Entanglement distillation protocols and number theory

H. Bombin and M. A. Martin-Delgado
Departamento de Física Teórica I, Universidad Complutense, 28040 Madrid, Spain
(Received 1 March 2005; published 13 September 2005)

We show that the analysis of entanglement distillation protocols for qudits of arbitrary dimension D benefits from applying basic concepts from number theory, since the set $\mathbf{Z}_D^n$ associated with Bell diagonal states is a module rather than a vector space. We find that a partition of $\mathbf{Z}_D^n$ into divisor classes characterizes the invariant properties of mixed Bell diagonal states under local permutations. We construct a very general class of recursion protocols by means of unitary operations implementing these local permutations. We study these distillation protocols depending on whether we use twirling operations in the intermediate steps or not, and we study them both analytically and numerically with Monte Carlo methods. In the absence of twirling operations, we construct extensions of the quantum privacy algorithms valid for secure communications with qudits of any dimension D. When D is a prime number, we show that distillation protocols are optimal both qualitatively and quantitatively.

DOI: 10.1103/PhysRevA.72.032313 PACS number(s): 03.67.Lx

I. INTRODUCTION

Quantum information theory (QIT) revolves around the concept of entanglement [1–4]. It is the product of combining the superposition principle of quantum mechanics with multipartite systems—described by the tensor product of Hilbert spaces. Entanglement is central to transmitting information in a quantum communication protocol or processing information in a quantum computation. There are two basic open problems in the study of entanglement: separability and distillability. Separability is concerned with two questions, namely whether a quantum state is factorizable, and if not, how much entanglement it contains. These questions are of great importance even in practice since entanglement amounts to interaction between two or more parties, and thus it demands more resources to establish an entangled state than a factorized one.

Likewise, distillability [5–11] is concerned with two questions: whether a quantum state is distillable, and if it is, how to devise an explicit protocol to extract entanglement out of the initial low entangled state. The main focus of our paper is on the construction of distillation protocols, rather than a direct analysis of the distillability issue.

The study of distillation protocols for qudits is justified since it is known that for mixed states of dimension higher than $2\times 2, 2\times 3$, neither a complete criterion for separability nor for distillability is known [9].

Separability and distillability are interconnected. Entanglement of a mixed state is a necessary condition for being distillable. However, it was quite a surprise to find that there exist states that, though they are entangled, cannot be distilled. They are the bound entangled states that are characterized by being positive partial transposition (PPT) entangled states [12–14]. This situation soon raised the question of whether non-PPT states were all distillable. Although there is not a conclusive answer, there is strong evidence that this is not the case since Werner states which are finite-n undistillable have been found [15,16].

In this paper, we study entanglement distillation protocols for qudits using the recursion method [5,6]. The mixed states to be distilled are diagonal Bell states of qudits, i.e., maximally entangled states, but they do not need to be tensor product of pairs of Bell states. Moreover, we can also distill nondiagonal states.

We make significant progress in the understanding of these protocols and find new efficient variants of them by using number theory. This number theory enters in the properties of the module $\mathbf{Z}_D^n$ that appears in the labeling of the Bell diagonal states of qudits. Local permutations acting on these states by means of unitary operations serve to construct generalized distillation protocols [10]. As a byproduct, we also introduce heterotropic states (38) as the invariant states under the group of local permutations.

As a result of this study, we find that qudit states with dimension D a prime number are qualitatively and quantitatively the best choices for quantum distillation protocols based on the recursion method. Qualitatively, these states are best since we show that for D not a prime number there appear disturbing attractor points in the space of fidelity parameters that deviate the distillation process from the desired fixed point that represents the maximally entangled state. This phenomenon is absent when D is a prime number and $\mathbf{Z}_D^n$ becomes a vector space [11]. Quantitatively, these states are best since we propose distillation protocols that when D is a prime number, they distill all states with fidelity bigger than $1/D$ without resorting to twirling operations.

We hereby summarize briefly some of our main results.

(i) We prove that the group of local permutations for qudits of arbitrary dimension D is the semidirect product of the group of translations and simplectic transformations. This structure plays an important role in the distillation protocols for qudits.

(ii) We simplify the problem of finding the best distillation protocol to that of finding the best set of coefficients of a certain polynomial constrained to the existence of a suitable vector space.

(iii) We introduce the concept of *joint performance parameter* η (65) that allows us the comparison of distillation protocols with different values of fidelity, probability of success, and number of Bell pairs used altogether. It is a figure

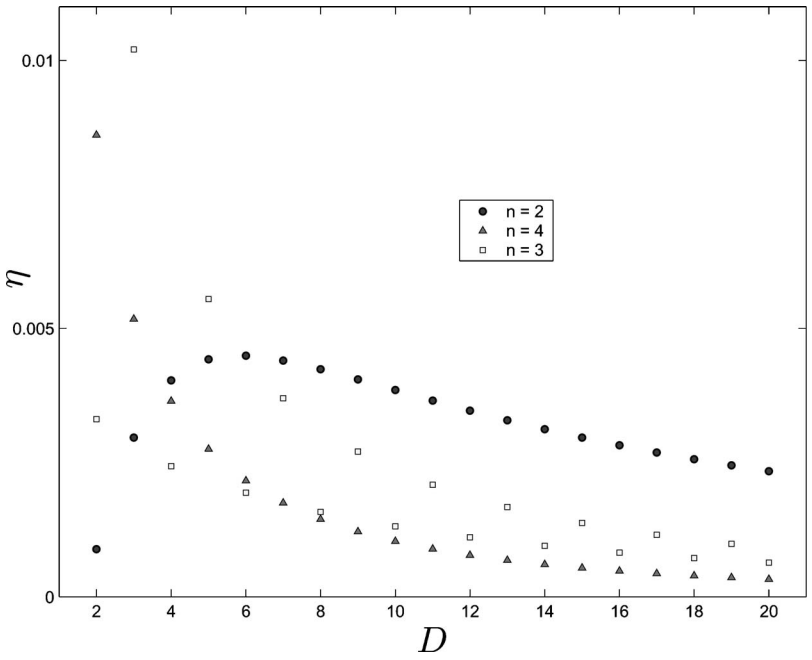

FIG. 1. Values of the coefficient η for the considered twirled-assisted protocols with $n=2,3,4$. Initial fidelity is close to $1/D$.

of merit for low fidelity states above the distillation threshold where the recursion method is specially suited for distillation, prior to switching a hashing or breeding method.

(iv) We analyze several distillation protocols assisted with twirling operations as the dimension D of qudits vary. We find that the best performance according to η is not achieved for qubits ($D=2$), but for qutrits ($D=3$) and $n=3$ input pairs of Bell states as shown in Fig. 1 and Fig. 2. We also find that it is not possible to improve η by indefinitely increasing the number of input pairs n.

(v) We propose a distillation protocol without resorting to twirling operations for $n=4$ input Bell states and $m=2$ output Bell states that is iterative and its yield is greatly improved: about four orders of magnitude with respect to the protocols based on twirling, even for states quite near the fixed point. This is shown in Fig. 5.

(vi) We propose and study an extension of the quantum privacy amplification protocols [7] that work for arbitrary dimension D.

(vii) We find indications of the existence of nondistillable NPPT states by studying the distillation capacities of protocols for several values of D, as shown in Fig. 7. By all means, the fact that some states are not distillable with a set of protocols does not necessarily mean that they are nondistillable.

This paper is organized as follows. In order to facilitate the reading and exposition of our results, we present the proofs of our theorems and technicalities of the numerical methods in independent Appendixes. Section II deals with the basic properties of diagonal Bell states for qudits and introduces a partition of the module $\mathbf{Z}_D^n$. Section III treats the group of local permutations acting on qudits in diagonal Bell

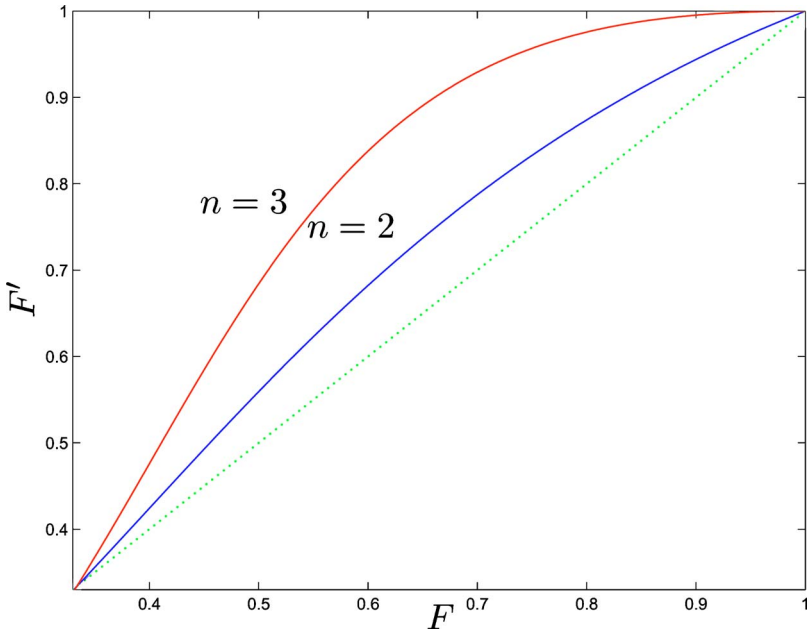

FIG. 2. (Color online) Evolution of the fidelity for the distillation protocols assisted by twirling when $n=2$, 3 for qudits with $D=3$.

states. Section IV describes the group of local permutations and the twirling operations associated with it. We characterize states that are invariant under these operations as heterotropic states. In Section V, we present all our distillation protocols based on local permutations of qudits in diagonal Bell states, both with and without resorting to twirling operations. To this end, we make extensive use of the theoretical results found in previous sections and devise numerical methods to analyze efficiently the properties of the proposed distillation protocols as different parameters such as D, F, n,m, etc. vary. Section VI is devoted to conclusions and future prospects.

II. BASIC PROPERTIES OF DIAGONAL BELL STATES FOR QUDITS

A. Bell states basic notation

The quantum systems we are going to consider are *qudits*, which are described by a Hilbert space of dimension $D \geqslant 2$, and finite. The elements of a given orthogonal basis can be denoted $|x\rangle$ with $x=0,\ldots,D-1$. This set of numbers is naturally identified with the elements of the set modulus,

$$\mathbf{Z}_D := \mathbf{Z}/D\mathbf{Z}, \tag{1}$$

and we shall informally use them as if they belonged to $\mathbf{Z}_D$. In general, whenever an element of $\mathbf{Z}_D$ appears in an expression, any integer in that expression must be understood to be mapped to $\mathbf{Z}_D$.

We consider two separate parties, Alice and Bob, each of them owning one of these systems. The entire Hilbert space is then $\mathcal{H}=\mathcal{H}_A \otimes \mathcal{H}_B$. A mixed state of the whole system is called *separable* when it can be expressed as a convex sum of *product states* [17],

$$\rho = \sum_i p_i |e_i\rangle\langle e_i| \otimes |f_i\rangle\langle f_i|, \quad |e_i\rangle \in \mathcal{H}_A,\ |f_i\rangle \in \mathcal{H}_B; \tag{2}$$

a state which is not separable is said to be *entangled*.

Elements of the *computational basis* of a pair of qudits shared by Alice and Bob are denoted as

$$|ij\rangle := |i\rangle \otimes |j\rangle, \quad i,j \in \mathbf{Z}_D. \tag{3}$$

To shorthen the notation, it is convenient to introduce the symbols

$$\underset{k}{\boldsymbol{\mathcal{S}}} := \frac{1}{\sqrt{D}} \sum_{k \in \mathbf{Z}_D}, \quad \delta(k) := \sqrt{D}\delta_{k,0}, \quad \varphi(k) := \mathrm{e}^{(2\pi \mathrm{i}/D)k}, \tag{4}$$

chosen so that $\mathcal{S}_k \varphi(ik) = \delta(i)$ and $\mathcal{S}_k \delta(i-k) f(k) = f(i)(i \in \mathbf{Z}_D), \forall f$. *Bell states* are defined as [18–21]

$$|ij\rangle_{\mathcal{B}} := \underset{k}{\boldsymbol{\mathcal{S}}} \varphi(ki) |k\ k-j\rangle, \quad i,j \in \mathbf{Z}_D. \tag{5}$$

Bell states are an example of maximally entangled states. In fact, any maximally entangled state can be identified with the $|00\rangle_{\mathcal{B}}$ state by suitably choosing the computational basis of each of the qudits, due to the Schmidt decomposition.

The *fidelity* of a mixed state ρ is defined as

$$F := \max_{\Psi} \langle \Psi | \rho | \Psi \rangle, \tag{6}$$

where the maximum is taken over the set of maximally entangled states. The aim of distillation protocols is to get maximally entangled pairs (fidelity 1) by means of local operations and classical communication (LOCC), starting with entangled states of fidelity lower than 1. Because of the previous comment, we will always suppose, without loss of generality, that the initial fidelity of the states to be distilled is equal to ${}_{\mathcal{B}}\langle 00|\rho|00\rangle_{\mathcal{B}}$, and the aim of our protocols will be to obtain distilled states as close as possible to this Bell state.

Of special interest are the mixtures of perfectly entangled states and white noise, known as *isotropic states*,

$$\rho_{\mathrm{iso}} := F|0\ 0\rangle_{\mathcal{B}}\langle 0\ 0| + \frac{1-F}{D^2-1}(1 - |0\ 0\rangle_{\mathcal{B}}\langle 0\ 0|), \tag{7}$$

where F is the fidelity of the state. These states are known to be entangled and distillable iff $F > 1/D$ [9].

The main interest of these states comes not only from their physical meaning, but also from the possibility of transforming any state in an isotropic one through a *twirling* operation. In general, the twirling consists in a random application of the elements of a certain group of unitary operations, say $\mathcal{U}$, to each of the systems in an ensemble. Namely, its action is

$$T_{\mathcal{U}}(\rho) := \int_{\mathcal{U}} dU U \rho U^{\dagger}. \tag{8}$$

The result of such an operator must be a sum over the states invariant under the action of the group. In the case of isotropic states, a suitable election is the set of transformations of the form $U \otimes U^*$.

When managing multiple shared pairs, vector notation is necessary; $\mathbf{k} \in \mathbf{Z}_D^n$ stands for $\mathbf{k}=(k_1,\ldots,k_n), k_i \in \mathbf{Z}_D$. A scalar product will be employed with its usual meaning. The generalization of the previous expressions is straightforward,

$$\underset{\mathbf{k}}{\boldsymbol{\mathcal{S}}} := \underset{k_1}{\boldsymbol{\mathcal{S}}} \cdots \underset{k_n}{\boldsymbol{\mathcal{S}}}, \quad \delta(\mathbf{k}) := \delta(k_1)\cdots\delta(k_n). \tag{9}$$

Again, $\mathcal{S}_{\mathbf{k}} \varphi(\mathbf{i}\cdot\mathbf{k}) = \delta(\mathbf{i})$ for any $\mathbf{i} \in \mathbf{Z}_D^n$. The computational basis and the Bell basis are

$$|\mathbf{i}\ \mathbf{j}\rangle := \bigotimes_{k=1}^{n} |i_k\ j_k\rangle,$$

$$|\mathbf{i}\ \mathbf{j}\rangle_{\mathcal{B}} := \bigotimes_{k=1}^{n} |i_k\ j_k\rangle_{\mathcal{B}} = \underset{\mathbf{k}}{\boldsymbol{\mathcal{S}}} \varphi(\mathbf{i}\cdot\mathbf{k}) |\mathbf{k}\ \mathbf{k}-\mathbf{j}\rangle, \tag{10}$$

with $\mathbf{i},\mathbf{j} \in \mathbf{Z}_D^n$. In order to simplify the notation, sometimes we will work with vectors over $\mathbf{Z}_D^{2n}$ and write states as $|\mathbf{x}\rangle$ in the place of $|\mathbf{i}j\rangle$, with

$$\mathbf{x} := (i_1,\ldots,i_n,j_1,\ldots,j_n). \tag{11}$$

B. A Partition of $\mathbf{Z}_D^n$ with divisor classes

In general, $\mathbf{Z}_D$ is not a field and thus $\mathbf{Z}_D^n$ is not a vector space but a module. We can still make use of some properties

associated with vector spaces and so we will abuse a bit of the term vector. For a detailed exposition, see Appendix A, but it is enough to know the following. The usual definition of linear independence makes sense, as one can demonstrate that any linearly independent set of vectors can be extended to a complete basis and also that a square matrix composed by such a set is invertible. A subspace is defined to be the set of linear combinations of a linearly independent set, and its dimension is the cardinality of these generators. Orthogonality poses no problem, since the set orthogonal to a subspace is a subspace, and it has the expected dimension.

Working with this pseudovector space $\mathbf{Z}_D^n$ requires care. Some vectors can be taken to the null vector by multiplying them by a nonzero number. For example, for $D=4$ we have $2\times(0,2)=(0,0)$. In order to classify this anomalous vectors, consider the set of divisors of D,

$$\mathrm{div}(D) \coloneqq \{d \in \mathbf{N}: d|D\}. \tag{12}$$

This set inherits the ordering of $\mathbf{N}$, and we shall use this property to introduce a suitable gcd function in $\mathbf{Z}_D$:

Definition II.1. For every $S \subset \mathbf{Z}_D$ we define the greatest common divisor of S, or $\gcd(S)$, to be the greatest $d \in \mathrm{div}(D)$ such that $(D/d)s=0$, $\forall s \in S$.

The nomenclature was chosen because for any $d \in \mathrm{div}(D)$ and $x \in \mathbf{Z}$ we have

$$d|x \Leftrightarrow D|\frac{D}{d}x \Leftrightarrow \frac{D}{d}x = 0 \pmod{D}, \tag{13}$$

and then for any set of integers X

$$\gcd(\bar{X}) = \max\{d \in \mathrm{div}(D): d|x \ \forall\, x \in X\}, \tag{14}$$

where $\bar{X}$ is the corresponding set in $\mathbf{Z}_D$.

Vectors over $\mathbf{Z}_D$ are n-tuples of elements in $\mathbf{Z}_D$, and so we extend the gcd function to act over $\mathbf{Z}_D^n$ in the natural way, that is, if $\mathbf{v}=(v_1,\ldots,v_n)$, $\gcd(\mathbf{v}) \coloneqq \gcd(\{v_1,\ldots,v_n\})$. Now we can consider an equivalence relation in $\mathbf{Z}_D^n$ governed by the equality under the gcd function. The corresponding partition consists in the sets

$$C_d(D,n) \coloneqq \{\mathbf{v} \in \mathbf{Z}_D^n : \gcd(\mathbf{v}) = d\},\ d \in \mathrm{div}(D). \tag{15}$$

The most important of these sets is $C_1(D,n)$, since it contains those vectors $\mathbf{v}$ for which $\{\mathbf{v}\}$ is linearly independent. Later we will need its cardinality when considering properties of local unitary operators acting on diagonal Bell states. Thus, it is useful to define

$$\phi_n(D) \coloneqq \begin{cases} 1 & \text{if } D = 1 \\ \mathcal{N}C_1(D,n) & \text{if } D > 1. \end{cases} \tag{16}$$

For the particular case of $n=1$, $\phi_1(x)$ corresponds to Euler's totient ϕ function [22]. Euler's ϕ function appears naturally in number theory since it gives for a natural number n, the cardinality of the set $\{m=1,\ldots,n-1:\gcd(m,n)=1\}$. That is, $\phi_1(n)$ is the total number of coprime integers (or totatives) below or equal to n. For example, there are eight totatives of 24, namely, $\{1,5,7,11,13,17,19,23\}$, thus $\phi_1(24)=8$. For $n \neq 1$, we have therefore introduced a generalization of Euler's totient function for elements in $\mathbf{Z}_D^n$. The following lemma tells us how to compute the cardinalities of the sets $C_d(D,n)$, which shall naturally arise in our analysis of distillation protocols.

TABLE I. Values of the generalized Euler's totient function $\phi_n(D)$ for several qudit dimensions D.

D	2	3	4	5	6
$\phi_1(D)$	1	2	2	4	2
$\phi_2(D)$	3	8	12	24	24
$\phi_3(D)$	7	26	56	124	182

Lemma II.2. For every $n \in \mathbf{N}$, $D \in \mathbf{N}-\{1\}$ and $d \in \mathrm{div}(D)$,

$$\text{(i)}\ \phi_n(D) = D^n \prod_{\substack{p|D \\ p \text{ prime}}} \frac{p^n-1}{p^n}, \tag{17}$$

$$\text{(ii)}\ \mathcal{N}C_d(D,n) = \phi_n\left(\frac{D}{d}\right), \tag{18}$$

$$\text{(iii)}\ \sum_{d' \in \mathrm{div}(D)} \phi_n(d') = D^n. \tag{19}$$

The proof of this lemma can be found in Appendix B. As an illustration, we list several values of $\phi_n(D)$ in Table I.

III. THE GROUP OF LOCAL PERMUTATIONS

The main constraint Alice and Bob have to face when they intend to distill qudits is that they can perform only local operations. If we consider only unitary operations, we are led to the group $\mathcal{U}_{\mathrm{loc}}$ of local unitary operations. Its elements are all of the form

$$U = U_{\mathrm{A}} \otimes U_{\mathrm{B}}. \tag{20}$$

In this section, we shall study the subgroup, $\mathcal{U}_{\mathrm{B\ loc}}$ defined as the group of local unitary operations which are closed over the space of Bell diagonal states, that is, states of the form

$$\rho^{(n)} = \sum_{\mathbf{x} \in \mathbf{Z}_D^{2n}} p_{\mathbf{x}}^{(n)} |\mathbf{x}\rangle_{\mathcal{B}}\langle\mathbf{x}|, \tag{21}$$

where the label (n) is a reminder that we are considering states of n pairs of qudits. The aim is to use the acquired knowledge to devise distillation protocols specially suited for these states.

More specifically, we analyze the group $\mathcal{U}_{\mathrm{B\ loc}}(D,n)$ of local unitary operators over the space spanned by the tensor product of n pairs of qudits of dimension D for which the image of a Bell diagonal state is another Bell diagonal state. The first we notice is that the result of applying such an operator over a pure Bell state is another Bell state (it cannot be the convex sum of several Bell states because it must remain pure). Since the mapping of Bell states must be one-to-one, the action of any $U \in \mathcal{U}_{\mathrm{B\ loc}}(D,n)$ involves a permutation of the Bell states,

TABLE II. Values of the number of elements of the group $\mathcal{P}_S(D,n)$ for several qudit dimensions D.

D	2	3	4	5	6
$\mathcal{NP}_S(D,1)$	6	24	48	120	144
$\mathcal{NP}_S(D,2)$	720	51840	737280	9.36×10^6	$\sim3.7\times10^7$
$\mathcal{NP}_S(D,3)$	1451520	$\sim9.2\times10^9$	$\sim3.0\times10^{12}$	$\sim9.1\times10^{13}$	$\sim1.3\times10^{16}$

$$U\rho U^\dagger = \sum_{\mathbf{x}\in\mathbf{Z}_D^{2n}} p_{\mathbf{x}}^{(n)} |\pi(\mathbf{x})\rangle_{\mathcal{B}}\langle\pi(\mathbf{x})|, \tag{22}$$

where $\pi:\mathbf{Z}_D^{2n}\to\mathbf{Z}_D^{2n}$ is a permutation. So we introduce $\mathcal{P}_{\text{loc}}(D,n)$, the *group of local permutations*, as the set of permutations over $\mathbf{Z}_D^{2n}$ implementable over Bell states by local (unitary) means.

Before stating the main result of this section, we shall define several groups. Consider the family of unitary operators $u_{\mathbf{x}}(\mathbf{x}\in\mathbf{Z}_D^{2n})$ over Bob's part of the system such that by definition

$$1\otimes u_{\mathbf{x}}^*|\mathbf{0}\rangle_{\mathcal{B}} := ||\mathbf{x}\rangle_{\mathcal{B}}, \tag{23}$$

where conjugation is taken with respect to the computational basis. With these operators at hand, we construct the group $\mathcal{U}_{\text{B inv}}(D,n)$ with the elements $U_{\mathbf{x}} := u_{\mathbf{x}}\otimes u_{\mathbf{x}}^*$. We claim that it is a subgroup of $\mathcal{U}_{\text{B loc}}(D,n)$. An explicit calculation shows that the action of its elements is

$$U_{\mathbf{x}}|\mathbf{y}\rangle_{\mathcal{B}} = \varphi(\mathbf{x}^t\Omega\mathbf{y})|\mathbf{y}\rangle_{\mathcal{B}}, \tag{24}$$

where $\Omega\in\mathbf{M}_{2n\times2n}(\mathbf{Z}_D)$ is

$$\Omega := \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}. \tag{25}$$

So the special feature of $\mathcal{U}_{\text{B inv}}$ is that for $\rho^{(n)}$ Bell diagonal $U_{\mathbf{x}}\rho^{(n)}U_{\mathbf{x}}^\dagger=\rho^{(n)}$, which means that its elements implement the identity permutation.

We also define two subgroups of the group of permutations over $\mathbf{Z}_D^{2n}$. The *translation group* $\mathcal{P}_T(D,n)$ contains the permutations of the form

$$\pi_{\mathbf{a}}(\mathbf{x}) = \mathbf{x}+\mathbf{a}, \tag{26}$$

with $\mathbf{a}\in\mathbf{Z}_D^{2n}$, and the *symplectic group* $\mathcal{P}_S(D,n)$ contains in turn those whose action is

$$\pi_M(\mathbf{x}) = M\mathbf{x}, \tag{27}$$

where $M\in\mathbf{M}_{2n\times2n}(\mathbf{Z}_D)$ is such that

$$M^t\Omega M = \Omega. \tag{28}$$

$\mathcal{P}_S(D,n)$ is a finite nonsimple group. A suitable generator set for this group is presented in Appendix C. Now, we are in a position to establish the following theorem that plays an important role in the distillation protocols for qudits to be devised later on.

Theorem III.1.

1. $\mathcal{P}_{\text{loc}}$ is the semidirect product of $\mathcal{P}_S$ and $\mathcal{P}_T$,

$$\mathcal{P}_{\text{loc}}(D,n) = \mathcal{P}_T(D,n)\ltimes\mathcal{P}_S(D,n). \tag{29}$$

2. Let h be the natural homomorphism from $\mathcal{U}_{\text{B loc}}$ onto $\mathcal{P}_{\text{loc}}$; then its kernel is

$$\ker h = \mathcal{U}_{\text{B inv}}(D,n)\otimes U(1), \tag{30}$$

where $U(1)$ denotes the global phase.

We prove this theorem in Appendix D. For qubits ($D=2$), part 1 of this theorem was proved in [10] using a mapping between Bell states and Pauli matrices. Our proof does not rely on this mapping, and being completely different, it becomes general enough so as to treat all qudits of dimension D on an equal footing.

$\mathcal{P}_S$ is specially well suited to construct distillation protocols, and so it is worth a closer study of its properties. There is another interesting way of writing (28); if we call $\mathbf{u}_i$ the first n rows (columns) of M and $\mathbf{v}_i$ the last n rows (columns), the condition can be rewritten in a canonical symplectic form,

$$\mathbf{u}_i^t\Omega\mathbf{u}_j = 0,$$

$$\mathbf{v}_i^t\Omega\mathbf{v}_j = 0,$$

$$\mathbf{u}_i^t\Omega\mathbf{v}_j = \delta_{ij}. \tag{31}$$

This point of view is especially useful when systematically constructing the elements of $\mathcal{P}_S$, thanks to the following result.

Theorem III.2.

1. Consider a linearly independent set of vectors $\{\mathbf{u}_1,\ldots,\mathbf{u}_r,\mathbf{v}_1,\ldots,\mathbf{v}_s,\mathbf{v}_{r+1},\ldots,\mathbf{v}_{r+t}\}\subset\mathbf{Z}_D^{2n}$ with $(s\leqslant r\leqslant n, s+t\leqslant n)$. If this set satisfies conditions (31), it is always possible to complete it while preserving them.

2. The cardinality of $\mathcal{P}_S$ is

$$\mathcal{NP}_S(D,n) = D^{n^2}\prod_{k=1}^{n}\phi_{2k}(D). \tag{32}$$

We prove this theorem in Appendix E. As an illustration, we list several values of $\mathcal{NP}_S(D,n)$ in Table II. Clearly numbers grow fast, which makes unfeasible any numerical investigation which requires going over the elements of the whole group even for not very large values of n. In any case, it can be helpful to have an algorithm which allows this task without the expense of storing the elements. Consider any suitable ordering over $\mathbf{Z}_D^{2n}$. Given an element of $\mathcal{P}_S(D,n)$, we can increase its last row according to this order until another element is reached. If this fails, the same can be done for the previous row, and so on and so forth. However, this is not very efficient, and we can do it better combining part one of theorem III.2 with lemma C.1. For example, for any of the

last n rows, we could substitute the search with the application of a suitable generator from lemma C.1. Then the additional information contained in the proof of theorem III.2 would guarantee that we were not forgetting any element of the group. Moreover, as we shall later see, typically we are only interested in some of the rows of the matrix, and then part 1 of the theorem is crucial since it allows us to ignore unimportant rows.

IV. TWIRLING WITH $\mathcal{U}_{\mathrm{B\ inv}}$ AND $\mathcal{P}_{\mathrm{loc}}$

We now explore the possibility of using the groups of the previous section with the twirling operator (8), which for finite groups is

$$T_{\mathcal{U}}(\rho) := \frac{1}{\mathcal{N}\mathcal{U}} \sum_{U \in \mathcal{U}} U \rho U^{\dagger}. \tag{33}$$

Consider now the group $\mathcal{U}_{\mathrm{B\ inv}}(D,n)$. From (24), it follows that

$$\sum_{\mathbf{z}} U_{\mathbf{z}} |\mathbf{x}\rangle_{\mathcal{B}} \langle \mathbf{y}| U_{\mathbf{z}}^{\dagger} = \delta(\mathbf{x}-\mathbf{y}) |\mathbf{x}\rangle_{\mathcal{B}} \langle \mathbf{x}|, \tag{34}$$

which means that the action of $T_{\mathcal{U}_{\mathrm{B\ inv}}}$ over a state leaves Bell diagonal elements invariant, whereas the off-diagonal components are sent to zero.

The group $\mathcal{P}_{\mathrm{S}}$ can also be successfully used in twirling operations. This asseveration, however, has no meaning by itself since $\mathcal{P}_{\mathrm{S}}$ is not a group of transformations over the n pairs of qudits. We have to choose any mapping $U:\mathcal{P}_{\mathrm{S}} \rightarrow \mathcal{U}_{\mathrm{B\ loc}}$ such that

$$U(\pi) |\mathbf{x}\rangle_{\mathcal{B}} \langle \mathbf{x}| U^{\dagger}(\pi) = |\pi(\mathbf{x})\rangle_{\mathcal{B}} \langle \pi(\mathbf{x})|. \tag{35}$$

There are many possible realizations for this mapping, and at least in general the image of the mapping is not a subgroup of $\mathcal{U}_{\mathrm{B\ loc}}$. However, it *is* a group when considered as a set of transformations over Bell diagonal states. Thus, for ρ Bell diagonal the following makes sense:

$$T_{\mathcal{P}_{\mathrm{S}}}(\rho) := \frac{1}{\mathcal{N}\mathcal{P}_{\mathrm{S}}} \sum_{\pi \in \mathcal{P}_{\mathrm{S}}} U(\pi) \rho U(\pi)^{\dagger}. \tag{36}$$

To perform the sum, we need to know which are the states invariant under the action of the group.

Theorem IV.1. For every $\mathbf{x},\mathbf{y} \in \mathbf{Z}_D^{2n}$, $\gcd(\mathbf{x}) = \gcd(\mathbf{y})$ if and only if there exists a permutation in $\mathcal{P}_{\mathrm{S}}(D,n)$ with associated matrix M such that $\mathbf{x} = M\mathbf{y}$.

Proof. The if direction follows from M being invertible, since then $dM\mathbf{i} = Md\mathbf{i} = 0$ iff $d\mathbf{i} = 0$. We now prove only the if direction. Let $d = \gcd(\mathbf{x})$ and consider any $\mathbf{u} \in \mathbf{Z}_D^{2n}$ such that $d\mathbf{u} = \mathbf{x}$ [note that $\gcd(\mathbf{u}) = 1$]. $\{\mathbf{u}\}$ is linearly independent, and so there exists a matrix M associated to a permutation in $\mathcal{P}_{\mathrm{S}}(D,n)$ having $\mathbf{u}$ as its first column (see theorem III.2). Then, if $\mathbf{v} = (d,0,\ldots,0)$ we have $\mathbf{x} = M_1\mathbf{v}$. The same reasoning is true for $\mathbf{y}$, giving $\mathbf{y} = M_2\mathbf{v}$ and thus $M = M_1 M_2^{-1}$. □

Now, let us recall the partition in $\mathbf{Z}_D^{2n}$ associated to the function gcd [see (15)]. We define the related states

$$\rho_d := \frac{1}{\mathcal{N}C_d} \sum_{\mathbf{x} \in C_d} |\mathbf{x}\rangle_{\mathcal{B}} \langle \mathbf{x}|. \tag{37}$$

These are the invariant states we were searching for. Thus, if ρ is Bell diagonal,

$$T_{\mathcal{P}_{\mathrm{S}}}(\rho) = \sum_{d \in \mathrm{div}(D)} \frac{\mathrm{Tr}(\rho_d \rho)}{\mathrm{Tr}(\rho_d \rho_d)} \rho_d. \tag{38}$$

Note that we have not taken into account the number of pairs involved, since it is unimportant. However, usually twirling operations are interesting for $n=1$. In this case, in analogy with isotropic states, we shall call heterotropic states those states invariant under (38). If ρ is not Bell diagonal, we can still obtain the same result with the operator

$$T_{\mathcal{U}_{\mathrm{B\ loc}} \times \mathcal{P}_{\mathrm{S}}}(\rho) := \frac{1}{\mathcal{N}\mathcal{U}_{\mathrm{B\ inv}}} \frac{1}{\mathcal{N}\mathcal{P}_{\mathrm{S}}} \sum_{\pi \in \mathcal{P}_{\mathrm{S}}} \sum_{U \in \mathcal{U}_{\mathrm{B\ inv}}} U(\pi) U \rho U^{\dagger} U(\pi)^{\dagger}. \tag{39}$$

As a corollary, if D is prime there are just two Bell diagonal invariant states,

$$\rho_1 = \frac{1}{D^{2n}-1} (1 - |0\rangle_{\mathcal{B}} \langle 0|), \tag{40}$$

$$\rho_D = |0\rangle_{\mathcal{B}} \langle 0|, \tag{41}$$

and thus the result of the twirling operation is an isotropic state, which is the simplest example of a heterotropic state.

V. PERMUTATION ASSISTED DISTILLATION

In the distillation protocols we consider, which are iterative, each iteration cycle can be decomposed in the following steps.

1. At start, Alice and Bob share n qudit pairs of dimension D and state matrix $\rho^{(n)}$.
2. They apply by local means one of the permutations $\pi_M \in \mathcal{P}_{\mathrm{S}}(D,n)$ in (27).
3. They measure the last n-m qudit pairs, both of them in their computational basis.
4. If the results of the measurement agree for each of the measured pairs, they keep the first m pairs (in the state $\rho^{(m)}$). Otherwise, they discard them.

In most situations, the initial n pairs are independent and have equal state matrices ρ. In these cases

$$\rho^{(n)} = \rho^{\otimes n}. \tag{42}$$

In general (for $m>1$) this does not guarantee that $\rho^{(m)}$ will be a product state, however, and thus it is preferable to consider the most general case.

The process can be performed several times in order to improve the entanglement progressively, but it is worth taking into account that a scheme of this kind with s steps and, say, $n=2$ and $m=1$, is equivalent to a single-step one with $n'=2^s$ and $m'=1$. It is enough to perform initially all the

permutations and afterward all the measurements *at the same time*. Although convergence properties are the same, the equivalence is not complete since from a practical point of view the step-by-step method will give a better yield. This is so because an undesired result in the measurement is more harmful in the second case, as more pairs must be discarded at once. An example clarifies this issue. Let us take a step-by-step case with $n=2$ and $m=1$, with a probability of success at the first step of P_1, and P_2 similarly for the second step. Then the yield in this case is $P_1P_2/2^2$. However, in the single-step protocol ($n'=2^2$, $m'=1$), there is a single probability of success given by $P_1^2P_2$, thus the corresponding yield is $P_1^2P_2/2^2$. Therefore, we see that in the single-step protocol there is an extra factor of $P_1<1$ that reduces its yield with respect to the step-by-step protocol, and this reduction gets amplified when considering a higher number of steps in the distillation.

In Appendix F, we derive an expression for the state of the remaining pairs of qudits after a successful measurement. It appears that the protocol is blind to nondiagonal states (in the Bell basis). So let us define

$$p_{\mathbf{x}}^{(n)} := {}_{\mathcal{B}}\langle \mathbf{x}|\rho^{(n)}|\mathbf{x}\rangle_{\mathcal{B}}, \quad \mathbf{x} \in \mathbf{Z}_D^{2n};$$

$$p_{\mathbf{x}}^{(m)} := {}_{\mathcal{B}}\langle \mathbf{x}|\rho^{(m)}|\mathbf{x}\rangle_{\mathcal{B}}, \quad \mathbf{x} \in \mathbf{Z}_D^{2m}.$$

If we call V_M the space generated by the last $n-m$ rows of M (the matrix associated to π_M), the probability of obtaining the desired measure is

$$P = \sum_{\mathbf{x}\in V_M^{\perp}} p_{\mathbf{x}}^{(n)}, \tag{43}$$

and the recurrence relation for the probabilities is

$$p_{\mathbf{x}}^{(m)} = \frac{1}{P}\sum_{\mathbf{y}\in V_M} p_{\Omega\mathbf{y}+M^{-1}\bar{\mathbf{x}}}^{(n)}, \tag{44}$$

where $\mathbf{x}\in\mathbf{Z}_D^{2m}$ and $\bar{\mathbf{x}}\in\mathbf{Z}_D^{2n}$ is

$$\bar{\mathbf{x}} := (x_1, \ldots, x_m, \underbrace{0, \ldots, 0}_{n-m}, x_{n+1}, \ldots, x_{n+m}, \underbrace{0, \ldots, 0}_{n-m}). \tag{45}$$

Note that since $M^{-1}=\Omega^t M^t \Omega$, rows $m+1$ to n (of M) do not take part in the expression, and therefore the protocol does not depend on them.

Consider the following family of heterotropic states:

$$\rho_d^{\mathrm{fix}} := \sum_{\mathbf{x}\in\mathbf{Z}_D^2} \frac{1}{Dd^2}\delta(d\mathbf{x})|\mathbf{x}\rangle_{\mathcal{B}}\langle\mathbf{x}|, \quad d\in \mathrm{div}(D). \tag{46}$$

From theorem IV.1, we know that

$$\delta(dM\mathbf{x}) = \delta(d\mathbf{x}). \tag{47}$$

Using this fact and Eq. (F4), one can readily check that for $\rho^{(n)}=\rho_d^{\mathrm{fix}\otimes n}$, Eq. (44) gives $\rho^{(m)}=\rho_d^{\mathrm{fix}\otimes m}$. Therefore, these heterotropic states are always fixed points of the protocol and candidates for attractors.

In the case of single-step protocols, we are only interested in the final joint fidelity (the probability of the state being $|\mathbf{0}\rangle_{\mathcal{B}}\langle\mathbf{0}|$),

$$F^{(m)} = \frac{1}{P}\sum_{\mathbf{x}\in V_M} p_{\Omega\mathbf{x}}^{(n)}, \tag{48}$$

where the label (m) reminds us that this is the joint probability of the m pairs being in the desired state. In general, the fidelity of each pair will be greater. If $m=1$, this distinction vanishes, and we will simply write F' instead of $F^{(1)}$. Equations (48) and (43) show how the effect of the entire process relies only on the set V_M, thereby reducing the search for efficient protocols according to part one of theorem III.2.

We now consider the Fourier transform of the probabilities,

$$p_{\tilde{\mathbf{x}}}^{(n)} := \sum_{\mathbf{x}\in\mathbf{Z}_D^{2n}} \varphi(\tilde{\mathbf{x}}\cdot\mathbf{x})p_{\mathbf{x}}^{(n)}, \quad \tilde{\mathbf{x}}\in\mathbf{Z}_D^{2n} \tag{49}$$

to obtain

$$P = D^{m-n}\sum_{\tilde{\mathbf{x}}\in V_M} p_{\tilde{\mathbf{x}}}^{(n)}, \tag{50}$$

where we have used

$$\sum_{\mathbf{v}\in V}\varphi(\mathbf{v}\cdot\mathbf{u}) := \begin{cases} 0 & \text{if } \mathbf{u}\notin V^{\perp} \\ \mathcal{N}V & \text{if } \mathbf{u}\in V^{\perp}, \end{cases} \tag{51}$$

with V being a subspace of $\mathbf{Z}_D^n$ and $\mathbf{u}\in\mathbf{Z}_D^n$. Gathering these results, we have

$$F^{(m)} = D^{n-m}\frac{\sum_{\mathbf{x}\in V_M} p_{\Omega\mathbf{x}}^{(n)}}{\sum_{\mathbf{x}\in V_M} p_{\tilde{\mathbf{x}}}^{(n)}}, \tag{52}$$

an expression for the final (joint) fidelity which can lighten its direct computation when (42) holds, since for this case we can define for $a, b\in\mathbf{Z}_D$

$$p_{ab} := {}_{\mathcal{B}}\langle ab|\rho|ab\rangle_{\mathcal{B}}, \quad p_{\tilde{a}\tilde{b}} := \sum_{a,b\in\mathbf{Z}_D}\varphi(\tilde{a}a+\tilde{b}b)p_{ab}, \tag{53}$$

and then

$$p_{\mathbf{x}}^{(n)} = \prod_{i=1}^{n} p_{x_i,x_{n+i}}, \quad p_{\tilde{\mathbf{x}}}^{(n)} = \prod_{i=1}^{n} p_{\tilde{x}_i,\tilde{x}_{n+i}}. \tag{54}$$

A. Twirling-assisted protocols

In order to understand better Eq. (44), we will adopt a useful simplification which generates by itself a whole family of distillation protocols. We consider only initial states for which the pairs are mutually independent and equal as in (42). Moreover, before each iteration we introduce any twirling operation which leads to an isotropic state while preserving the fidelity, as in (7). This way, the evolution of the state through the distillation protocol is described entirely by a single parameter, the fidelity F.

We would like to evaluate (54). We need

$$p_{ab} = F\delta_{ab} + \frac{1-F}{D^2-1}(1-\delta_{ab}), \tag{55}$$

$$p_{\bar{a}\bar{b}} = \delta_{\bar{a}\bar{b}} + \frac{D^2F - 1}{D^2 - 1}(1 - \delta_{\bar{a}\bar{b}}). \tag{56}$$

Defining $c_1(F) := (1-F)/[F(D^2-1)]$ and $c_2(F) := (D^2F - 1)/(D^2-1)$, this implies that for any $\mathbf{x}$ we have $p_{\mathbf{x}}^{(n)} = F^n c_1(F)^s$ and $p_{\bar{\mathbf{x}}}^{(n)} = c_2(F)^s$, where $n-s$ is the number of occurrences of p_{00} in the product (that is, $\mathcal{N}\{r=1,\ldots,n : x_r = x_{n+r}=0\}$). Since $p_{\Omega\mathbf{x}}^{(n)} = p_{\mathbf{x}}^{(n)}$, we can write

$$F^{(m)} = D^{n-m}F^n \frac{\chi[c_1(F)]}{\chi[c_2(F)]} \tag{57}$$

and

$$P = D^{m-n}\chi[c_2(F)], \tag{58}$$

where $\chi(x) = \sum_{s=0}^{n} \lambda_s x^s$ is a polynomial with coefficients defined by

$$\lambda_s := \mathcal{N}\{\mathbf{x} \in V_m : n - s = \mathcal{N}\{r = 1, \ldots, n : x_r = x_{n+r} = 0\}\}. \tag{59}$$

This definition is not very useful when trying to construct a suitable V_M for a given χ, but we can do better. Let V_r be the set of linear combinations of columns r and $n+r$ of the matrix formed by the last $n-m$ rows of M [or any other matrix of size $(n-m)\times 2n$ such that its rows span V_M]. If V_r is a subspace of $\mathbf{Z}_D^{n-m}$ for every r, which indeed is always the case for D prime, we can rewrite (59) as

$$\lambda_s = \mathcal{N}\{\mathbf{v} \in \mathbf{Z}_D^{n-m} : n - s = \mathcal{N}\{r = 1, \ldots, n : \mathbf{v} \in V_r^{\perp}\}\}. \tag{60}$$

We remark that χ depends only on V_M, which is a subspace of $\mathbf{Z}_D^{2n}$ of dimension $n-m$ constrained only by

$$\mathbf{u}^t \Omega \mathbf{v} = 0 \quad \forall\, \mathbf{u}, \mathbf{v} \in V_M. \tag{61}$$

It is apparent that $\chi(1) = D^{n-m}$ and $\chi(0) = 1$. For $m=1$, (57) becomes the recurrence relation $F' = F'(F)$, and then among the fixed points are $F=1$ (perfect entanglement), $F = 1/D$ (maximum fidelity for separable states) and $F = 1/D^2$ (pure noise).

Therefore, we have reduced the problem of finding the best protocol to that of finding the best coefficients for the polynomial, constrained to the existence of a suitable vector space. In the next subsection, we survey this issue for several values of n, but previously a small consideration is worthwhile. For the identity permutation (which of course is completely useless), the coefficients of the polynomial are $\lambda_s = \binom{n-m}{s}(D-1)^s$ and therefore

$$P = \left(\frac{1 + (D-1)F}{D}\right)^{n-m}. \tag{62}$$

One expects the probability of a useful protocol to be less than this, but then the decay is at least exponential with respect to an increase in $n-m$. This is an early advisory that considering progressively larger values of n need not be better.

1. Low fidelity states

Now we will concentrate on low fidelity states near $F = 1/D$. Since hashing and breeding protocols are available for high fidelities [6,11], one reason for studying this range is that it is the natural testing ground for the class of protocols we are analyzing. It is interesting also because we can develop a method to compare in a simple manner protocols with different n.

We start by discarding protocols with $m > 1$. In order to see why this is reasonable, let us consider the individual fidelities of each of the resulting m pairs, say F_i, $i = 1, \ldots, m$. For isotropic states, $F \leqslant 1/D$ is equivalent to separability, and thus $F_i(1/D) \leqslant 1/D$. Since $F^{(m)}(1/D) = 1/D^m$, for uncorrelated pairs we have $F_i = 1/D$. However, for correlated pairs in general (although not necessarily) the individual fidelities will be less than $1/D$, making the algorithm useless near the point of interest. We will see later how protocols with $m > 1$ can be fruitfully used.

A problem arising when comparing different protocols is that several factors take part at the same time, making it difficult to balance them in a simple manner. In our case, we have to take into account the probability of obtaining the right measure P, the number of pairs used n, and the output fidelity F'. We will now see, however, that restricting our attention to low fidelity states allows us to introduce a single coefficient, which makes possible the comparison. We shall call this coefficient the *joint performance* η of a distillation protocol and it is constructed as follows.

Let us consider an isotropic state of fidelity $(1/D) + \epsilon$. After q steps of the protocol, at the lowest order in ϵ, the state will have a fidelity $(1/D) + F_1^q \epsilon$ and the yield will be $(P_0/n)^q$, with

$$F_1 := \left.\frac{dF'}{dF}\right|_{F=1/D} = n - \frac{2D}{D^2-1}\left[\frac{d}{dx}\log[\chi(x)]\right]_{x=1/(D+1)}, \tag{63}$$

$$P_0 := \left. P \right|_{F=1/D} = D^{1-n}\chi\left(\frac{1}{D+1}\right). \tag{64}$$

We will assume $F_1 > 1$, since the protocol must be meaningful. The yield after amplifying ϵ by a factor t is $\eta^{\log(t)}$, and thus it is justified to introduce the coefficient

$$\eta := \exp\left(\frac{\log(P_0) - \log(n)}{\log(F_1)}\right). \tag{65}$$

As $F_1 < n$ and $P_0 \leqslant 1$, then $\eta < e^{-1}$.

We are ready to compare several protocols, which we shall do progressively increasing the number of discarded pairs.

(i) $n=2$. When Eq. (60) applies, there are just two possibilities. One corresponds to the identity permutation and the other is

$$\chi(x) = 1 + (D-1)x^2. \tag{66}$$

This corresponds to the original distillation protocol discussed in [6], as we expected. If D is not prime, there are other possibilities, but η is not greater for them.

(ii) $n=3$. We must distinguish two cases. If D is odd, the best value of η is attained with

$$\chi(x) = 1 + (D^2 - 1)x^3. \tag{67}$$

If D is even, however, the best option is

$$\chi(x) = 1 + (D - 1)x^2 + (D^2 - D)x^3. \tag{68}$$

The difference is due to the impossibility of constructing a suitable V_M in the second case, as we now show. Let $\{\mathbf{u},\mathbf{v}\}$ be a basis of V_M. Then, condition (61) is equivalent to

$$\sum_{i=1}^{3} \begin{vmatrix} u_i & u_{3+i} \\ v_i & v_{3+i} \end{vmatrix} = 0. \tag{69}$$

In order to obtain Eq. (67), the determinants appearing in the sum should have an invertible value [see Eq. (60)], but then they cannot sum up 0 if D is even.

(iii) $n=4$. In this case, we have found that the best polynomial is

$$\chi(x) = 1 + 4(D - 1)x^3 + (D^3 - 4D + 3)x^4. \tag{70}$$

As an example of realizing this case, set $V_M=\mathrm{Lin}\{\mathbf{u},\mathbf{v},\mathbf{w}\}$ with

$$\mathbf{u} = (1,0,0,1,0,1,0,0), \tag{71}$$

$$\mathbf{v} = (0,1,0,0,1,0,1,0), \tag{72}$$

$$\mathbf{w} = (1,1,-1,0,0,0,0,1). \tag{73}$$

Figure 1 displays the values of η for these protocols and several dimensions. Note the bad performance of the case $n=2$ for qubits, which is in fact the most important of all. On the other hand, qutrits ($D=3$) obtain the best yield among the proposed protocols, thanks to the advantage of odd dimensionality (for $n=3$). In connection with this, see also Fig. 2.

One could ask whether further improvement on η is possible by means of increasing n. Figure 1 suggests that this is not the case, at least for $D>2$. Exploration shows that nothing is gained in the case of qubits either. This result is an indication of the futility of increasing n with the aim of improving performance within the context of the current protocol. In [10] it is claimed that protocols with higher n should improve the yield, but apparently they do not take into account the (strong) reduction in the probability as n increases. This idea clarifies Fig. 1, since from Eq. (62) we expect $P_0 < 2^{n-1}(D+1)^{1-n}$ and then the reduction of the probability with n is more dramatic as D increases, whereas the performance gain from F_1 is at most linear ($F_1<n$).

2. Protocols with $m>1$

When considering states of higher fidelity, an important advantage of the proposed protocol for $n=3$ and D odd is that the derivative of $F'(F)$ vanishes for $F=1$, a qualitative difference with the $n=2$ case (see Fig. 2). This is important for states close to the Bell state, since a fidelity $1-\epsilon$ is mapped to $1-O(\epsilon^2)$. We now show how the ratio $n/m=2$ can be preserved while this desirable characteristic is added.

Using the definitions in lemma C.1, consider the permutation

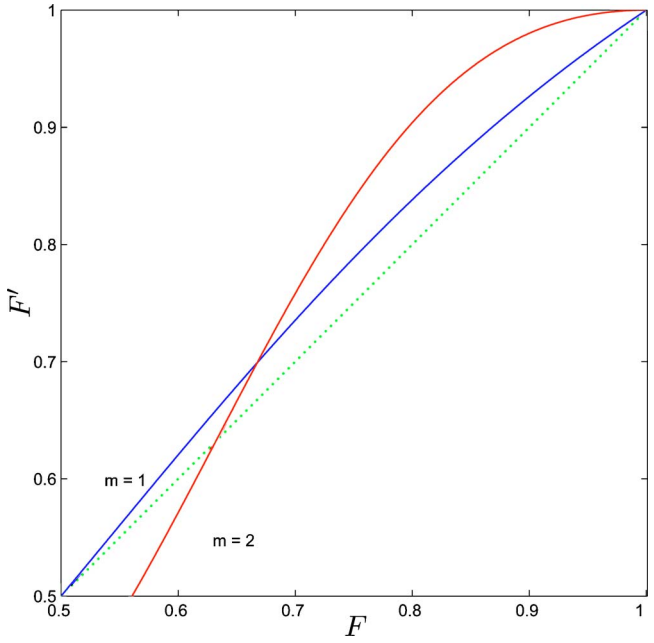

FIG. 3. (Color online) Evolution of fidelity through the proposed $n=2m$ twirled-assisted protocols for $D=2$.

$$\pi_{++}^{kl} := (\pi_{+}^{kl} \circ \pi_{+}^{l} \circ \pi_{+}^{k})^{-1} \circ \pi_{+}^{k} \circ \pi_{+}^{kl}. \tag{74}$$

The action of π_{++}^{12} is

$$\mu(\mathbf{i},\mathbf{j}) = \mathbf{i}, \tag{75}$$

$$\nu(\mathbf{i},\mathbf{j}) = (j_1 + i_2, j_2 + i_1, j_3, \ldots, j_n). \tag{76}$$

We propose for $n=4$ and $m=2$ a protocol in which the permutation π_M of step 2 is

$$\pi_{+}^{13} \circ \pi_{+}^{24} \circ \pi_{++}^{14} \circ \pi_{++}^{23}. \tag{77}$$

This permutation yields

$$\chi(x) = 1 + (D^2 - 1)x^4. \tag{78}$$

The resulting two pairs of qudits will be correlated, thereby providing us with a good scenario for hashing, and one could consider an iterative protocol in which the basic units were pairs of pairs of qudits (instead of pairs). We shall keep things simple by choosing $D=2$ and considering the partial traces of each of the pairs (which are equal due to symmetry of the permutation) in order to obtain the individual fidelity,

$$F' = \frac{F^4}{P}[1 + 4c_2(F)^2 + 4c_2(F)^3 + 7c_2(F)^4]. \tag{79}$$

The results for $D=2$ are displayed in Figs. 3 and 4. The yield Υ of Fig. 4 is calculated step by step through the following recursion relation:

$$\Upsilon_k = P_k \frac{m}{n} \Upsilon_{k-1}, \quad \Upsilon_0 = 1. \tag{80}$$

Regarding the $n=2$ case, the yield is greatly increased (four orders of magnitude) even for states quite near to the fixed point. The drawback is the impossibility of distillation for

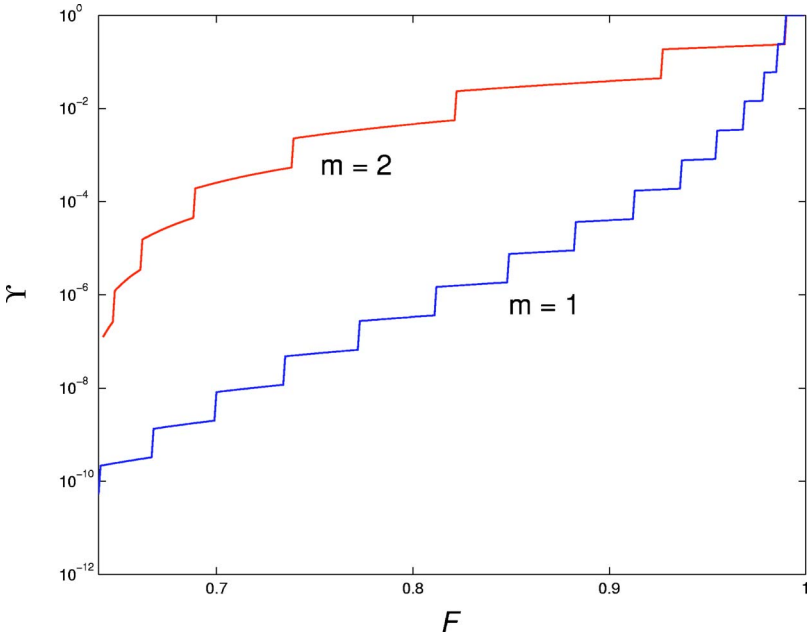

FIG. 4. (Color online) Yield Υ (80) after reaching a fidelity at least 0.99 through the proposed $n=2m$ twirled-assisted protocols for $D=2$. F is the initial fidelity.

$F \lesssim 0.64$, but let us recall that this is below the lowest fidelity distillable with a hashing method, $F \approx 0.81$ [6]. The conclusion then is that one should consider this kind of protocol in the latest steps prior to hashing.

B. Protocols without twirling

The use of twirling involves losing entanglement, and thus the protocols we have considered so far are a good starting point toward more sophisticated ones in which a careful selection of the permutations avoids the use of twirling techniques and allows for the distillation of states with fidelity less than $1/D$, if $D>2$.

1. Quantum privacy amplification

This idea was first explored (for qubits) in [7], where a quantum privacy amplification scenario was considered. In this situation, the state to be distilled is the average over an ensemble, not necessarily known, and so the permutations must work well in general. We shall now generalize this algorithm to qudits, guided by the main role the vector space V_M plays, as introduced in the beginning of Sec. V. The proposed generalization is an iteration with $n=2$ and $m=1$ as in the original case. It consists of an alternated application of two permutations,

$$\pi_+^{12} \circ \pi_+^1 \circ \pi_+^2,$$

$$\pi_+^{12} \circ (\pi_+^1 \circ \pi_+^2)^{-1}. \tag{81}$$

The (relative) simplicity of these operations is a first interesting point of the algorithm. The choice follows from the intention to preserve the form of V_M with respect to the known case $D=2$, as the number of iterations grows. For s iterations, M in this case is the $2^{s+1} \times 2^{s+1}$ matrix which would follow by considering the process as a single iteration, with a unique permutation and a unique measurement.

Although the two permutations alternate, it is possible to give a single recursion relation for every iteration cycle. To this end, let us introduce the elements of an alternative Bell basis as

$$|\mathbf{i}\ \mathbf{j}\rangle_{\mathcal{B}'} := |\mathbf{i}-\mathbf{j}\rangle_{\mathcal{B}}. \tag{82}$$

Then, in order to achieve this, it is enough to change the Bell basis to (82) after the first cycle, switch to the original Bell basis after the second, change again to (82) after the third one, and so on. with this little trick, we get the following recursion relation:

$$p'_{ij} = \frac{1}{P} \sum_{k \in \mathbf{Z}_D} p_{i+k,-i-j-k} p_{k,j-k}, \tag{83}$$

$$P = \frac{1}{D} \sum_{\tilde{k} \in \mathbf{Z}_D} p_{\tilde{k}\tilde{k}}^2. \tag{84}$$

It is interesting to note that the permutation that can switch between the two bases is not achievable by local means. If this were so, we could avoid the use of two different permutations and still get the same recursion relation, but unfortunately it is not the case.

Figure 5 shows the yield of this protocol compared to the equivalent protocol of the previous Sec. V A. The improvement is clear. As D increases, the results are less spectacular, however. An important detail is that now the distillable states are not simply those for which fidelity is greater than $1/D$. As one can check in Fig. 6, some states over this point are not distilled whereas other states beneath it are. Qubits are the only ones behaving as expected: the total volume of Bell states with $F>\frac{1}{2}$ is distilled, while those below it are undistilled. In any case, the normalized volume of states showing bad behavior, i.e., those with $F>1/D$ which are not distilled, is small if D is prime. This is not so for nonprime D's because, as discussed in Sec. V, new fixed states emerge for

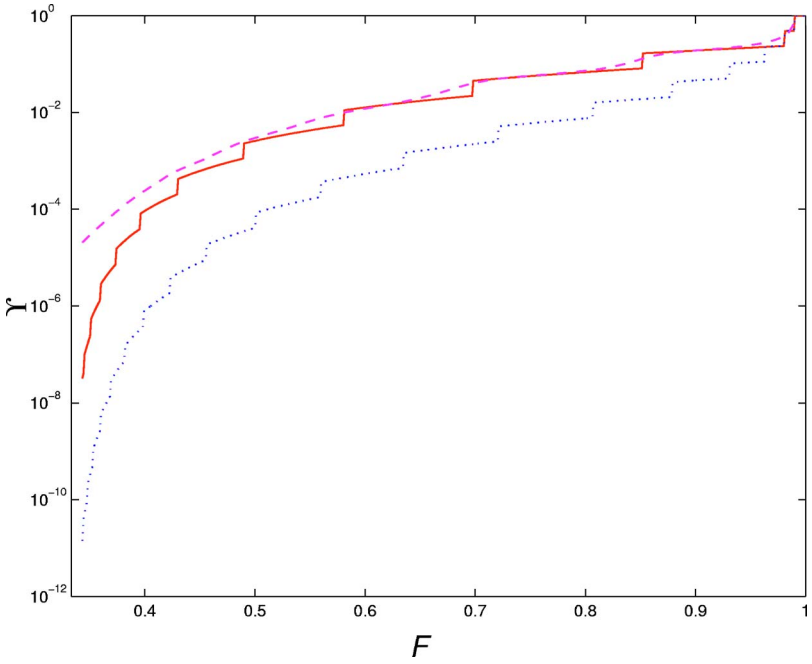

FIG. 5. (Color online) Yield Υ (80) after reaching a fidelity of at least 0.99 through the protocol proposed in the text: for isotropic states (solid line) and as a mean over Bell diagonal states (dashed line) compared to the same yield using the twirling protocol for $n=2$ (dotted line). The case under study is qutrits ($D=3$) and the mean refers to the measure discussed in Appendix G with Monte Carlo. F is the initial fidelity.

composite numbers creating undesirable attractors. We find that these attractors are especially harmful for states near to heterotropic states. As a corollary, we show that the permutational approach to distillation is more suited to prime D's.

2. Distillability

When the initial state is known, we can make use of this information to improve the distillation by selecting at each step the most convenient permutation. Then the question is whether the protocols we are managing are able to distill any distillable state. As we lack a working algorithm to decide whether a given state is distillable, we will compare the normalized volume of distilled states to that of NPPT states (states with a nonpositive partial transpose), since belonging to this set is a necessary condition for distillability.

We have chosen the following protocol with $n=2$ and $m=1$: At each step, one of the elements of $\mathcal{P}_S(D,1)$ is applied to both pairs of qudits before the permutation p_+^{12}. The element is chosen so as to give the best fidelity after the (correct) measurement. This does not necessarily lead to an optimal strategy.

Figure 7 shows the distillation capacities for $D=3$. In general, for D prime the behavior is good since all states known to be distillable, i.e., those for which the fidelity is more than $1/D$, happen to be distillable with our protocol. More precisely, we have not found computationally any counterexample of this fact. Not all the NPPT states are distilled. This is perhaps another indication of the existence of nondistillable

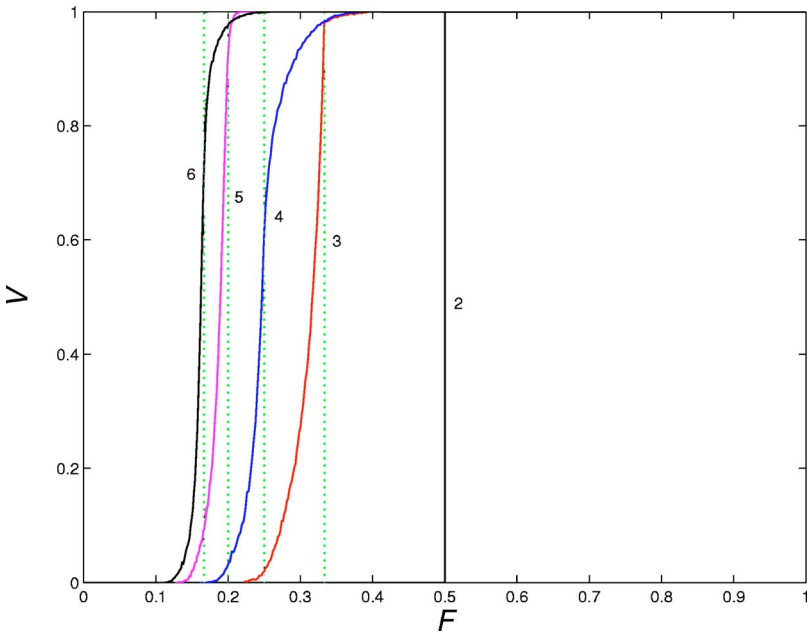

FIG. 6. (Color online) Normalized volume V of distilled Bell diagonal states for the protocol under study, given by permutations (81) with $D=2,3,4,5,6$. The measure is described in Appendix G and uses Monte Carlo. F is the initial fidelity.

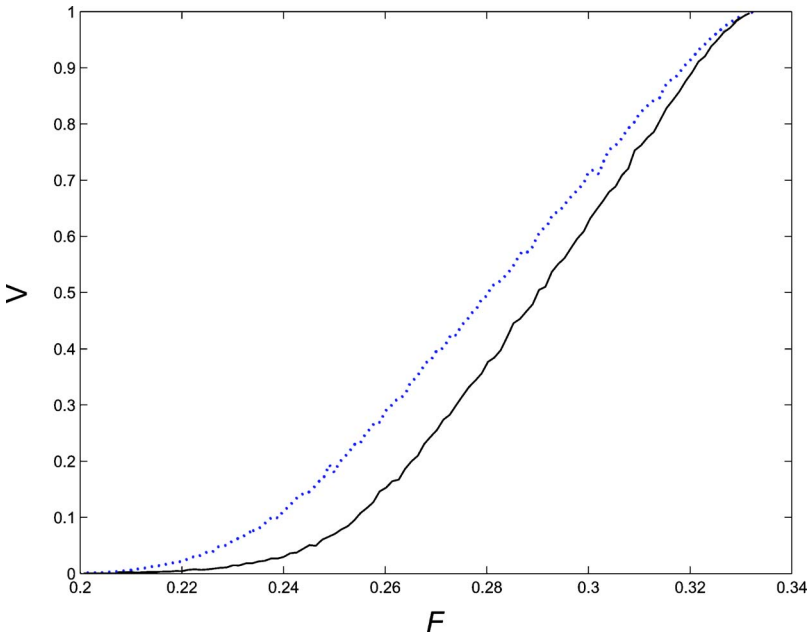

FIG. 7. (Color online) Normalized volume V of distilled Bell diagonal states compared to that of NPPT states for $D=3$. The metric and the measuring algorithm are discussed in Appendix G and uses Monte Carlo. F is the state fidelity.

NPPT states. In the case of composite numbers, the algorithm performs much worse.

VI. CONCLUSIONS AND PROSPECTS

We have shown that the study of entanglement distillation protocols based on the recursion method [5] benefits greatly from the application of basic number theory concepts when the set $\mathbf{Z}_D^n$ associated to qudits of arbitrary dimensions D is a module and not a vector space. In particular, we have found that a partition of $\mathbf{Z}_D^n$ into divisor classes is very useful to characterize the invariant properties of mixed Bell diagonal states under unitary groups that implement local permutations. These permutations, in turn, are used in very general distillation protocols based on the recursion method.

We have proposed and study a variety of distillation protocols that fall into two classes depending on whether we use twirling operations or not at intermediate steps of the protocols. When the twirling operations are absent, our distillation protocols amount to extensions of the quantum privacy amplification protocols [7] valid for arbitrary qudit dimensions D. This is very interesting and relevant for quantum communications with arbitrary large alphabets since they remain secure and operative even in the presence of quantum noisy channels.

These properties obtained from number theory are not only useful in the analytical understanding of the protocols, but also facilitate the construction of numerical methods for their study using Monte Carlo. In particular, we have characterized how the distillation protocols based on the recursion method and local permutations are qualitatively and quantitatively optimal when the dimension of the qudit states D is a prime number. We leave open the problem of how to construct better distillation protocols when D is not a prime number, and in this regard the use of the heterotropic states introduced here is a promising tool.

ACKNOWLEDGMENTS

We acknowledge financial support from the EJ-GV (H.B.) and DGS grant under Contract No. BFM 2003-05316-C02-01 (M.A.M-D.).

APPENDIX A: PROPERTIES OF THE MODULE $\mathbf{Z}_D^n$

In this appendix, we show how many ideas from genuine vector spaces can be adapted to the module $\mathbf{Z}_D^n$. First of all, we say that an element of $s \in \mathbf{Z}_D$ is *invertible* if there exists $s' \in \mathbf{Z}_D$ such that $ss'=1$. If $x \in \mathbf{N}$ is a representant of s, this is equivalent to $\gcd(x,D)=1$. When D is not prime, noninvertible elements other than zero exist (they are multiples of proper divisors of D) and we need to introduce a work around in the *Gaussian elimination* method, as we shall explain now.

Suppose we are given an element of $\mathbf{Z}_D^2$, say (x,y), and we are asked to get $x=0$ using two elementary transformations,

$$(x,y) \xrightarrow{\mathcal{O}_1} (x,x+y), \quad (x,y) \xrightarrow{\mathcal{O}_2} (x+y,y). \tag{A1}$$

The algorithm turns out to be quite simple. Consider for a moment the arbitrary ordering in $\mathbf{Z}_D$, $0<1<\cdots<D-1$. At each step, if $x \leqslant y$, use $\mathcal{O}_1$ to get $0 \leqslant y < x$; proceed inversely on the contrary. Clearly, $x=0$ or $y=0$ is reached in a finite number of steps. If $y=0$, just apply $\mathcal{O}_1$ once and $\mathcal{O}_2$ $D-1$ times.

We shall use Gaussian elimination, with the aid of the above trick, to convert any matrix M of size $p \times q$ into a very simple one. Suppose $p \leqslant q$; the converse case is similar. Summing one row to another amounts to taking the product (from the left) with a $p \times p$ invertible matrix. The same is true for columns (from the right, $q \times q$). Using these elementary operations, we can obtain

$$C = AMB, \quad C = [D\ 0], \tag{A2}$$

where A and B are invertible and D is a diagonal matrix.

With this tool at hand, we are ready to start with our analysis. We adopt the usual definition of linear independence for a finite subset of $\mathbf{Z}_D^n$, but the following is more surprising.

Definition A.1. Consider any $M \in \mathbf{M}_{p\times q}(\mathbf{Z}_D)$ and let S_r be the set of all $r\times r$ minors of M, $r \in R := \{1,\ldots,\min(p,q)\}$. The rank of M is defined as

$$\mathrm{rank}(M) := \max\{0\} \cup \{r \in R : \gcd(S_r) = 1\}. \tag{A3}$$

The rank of a matrix does not vary when we apply the elementary operations discussed above [see (14) and take into account that $d|x$ and $d|y$ iff $d|x$ and $d|x+y$]. As expected, a square matrix is invertible iff its rank is maximal. The following statement clarifies this strange definition.

Proposition A.2. The rows of a matrix $M \in \mathbf{M}_{p\times n}(\mathbf{Z}_D)$ form a linear independent (LI) set iff rank $(M)=p$.

Proof. Recall decomposition (A2) for M (we will use directly the notation there) and consider any $\mathbf{v} \in \mathbf{Z}_D^p$,

$$\mathbf{v}^t M = 0 \Leftrightarrow \mathbf{v}^t A^{-1} C B^{-1} = 0 \Leftrightarrow \mathbf{v}'^t C = 0,$$

where $\mathbf{v}=A^t\mathbf{v}'$. This shows that the rows of M form a LI set iff the rows of C do. On the other hand, $\mathrm{rank}(C)=\mathrm{rank}(M)$, and for the matrix C the statement is trivial. □

Given a set $S=\{\mathbf{v}_1,\ldots,\mathbf{v}_k\} \subset \mathbf{Z}_D^n$, we will denote by Lin S the subset of $\mathbf{Z}_D^n$ spanned by the elements of S, that is, the set of linear combinations of the vectors in S. Clearly, if S is LI, $\mathcal{N}\mathrm{Lin}\ S=D^k$, and so $k\leqslant n$. Not surprisingly, any LI set which spans $\mathbf{Z}_D$ will be called a basis of $\mathbf{Z}_D^n$. The usual definition of subspace does not work, however, and we introduce in its place the following.

Definition A.3. A set $V \subset \mathbf{Z}_D^n$ is said to be a subspace of $\mathbf{Z}_D^n$ if $V=\{0\}$ or if there exists a set $S=\{\mathbf{v}_1,\ldots,\mathbf{v}_k\} \subset \mathbf{Z}_D^n$ such that it is LI and Lin $S=V$. Such a set is called a basis of V, and its cardinality is the dimension of V (dim V).

Dimension is well defined since $\mathcal{N}V=D^{\sharp S}$ forbids the possibility of two bases of different cardinality.

Proposition A.4. Given a set $S=\{\mathbf{v}_1,\ldots,\mathbf{v}_k\} \subset \mathbf{Z}_D^n$ which is linearly independent, there exists a set $S'=\{\mathbf{v}_{k+1},\ldots,\mathbf{v}_n\}$ such that $S \cup S'$ is a basis of $\mathbf{Z}_D$.

Proof. Let M be a $k\times n$ matrix such that its rows are the elements of S. We recall (A2) but rewrite it in terms of $n \times n$ square matrices,

$$\begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} M \\ 0 \end{bmatrix} B. \tag{A4}$$

Now consider the following:

$$\begin{bmatrix} M \\ M' \end{bmatrix} = \begin{bmatrix} A^{-1} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} D & 0 \\ 0 & 1 \end{bmatrix} B^{-1}. \tag{A5}$$

It is enough to construct S' with the rows of M'. □

Corollary A.5. Given a subspace V of dimension d and a set $S=\{\mathbf{v}_1,\ldots,\mathbf{v}_k\} \subset V$ which is linearly independent, there exists a set $S'=\{\mathbf{v}_{k+1},\ldots,\mathbf{v}_d\}$ such that $S \cup S'$ is a basis of V.

Proof. Select any basis of V and consider the components of vectors with respect to that basis as elements of $\mathbf{Z}_D^d$. □

We adopt the usual definition and notations for the scalar product and orthogonality.

Proposition A.6. The subset of $\mathbf{Z}_D^n$ orthogonal to a subspace V (that is, $V^\perp$) is a subspace. Moreover, $\dim V + \dim V^\perp = n$.

Proof. Let $d=\dim V$ and let M be a $d\times n$ matrix such that its rows form a basis of V. We use decomposition (A2) (again). For any $\mathbf{v} \in \mathbf{Z}_D^n$,

$$M\mathbf{v} = 0 \Leftrightarrow A^{-1} C B^{-1} \mathbf{v} = 0 \Leftrightarrow C(\mathbf{v}') = 0, \tag{A6}$$

with $\mathbf{v}=B\mathbf{v}'$. The set of the $\mathbf{v}'$'s verifying the equation is clearly a subspace of the expected dimension. □

APPENDIX B: CARDINALITY OF $C_d(D,n)$: GENERALIZED EULER'S TOTIENT FUNCTION

This appendix is devoted to the proof of Lemma II.2. We start with part 2, which can be rewritten as

$$\mathcal{N}C_1(D,n) = \mathcal{N}C_d(dD,n) \quad \forall\, D \geqslant 2, n \geqslant 1, d \geqslant 1. \tag{B1}$$

This is equivalent to the existence of a one-to-one mapping from $C_1(D,n)$ onto $C_d(dD,n)$. Consider the mapping

$$\mu : \mathbf{Z}^n \to \mathbf{Z}^n, \tag{B2}$$

$$\mathbf{v} \to d\mathbf{v}. \tag{B3}$$

μ induces a mapping $\bar{\mu} : \mathbf{Z}_D^n \to \mathbf{Z}_{dD}^n$, which is well defined and one-to-one because $x=y \pmod D \Leftrightarrow dx=dy \pmod{dD}$. From (14) we learn that $\forall \mathbf{v} \in \mathbf{Z}^n$,

$$d | \gcd(\dot{\bar{\mu}}(\bar{\mathbf{v}})),$$

where $\bar{\mathbf{v}}$ is the result of mapping $\mathbf{v}$ in $\mathbf{Z}_D^n$. Since for any $x \in \mathbf{Z}$ and $d' \in \mathrm{div}(D)$ we have $d'|x \Leftrightarrow d'd|dx$, it follows that

$$d\ \gcd(\bar{\mathbf{v}}) = \gcd(\bar{\mu}(\bar{\mathbf{v}})), \tag{B4}$$

which implies that $C_1(D,n)$ is mapped into $C_d(dD,n)$. Since for any element of $C_d(dD,n)$ there exists a suitable $\mathbf{v}$, the mapping is onto ■.

Now proving part 1 of the lemma is easy. Start with

$$\begin{aligned} \phi_n(D) &= D^n - \sum_{d \in \mathrm{div}(D)-\{1\}} \mathcal{N}C_d(D,n) \\ &= D^n - \sum_{d \in \mathrm{div}(D)-\{1\}} \phi_n\left(\frac{D}{d}\right). \end{aligned} \tag{B5}$$

Changing the index, we get a beautiful recursive relation,

$$\phi_n(D) = D^n - \sum_{d \in \mathrm{div}(D)-\{D\}} \phi_n(d). \tag{B6}$$

With some algebra on this expression it is possible to show that $\phi_n(D)$ is a multiplicative function: A function $f : \mathbf{N} \to \mathbf{N}$ is said to be multiplicative if $f(nm)=f(n)f(m)\ \forall\, n,m \in \mathbf{N}$ such that $\gcd(m,n)=1$. Thus, we only have to solve the recursion for $D=p^q$, p prime, but this poses no difficulty,

$$\phi_n(p^q) = p^{nq} - p^{(n-1)q}, \tag{B7}$$

from which (17) follows.

Part 3 of the lemma is merely the recursion relation just constructed.

APPENDIX C: GENERATORS OF $\mathcal{P}_S$

In order to study $\mathcal{P}_S$, it is preferable to consider its elements as permutations over $\mathbf{Z}_D^n \times \mathbf{Z}_D^n$, and so we change the notation

$$\mathbf{x} \to \pi(\mathbf{x}), \quad \mathbf{x} \in \mathbf{Z}_D^{2n} \tag{C1}$$

for

$$(\mathbf{i},\mathbf{j}) \to (\mu(\mathbf{i},\mathbf{j}), \nu(\mathbf{i},\mathbf{j})), \quad \mathbf{i},\mathbf{j} \in \mathbf{Z}_D^n, \tag{C2}$$

where the correspondence is the same as in (11).

Lemma C.1. $\mathcal{P}_S$ is generated by its following elements:

$$\pi_+^1, \quad \text{with } \mu(\mathbf{i},\mathbf{j}) = \mathbf{i},$$

$$\nu(\mathbf{i},\mathbf{j}) = (i_1 + j_1, j_2, \ldots, j_n);$$

$$\pi_{\text{xch}}^1, \quad \text{with } \mu(\mathbf{i},\mathbf{j}) = (j_1, i_2, \ldots, i_n),$$

$$\nu(\mathbf{i},\mathbf{j}) = (-i_1, j_2, \ldots, j_n);$$

$$\pi_+^{12}, \quad \text{with } \mu(\mathbf{i},\mathbf{j}) = (i_1 + i_2, i_2, \ldots, i_n),$$

$$\nu(\mathbf{i},\mathbf{j}) = (j_1, j_2 - j_1, \ldots, j_n);$$

$$\pi_{\text{swap}}^{lm}, \quad \text{with } \mu(\mathbf{i},\mathbf{j}) = (\ldots, i_{l-1}, i_m, i_{l+1}, \ldots, i_{m-1}, i_l, i_{m+1}, \ldots),$$

$$\nu(\mathbf{i},\mathbf{j}) = (\ldots, j_{l-1}, j_m, j_{l+1}, \ldots, j_{m-1}, j_l, j_{m+1}, \ldots), \quad l,m = 1, \ldots, n.$$

Proof. In order to prove this, first let us define

$$\pi_+^i := \pi_{\text{swap}}^{1i} \circ \pi_+^1 \circ \pi_{\text{swap}}^{1i}, \tag{C3}$$

$$\pi_{\text{xch}}^i := \pi_{\text{swap}}^{1i} \circ \pi_{\text{xch}}^1 \circ \pi_{\text{swap}}^{1i}, \tag{C4}$$

$$\pi_+^{ij} := \pi_{\text{swap}}^{1i} \circ \pi_{\text{swap}}^{2j} \circ \pi_+^{12} \circ \pi_{\text{swap}}^{1i} \circ \pi_{\text{swap}}^{2j}, \tag{C5}$$

with $i,j = 1, \ldots, n$. Consider any $p \in \mathcal{P}_S$; our goal is to act from the left and from the right with these permutations until we get the identity which is equivalent to the statement of the lemma. We shall use the matrix representation of the permutations, and the process is a suitable Gaussian elimination similar to the one used in Appendix A. The difference is that now we cannot perform freely any sum of lines or columns, but only those which have associated a permutation in the above set.

To work around this problem, in place of (A1) we consider

$$(x,y) \xrightarrow{\mathcal{O}_1} (x, x+y), \quad (x,y) \xrightarrow{\mathcal{O}_3(e)} (ey, x), \quad e = \pm 1. \tag{C6}$$

Since $\mathcal{O}_2$ can be constructed suitably combining $\mathcal{O}_1$ and $\mathcal{O}_3(-1)$, only the case $e=1$ is really different, but adapting the algorithm is straightforward. The point is that π_+^i and π_+^{ij} can be attached to $\mathcal{O}_1$, π_{xch}^i to $\mathcal{O}_3(-1)$, and π_{swap}^{ij} to $\mathcal{O}_3(1)$, with care in the case of π_+^{ij} and π_{swap}^{ij} for their additional effects.

To perform the elimination in an element of $\mathcal{P}_S$ with associated matrix M, start working over the first column (permutations act thereby from the left). Using p_+^i and p_{ex}^i, make zero the elements $M_{n+1,1}$ to $M_{2n,1}$, and then use p_+^{1i} and p_{swap}^{1i} until just M_{11} is nonzero in the first column. The process must be repeated for the first row (this time permutations act from the right). Now let us deal with the second column, first making zero the elements $M_{n+2,2}$ to $M_{2n,2}$, and afterward the elements $M_{3,2}$ to $M_{n,2}$. Do the same for the second row. The process must be carried out for the first n rows and columns, until we get something of the form

$$\begin{bmatrix} D & T_1 \\ T_2 & M \end{bmatrix}, \tag{C7}$$

with D diagonal, T_1 lower triangular, and T_2 upper triangular. But in fact applying the condition (28) forces $T_1 = T_2 = 0$ and $M = D^{-1}$. Now we note that

$$\begin{bmatrix} D & 0 \\ 0 & D^{-1} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -D & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ D-1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}. \tag{C8}$$

Since the matrices in the right side are trivially constructed with the given set of generators, this ends the proof. □

APPENDIX D: PROOF OF THEOREM III.1

(This proof uses the notation and results of Appendix C.)

(1) Since $\mathcal{P}_T$ is invariant in the product group trivially, we prove both sides of the inclusion, starting with $\mathcal{P}_{\text{loc}} \subset \mathcal{P}_T \ltimes \mathcal{P}_S$.

Later we define local unitary operators implementing $\mathcal{P}_T$ [see (D7)], and so we just bother about those $U \in \mathcal{U}_{B\ \text{loc}}(D,n)$ that leave $|0\ 0\rangle_{\mathcal{B}}\langle 0\ 0|$ invariant. Moreover, as the global phase is unimportant, we select for the analysis those operators for which $U|0\ 0\rangle_{\mathcal{B}} = |0\ 0\rangle_{\mathcal{B}}$. But U is local, and then these constraints are equivalent to $U = U_A \otimes U_A^*$ (conjugation with respect to the computational basis). We can thus write

$$U|\mathbf{ij}\rangle = \sum_{\mathbf{kl}} \mathcal{S} A_{\mathbf{ki}} A_{\mathbf{lj}}^* |\mathbf{kl}\rangle, \quad \sum_{\mathbf{k}} \mathcal{S} A_{\mathbf{ik}} A_{\mathbf{jk}}^* = \delta(\mathbf{i}-\mathbf{j}) \tag{D1}$$

with A a unitary matrix on a single party (Alice or Bob). Using (5), we get

$$U|\mathbf{ij}\rangle_{\mathcal{B}} = \sum_{\mathbf{klmn}} \mathcal{S}\, \varphi(\mathbf{k}\cdot\mathbf{i} - \mathbf{m}\cdot\mathbf{l}) A_{\mathbf{lk}} A_{\mathbf{l-n,k-j}}^* |\mathbf{mn}\rangle_{\mathcal{B}}. \tag{D2}$$

On the other hand, the action of U involves a permutation over Bell states,

$$U|\mathbf{ij}\rangle_{\mathcal{B}} = \phi(\mathbf{i},\mathbf{j}) |\mu(\mathbf{i},\mathbf{j})\, \nu(\mathbf{i},\mathbf{j})\rangle_{\mathcal{B}}. \tag{D3}$$

Identifying both expressions (Bell states are orthogonal),

$$\phi(\mathbf{i},\mathbf{j})\,\delta(\mathbf{m}-\mu(\mathbf{i},\mathbf{j}))\,\delta(\mathbf{n}-\nu(\mathbf{i},\mathbf{j})) = \sum_{\mathbf{kl}} \mathcal{S}\varphi(\mathbf{k}\cdot\mathbf{i}-\mathbf{l}\cdot\mathbf{m}) A_{\mathbf{lk}} A^{*}_{\mathbf{l-n,k-j}}. \quad \text{(D4)}$$

Act on both sides of this equation with the operator $\mathcal{S}_{\mathbf{mn}}\varphi(\mathbf{r}\cdot\mathbf{m})A_{\mathbf{r-n,s-j}}$ to obtain

$$A_{\mathbf{r}-\nu(\mathbf{i},\mathbf{j}),\mathbf{s-j}} = \varphi(\mathbf{s}\cdot\mathbf{i}-\mathbf{r}\cdot\mu(\mathbf{i},\mathbf{j}))\,\phi(\mathbf{i},\mathbf{j})^{*} A_{\mathbf{r,s}}. \quad \text{(D5)}$$

Choose any $\mathbf{r},\mathbf{s}$ such that $A_{\mathbf{rs}} \neq 0$, interpret this equation as a recurrence relation, and consider the commutative diagram

$$\begin{array}{ccc} A_{\mathbf{rs}} & \rightarrow & A_{\mathbf{r}-\nu(\mathbf{i},\mathbf{j}),\mathbf{s-j}} \\ \downarrow & & \downarrow \\ A_{\mathbf{r}-\nu(\mathbf{i}',\mathbf{j}'),\mathbf{s-j}'} & \rightarrow & A_{\mathbf{r}-\nu(\mathbf{i},\mathbf{j})-\nu(\mathbf{i}',\mathbf{j}'),\mathbf{s-j}-\mathbf{j}'}. \end{array}$$

Switching again to the $\mathbf{Z}_D^{2n}$ notation, the commutation condition is

$$\mathbf{x}^t \Omega \mathbf{x}' = \pi(\mathbf{x})^t \Omega\, \pi(\mathbf{x}') \quad \text{(D6)}$$

for any $\mathbf{x},\mathbf{x}' \in \mathbf{Z}_D^{2n}$. Thus, $\pi \in \mathcal{P}_{\mathrm{S}}$.

We now show that $\mathcal{P}_{\mathrm{T}} \ltimes \mathcal{P}_{\mathrm{S}} \subset \mathcal{P}_{\mathrm{loc}}$. Thanks to lemma C.1, it is enough to construct a few permutations by means of unitary local operators. We start with translations. Choosing

$$U_{\mathrm{A}}|\mathbf{i}\rangle = \varphi(\mathbf{i}\cdot\mathbf{a})|\mathbf{i}\rangle, \quad U_{\mathrm{B}}|\mathbf{i}\rangle = |\mathbf{i}-\mathbf{b}\rangle, \quad \text{(D7)}$$

with $\mathbf{a},\mathbf{b} \in \mathbf{Z}_D^n$, the effect is

$$\mu(\mathbf{i},\mathbf{j}) = \mathbf{i}+\mathbf{a}, \quad \nu(\mathbf{i},\mathbf{j}) = \mathbf{j}+\mathbf{b}.$$

This is not the only subgroup easily generated. We have also that

$$U_{\mathrm{A}}|\mathbf{i}\rangle = |S\mathbf{i}\rangle, \quad U_{\mathrm{B}}|\mathbf{i}\rangle = |S\mathbf{i}\rangle, \quad \text{(D8)}$$

with $S \in \mathbf{Z}_D^n \times \mathbf{Z}_D^n$ and invertible give

$$\mu(\mathbf{i},\mathbf{j}) = (S^t)^{-1}\mathbf{i}, \quad \nu(\mathbf{i},\mathbf{j}) = S\mathbf{j}.$$

Both of these results can be checked with a few manipulations in (10). Since π_{swap}^{lm} is physically trivial and π_{+}^{lm} is contained in the last construction, the only permutations we have not still covered are π_{+}^{1} and π_{xch}^{1}, but as these involve only the first pair of qudits we can fix $n=1$ in (D5) and try the ansatz $A_{ij}=\varphi(\eta(i,j))$. Working modulo D, this results in

$$\eta(r-\nu(i,j),s-j) = \eta(r,s) + si - r\mu(i,j) - \tilde{\phi}(i,j), \quad \text{(D9)}$$

where $\varphi(\tilde{\phi}(i,j)) := \phi(i,j)$. Solutions to this equation require η to be a second-order polynomial, limiting the permutation to

$$\mu(i,j) = aj + b\nu(i,j), \quad \nu(i,j) = -a^{-1}(i+cj), \quad \text{(D10)}$$

where $a,b,c \in \mathbf{Z}_D$, a invertible. A compatible choice for η is

$$\eta(i,j) = aij + \frac{b}{2}i^2 + \frac{c}{2}j^2. \quad \text{(D11)}$$

The permutations we were searching for belong to the set of (D10).

(2) In order to prove the second part of the theorem, it is enough to analyze which are the realizations of the identity permutation. Going back to (D5) and fixing $\mu(\mathbf{i},\mathbf{j})=\mathbf{i}$ and $\nu(\mathbf{i},\mathbf{j})=\mathbf{j}$, we find the equation

$$A_{\mathbf{r-j,s-j}} = \varphi(\mathbf{s}\cdot\mathbf{i}-\mathbf{r}\cdot\mathbf{i})\,\phi(\mathbf{i},\mathbf{j})^{*} A_{\mathbf{r,s}}. \quad \text{(D12)}$$

Modulo a global phase (30), the solutions are exactly of the form

$$\phi(\mathbf{i},\mathbf{j}) = \varphi(\mathbf{a}\cdot\mathbf{i}+\mathbf{b}\cdot\mathbf{j}), \quad \text{(D13)}$$

$$A_{\mathbf{rs}} = \varphi(\mathbf{b}\cdot\mathbf{s})\,\delta(\mathbf{s}-\mathbf{r}-\mathbf{a}), \quad \text{(D14)}$$

where $\mathbf{a},\mathbf{b} \in \mathbf{Z}_D^n$. But this is $U_{\mathbf{x}}$ in (30) with

$$\mathbf{x} = (b_1,\ldots,b_n,-a_1,\ldots,-a_n). \quad \text{(D15)}$$

APPENDIX E: ORDER OF $\mathcal{P}_{\mathrm{S}}$

In this Appendix, we offer a proof of Theorem III.2.

We first note that, except for a sign, $\mathbf{u}_i$ and $\mathbf{v}_i$ play interchangeable roles. Thus it is enough to consider a case with $t=0$ (if $t \geqslant 1$, suitable exchanges between $\mathbf{u}$'s and $\mathbf{v}$'s and sign adjustments will be enough). We shall consider two cases separately, depending on whether $r<n$. In both cases the target is to find out in how many ways a new vector can be included in the set. Such a vector must fulfill (31) and be linearly independent with respect to the initial set.

Suppose $r<n$. We would like to know how many vectors can take the role of $\mathbf{u}_{r+1}$. Let $S=\{\mathbf{u}_1,\ldots,\mathbf{u}_r,\mathbf{v}_1,\ldots,\mathbf{v}_s\}$ and $V=\mathrm{Lin}\{\Omega\mathbf{v}:\mathbf{v}\in S\}$. From (31) we have $\mathbf{u}_{r+1} \in V^{\perp}$ (and no further conditions), so let $S'=\{\mathbf{u}_{s+1},\ldots,\mathbf{u}_r,\mathbf{w}_1,\ldots,\mathbf{w}_{2(n-r)}\}$ be a basis of $V^{\perp}$. We claim that $S \cup S'$ is LI, that is, the equation

$$\sum_{i=1}^{r-s} a_i \mathbf{u}_{s+j} + \sum_{i=1}^{2(n-r)} b_i \mathbf{w}_i + \sum_{i=1}^{s} (c_i \mathbf{u}_i + d_i \mathbf{v}_i) = 0$$

holds only if all the scalars are zero. This is so because taking the scalar product with $\Omega\mathbf{u}_k (k \leqslant s)$ we get $d_k=0$; using $\Omega\mathbf{v}_k$, $c_k=0$. The rest of the scalars must be zero because S' is LI. Therefore, there are $D^{r-s}\phi_{2(n-r)}(D)$ suitable vectors, since we can choose any combination of the form

$$\mathbf{u}_{r+1} = \sum_{i=1}^{r-s} a_i \mathbf{u}_{s+j} + \sum_{i=1}^{2(n-r)} b_i \mathbf{w}_i$$

for which $\gcd(\{b_1,\ldots,b_{2(n-r)}\})=1$ [this is why the factor $\phi_{2(n-r)}(D)$ appears, see (16)].

Now suppose $r=n$, $s<n$. We pursue $\mathbf{v}_{s+1}$. Let $S=\{\mathbf{u}_1,\ldots,\mathbf{u}_s,\mathbf{u}_{s+2},\ldots,\mathbf{u}_n,\mathbf{v}_1,\ldots,\mathbf{v}_s\}$ and $V=\mathrm{Lin}\{\Omega\mathbf{v}:\mathbf{v}\in S\}$. From (31) we have $\mathbf{v}_{s+1} \in V^{\perp}$ and $\mathbf{u}_{s+1}^t \Omega \mathbf{v}_{s+1}=1$. Let $S'=\{\mathbf{u}_{s+1},\ldots,\mathbf{u}_n,\mathbf{w}\}$ be a base of $V^{\perp}$ and let $s := \mathbf{u}_{s+1}^t \Omega \mathbf{w}$. We first show that s is invertible. Let $V'=\mathrm{Lin}\{\Omega\mathbf{u}_1,\ldots,\Omega\mathbf{u}_n,\Omega\mathbf{v}_1,\ldots,\Omega\mathbf{v}_s\}$. If s is noninvertible, choose any $k \neq 0$ such that $ks=0$. Then $k\mathbf{u}_{s+1}^t\Omega\mathbf{w}=0$ implies $k\mathbf{w} \in V'^{\perp}=\mathrm{Lin}\{\mathbf{u}_{s+1},\ldots,\mathbf{u}_n\}$, but this is not possible because S' is linearly independent. We now show that $S \cup \{\mathbf{u}_{s+1},\mathbf{w}\}$ is LI. If it is not, then $\mathbf{w} \in \mathrm{Lin}(S \cup \{\mathbf{u}_{s+1}\})$, but this in turn implies $\mathbf{u}_{s+1}^t \Omega \mathbf{w}=0$, which again is false. Therefore, there are

D^{n-s} suitable vectors, since we can choose any of the following combinations:

$$\mathbf{v}_{s+1} = s^{-1}\mathbf{w} + \sum_{i=1}^{n-s} c_i \mathbf{u}_{s+i}.$$

With this, part 1 of the theorem is proved. For part 2, it only remains to count. There are $\phi_{2n}(D)$ possible values for $\mathbf{u}_1$. If $\mathbf{u}_1$ is fixed, there are $D\phi_{2(n-1)}(D)$ possible elections for $\mathbf{u}_2$. Continuing this way, one gets the desired result.

APPENDIX F: RECURSION RELATIONS FOR DISTILLATION PROTOCOLS

In this appendix, we derive an expression for the final state of the remaining pairs of qudits when the procedure of Sec. V has been successfully performed. We will use the same notation found there.

So let us define for $\mathbf{x},\mathbf{y} \in \mathbf{Z}_D^{2n}$

$$\rho_{\mathbf{xy}}^{(n)} := {}_{\mathcal{B}}\langle \mathbf{x}|\rho^{(n)}|\mathbf{y}\rangle_{\mathcal{B}}, \tag{F1}$$

from which $p_{\mathbf{x}} = \rho_{\mathbf{xx}}^{(n)}$. After the permutation with associated matrix M and phase function ϕ, the state is

$$\rho^{(n)\prime} := \sum_{\mathbf{x},\mathbf{y}\in\mathbf{Z}_D^{2n}} \phi(\mathbf{x})\phi^*(\mathbf{y})\rho_{\mathbf{xy}}^{(n)} |M\mathbf{x}\rangle_{\mathcal{B}}\langle M\mathbf{y}|. \tag{F2}$$

The measurement is performed in the computational basis (for the last $n-m$ pairs), and the rest of the pairs are kept only if this measurement coincides for each of the measured pairs (if Alice measures $|3\rangle$, so does Bob for the corresponding qudit). Going back to (5), this means that j is zero for each of the pairs. Therefore, after the measurement and taking the partial trace over the measured pairs, the state of the first m pairs is

$$\rho^{(m)} = \frac{1}{P} \sum_{\mathbf{k}\in\mathbf{Z}_D^{n-m}} {}_{\mathcal{B}}\langle \mathbf{k0}|\rho^{(n)\prime}|\mathbf{k0}\rangle_{\mathcal{B}}, \tag{F3}$$

where the Bell states must be understood to belong to the space of the last $n-m$ pairs and P is the probability of having obtained the suitable measurement. Calculating it amounts to taking the total trace,

$$P = \sum_{\mathbf{x}\in\mathbf{Z}_D^{2m}} \sum_{\mathbf{k}\in\mathbf{Z}_D^{n-m}} ({}_{\mathcal{B}}\langle \mathbf{x}| \otimes {}_{\mathcal{B}}\langle \mathbf{k0}|)\rho^{(n)\prime}(|\mathbf{x}\rangle_{\mathcal{B}} \otimes |\mathbf{k0}\rangle_{\mathcal{B}}).$$

Inserting definition (F2), we get (43).

The state of the system before the measurement can also be expressed,

$$\rho^{(n)\prime} = \sum_{\mathbf{x},\mathbf{y}\in\mathbf{Z}_D^{2n}} \phi(M^{-1}\mathbf{x})\phi^*(M^{-1}\mathbf{y})\rho^{(n)}_{M^{-1}\mathbf{x},M^{-1}\mathbf{y}} |\mathbf{x}\rangle_{\mathcal{B}}\langle \mathbf{y}|.$$

Inserting this expression in (F3),

$${}_{\mathcal{B}}\langle \mathbf{x}|\rho^{(m)}|\mathbf{y}\rangle_{\mathcal{B}} = \frac{1}{P} \sum_{\mathbf{k}\in\mathbf{Z}_D^{n-m}} \phi(M^{-1}(\hat{\mathbf{k}}+\bar{\mathbf{x}}))\phi^*(M^{-1}(\hat{\mathbf{k}}+\bar{\mathbf{y}}))\rho^{(n)}_{M^{-1}(\hat{\mathbf{k}}+\bar{\mathbf{x}}),M^{-1}(\hat{\mathbf{k}}+\bar{\mathbf{y}})}, \tag{F4}$$

where $\mathbf{x},\mathbf{y} \in \mathbf{Z}_D^{2m}$, $\bar{\mathbf{x}}$ and $\bar{\mathbf{y}}$ are defined as in (45), and

$$\hat{\mathbf{k}} := (\underbrace{0,\ldots,0}_{m},k_1,\ldots,k_{n-m},\underbrace{0,\ldots,0}_{n}).$$

With the definition for V_M given in Sec. V, we have

$${}_{\mathcal{B}}\langle \mathbf{x}|\rho^{(m)}|\mathbf{y}\rangle_{\mathcal{B}} = \frac{1}{P} \sum_{\mathbf{z}\in V_M} \phi(\Omega\mathbf{z}+M^{-1}\bar{\mathbf{x}})\phi^*(\Omega\mathbf{z}+M^{-1}\bar{\mathbf{y}})\rho^{(n)}_{\Omega\mathbf{z}+M^{-1}\bar{\mathbf{x}},\Omega\mathbf{z}+M^{-1}\bar{\mathbf{y}}}.$$

Equation (44) follows setting $\mathbf{x}=\mathbf{y}$.

APPENDIX G: MONTE CARLO MEASURING

We introduce a suitable metric in the space of Bell diagonal states in order to perform several measures. For simplicity, we have chosen the metric induced by mapping physical states into Euclidean space taking the eigenvalues as coordinates.

We have chosen a Monte Carlo approach to perform the measurements. This approach consists in randomly generating points of the space according to the measure on that space, and counting how many of them are inside the measured set.

In our case, numerically implementing such a measure is not difficult if fidelity is not low. Consider a Bell diagonal state of fidelity F. There are D^2-1 free coordinates (eigenvalues) λ_i subject to the constraints

$$0 \leqslant \lambda_i \leqslant F, \quad \sum_i \lambda_i = 1 - F. \tag{G1}$$

The random generation is achieved as follows. We take for each point D^2-2 real random variables x_i uniformly distributed in $[0,1]$ $(i=1,\ldots,D^2-2)$. Defining $x_0 := 0$ and $x_{D^2-1} := 1$, we set $\lambda_i = (1-F)(x_{i+1}-x_i)$. If $\lambda_i > F$ for any i, we simply discard the point. Otherwise, it belongs to the space of interest. Then it is checked whether it belongs to the measured set by running the proper algorithm. For example, if we are checking distillability through a given protocol, this is the moment were the protocol is numerically simulated until the point converges.

[1] M. Lewenstein, D. Bruss, J. I. Cirac, B. Kraus, M. Kus, J. Samsonowicz, A. Sanpera, and R. Tarrach, J. Mod. Opt. **47**, 2841 (2000).

[2] D. Bruss, J. I. Cirac, P. Horodecki, F. Hulpke, B. Kraus, M. Lewenstein, and A. Sanpera, J. Mod. Opt. **49**, 1399 (2002).

[3] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).

[4] A. Galindo and M. A. Martin-Delgado, Rev. Mod. Phys. **74**, 347 (2002).
[5] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, Phys. Rev. Lett. **76**, 722 (1996).
[6] Ch. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Phys. Rev. A **54**, 3824 (1996).
[7] D. Deutsch, A. Ekert, R. Jozsa, C. Macchiavello, S. Popescu, and A. Sanpera, Phys. Rev. Lett. **77**, 2818 (1996).
[8] C. Macchiavello, Phys. Lett. A **246**, 385 (1998).
[9] M. Horodecki and P. Horodecki, Phys. Rev. A **59**, 4206 (1999).
[10] J. Dehaene, M. Van den Nest, B. de Moor, and F. Verstraete, Phys. Rev. A **67**, 022310 (2003).
[11] K. G. H. Vollbrecht and M. M. Wolf, Phys. Rev. A **67**, 012303 (2003).
[12] P. Horodecki, Phys. Lett. A **232**, 333 (1997).
[13] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. Lett. **80**, 5239 (1998).
[14] P. Horodecki, M. Horodecki, and R. Horodecki, Phys. Rev. Lett. **82**, 1056 (1999).
[15] W. Dür, J. I. Cirac, M. Lewenstein, and D. Bruss, Phys. Rev. A **61**, 062313 (2000).
[16] D. P. DiVincenzo, P. W. Shor, J. A. Smolin, B. M. Terhal, and A. V. Thapliyal, Phys. Rev. A **61**, 062312 (2000).
[17] R. F. Werner, Phys. Rev. A **40**, 4277 (1989).
[18] G. Alber, A. Delgado, N. Gisin, and I. Jex, e-print quant-ph/000802.
[19] G. Alber, A. Delgado, N. Gisin, and I. Jex, J. Phys. A **34**, 8821 (2001).
[20] M. A. Martin-Delgado and M. Navascues, Eur. Phys. J. D **27**, 169 (2003).
[21] M. A. Martin-Delgado and M. Navascues, Phys. Rev. A **68**, 012322 (2003).
[22] J. H. Conway and R. K. Guy, Euler's Totient Numbers, in *The Book of Numbers* (Springer-Verlag, New York, 1996), pp. 154–156.