Article

Optimal High-Dimensional Entanglement Concentration for Pure Bipartite Systems

Lukas Palma Torres [1], **Miguel Ángel Solís-Prosser** [1,*], **Omar Jiménez** [2], **Esteban S. Gómez** [3,4] **and Aldo Delgado** [3,4]

1 Departamento de Ciencias Físicas, Facultad de Ingeniería y Ciencias, Universidad de La Frontera, Avenida Francisco Salazar 01145, Temuco 4811230, Chile; l.palma03@ufromail.cl
2 Centro de Óptica e Información Cuántica, Facultad de Ciencias, Universidad Mayor, Camino La Pirámide 5750, Huechuraba, Santiago 8580745, Chile; omar.jimenez@umayor.cl
3 Departamento de Física, Universidad de Concepción, Casilla 160-C, Concepción 4070043, Chile; estesepulveda@udec.cl (E.S.G.); aldelgado@udec.cl (A.D.)
4 Millennium Institute for Research in Optics, Universidad de Concepción, Casilla 160-C, Concepción 4070043, Chile
* Correspondence: miguel.solis@ufrontera.cl

Abstract: Considering pure quantum states, entanglement concentration is the procedure where, from N copies of a partially entangled state, a single state with higher entanglement can be obtained. Obtaining a maximally entangled state is possible for $N = 1$. However, the associated success probability can be extremely low when increasing the system's dimensionality. In this work, we study two methods to achieve a probabilistic entanglement concentration for bipartite quantum systems with a large dimensionality for $N = 1$, regarding a reasonably good probability of success at the expense of having a non-maximal entanglement. Firstly, we define an efficiency function $\mathcal{Q}$ considering a tradeoff between the amount of entanglement (quantified by the I-Concurrence) of the final state after the concentration procedure and its success probability, which leads to solving a quadratic optimization problem. We found an analytical solution, ensuring that an optimal scheme for entanglement concentration can always be found in terms of $\mathcal{Q}$. Finally, a second method was explored, which is based on fixing the success probability and searching for the maximum amount of entanglement attainable. Both ways resemble the Procrustean method applied to a subset of the most significant Schmidt coefficients but obtaining non-maximally entangled states.

Keywords: entanglement concentration; Schmidt number

1. Introduction

Quantum entanglement is the most known, remarkable, and useful quantum resource in the quantum information (QI) theory [1] as it underlies several QI protocols, such as dense coding [2], entanglement swapping [3], quantum teleportation [4], and quantum cryptography [5]. For instance, in the bipartite scenario, two users who want to communicate—usually called Alice and Bob—can share an entangled state [6]. In this case, the ability to transmit information encoded in the state shared by Alice and Bob depends on the amount of entanglement [7,8]. Moreover, the most favorable case for faithful communication is when Alice and Bob share a maximally pure entangled state (MES) [9]. However, even if it was the initial state, the quantum noisy channel used to send the information will produce a loss of correlations in the MES [10]. Moreover, the quantum operations needed to carry out a particular quantum application are performed imperfectly due to the experimental errors, yielding to fidelities of less than one [11].

In such cases where they have access only to a partially entangled state ρ, it is desirable to access a channel that allows a more faithful way to send quantum information. One solution is to implement protocols to increase the amount of entanglement [12,13]. These

protocols are known as entanglement purification or entanglement distillation [14–16] and entanglement concentration [17]. These methods are based on the fact that local operations and classical communication between Alice and Bob cannot increase, on average, the amount of entanglement in the initially entangled pairs [18].

In the case of entanglement purification, the goal is to increase the purity and the entanglement in the initial state ρ, but under the cost of reducing the number of initial copies available. It can only be implemented successfully in a probabilistic way [14]. Moreover, an experimental realization of entanglement purification was carried out for mixed states of polarization-entangled photons using linear optics [19].

In the entanglement concentration, the process considers the cases where the initial partially entangled state is pure [20,21]. Indeed, there are two ways to implement entanglement concentration: the Procrustean method and the Schmidt projection method [17,20,21]. The Procrustean method is easier to implement than the Schmidt projection method because the initial partially entangled state is known. The entanglement concentration procedure is carried out by local filtering onto individual pairs of the initial state [17]. In the Schmidt method, however, the process of entanglement concentration is implemented in at least two unknown partially entangled states through collective simultaneous measurements onto the particles [22]. Thus, schemes for carrying out the entanglement concentration have been proposed for the Procrustean [23] and the Schmidt method [24,25]. Moreover, its experimental implementation has been achieved in the case of the Procrustean method [26] and for the Schmidt method [22] using partially polarization-entangled photons.

The entanglement concentration can also be classified as deterministic [12,27,28] as well as probabilistic [11,13,20,29]. In the deterministic case, the process has a probability equal to one to be successfully implemented in the regimes of few copies or in the asymptotic limit of infinite copies [30]. In this scheme, the quantum circuits to carry out deterministic entanglement concentration have been proposed [31] and recent experimental efforts demonstrate its feasibility [32–35]. In these experimental works, the copies are replaced by additional degrees of freedom of the same pair of photons, which improves the possibility of short-term implementations of entanglement distillation for technological purposes. On the other hand, in the probabilistic entanglement concentration, the process is achieved with a probability of less than one and has been experimentally implemented [36]. Moreover, the relation in the asymptotic limit between the entanglement concentration in a deterministic and probabilistic way was studied [30]. They found that these methods are equivalent considering many copies of the initial state: the error probability for the probabilistic method goes to zero quickly with the number of copies. In addition, the entanglement concentration is generally studied considering two entangled quantum states, but has also been studied for the case of tripartite correlated systems [37,38].

In this work, we studied the probabilistic entanglement concentration in the bipartite scenario of a pure two-qudit (D-dimensional) state. Considering a large dimensionality ($D \gg 2$), we study two methods to achieve entanglement concentration regarding a reasonably good probability of success at the expense of having a non-maximal entanglement. At first glance, we consider a tradeoff between the amount of entanglement of the state after the concentration procedure and its success probability, quantified by the payoff function $\mathcal{Q}$. This figure of merit leads to analytically solving a quadratic optimization problem, ensuring that an optimal scheme for entanglement concentration can always be found in terms of $\mathcal{Q}$. Then, a second method was studied, where we fixed the success probability and searched for the maximum amount of entanglement attainable in this case. We found that both ways resemble the Procrustean method applied to a subset of the most significant Schmidt coefficients without the constraint of obtaining a MES. We envisage the usefulness of these methods in entanglement-based quantum communication and also for device-independent protocols where high-dimensional entangled states are required with a certain amount of entanglement, such as randomness certification, expansion and self-testing [39–41].

2. Revisiting Entanglement Concentration

Throughout this work, we will limit ourselves to the case of entanglement concentration from a single copy of a two-qudit non-maximally entangled pure state. This state will be given by

$$|\Phi\rangle_{12} = \sum_{m=1}^{D} a_m |m\rangle_1 |m\rangle_2, \tag{1}$$

where a_m are positive coefficients such that $\sum_m a_m^2 = 1$. The set of states $\{|m\rangle_1|m\rangle_2\}_{m=1}^{D}$ can be regarded as the Schmidt basis for the entangled state $|\Phi\rangle_{12}$, and, therefore, a_m will be the respective Schmidt coefficients. In order to quantify the entanglement conveyed by $|\Phi\rangle_{12}$, the I-Concurrence [42] can be used, which is given by

$$\mathcal{C}(|\Phi\rangle_{12}) = \sqrt{\frac{D}{D-1}\left(1 - \mathrm{tr}(\rho_1^2)\right)}$$

$$= \sqrt{\frac{D}{D-1}\left(1 - \sum_{m=1}^{D} a_m^4\right)}, \tag{2}$$

where ρ_1 is the reduced density matrix of one of the qudits. This function fulfills the necessary conditions an entanglement measure needs to satisfy [43]. Its minimum value is zero, and its maximum is one, which arises when $|\Phi\rangle_{12}$ is a product state and a maximally entangled state, respectively. This document will refer to $\mathcal{C}$ simply as entanglement. Another function widely used to assess entanglement is the Schmidt number [44–51], defined as

$$K(|\Phi\rangle_{12}) = \frac{1}{\mathrm{tr}(\rho_1^2)} = \left[\sum_{m=1}^{D} a_m^4\right]^{-1}. \tag{3}$$

It is straightforward to see that $\mathcal{C}(|\Phi\rangle_{12})$ and $K(|\Phi\rangle_{12})$ are closely related, as both depend on $\mathrm{tr}(\rho_1^2)$.

As we mentioned above, it is well known the correlated state given in Equation (1) can have its entanglement increased through an entanglement concentration procedure [7,13,30,52,53]. This process is, in general, a probabilistic one [54]. We will follow the next approach to show the concentration scheme. Assuming we have an ancillary qubit initially prepared in state $|0\rangle_a$, it can be used for concentration through a unitary bipartite operation U_{a1} acting over the ancilla and one of the qudits. Let

$$U_{a1} \otimes \mathbb{I}_2 |0\rangle_a |\Phi\rangle_{12} = |0\rangle_a A_\mathrm{S} |\Phi\rangle_{12} + |1\rangle_a A_\mathrm{F} |\Phi\rangle_{12}, \tag{4}$$

where $|\mu\rangle_a$ is the state of the ancilla which flags whether concentration was accomplished ($\mu = 0$) or not ($\mu = 1$). A_S and A_F are Kraus operators acting on qudit 1, modifying the entangled state in each case. A measurement on the ancilla announces if we succeeded. Through this work, we will be only concerned with the successful cases, whose study can be simplified considering $A_\mathrm{S}|\Phi\rangle_{12}$ only. Without loss of generality, we may write

$$A_\mathrm{S}|\Phi\rangle_{12} = \sqrt{p_\mathrm{S}}|\Psi\rangle_{12}, \tag{5}$$

where p_S is the probability of success for the concentration procedure, and $|\Psi\rangle_{12}$ is the resulting state; therefore, we obtain $\mathcal{C}(|\Psi\rangle_{12}) > \mathcal{C}(|\Phi\rangle_{12})$. If the intention is to obtain a MES, it is known that $p_\mathrm{S} = D a_\mathrm{min}^2$, where $a_\mathrm{min}^2 = \min\{|a_m|^2\}$ [7,53]. This probability, however, may adopt very small values if the Schmidt coefficients exhibit large differences among them, rendering the procedure inefficient.

Alternatively, one may increase the success probability at expense of having a partially entangled state as result. In Ref. [20], Vidal studied the case of transforming Schmidt

coefficients $\{a_m\}$ onto a given set $\{b_m\}$ and showed the optimal probability of success for such map. In this way, one may choose the b_m coefficients in such a way that the success probability is *good enough* at the same time the entanglement is increased. Another possibility is to set the resulting state $|\Psi\rangle_{12}$ as a maximally entangled one *for a subspace* of dimension $N \leqslant D$, which is analogous to a Procrustean method (i.e., cutting off extra probabilities from a given reference value [14]) applied only on a subset of the original Schmidt coefficients [13]. Both approaches, however, force one to constrain the final state to be a given one. Thus, the problem contains D arbitrary parameters b_m, and one has to search thoroughly for a convenient combination of the b_m.

A possible way to decrease the number of free parameters is to use the Kraus operator $A_{\mathrm{S}}(\xi)$ given in Ref. [55]. This approach allows to interpolate between the initial Schmidt coefficients (a_m) and the ones from a maximally entangled state ($1/\sqrt{D}$) using a single parameter ξ. Thus, we may transform $a_m \to b_m(\xi)$, where $0 \leqslant \xi \leqslant 1$, and

$$b_m^2(\xi) = a_m^2 + \left(\frac{1}{D} - a_m^2 \right) \xi.$$ (6)

It can be seen that Equation (6) shows a transformation that preserves the norm of the new state and represents a linear interpolation for the squares of the Schmidt coefficients. Moreover, the success probability is $p(\xi) = \left[1 - \xi + \xi/(Da_{\min}^2) \right]^{-1}$ [55]. This method, although straightforward to understand, leads to little improvement in terms of success probabilities. For instance, Figure 1 evidences that even a little improvement in any of the functions used to assess entanglement is achieved at the expense of a substantial drop in the success probability. This figure also evidences that the I-Concurrence, although simple to work with because it is not a rational function, is not good for graphical assessment since even initial I-Concurrence (see $\xi = 0$) exhibits values close to one. Instead, the Schmidt number is not simple to work with due to its inverse dependence on $\mathrm{tr}(\rho_1^2)$ but makes graphical evaluation uncomplicated.
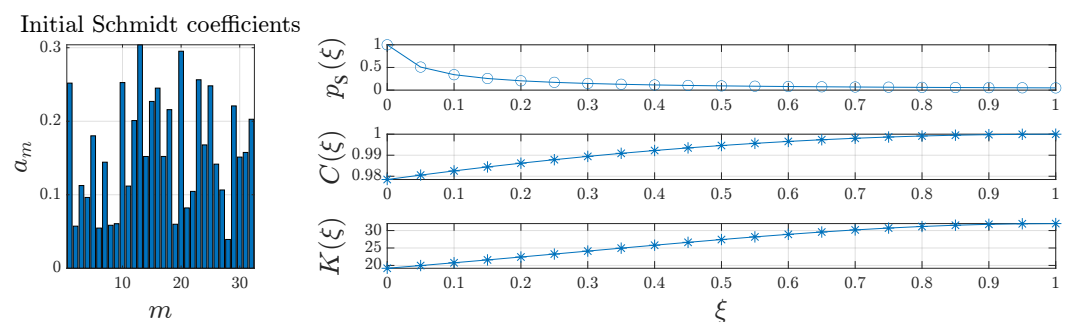
Figure 1. Example of entanglement concentration for $D = 32$ by using linear interpolation for the squares of the Schmidt coefficients.

These previous attempts lead us to question whether a method can obtain a *reasonable* increment in entanglement with a non-negligible success probability without imposing constraints on the final state beforehand. The next sections will address this question.

3. Towards Efficient Entanglement Concentration

Here, we shall propose and analyze a more efficient method for entanglement concentration from a single copy of a partially entangled pure state. Let us define a parameterized Kraus operator $A_{\mathrm{S}}(\vec{z})$ being applied on one of the qudits. This operator can be written as

$$A_{\mathrm{S}}(\vec{z}) = \sum_{m=1}^{D} z_m |m\rangle\langle m|,$$ (7)

so its action on the two-qudit system after successful concentration will be

$$A_{\mathrm{S}}(\vec{z})|\Phi\rangle_{12} = \sum_{m=1}^{D} z_m a_m |m\rangle_1 |m\rangle_2. \tag{8}$$

Thus, keeping Equation (5) in mind, the post-concentration state and its probability of success are

$$|\Psi(\vec{z})\rangle_{12} = \frac{1}{\sqrt{p_{\mathrm{S}}(\vec{z})}} \sum_{m=1}^{D} a_m z_m |m\rangle_1 |m\rangle_2, \tag{9}$$

$$p_{\mathrm{S}}(\vec{z}) = \sum_{m=1}^{D} a_m^2 |z_m|^2, \tag{10}$$

respectively. Since $p_{\mathrm{S}}(\vec{z})$ must not exceed one, it is mandatory to impose $|z_m| \leqslant 1$. The reduced density matrix for one of the subsystems shall be

$$\rho_1(\vec{z}) = \frac{1}{p_{\mathrm{S}}(\vec{z})} \sum_{m=1}^{D} a_m^2 |z_m|^2 |m\rangle\langle m|. \tag{11}$$

I-Concurrence and Schmidt number, as function of $\vec{z}$, will be given by

$$\mathcal{C}(\vec{z}) = \sqrt{\frac{D}{D-1}\left(1 - \sum_{m=1}^{D} \frac{a_m^4 |z_m|^4}{p_{\mathrm{S}}^2(\vec{z})}\right)}, \tag{12}$$

$$K(\vec{z}) = \frac{p_{\mathrm{S}}^2(\vec{z})}{\sum_m a_m^4 |z_m|^4}. \tag{13}$$

Let us now define a quantity $\mathcal{Q}(\vec{z})$ aimed to assess the efficiency of the concentration procedure considering a trade-off between the probability of success and the increment in entanglement. A Kraus operator that maximizes this efficiency will be pursued. A choice, although not unique at all, might be $p_{\mathrm{S}}(\vec{z})\mathcal{C}(\vec{z})$. Maximizing it will be equivalent to maximizing its square, $[p_{\mathrm{S}}(\vec{z})\mathcal{C}(\vec{z})]^2$, which should be a simpler procedure since the square root we can see in Equation (12) will not be present. However, $[p_{\mathrm{S}}(\vec{z})\mathcal{C}(\vec{z})]^2$ has its maximum when $z_m = 1$, $\forall\, m$, which means state $|\Phi\rangle_{12}$ will be kept unaltered (This will be proven in Appendix A). Instead, we may try with the difference between $\mathcal{C}^2(\vec{z})$ and a constant reference level for the I-Concurrence ($\mathcal{C}_{\mathrm{REF}}^2$). This reference level could be, for instance, the initial value $\mathcal{C}_{\mathrm{INIT}} = \mathcal{C}(|\Phi\rangle_{12})$. Let us try by defining an efficiency function such as

$$\mathcal{Q}(\vec{z}) = p_{\mathrm{S}}^2(\vec{z})\left(\mathcal{C}^2(\vec{z}) - \mathcal{C}_{\mathrm{REF}}^2\right). \tag{14}$$

Equations (10) and (12) allow us to transform Equation (14) into

$$\mathcal{Q}(\vec{z}) = \frac{D}{D-1} \sum_{m,n=1}^{D} |z_m|^2 a_m^2 (\mathscr{P}_{\mathrm{REF}} - \delta_{mn}) a_n^2 |z_n|^2, \tag{15}$$

$$\mathscr{P}_{\mathrm{REF}} = 1 - \frac{D-1}{D}\mathcal{C}_{\mathrm{REF}}^2, \tag{16}$$

where $\mathscr{P}_{\mathrm{REF}}$ has been defined for mathematical convenience. It ranges from $1/D$ to 1, and it can be interpreted as a reference value for the purity of a reduced density matrix, as it can be seen from Equation (2). Another interpretation, as can be seen from Equation (3), is $\mathscr{P}_{\mathrm{REF}} = 1/\mathcal{K}_{\mathrm{REF}}$, where $\mathcal{K}_{\mathrm{REF}}$ is a reference value for the Schmidt number. A careful observation of Equation (15) leads us to infer that the problem of efficient entanglement concentration, in the form it has been described in this document, can be rewritten as a quadratic optimization problem given by

$$\max_{\vec{y}} \mathcal{Q}(\vec{y}) = \frac{D}{D-1} \vec{y}^{\mathsf{T}} \mathbf{H} \, \vec{y}, \tag{17a}$$

$$\text{subject to} \quad 0 \leqslant y_m \leqslant 1, \tag{17b}$$

where

$$y_m = |z_m|^2, \tag{17c}$$

$$[\mathbf{H}]_{m,n} = (\mathscr{P}_{\mathrm{REF}} - \delta_{mn}) a_m^2 a_n^2. \tag{17d}$$

Therefore, the problem of efficient entanglement concentration for a single pair of entangled qudits can be written as the quadratic optimization problem described in Equations (17a)–(17d), with the optimization variables y_m lying in a unit hypercube. Finally, without loss of generality, we may choose the positive root of $z_m = \sqrt{y_m}$. Note that the presence of $\mathcal{C}_{\mathrm{REF}}$ forces the optimization to look for a solution $\vec{y}_{\mathrm{OPT}}$ such that $\mathcal{C}(\vec{y}_{\mathrm{OPT}}) \geqslant \mathcal{C}_{\mathrm{REF}}$. Otherwise, function $\mathcal{Q}(\vec{y}_{\mathrm{OPT}})$ would adopt a negative value [see Equation (14)] and, therefore, it will not represent a maximum. For this reason, we can assure that $\mathcal{C}_{\mathrm{REF}} \geqslant \mathcal{C}_{\mathrm{INIT}}$ forces entanglement concentration. In an extreme case, $\mathcal{C}_{\mathrm{REF}} = 1$ means that the reference level is equal to the maximum possible value I-Concurrence can adopt. Therefore, $\mathcal{Q}(\vec{y})$ will adopt a negative value *unless* the final entanglement is also equal to one, for which $\mathcal{Q} = 0$. This is the standard entanglement concentration procedure. On the other hand, $\mathcal{C}_{\mathrm{REF}}$ could be slightly smaller than $\mathcal{C}_{\mathrm{INIT}}$ and, still, entanglement concentration may occur, as it will be shown in Section 4.1. For this problem, the square of the I-Concurrence has been used also because both numerical and analytical solutions are accessible. For graphical purposes, as it was already seen in Figure 1, the Schmidt number shall be used. Moreover, the Schmidt number provides an estimation of the number of relevant Schmidt modes involved [45].

We must add that the Kraus operator defined in Equation (7) is diagonal in the Schmidt basis. We may have started by a general Kraus operator, instead of a diagonal one. However, Appendix B shows it suffices to look for diagonal operators.

4. Solving the Problem

4.1. Numerical Hints

Figure 2 shows the results of numerical resolution of the aforementioned optimization problem for a given set of $D = 16$ Schmidt coefficients a_m^2, randomly chosen, and sorted decreasingly in order to ease observation. For this example, we tested four possible values of $\mathcal{C}_{\mathrm{REF}}^2$, given by (i) $\mathcal{C}_{\mathrm{INIT}}^2/2$, much smaller than the initial entanglement; (ii) $0.98\mathcal{C}_{\mathrm{INIT}}^2$, slightly smaller than the initial entanglement; (iii) average value between $\mathcal{C}_{\mathrm{INIT}}$ and 1, a significant increase in entanglement; and (iv) $\mathcal{C}_{\mathrm{REF}}^2 = 1$, the maximum possible value for $\mathcal{C}_{\mathrm{REF}}^2$. The optimization was performed using the function QUADPROG of Matlab R2022b. Since this is a non-convex problem with constant bounds only, the algorithm "trust-region-reflective" was used since it was the best suited for our optimization problem [56].

The plots show the original Schmidt coefficients (cyan) and the non-normalized coefficients after concentration (dark red). A pattern is evident. For small values of $\mathcal{C}_{\mathrm{REF}}^2$, keeping the state as it is seems to be the best option in terms of efficiency. As $\mathcal{C}_{\mathrm{REF}}^2$ increases, the solutions of the optimization problem suggest one to use a Procrustean method on the n largest Schmidt coefficients, where n increases according $\mathcal{C}_{\mathrm{REF}}^2$ moves closer to one. This is analogous to entanglement concentration on a subspace of the bipartite Hilbert space as the one proposed in Ref. [13], although we have not required the final state to be fixed to a given one. Finally, $\mathcal{C}_{\mathrm{REF}}^2 = 1$ represents the ideal entanglement concentration context, in which the resulting state exhibits the maximal entanglement possible. The optimization problem shows the correct result, which consists of uniforming all post-concentration Schmidt coefficients.
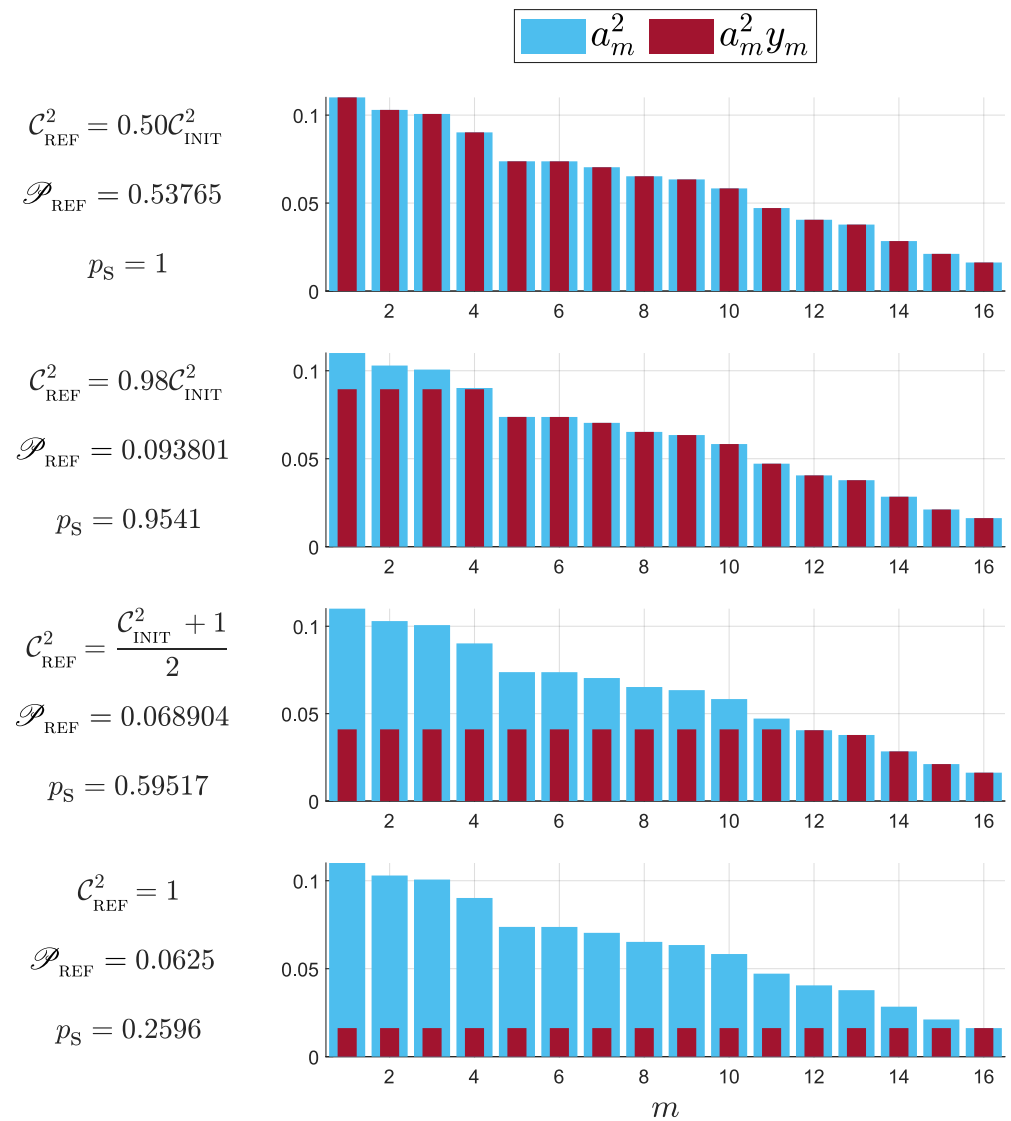
Figure 2. Numerical example of resolution of the quadratic optimization problem [Equation (17)] for dimension $D = 16$, using 4 different values of $\mathcal{C}_{\mathrm{REF}}^2$. Bars show the original Schmidt coefficients (cyan) and the non-normalized coefficients after concentration (dark red). Their respective values of $\mathscr{P}_{\mathrm{REF}}$ and probabilities of success p_{S} are also shown.

Although Figure 2 shows a single set of initial Schmidt coefficients, the same pattern is observed for other states in any dimension $D > 2$. In the following, we shall prove why the Procrustean method on a subspace is the most efficient method, according to our figures of merit.

4.2. Analytical Results

One of the goals of this work is to find the analytical solution of the optimization problem of Equation (17). The details of the proof will be shown in the next subsections. The procedure can be summarized as follows:

1. If $\mathscr{P}_{\mathrm{REF}} = 1/D$ (minimum attainable value, equivalent to $\mathcal{C}_{\mathrm{REF}} = 1$), it means we are pursuing a standard entanglement concentration using all Schmidt coefficients. Then, perform concentration using $z_m = a_{\min}/a_m$. Otherwise, follow Steps 2–8.
2. Sort the Schmidt coefficients in decreasing order. Let us label these sorted coefficients as a_m.
3. Define a vector $\vec{\beta}$ such that $\beta_n = 1 - \sum_{m=1}^{n} \mathsf{a}_m^2$, for $n = 1, \ldots, D$.
4. Define a vector $\vec{\alpha}$ such that $\alpha_n = \mathscr{P}_{\mathrm{REF}} \beta_n / (1 - n \mathscr{P}_{\mathrm{REF}})$.

5. Find the largest value of n that allow both $\alpha_n \leqslant a_n^2$ and $n < 1/\mathscr{P}_{\mathrm{REF}}$ to be simultaneously satisfied. Let us label this value as n_{OPT}.
6. Define $\vec{x}$ such that

$$
x_m = \begin{cases} \alpha_{n_{\mathrm{OPT}}} & , \text{for } m = 1, \ldots, n_{\mathrm{OPT}}, \\ 1 & , \text{for } m = n_{\mathrm{OPT}} + 1, \ldots, D. \end{cases}
$$

7. Define $y_m = x_m/a_m^2$. Afterwards, sort the y_m using the inverse of the sorting operation described in Step 1. These sorted values will be the y_m that solve the optimization problem of Equation (17).
8. Define $z_m = \sqrt{y_m}$. These values are the ones needed to construct the Kraus operator of Equation (7).

Sections 4.2.1–4.2.7 hereunder shall detail the underlying reasoning for the algorithm shown above.

4.2.1. Redefining the Optimization Problem

In order to prove the solution detailed above, we shall define $x_m = a_m^2 y_m = a_m^2 |z_m|^2$. This allows us to write the optimization problem [Equation (17)], up to a proportionality constant, in a simpler way:

$$
\begin{cases} \max_{\vec{x}} \quad Q(\vec{x}) = \mathscr{P}_{\mathrm{REF}} \left(\sum_{m=1}^{D} x_m \right)^2 - \sum_{m=1}^{D} x_m^2, \\ \text{s. t.} \quad 0 \leqslant x_m \leqslant a_m^2. \end{cases} \tag{18}
$$

These new variables x_m are the ones plotted in Figure 2 using dark red bars. Thus, the x_m will provide an idea about the post-concentration Schmidt coefficients.

The domain is no longer the unit hypercube, but an orthotope whose vertices have coordinates components equal to zero and a_m^2. Thus, every x_m has three options: (i) having a fixed value equal to zero, (ii) having a fixed value equal to a_m^2, and (iii) having a variable value between zero and a_m^2. These options had to be taken into account in order to find all critical points.

4.2.2. Finding Critical Points

For starters, we shall define set of indices according to the aforementioned options:
1. $\mathcal{Z} = \{ j : x_j = 0 \}$;
2. $\mathcal{O} = \{ k : x_k = a_k^2 \}$;
3. $\mathcal{I} = \{ \ell : 0 < x_\ell < a_\ell^2 \}$.

The symbols $\mathcal{Z}$, $\mathcal{O}$, and $\mathcal{I}$ stand for *zero*, *outer*, and *inner*, respectively. In this way, any summation can be written as $\sum_m = \sum_{j \in \mathcal{Z}} + \sum_{k \in \mathcal{O}} + \sum_{\ell \in \mathcal{I}}$. There exist 3^D configurations for $(\mathcal{Z}, \mathcal{O}, \mathcal{I})$. If we label each of those 3^D combinations by using the index μ, then we can define function $Q_\mu(\vec{x})$ as the function $Q(\vec{x})$ for the μth configuration. Explicitly,

$$
Q_\mu(\vec{x}) = \mathscr{P}_{\mathrm{REF}} \left(\sum_{k \in \mathcal{O}_\mu} a_k^2 + \sum_{\ell \in \mathcal{I}_\mu} x_\ell \right)^2 - \sum_{k \in \mathcal{O}_\mu} a_k^4 - \sum_{\ell \in \mathcal{I}_\mu} x_\ell^2. \tag{19}
$$

By imposing $\partial_{x_r} Q_\mu(\vec{x}) = 0$, we can find the critical points of $Q_\mu(\vec{x})$. Consequently,

$$
x_r = \mathscr{P}_{\mathrm{REF}} \left(\sum_{k \in \mathcal{O}_\mu} a_k^2 + \sum_{\ell \in \mathcal{I}_\mu} x_\ell \right), \quad r \in \mathcal{I}_\mu. \tag{20}
$$

This means that as long as x_r is not fixed in either 0 or a_r^2, the optimal solution is such that those x_r all adopt the same value. Let us define some additional ancillary parameters,

$$\beta_\mu = \sum_{k \in \mathcal{O}_\mu} a_k^2, \qquad \gamma_\mu = \sum_{k \in \mathcal{O}_\mu} a_k^4, \qquad n_\mu = |\mathcal{I}_\mu|, \tag{21}$$

n_μ being the number of free parameters x_ℓ. With these definitions, we can now assert that $x_\ell = \alpha_\mu$ is the critical point for the μth configuration, where

$$x_\ell = \alpha_\mu = \frac{\mathscr{P}_{\mathrm{REF}} \beta_\mu}{1 - \mathscr{P}_{\mathrm{REF}} n_\mu}, \qquad \forall \, \ell \in \mathcal{I}_\mu. \tag{22}$$

Consequently, if $\mathbb{Q}_\mu$ is the value of $Q_\mu(\vec{x})$ evaluated at the μth critical point, then

$$\begin{aligned}
\mathbb{Q}_\mu &= \mathscr{P}_{\mathrm{REF}}(\beta_\mu + n_\mu \alpha_\mu)^2 - \gamma_\mu - n_\mu \alpha_\mu^2 \\
&= \alpha_\mu \beta_\mu - \gamma_\mu.
\end{aligned} \tag{23}$$

The fact that $x_\ell = \alpha_\mu$ means that, for every $\ell \in \mathcal{I}_\mu$, coefficients a_ℓ^2 will be transformed into α_μ as consequence of the concentration procedure. This is, precisely, the Procrustean method applied on a n_μ-dimensional subset of the coefficients $\{a_m\}$.

It is worth mentioning that Equation (22) contains the implicit assumption $\mathscr{P}_{\mathrm{REF}} \neq 1/n_\mu$, which raises questions regarding the case $\mathscr{P}_{\mathrm{REF}} = 1/n_\mu$. If that were the case, trying to solve Equation (20) leads us to conclude $\beta_\mu = 0$ and, equivalently, $\mathcal{O}_\mu = \varnothing$. In turn, this implies $Q_\mu(\vec{x}) = 0$. Nevertheless, we may see from the original definition of $\mathcal{Q}(\vec{z})$ [Equation (14)] that the only possible way in which $Q_\mu(\vec{x}) = 0$ represents a maximum occurs when $\mathcal{C}_{\mathrm{REF}}^2 = 1$ and $\mathcal{C}^2(\vec{z}) = 1$ simultaneously, i.e., $\mathscr{P}_{\mathrm{REF}} = 1/D$ has been set and the resulting state is a D-dimensional maximally entangled state.

4.2.3. Upper Bounds for n_μ

The Hessian matrix has components given by

$$\partial_{x_s} \partial_{x_r} Q_\mu(\vec{x}) = 2(\mathscr{P}_{\mathrm{REF}} - \delta_{rs}). \tag{24}$$

It can be shown that $\mathbb{Q}_\mu$ will represent a local maximum for the μth configuration provided, $(1 - n_\mu \mathscr{P}_{\mathrm{REF}}) > 0$, since this condition ensures Hessian matrix to be negative-definite. In other words,

$$n_\mu < \frac{1}{\mathscr{P}_{\mathrm{REF}}}. \tag{25}$$

Thus, some configurations $(\mathcal{Z}_\mu, \mathcal{O}_\mu, \mathcal{I}_\mu)$ can be immediately discarded if n_μ exceeds this bound.

4.2.4. Eliminating Zeros

Let us start by analyzing the effect of zeros by comparing a given $\mathbb{Q}_\mu$—for which $x_r = 0$—with the value of $Q_{\mu'}(\vec{x})$ when $x_r = \delta \gtrsim 0$. Using Equation (19), we have that

$$\mathbb{Q}_\mu \Big|_{x_r=0} = \mathscr{P}_{\mathrm{REF}}(\beta_\mu + n_\mu \alpha_\mu)^2 - \gamma_\mu - n_\mu \alpha_\mu^2, \tag{26}$$

$$Q_{\mu'} \Big|_{x_r \to \delta} = \mathscr{P}_{\mathrm{REF}}(\beta_\mu + n_\mu \alpha_\mu + \delta)^2 - \gamma_\mu - n_\mu \alpha_\mu^2 - \delta^2, \tag{27}$$

which, in turn, leads us to

$$Q_{\mu'} \Big|_{x_r \to \delta} - \mathbb{Q}_\mu \Big|_{x_r=0} = 2\mathscr{P}_{\mathrm{REF}}(\beta_\mu + n_\mu \alpha_\mu)\delta + \mathcal{O}(\delta^2) > 0. \tag{28}$$

We can see that $Q_{\mu'}$ actually grows if x_r moves away from zero within its neighborhood. This means that every configuration containing a null value on *any* of its x_m cannot represent a maximum since all neighboring points have higher values for $Q(\vec{x})$. Therefore, the solution we are looking for is such that $\mathcal{Z}_\mu = \emptyset$. The number of remaining configurations is now less than 2^D.

4.2.5. Optimal n Will Be the Largest Possible

We are left with the options $x_m \in \{\alpha_\mu, a_m^2\}$. We know that the μth critical point is such that $x_\ell = \alpha_\mu$, $\forall \ell \in \mathcal{I}_\mu$. Since $\vec{x}$ still belongs to the orthotope, an additional condition arises: $\alpha_\mu \leqslant a_\ell^2$, $\forall \ell \in \mathcal{I}_\mu$.

Let us now compare two solutions $\mathbb{Q}_\lambda$ and $\mathbb{Q}_\nu$, whose critical points differ only in one term x_r, so $r \in \mathcal{O}_\lambda$ and $r \in \mathcal{I}_\nu$. Thus, by using Equations (21)–(23), we have that

$$\beta_\nu = \beta_\lambda - a_r^2, \tag{29}$$

$$\gamma_\nu = \gamma_\lambda - a_r^4, \tag{30}$$

$$n_\nu = n_\lambda + 1, \tag{31}$$

$$\alpha_\nu = \frac{\mathscr{P}_{\text{REF}}(\beta_\lambda - a_r^2)}{1 - (n_\lambda + 1)\mathscr{P}_{\text{REF}}}, \tag{32}$$

$$\mathbb{Q}_\lambda = \alpha_\lambda \beta_\lambda - \gamma_\lambda, \tag{33}$$

$$\mathbb{Q}_\nu = \alpha_\nu \beta_\nu - \gamma_\nu. \tag{34}$$

Consequently,

$$\mathbb{Q}_\lambda - \mathbb{Q}_\nu = -\frac{\left(\mathscr{P}_{\text{REF}}\beta_\lambda - (1 - n_\lambda \mathscr{P}_{\text{REF}})a_r^2\right)^2}{(1 - n_\lambda \mathscr{P}_{\text{REF}})(1 - (n_\lambda + 1)\mathscr{P}_{\text{REF}})} < 0. \tag{35}$$

Therefore, a better solution is obtained when r belongs to $\mathcal{I}_\nu$ over $\mathcal{O}_\lambda$, provided that the constraints are fulfilled. In simpler words, the best of the $\{n_\mu\}$ will be the largest possible within the conditions $n_\mu < 1/\mathscr{P}_{\text{REF}}$ and $\alpha_\mu \leqslant a_\ell^2$, $\forall \ell \in \mathcal{I}_\mu$.

4.2.6. Sorting Preference

For the following comparison, it will be helpful to define two sets $\mathcal{O}_0$ and $\mathcal{I}_0$. We will center our attention on two values x_r and x_s. Now, let us compare two solutions $\mathbb{Q}_\rho$ and $\mathbb{Q}_\sigma$ that satisfy

$$n_\rho = n_\sigma = n, \tag{36}$$

$$\mathcal{I}_\rho = \mathcal{I}_0 \cup \{r\}, \qquad \mathcal{I}_\sigma = \mathcal{I}_0 \cup \{s\}, \tag{37}$$

$$\mathcal{O}_\rho = \mathcal{O}_0 \cup \{s\}, \qquad \mathcal{O}_\sigma = \mathcal{O}_0 \cup \{r\}. \tag{38}$$

Thus, $\mathcal{I}_\rho$ and $\mathcal{I}_\sigma$ have $n - 1$ elements in common, whereas $\mathcal{O}_\rho$ and $\mathcal{O}_\sigma$ have $D - n - 1$ elements in common. Consequently,

$$\beta_\rho = \beta_0 + a_s^2, \qquad \gamma_\rho = \gamma_0 + a_s^4, \tag{39}$$

$$\beta_\sigma = \beta_0 + a_r^2, \qquad \gamma_\sigma = \gamma_0 + a_r^4, \tag{40}$$

where $\beta_0 = \sum_{k \in \mathcal{O}_0} a_k^2$ and $\gamma_0 = \sum_{k \in \mathcal{O}_0} a_k^4$. For the following, we shall assume $a_r > a_s$. Now, since both $\mathbb{Q}_\rho$ and $\mathbb{Q}_\sigma$ are admissible solutions, it *must* happen that $\alpha_\rho \leqslant a_r^2$ and $\alpha_\sigma \leqslant a_s^2$ as consequence of Equations (18), (22), and (37). This means

$$t(\beta_0 + a_s^2) \leqslant a_r^2, \quad \text{and} \quad t(\beta_0 + a_r^2) \leqslant a_s^2, \tag{41}$$

where $t = \mathscr{P}_{\text{REF}}/(1 - n\mathscr{P}_{\text{REF}})$ is a positive parameter. If we add these two inequalities, we obtain

$$(a_r^2 + a_s^2)(1 - t) - 2t\beta_0 \geqslant 0. \tag{42}$$

The difference between the solutions $\mathbb{Q}_\rho$ and $\mathbb{Q}_\sigma$ is

$$\begin{aligned}
\Delta\mathbb{Q} &= \mathbb{Q}_\rho - \mathbb{Q}_\sigma \\
&= \left(a_r^2 - a_s^2\right)\left((1 - t)(a_r^2 + a_s^2) - 2t\beta_0\right).
\end{aligned} \tag{43}$$

Since $a_r > a_s$ was assumed and the inequality of Equation (42) was obtained, it can be assured that $\mathbb{Q}_\rho > \mathbb{Q}_\sigma$. Now, let us remember that $\mathbb{Q}_\rho$ is the solution in which $x_r = \alpha_\rho$ and $x_s = a_s^2$. This means it is better to cut off coefficient a_r (the larger one) over a_s.

Since we already know (see Section 4.2.5) that n must be the largest possible within the constraints $n < 1/\mathscr{P}_{\mathrm{REF}}$ and $\alpha_\mu \leqslant a_\ell^2$, $\forall \ell \in \mathcal{I}_\mu$, we must compare now all the solutions $\mathbb{Q}_\mu$ such that n_μ is equal to that optimal value of n. According to the computations of this section, the most efficient concentration scheme will consist in cutting off the n largest Schmidt coefficients, which is in complete agreement with the results shown in Figure 2.

4.2.7. How to Construct the Optimal Concentration Scheme

In summary, we know now that if $\mathcal{C}_{\mathrm{REF}} = 1$ (equivalently, $\mathscr{P}_{\mathrm{REF}} = 1/D$), then the optimal solution corresponds to a entanglement concentration procedure that yields a D-dimensional maximally entangled state. On the other hand, if $\mathcal{C}_{\mathrm{REF}} < 1$ (equivalently, $\mathscr{P}_{\mathrm{REF}} > 1/D$), we have shown that the optimal solution (i) does not contain zeros, (ii) it has values either given by $x_m = a_m^2$ (i.e., keep a_m as they are) or by $x_m = \alpha_\mu$ (i.e., crop coefficients a_m to a given value α_μ), (iii) the n largest Schmidt coefficients are to be cropped, and (iv) n must be as large as possible within constraints given by $n < 1/\mathscr{P}_{\mathrm{REF}}$ and $\alpha_\mu \leqslant a_m^2$. Once the optimal x_m are found, we may compute the corresponding y_m and z_m. These rules gave rise to the algorithm described at the beginning of Section 4.2. Moreover, we performed thousands of numerical simulations, ranging from $D = 32$ to $D = 1024$, that confirmed such an algorithm actually provides the optimal solution. Figure 3 shows a sample of those simulations for $D = 1024$, depicting relative differences between the results from numerical optimization ($\vec{y}_{\mathrm{num}}$ and $Q(\vec{y}_{\mathrm{num}})$) and the ones from the algorithm proposed in this section ($\vec{y}_{\mathrm{alg}}$ and $Q(\vec{y}_{\mathrm{alg}})$) for 100 values of $\mathscr{P}_{\mathrm{REF}}$. These relative differences are computed as

$$\Delta y_{\mathrm{relative}} = \frac{1}{D} \sum_{m=1}^{D} \left| \frac{(\vec{y}_{\mathrm{num}})_m - \left(\vec{y}_{\mathrm{alg}}\right)_m}{(\vec{y}_{\mathrm{num}})_m} \right|, \tag{44}$$

$$\Delta \mathcal{Q}_{\mathrm{relative}} = \left| \frac{\mathcal{Q}(\vec{y}_{\mathrm{num}}) - \mathcal{Q}(\vec{y}_{\mathrm{alg}})}{\mathcal{Q}(\vec{y}_{\mathrm{num}})} \right|. \tag{45}$$

The initial Schmidt coefficients were computed from a randomly-generated $D \times D$ entangled state. As the data of Figure 3 show, relative differences between the two solutions being compared are negligible, thus demonstrating the adequateness of the proposed algorithm. Discrepancies can be explained as a consequence of floating-point computation precision.
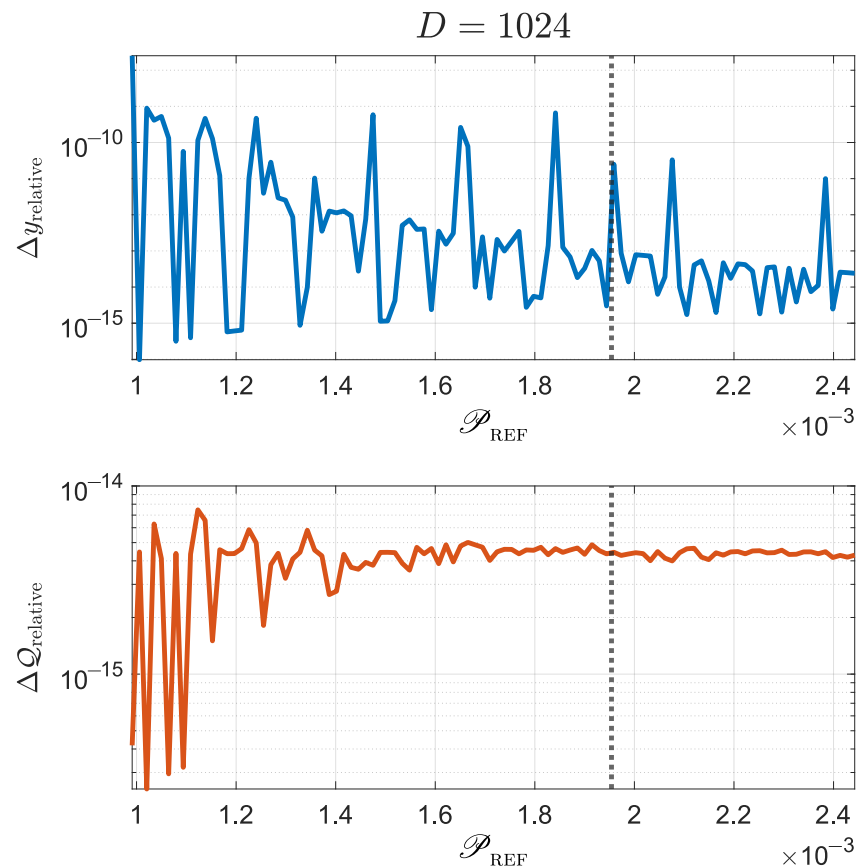
Figure 3. Comparison between results obtained through numerical optimization ($\mathcal{Q}(\vec{y}_{\mathrm{num}})$) and the ones obtained by using the algorithm introduced at the beginning of Section 4.2 ($\mathcal{Q}(\vec{y}_{\mathrm{alg}})$). Relative differences for are shown for 100 values of $\mathscr{P}_{\mathrm{REF}}$. The vertical dotted line indicates the initial value of the purity of the reduced density matrix, i.e., $\mathscr{P}_{\mathrm{REF}} = \mathscr{P}_{\mathrm{INIT}}$. See the main text for details about the computation of these relative differences.

After efficiency optimization, one should evaluate whether practical advantages were obtained from it. Figure 4 shows the probability of success and Schmidt number for the same optimizations carried out for Figure 3. The initial state had a Schmidt number $K_{\mathrm{INIT}} \approx 512$. Raising this number to its maximum (i.e., $K = 1024$) can be done with a probability of success $p_{\mathrm{S}} = D a_{\mathrm{min}}^2 \sim 10^{-7}$ (not shown in the graphs in order to ease observation). However, non-maximal Schmidt numbers can be obtained with much better probabilities. For instance, $\mathscr{P}_{\mathrm{REF}} \approx 1.15 \times 10^{-3}$ allows one to achieve a considerable Schmidt number ($K = 900$) with a success probability $p_{\mathrm{S}} = 11\%$. Although $\mathscr{P}_{\mathrm{REF}} \approx 1.15 \times 10^{-3}$ seems to be a non-trivial number of uncertain origin, we may notice that $1/\mathscr{P}_{\mathrm{REF}} \sim 868$. Thus, an acceptable method to estimate the necessary value of $\mathscr{P}_{\mathrm{REF}}$ consists in setting a minimum desirable Schmidt number $\mathcal{K}_{\mathrm{MIN}}$, define a slightly smaller threshold number $\mathcal{K}_{\mathrm{THR}} < \mathcal{K}_{\mathrm{MIN}}$, and computing $\mathscr{P}_{\mathrm{REF}} = 1/\mathcal{K}_{\mathrm{THR}}$.

It is worth mentioning that the solution described in this section closely resembles the entanglement concentration procedure described in Ref. [13], which was also graphically explained in Ref. [30]. However, we did not set the final state to a fixed one in our formulation. Instead, we defined a single figure of merit to be interpreted as efficiency, and its optimization suggested performing entanglement concentration on the subspace of the largest Schmidt coefficients.

$$D = 1024$$
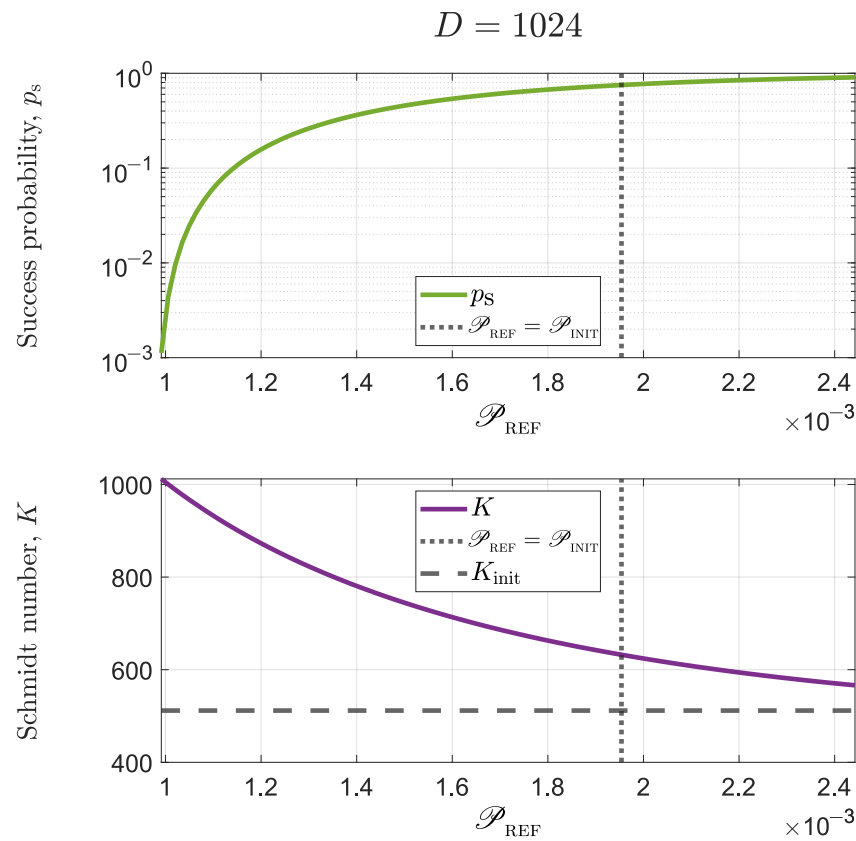
Figure 4. Success probability and Schmidt number for the same state and optimizations used in Figure 3. The vertical dotted line indicates the initial value of the purity of the reduced density matrix, i.e., $\mathscr{P}_{\text{REF}} = \mathscr{P}_{\text{INIT}}$ and the horizontal dashed line shows the initial Schmidt number. Keep in mind that larger values of $\mathscr{P}_{\text{REF}}$ mean smaller values of $\mathcal{C}_{\text{REF}}$.

5. Entanglement Concentration with Fixed Probability of Success

An alternative way to solve the problem of efficient entanglement concentration is by setting the success probability to a fixed value p_{FIX} and inquiring about the largest entanglement that can be extracted. As it can be seen from Equations (12) and (13), this question reduces to minimization of the purity of the reduced density matrix, as

$$\min_{\vec{y}} \ \mathcal{P}(\vec{y}) = \left[\frac{1}{p_{\text{S}}(\vec{y})} \sum_{m=1}^{D} a_m^4 y_m^2 \right],$$

$$\text{subject to } 0 \leqslant y_m \leqslant 1 \quad \text{and} \quad \sum_{m=1}^{D} a_m^2 y_m = p_{\text{FIX}}, \tag{46}$$

where we have already used $y_m = |z_m|^2$. As we have imposed $p_{\text{S}}(\vec{y}) = p_{\text{FIX}}$, the optimization reduces to optimize $\sum_m a_m^4 y_m^2$. As in the previous section, we shall resort to $x_m = y_m^2$ and the sets of indices $\mathcal{Z}_\mu$, $\mathcal{O}_\mu$, and $\mathcal{I}_\mu$. Using the x_m, we are left to optimize $\sum_m x_m^2$, and the constraint of fixed probability can be rewritten as $\sum_m x_m = p_{\text{FIX}}$, which also allows us to write one of the variables in terms of the others. Let

$$x_\vartheta = p_{\text{FIX}} - \sum_{m \neq \vartheta} x_m. \tag{47}$$

Then, the minimization of the purity can be rewritten as

$$\text{minimize}(p_{\text{FIX}}\mathcal{P}(\vec{x})) = \sum_{m\neq\vartheta} x_m^2 + \left(p_{\text{FIX}} - \sum_{m\neq\vartheta} x_m\right)^2$$

$$= \sum_{k\in\mathcal{O}_\mu} a_k^4 + \sum_{\substack{\ell\in\mathcal{I}_\mu \\ \ell\neq\vartheta}} x_\ell^2 + \left(\sum_{k\in\mathcal{O}_\mu} a_k^2 + \sum_{\substack{\ell\in\mathcal{I}_\mu \\ \ell\neq\vartheta}} x_\ell\right)^2. \tag{48}$$

Critical points are found by setting $\partial(p_{\text{FIX}}\mathcal{P}(\vec{x}))/\partial x_r = 0$, with $r \in \mathcal{I}_\mu$ and $r \neq \vartheta$. This leads us to $x_r = \kappa_\mu$, where

$$\kappa_\mu = \frac{p_{\text{FIX}} - \beta_\mu}{n_\mu}. \tag{49}$$

In turn, Equation (47) implies that $x_\vartheta = \kappa_\mu$ as well. Thus, we obtained solutions given by either $x_m = a_m^2$, $x_m = 0$, or $x_m = \kappa_\mu$, which is the exact behavior exhibited by the x_m from Section 4 up to a change from α_μ to κ_μ. The same analysis performed in Sections 4.2.4–4.2.7 can be applied here. The conclusions are very similar: (i) the optimal values of x_m are different from zero, (ii) if n is the number of variables x_m being equal to κ_μ, then n must be as large as possible within the constraint $0 \leqslant \kappa \leqslant a_\ell^2$, and (iii) the n largest Schmidt coefficients are cut off. Thus, an algorithm can be constructed as follows:

1. Sort the Schmidt coefficients in decreasing order. Let us label these sorted coefficients as a_m.
2. Define a vector $\vec{\beta}$ such that $\beta_n = 1 - \sum_{m=1}^{n} \mathsf{a}_m^2$, for $n = 1, \ldots, D$.
3. Define a vector $\vec{\kappa}$ such that $\kappa_n = (p_{\text{FIX}} - \beta_n)/n$.
4. Find the largest value of n such that $\kappa_n \geqslant 0$ and $\kappa_n < \mathsf{a}_n^2$ are simultaneously satisfied. Let us label this value as n_{OPT}.
5. Define $\vec{\mathsf{x}}$ such that

$$\mathsf{x}_m = \begin{cases} \kappa_{n_{\text{OPT}}} & \text{, for } m = 1, \ldots, n_{\text{OPT}}, \\ 1 & \text{, for } m = n_{\text{OPT}} + 1, \ldots, D. \end{cases}$$

6. Define $\mathsf{y}_m = \mathsf{x}_m/\mathsf{a}_m^2$. Afterwards, sort the y_m using the inverse of the sorting operation described in Step 1. These sorted values will be the y_m that solve the optimization problem of Equation (17).
7. Define $z_m = \sqrt{y_m}$. These values are the ones needed to construct the Kraus operator of Equation (7).

As it can be seen, the solutions obtained for this problem are completely analogous to the ones of the previous section. The advantage of this approach lies in the fact that $\mathcal{P}(\vec{x})$ appears in both I-Concurrence and Schmidt number. Thus, it is a favorable way to increase the Schmidt number without introducing nontrivial mathematical complications. Once more, this result represents a Procrustean method applied on a subspace, although only one parameter has been fixed (p_{FIX}) instead of a whole state.

6. Conclusions

In summary, we have studied entanglement concentration from a single copy of a two-qudit entangled state in terms of efficiency. As the ideal procedure—obtaining a maximally entangled state—is extremely inefficient in terms of probability, we studied the possibility of concentrating a fair enough amount of entanglement and, simultaneously, increment the success probability. Two methods were analyzed. For the first one, a function $\mathcal{Q}(\vec{y})$ was defined in order to quantify efficiency as the product of success probability and entanglement increment. This function allows one to introduce a parameter $\mathscr{P}_{\text{REF}}$, which is loosely related to a minimal entanglement amount intended to extract. The other one consisted of fixing the success probability to a given value and finding the maximal

entanglement it can be extracted under the constraint herein. We found that, for both cases, the solution resembles a Procrustean method applied on a subset of the largest Schmidt coefficients. Such application of the Procrustean method has been already studied in the literature under the assumption that the final state *must* be a n-dimensional maximally entangled state, with $n < D$. Therefore, n constraints are implicitly assumed. Instead, this work does not impose constraints on the final state. In the first method, the Procrustean method results as consequence of a quadratic optimization problem. In the second one, it emerges after optimizing entanglement and using a single constraint.

We anticipate that this work may be useful for understanding how to concentrate entanglement efficiently in very large dimensions. As entanglement is a resource underlying many protocols in Quantum Information Science, we believe many people in the Quantum Information community may benefit from these findings.

Appendix A. Why Is It Necessary to Add a Difference?

In Section 3, we asserted that $[p(\vec{z})\mathcal{C}(\vec{z})]^2$ has its maximum when $z_m = 1$, $\forall\, m$. This means to keep the original state unaltered, without making any attempt to concentrate entanglement. In order to prove it, let us remember Equations (10) and (12). We may observe that

$$
\begin{aligned}
\delta &= [p_{\mathrm{S}}(\vec{z})\mathcal{C}(\vec{z})]^2\Big|_{z_m=1} - [p_{\mathrm{S}}(\vec{z})\mathcal{C}(\vec{z})]^2 \\
&= \frac{D}{D-1}\left(p_{\mathrm{S}}^2(\vec{z}) - \sum_{m=1}^{D} a_m^4 |z_m|^4 \right)\Bigg|_{\vec{z}}^{z_m=1} \\
&= \frac{D}{D-1}\sum_{m,n=1}^{D}(1-\delta_{mn})\left(1-|z_m|^2|z_n|^2\right)a_m^2 a_n^2 \\
&\geqslant 0,
\end{aligned}
$$

because $|z_m| \leqslant 1$. Thus, straight optimization of $p^2(\vec{z})\mathcal{C}^2(\vec{z})$ will suggest to *do nothing* and, instead, keep entanglement as it is. For this reason, it is necessary to add a reference level for entanglement. In other words, it is better to optimize $p^2(\vec{z})\big[\mathcal{C}^2(\vec{z}) - \mathcal{C}_{\mathrm{REF}}^2\big]$ rather than maximizing solely $p^2(\vec{z})\mathcal{C}^2(\vec{z})$ in order to actually increment entanglement.

Appendix B. Why Does a Diagonal Kraus Operator Suffice?

In Equation (7), we assumed $A_{\mathrm{S}}(\vec{z})$ to be diagonal in the $\{|m\rangle\}$ basis. This section will show why nondiagonal terms do not increase efficiency. Let us redefine A_{S} to be a general operator with components ζ_{mn}. We will add an additional definition. Let $\Pi(\zeta) = A_{\mathrm{S}}^\dagger A_{\mathrm{S}}$ be a positive operator whose matrix components are $\pi_{mn} = \sum_j \zeta_{jm}^* \zeta_{jn}$ and satisfy $\pi_{mn}^* = \pi_{nm}$ and $\pi_{jj} \geqslant 0$. If Π is known, then $A_{\mathrm{S}} = U\sqrt{\Pi}$, where U is an arbitrary unitary operator

whose explicit form depends on experimental details about the physical implementation of A_{S}

Now, considering that $A_{\mathrm{S}} = U\sqrt{\Pi}$, Equations (9) and (10) become

$$|\Psi(\zeta)\rangle_{12} = (U \otimes \mathbb{I}) \sum_{m=1}^{D} \frac{a_m}{\sqrt{p_{\mathrm{S}}(\zeta)}} \sqrt{\Pi}|m\rangle_1|m\rangle_2,$$

$$p_{\mathrm{S}}(\zeta) = \sum_{m=1}^{D} \pi_{mm} a_m^2,$$

and the efficiency function is written as

$$\mathcal{Q}(\zeta) = \frac{D}{D-1}\left[\mathscr{P}_{\mathrm{REF}}\left(\sum_{m=1}^{D} a_m^2 \pi_{mm}\right)^2 \right.$$
$$\left. - \sum_{m=1}^{D} a_m^4 \pi_{mm}^2 - \sum_{m \neq n}^{D} a_m^2 a_n^2 |\pi_{mn}|^2 \right] \tag{A1}$$

It can be seen that $\mathcal{Q}(\zeta)$ does not depend on U. In addition, the only positive term on the RHS of Equation (A1) depends on the diagonal components π_{mm}, whereas nondiagonal components only diminish the efficiency. Consequently, the optimal operator Π *must* be diagonal. This last condition can be satisfied, although not uniquely, by imposing A_{S} to be diagonal, so Equation (7) suffices to find the adequate operation to optimize the function $\mathcal{Q}$.

References

1. Horodecki, R.; Horodecki, P.; Horodecki, M.; Horodecki, K. Quantum Entanglement. *Rev. Mod. Phys.* **2009**, *81*, 865–942. [CrossRef]
2. Bennett, C.H. Quantum Cryptography Using Any Two Nonorthogonal States. *Phys. Rev. Lett.* **1992**, *68*, 3121–3124. [CrossRef] [PubMed]
3. Żukowski, M.; Zeilinger, A.; Horne, M.A.; Ekert, A.K. "Event-ready-detectors" Bell Experiment via Entanglement Swapping. *Phys. Rev. Lett.* **1993**, *71*, 4287–4290. [CrossRef] [PubMed]
4. Bennett, C.H.; Brassard, G.; Crépeau, C.; Jozsa, R.; Peres, A.; Wootters, W.K. Teleporting an Unknown Quantum State via Dual Classical and Einstein-Podolsky-Rosen Channels. *Phys. Rev. Lett.* **1993**, *70*, 1895–1899. [CrossRef]
5. Ekert, A.K. Quantum Cryptography Based on Bell's Theorem. *Phys. Rev. Lett.* **1991**, *67*, 661–663. [CrossRef]
6. Einstein, A.; Podolsky, B.; Rosen, N. Can Quantum-Mechanical Description of Physical Reality Be Considered Complete? *Phys. Rev.* **1935**, *47*, 777–780. [CrossRef]
7. Neves, L.; Solís-Prosser, M.A.; Delgado, A.; Jiménez, O. Quantum Teleportation via Maximum-Confidence Quantum Measurements. *Phys. Rev. A* **2012**, *85*, 062322. [CrossRef]
8. Solís-Prosser, M.A.; Delgado, A.; Jiménez, O.; Neves, L. Deterministic and Probabilistic Entanglement Swapping of Nonmaximally Entangled States Assisted by Optimal Quantum State Discrimination. *Phys. Rev. A* **2014**, *89*, 012337. [CrossRef]
9. Holevo, A.S.; Giovannetti, V. Quantum Channels and Their Entropic Characteristics. *Rep. Prog. Phys.* **2012**, *75*, 046001. [CrossRef]
10. Wilde, M.M. *Quantum Information Theory*, 2nd ed.; Cambridge University Press: Cambridge, UK, 2017. [CrossRef]
11. Yang, J.; Bao, X.H.; Zhang, H.; Chen, S.; Peng, C.Z.; Chen, Z.B.; Pan, J.W. Experimental Quantum Teleportation and Multiphoton Entanglement via Interfering Narrowband Photon Sources. *Phys. Rev. A* **2009**, *80*, 042321. [CrossRef]
12. Nielsen, M.A. Conditions for a Class of Entanglement Transformations. *Phys. Rev. Lett.* **1999**, *83*, 436–439. [CrossRef]
13. Lo, H.K.; Popescu, S. Concentrating Entanglement by Local Actions: Beyond Mean Values. *Phys. Rev. A* **2001**, *63*, 022301. [CrossRef]
14. Bennett, C.H.; Bernstein, H.J.; Popescu, S.; Schumacher, B. Concentrating Partial Entanglement by Local Operations. *Phys. Rev. A* **1996**, *53*, 2046–2052. [CrossRef] [PubMed]
15. Deutsch, D.; Ekert, A.; Jozsa, R.; Macchiavello, C.; Popescu, S.; Sanpera, A. Quantum Privacy Amplification and the Security of Quantum Cryptography over Noisy Channels. *Phys. Rev. Lett.* **1996**, *77*, 2818–2821. [CrossRef]
16. Pan, J.W.; Simon, C.; Brukner, Č.; Zeilinger, A. Entanglement Purification for Quantum Communication. *Nature* **2001**, *410*, 1067–1070. [CrossRef]
17. Bennett, C.H.; Brassard, G.; Popescu, S.; Schumacher, B.; Smolin, J.A.; Wootters, W.K. Purification of Noisy Entanglement and Faithful Teleportation via Noisy Channels. *Phys. Rev. Lett.* **1996**, *76*, 722–725. [CrossRef]
18. Plenio, M.; Virmani, S. An Introduction to Entanglement Measures. *Quantum Inf. Comput.* **2007**, *7*, 1–51. [CrossRef]

19. Pan, J.W.; Gasparoni, S.; Ursin, R.; Weihs, G.; Zeilinger, A. Experimental Entanglement Purification of Arbitrary Unknown States. *Nature* **2003**, *423*, 417–422. [CrossRef] [PubMed]
20. Vidal, G. Entanglement of Pure States for a Single Copy. *Phys. Rev. Lett.* **1999**, *83*, 1046–1049. [CrossRef]
21. Hardy, L. Method of Areas for Manipulating the Entanglement Properties of One Copy of a Two-Particle Pure Entangled State. *Phys. Rev. A* **1999**, *60*, 1912–1923. [CrossRef]
22. Zhao, Z.; Yang, T.; Chen, Y.A.; Zhang, A.N.; Pan, J.W. Experimental Realization of Entanglement Concentration and a Quantum Repeater. *Phys. Rev. Lett.* **2003**, *90*, 207901. [CrossRef]
23. Thew, R.T.; Munro, W.J. Entanglement Manipulation and Concentration. *Phys. Rev. A* **2001**, *63*, 030302. [CrossRef]
24. Yamamoto, T.; Koashi, M.; Imoto, N. Concentration and Purification Scheme for Two Partially Entangled Photon Pairs. *Phys. Rev. A* **2001**, *64*, 012304. [CrossRef]
25. Zhao, Z.; Pan, J.W.; Zhan, M.S. Practical Scheme for Entanglement Concentration. *Phys. Rev. A* **2001**, *64*, 014301. [CrossRef]
26. Kwiat, P.G.; Barraza-Lopez, S.; Stefanov, A.; Gisin, N. Experimental Entanglement Distillation and 'Hidden' Non-Locality. *Nature* **2001**, *409*, 1014–1017. [CrossRef] [PubMed]
27. Morikoshi, F. Recovery of Entanglement Lost in Entanglement Manipulation. *Phys. Rev. Lett.* **2000**, *84*, 3189–3192. [CrossRef] [PubMed]
28. Morikoshi, F.; Koashi, M. Deterministic Entanglement Concentration. *Phys. Rev. A* **2001**, *64*, 022316. [CrossRef]
29. Chefles, A. Unambiguous Discrimination between Linearly Independent Quantum States. *Phys. Lett. A* **1998**, *239*, 339–347. [CrossRef]
30. Hayashi, M.; Koashi, M.; Matsumoto, K.; Morikoshi, F.; Winter, A. Error Exponents for Entanglement Concentration. *J. Phys. A Math. Gen.* **2003**, *36*, 527–553. [CrossRef]
31. Gu, Y.J.; Li, W.D.; Guo, G.C. Protocol and Quantum Circuits for Realizing Deterministic Entanglement Concentration. *Phys. Rev. A* **2006**, *73*, 022321. [CrossRef]
32. Hu, X.M.; Huang, C.X.; Sheng, Y.B.; Zhou, L.; Liu, B.H.; Guo, Y.; Zhang, C.; Xing, W.B.; Huang, Y.F.; Li, C.F.; et al. Long-Distance Entanglement Purification for Quantum Communication. *Phys. Rev. Lett.* **2021**, *126*, 010503. [CrossRef] [PubMed]
33. Ecker, S.; Sohr, P.; Bulla, L.; Huber, M.; Bohmann, M.; Ursin, R. Experimental Single-Copy Entanglement Distillation. *Phys. Rev. Lett.* **2021**, *127*, 040506. [CrossRef]
34. Ecker, S.; Sohr, P.; Bulla, L.; Ursin, R.; Bohmann, M. Remotely Establishing Polarization Entanglement Over Noisy Polarization Channels. *Phys. Rev. Appl.* **2022**, *17*, 034009. [CrossRef]
35. Huang, C.X.; Hu, X.M.; Liu, B.H.; Zhou, L.; Sheng, Y.B.; Li, C.F.; Guo, G.C. Experimental One-Step Deterministic Polarization Entanglement Purification. *Sci. Bull.* **2022**, *67*, 593–597. [CrossRef] [PubMed]
36. Marques, B.; Matoso, A.A.; Pimenta, W.M.; Gutiérrez-Esparza, A.J.; Lima, G.; Neves, L.; Delgado, A.; Saavedra, C.; Pádua, S. Optimal Entanglement Concentration for Photonic Qutrits Encoded in Path Variables. *Phys. Rev. A* **2013**, *87*, 052327. [CrossRef]
37. Smolin, J.A.; Verstraete, F.; Winter, A. Entanglement of Assistance and Multipartite State Distillation. *Phys. Rev. A* **2005**, *72*, 052317. [CrossRef]
38. Groisman, B.; Linden, N.; Popescu, S. Entanglement Concentration of Three-Partite States. *Phys. Rev. A* **2005**, *72*, 062322. [CrossRef]
39. Gómez, S.; Mattar, A.; Gómez, E.S.; Cavalcanti, D.; Farías, O.J.; Acín, A.; Lima, G. Experimental nonlocality-based randomness generation with nonprojective measurements. *Phys. Rev. A* **2018**, *97*, 040102. [CrossRef]
40. Gómez, S.; Mattar, A.; Machuca, I.; Gómez, E.S.; Cavalcanti, D.; Farías, O.J.; Acín, A.; Lima, G. Experimental investigation of partially entangled states for device-independent randomness generation and self-testing protocols. *Phys. Rev. A* **2019**, *99*, 032108. [CrossRef]
41. Martínez, D.; Gómez, E.S.; Cariñe, J.; Pereira, L.; Delgado, A.; Walborn, S.P.; Tavakoli, A.; Lima, G. Certification of a non-projective qudit measurement using multiport beamsplitters. *Nat. Phys.* **2023**, *19*, 190–195. [CrossRef]
42. Rungta, P.; Bužek, V.; Caves, C.M.; Hillery, M.; Milburn, G.J. Universal State Inversion and Concurrence in Arbitrary Dimensions. *Phys. Rev. A* **2001**, *64*, 042315. [CrossRef]
43. Vedral, V.; Plenio, M.B.; Rippin, M.A.; Knight, P.L. Quantifying Entanglement. *Phys. Rev. Lett.* **1997**, *78*, 2275–2279. [CrossRef]
44. Grobe, R.; Rzazewski, K.; Eberly, J.H. Measure of Electron-Electron Correlation in Atomic Physics. *J. Phys. B At. Mol. Opt. Phys.* **1994**, *27*, L503–L508. [CrossRef]
45. Law, C.K.; Eberly, J.H. Analysis and Interpretation of High Transverse Entanglement in Optical Parametric Down Conversion. *Phys. Rev. Lett.* **2004**, *92*, 127903. [CrossRef]
46. Fedorov, M.V.; Efremov, M.A.; Kazakov, A.E.; Chan, K.W.; Law, C.K.; Eberly, J.H. Packet Narrowing and Quantum Entanglement in Photoionization and Photodissociation. *Phys. Rev. A* **2004**, *69*, 052117. [CrossRef]
47. Brida, G.; Caricato, V.; Fedorov, M.V.; Genovese, M.; Gramegna, M.; Kulik, S.P. Characterization of Spectral Entanglement of Spontaneous Parametric-down Conversion Biphotons in Femtosecond Pulsed Regime. *EPL (Europhys. Lett.)* **2009**, *87*, 64003. [CrossRef]
48. Di Lorenzo Pires, H.; Monken, C.H.; van Exter, M.P. Direct Measurement of Transverse-Mode Entanglement in Two-Photon States. *Phys. Rev. A* **2009**, *80*, 022307. [CrossRef]
49. Straupe, S.S.; Ivanov, D.P.; Kalinkin, A.A.; Bobrov, I.B.; Kulik, S.P. Angular Schmidt Modes in Spontaneous Parametric Down-Conversion. *Phys. Rev. A* **2011**, *83*, 060302. [CrossRef]

50. Just, F.; Cavanna, A.; Chekhova, M.V.; Leuchs, G. Transverse Entanglement of Biphotons. *New J. Phys.* **2013**, *15*, 083015. [CrossRef]
51. Gómez, E.S.; Riquelme, P.; Solís-Prosser, M.A.; González, P.; Ortega, E.; Xavier, G.B.; Lima, G. Tunable Entanglement Distillation of Spatially Correlated Down-Converted Photons. *Opt. Express* **2018**, *26*, 13961. [CrossRef]
52. Gu, Y.J.; Zheng, Y.Z.; Guo, G.C. Conclusive Teleportation and Entanglement Concentration. *Phys. Lett. A* **2002**, *296*, 157–160. [CrossRef]
53. Yang, M.; Delgado, A.; Roa, L.; Saavedra, C. Entanglement Concentration for Non-Maximally Entangled States of Qudits. *Opt. Commun.* **2009**, *282*, 1482–1487. [CrossRef]
54. Vidal, G.; Jonathan, D.; Nielsen, M.A. Approximate Transformations and Robust Manipulation of Bipartite Pure-State Entanglement. *Phys. Rev. A* **2000**, *62*, 012304. [CrossRef]
55. Solís-Prosser, M.A.; Delgado, A.; Jiménez, O.; Neves, L. Parametric Separation of Symmetric Pure Quantum States. *Phys. Rev. A* **2016**, *93*, 012337. [CrossRef]
56. The MathWorks Inc. Quadratic Programming-MATLAB Quadprog Documentation. Available online: https://www.mathworks.com/help/optim/ug/quadprog.html (accessed on 9 April 2023).