{i-1}\alpha) \right] S_-^{(k)} S_-^{(k-1)} \dots S_-^{(1)} = R_B^y(-\pi/2 + \alpha) S_-^{(k)} S_-^{(k-1)} \dots S_-^{(1)}. \quad (\text{A.6})$$

- 
- [1] Here *joint* does not necessary mean *entangled*.  
[2] A. Chefles, Physical Review A **69**, 050307 (2004).  
[3] S. De Rinaldis, Physical Review A **70**, 022309 (2004).  
[4] J. Walgate, A. J. Short, L. Hardy, and V. Vedral, Physical Review Letters **85**, 4972 (2000).  
[5] A. Chefles, Physics Letters A **239**, 339 (1998).  
[6] C. H. Bennett, D. P. DiVincenzo, C. A. Fuchs, T. Mor, E. Rains, P. W. Shor, J. A. Smolin, and W. K. Wootters, Phys. Rev. A **59**, 1070 (1999).  
[7] Y. Aharonov and D. Z. Albert, Physical Review D **21**, 3316 (1980).  
[8] B. Groisman and B. Reznik, Phys. Rev. A **66**, 022110 (2002).  
[9] L. Vaidman, Phys. Rev. Lett. **90**, 010402 (2003).

- [10] S. R. Clark, A. J. Connor, D. Jaksch, and S. Popescu, *New Journal of Physics* **12**, 083034 (2010).
- [11] S. Beigi and R. König, *New Journal of Physics* **13**, 093036 (2011).
- [12] H. Buhrman, N. Chandran, S. Fehr, R. Gelles, V. Goyal, R. Ostrovsky, and C. Schaffner, *SIAM Journal on Computing* **43**, 150 (2014).
- [13] In [10] authors derive finite expectation value for entanglement consumption, but infinite amount must be available.
- [14] M. Horodecki, P. Horodecki, R. Horodecki, J. Oppenheim, A. Sen(De), U. Sen, and B. Synak-Radtke, *Physical Review A* **71**, 062307 (2005).
- [15] M. Dall’Arno, G. M. D’Ariano, and M. F. Sacchi, *Physical Review A* **83**, 062304 (2011).
- [16] B. Groisman, D. Kenigsberg, and T. Mor, <http://arxiv.org/abs/quant-ph/0703103> (2007).
- [17] H. Ollivier and W. H. Zurek, *Physical Review Letters* **88**, 017901 (2001).
- [18] B. Reznik, Y. Aharonov, and B. Groisman, *Physical Review A* **65**, 032312 (2002).
- [19] L. Vaidman and Z. Mitrani, *Physical Review Letters* **92**, 217902 (2004).
- [20] B. Groisman and L. Vaidman, *Journal of Physics A: Mathematical and General* **34**, 6881 (2001).
- [21] We could arrive to the same conclusion based on the fact that the corresponding transformation on Bob’s side cannot be performed by unitary on  $B$ .
- [22] In principle,  $V_{A\tilde{a}_m}$  may be decomposed into a product of two unitaries: some general unitary  $\tilde{V}_{A\tilde{a}_m}$  and a permutation unitary  $T_{A\tilde{a}_m}$ . The latter always exist for any  $\tilde{V}_{A\tilde{a}_m}$  as it preserves orthogonality of bit strings in blocks  $C_0$  and  $C_1$ . Thus, every unitary acting on  $A\tilde{a}_m$  can be cast in the form (31).
- [23] It is known that any two pure orthogonal states, entangled or product, can be perfectly distinguished by LOCC [4].
- [24] R. F. Werner, *Physical Review A* **40**, 4277 (1989).