Linear-optics realization of channels for single-photon multimode qudits

Marco Piani,[1] David Pitkanen,[1] Rainer Kaltenbaek,[1,2] and Norbert Lütkenhaus[1]

[1]*Institute for Quantum Computing & Department of Physics and Astronomy, University of Waterloo, 200 University Avenue West, Waterloo, Ontario, Canada N2L 3G1*

[2]*Vienna Center for Quantum Science and Technology, Faculty of Physics, University of Vienna, Boltzmanngasse 5, A-1090 Vienna, Austria*

We propose and theoretically study a method for the stochastic realization of arbitrary quantum channels on multimode single-photon qudits. In order for our method to be undemanding in its implementation, we restrict our analysis to linear-optical techniques, vacuum ancillary states, and nonadaptive schemes, but we allow for random switching between different optical networks. With our method it is possible to deterministically implement random-unitary channels and to stochastically implement general channels. We provide an expression for the optimal probability of success of our scheme and calculate this quantity for specific examples such as the qubit amplitude-damping channel. The success probability is shown to be related to the entanglement properties of the Choi-Jamiołkowski state isomorphic to the channel.

I. INTRODUCTION

The most general transformation a quantum state can undergo is described by a quantum channel. For example, it may correspond to a controlled manipulation of a quantum system for some final aim—as in quantum-information-processing protocols [1]—or it may represent an unwanted interaction with the environment. While in the first case implementing the respective quantum channel is of direct practical interest, in the second case one may still be interested in the implementation of the channel for the sake of understanding the role of noise, and how to counteract it, in real-world implementations of quantum-information-processing protocols. It is worth remarking that striking effects in quantum-information processing (QIP), e.g., the superactivation of the quantum capacity of channels [2], involve nontrivial noisy channels.

Quantum optics is one of the best established physical architectures for QIP [3–7]. It has the advantage that the carriers of information—photons—interact naturally weakly with the environment, so that *real* noise is low. This makes *simulating* noise possible in a very controlled way. The workhorse of optical experiments is the manipulation via linear-optical elements, such as beam splitters and phase shifters. The linear-optics realization of channels has been investigated in general [8] as well as in specific cases; for example, random-unitary channels are common in experiments on decoherence-free and unitarily recoverable subspaces [9,10] and in the realization of mixed states [11,12]. The simplest nontrivial example of a channel that is not random unitary is perhaps given by the qubit amplitude-damping channel [1]. The counting statistics of this channel have been simulated using linear optics [13], and a stochastic linear-optical implementation with a fixed success probability of 50%, independent of the value of the damping parameter, has been suggested in [14] and realized in [15].

He *et al.* [8] proposed a way to implement an arbitrary single Kraus operator on a single-photon qudit. They proceeded to show how it is then possible to implement a generalized measurement, that is, a POVM. He *et al.* argued that their method allows for the stochastic realization of a general channel, but did not address the issue of the probability of success of such a realization. Actually, the scheme proposed in [8] would allow for the *deterministic* realization of an

arbitrary channel, if multimode nondemolition measurements were allowed. Unfortunately, such measurements are experimentally challenging and not amenable to a linear-optics realization.

In this article, we propose a linear-optics scheme for the stochastic exact realization of an arbitrary channel for single-photon multimode qudits. In contrast to He *et al.*, we prove that, under constraints motivated by the ease of experimental realization, our scheme achieves an optimal probability of success. An interesting result is that such a success probability is related to the entanglement properties of the Choi-Jamiołkowski state isomorphic to the channel [16,17]. This connection allows us to apply results in entanglement theory [18] to the quite different problem of channel realization.

Our results provide an optimal strategy for the realization of arbitrary channels, an important building block in experimental studies of QIP. In the specific case of the qubit amplitude-damping channel, our scheme provides a significantly higher efficiency than alternative schemes [14] without leaving the subspace of the encoding of the input state. In contrast to [13], this allows us to further process the output of the channel.

The paper is structured as follows. In Sec. II, we provide definitions, fixing both the framework and the notation. In Sec. III, we illustrate in detail the problem we consider, that is, the realization of a quantum channel with a fixed set of tools. In Sec. IV, we provide a scheme to realize any channel perfectly albeit only stochastically. In Sec. V, we relate the optimal success probability of the method proposed to the entanglement properties of the Choi-Jamiołkowski state isomorphic to the channel of interest. In Sec. VI, we use this relation to provide bounds on the probability of success, both in the specific case of qubits, for which we are able to give analytic bounds, and qudits. In Sec. VII, we apply our technique to two examples, one being the qubit amplitude-damping channel. Finally, we conclude and discuss possible future venues to investigate.

II. DEFINITIONS AND FRAMEWORK

The state of a quantum system may change over time due to some internal dynamics, to an interaction with its environment,

or to a measurement performed on it by an observer. Any physical transformation a quantum system can experience can be modeled as a quantum channel $\Lambda : \rho_{\text{in}} \mapsto \rho_{\text{out}}$. Every channel acting on a system S admits a dilation, which means that it can be realized as some unitary interaction with an ancilla E, which is subsequently discarded:

$$\Lambda[\rho_S] = \text{Tr}_E(U_{SE}\,\rho_S \otimes \sigma_E\,U_{SE}^\dagger),$$

with σ_E the initial state of the ancilla [1]. More abstractly a quantum channel can be defined as a completely positive trace-preserving linear map. Each channel can be represented in the form $\Lambda[\rho] = \sum_i A_i \rho A_i^\dagger$, where $\{A_i\}$ is a set of *Kraus operators* fulfilling the trace-preserving condition, $\sum_i A_i^\dagger A_i = I$. The Kraus representation of a channel is not unique. For instance, if $\{A_i\}$ forms a *Kraus decomposition* of a channel Λ, the relation $B_i = \sum_j u_{ij} A_j$, assuming u_{ij} are the elements of a unitary matrix, will define a new decomposition $\{B_i\}$ for Λ [1].

We will frequently find the notion of operator norm useful in our discussions of quantum channels. Since we will always work in finite dimensions, the operator norm $\|A\|_\infty$ of A corresponds to the largest singular value of A. An operator is an admissible Kraus operator—that is, it can be considered as part of some valid Kraus-operator set—as long as $\|A\|_\infty \leqslant 1$. Any set of linear operators that satisfy the completion relation $\sum_{i=1}^k A_i^\dagger A_i = I$ will constitute a valid quantum channel.

In this paper, we will be interested in optical quantum systems. Each mode of an optical system is associated to a basis of Fock states $|n\rangle$, where $n = 0,1,2\ldots$ denotes the number of photons in the mode. The creation and annihilation operators, $a^\dagger$ and a, respectively, provide a convenient notational framework for describing Fock states because of the relations $a|n\rangle = \sqrt{n}|n-1\rangle$, $a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle$, so that $|n\rangle = (a^\dagger)^n/\sqrt{n!}|0\rangle$. These operators have commutation relations $[a_i,a_j^\dagger] = \delta_{ij}$, $[a_i^\dagger,a_j^\dagger] = 0$, and $[a_i,a_j] = 0$, where the indices i and j denote the optical mode and δ_{ij} is the Kronecker delta.

The most common optical elements that are used in experiments for the manipulation of optical modes are beam splitters and phase shifters. Optical networks that are composed only of instances of these two elements are referred to as passive linear devices. Linear (quantum) optics is the part of quantum optics that, apart from the initial generation of entangled photon pairs and single-photon detection, deals only with passive linear devices [3]. Any unitary transformation U acting on d optical modes and preserving the total photon number can conveniently be described by the way it transforms the creation operators of the modes:

$$a_i^{\text{out}\dagger} = \sum_j u_{ij} a_j^{\text{in}\dagger}, \tag{1}$$

where u_{ij} are the elements of a unitary matrix. A transformation can be realized by linear optics if and only if it is of this kind [19].

A phase shifter is an optical element that acts on a single mode as $U a^\dagger U^\dagger = e^{i\phi}a^\dagger$. A beam splitter acts on two optical modes at a time and can be described by

$$(u_{ij}) = \begin{pmatrix} \cos\theta & -e^{i\phi}\sin\theta \\ e^{-i\phi}\sin\theta & \cos\theta \end{pmatrix}. \tag{2}$$

Any unitary that acts on d modes preserves the total photon number if and only if it can be implemented using these two devices [19].

III. THE PROBLEM

A qudit can be encoded by using one photon in d optical modes. An arbitrary logical state can be written as $|\psi\rangle = \sum_{i=1}^d \psi_i |i_L\rangle$, with a logical basis $\{|i_L\rangle\}_{i=1}^d$, where $|i_L\rangle = a_i^\dagger|0\rangle$. We call this kind of encoding a *d-rail encoding*. This encoding is convenient when the interactions are limited to linear optics, because any unitary operation (1) can be performed on the creation operators using linear optics, and under this encoding the basis states of a single qudit and the creation operators transform identically.

We are interested in the simulation of an arbitrary quantum channel Λ that acts on a qudit, using only passive linear optics. What we want is a realization of Λ on the d-rail qudit, such that the logical subspace—the encoding—is mapped onto itself. This allows for further processing of the output of the channel.

We will refer to the channel to be realized as the *logical channel*, to distinguish it from physical channels that evolve the state of the modes without necessarily preserving the logical subspace.

As we noted earlier, we can always represent a channel in the form of a dilation where the channel is realized via the unitary interaction of the system with ancillary modes. We will limit ourselves to linear-optics evolution. For the sake of the ease of experimental implementation, we will assume several other reasonable restrictions: (i) to limit the number of photons that need to be generated, we only introduce ancillary modes that are initially in the vacuum state; (ii) in order to prevent the necessity of using expensive feed-forward mechanisms (Pockels cells and high-speed high-voltage switches—see, e.g., [7,20]), we do not allow adaptive schemes; (iii) we will restrict ourselves to photon-number measurements, although it will actually turn out that commonly used threshold detectors suffice.

When we consider the dilation representation of a channel, we can imagine that the final trace over the ancillary space corresponds to a measurement of the ancilla, whose result is discarded. If we assume that the ancilla starts in the (vacuum) state $|0\rangle_E$, then we have

$$\begin{aligned} \Lambda[\rho_S] &= \text{Tr}_E(U_{SE}\rho_S \otimes \sigma_E U_{SE}^\dagger) \\ &= \sum_k \text{Tr}_E(U_{SE}\rho_S \otimes \sigma_E U_{SE}^\dagger M_E^k) \\ &= \sum_{jk} \left(\langle j|_E \sqrt{M_k^E}\,U_{SE}|0\rangle_E\right)\rho_S\left(\langle j|_E \sqrt{M_k^E}\,U_{SE}|0\rangle_E\right)^\dagger \end{aligned}$$

with $\{M_k\}$, $M_k \geqslant 0$, $\sum_k M_k = I$ a POVM on the ancilla system E, and $\{|j\rangle\}$ an orthonormal basis for E. With our constraints—vacuum input ancillas and linear-optics evolution—measuring the vacuum on the output ancillas is the only result that leaves the system within the encoding. This can be seen easily considering the action of the linear-optics unitary $U_{SE} = U_{LO}$ on initial states $|i_L\rangle|0\rangle_E$, $i = 1,\ldots,d$. We will consider $d + e$

modes, with the first d used for the encoding, and the remaining e constituting the ancilla system E. Then we have

$$U_{LO}|i_L\rangle_S|0\rangle_E = U_{LO} a_i^\dagger |0\rangle_S |0\rangle_E = \sum_{j=1}^{d+e} u_{ij} a_j^\dagger |0\rangle_S |0\rangle_E$$

$$= \left(\sum_{j=1}^{d} u_{ij}|j_L\rangle \right) |0\rangle_E$$

$$+ |0\rangle_S \left(\sum_{j=d+1}^{d+e} u_{ij} a_j^\dagger |0\rangle_E \right).$$

From this expression it is evident that if we perform a photon-number measurement on the output ancillary modes and we obtain a result different from the vacuum, then the encoding is lost. The reason for this is that linear optics preserves the photon number, and the initial state of the system $|i_L\rangle|0\rangle_E$ has only one photon in it. If the photon is measured in the ancilla, then the initial state $|i_L\rangle$ will be mapped out of the encoding to the vacuum, independently of which output ancilla mode the photon is measured in.

Therefore, under the constraints that we have imposed, the only logical channels that can be realized deterministically must have a single Kraus operator. Such channels are necessarily unitary transformations, as can be seen by the trace-preservation condition $A^\dagger A = I$.

IV. THE SOLUTION: STOCHASTIC IMPLEMENTATION

In this section, we will first see that *any* single logical Kraus operator (i.e., any Kraus operator of the logical channel) can be realized stochastically. Later we will introduce a further resource, randomness, and the ability to switch—according to such randomness—among different optical networks, and we will show that then any logical channel can be realized, albeit only stochastically.

A. Implementation of a logical Kraus operator

For any logical Kraus operator A that we want to apply to the input state, it is possible to construct an optical network such that A will correspond to the transformation of the logical state if the output ancillary modes are detected to be in the vacuum state, given that they were in the vacuum state before the channel. Every Kraus operator has a singular value decomposition $A = VSU$, where U and V are unitaries and the matrix S is positive and diagonal, with diagonal elements $0 \leqslant s_i \leqslant 1$ that correspond to the singular values of A. As unitary rotations can be realized deterministically on the encoding, in order to prove that A can be realized under our constraints, it is sufficient to prove that any diagonal matrix S can be realized (see Fig. 1).

This is proven possible by considering the action (2) of a beam splitter on two modes. If the first mode, with creation operator $a^\dagger$, belongs to the encoding and the second mode is an ancilla—which means it starts in the vacuum—then the transformation that results when the vacuum is measured on the ancilla state effectively realizes the mapping $a^\dagger \mapsto \cos\theta a^\dagger$. Since the angle θ is arbitrary, we can simply implement any
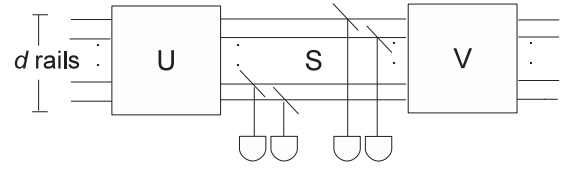

FIG. 1. The diagram describes an optical circuit for a channel that would realize the Kraus operator $A = VSU$ as well as a Kraus operator that would map the encoding to the vacuum. The boxes represent optical arrays that perform the unitary that labels them. The S transformation then consists of a set of beam splitters, one for each mode, whose transmission coefficients are matched to the singular values of the matrix S.

diagonal logical Kraus operator S by using d ancillary modes and d beam splitters, choosing the angles θ_i such that $s_i = \cos\theta_i$.

A proposal for a similar implementation of a single Kraus operator can be found in [8] and [21].

B. Perfect but stochastic implementation of an arbitrary logical channel

A logical channel Λ that we may want to apply on the encoding will in general have a Kraus decomposition $\{A_i\}_{i=1}^n$, with $n \geqslant 1$. Therefore, by using a fixed linear-optical network in the framework defined in Sec. III it will not be possible in general to simulate the channel, as only one logical Kraus operator can be realized per fixed optical network.

We will circumvent this problem by realizing individually the various Kraus operators A_i, $i = 1, \ldots, n$, in this way being able to preserve the encoding for each A_i. Roughly speaking, by randomly applying the different Kraus operators the logical channel Λ will be realized. Of course, this is possible only by allowing the linear-optical network to change. We will introduce the possibility of switching among various optical networks—one for each A_i—according to a probability distribution $\{p_i\}$. Each fixed optical network that we will introduce to realize the Kraus operator A_i will itself correspond to a quantum channel Γ_i (see Fig. 2). This "average realization" of the logical channel will anyway be stochastic, because in the implementation of any A_i that is not unitary there will necessarily be a finite probability of ending up outside the encoding, which corresponds to finding the input photon in the output ancillary modes.

One important point is that, given the additional degree of freedom due to the choice of the probability distribution $\{p_i\}$, it is possible to consider the realization of a rescaled version $\tilde{A}_i$ of A_i rather than exactly A_i. Of course each $\tilde{A}_i$ must be a valid Kraus operator, i.e., $\|\tilde{A}_i\|_\infty \leqslant 1$. We will use this rescaling degree of freedom to maximize the success probability for the realization of the channel.

If we postselect on finding the output ancillary modes in the vacuum state, and if we choose the probability distribution $\{p_i\}$ and the $\tilde{A}_i$ operators such that

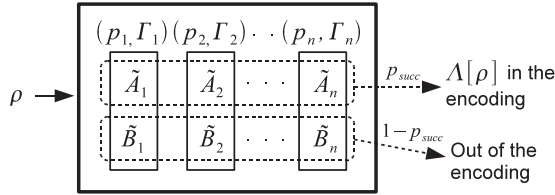$$\sqrt{p_i}\tilde{A}_i = \sqrt{p_{\text{succ}}}A_i, \tag{3}$$

FIG. 2. Pictorial representation of our scheme. Each solid rectangle represents a channel. The most external box is a mixture of the n channels Γ_i inside of it. Each of these inner channels corresponds to a linear-optics setup and for our scope its action on the encoding can be completely described without loss of generality by two Kraus operators, $\tilde{A}_i$ and $\tilde{B}_i$. Each $\tilde{A}_i$ preserves the d-rail encoding, while the Kraus operators $\tilde{B}_i$ map an encoded state out of the encoding. If the condition $\sqrt{p_i}\tilde{A}_i = \sqrt{p_{\text{succ}}}A_i$, for all i, is met, the overall result of randomly switching among the channels Γ_i according to the probability distribution $\{p_i\}$ is that of realizing the target logical channel Λ with probability p_{succ} (independent of the input ρ).

for all i and for some $0 \leqslant p_{\text{succ}} \leqslant 1$, then the logical input state ρ will be mapped into the (unnormalized) logical state

$$\sum_i p_i \tilde{A}_i \rho \tilde{A}_i^\dagger = p_{\text{succ}} \sum_i A_i \rho A_i^\dagger.$$

This will happen with probability

$$\text{Tr}\left(\sum_i p_i \tilde{A}_i \rho \tilde{A}_i^\dagger\right) = \sum_i p_i \text{Tr}(\tilde{A}_i \rho \tilde{A}_i^\dagger) = p_{\text{succ}},$$

with:

(1) $\tilde{A}_i$ proportional to the logical Kraus operator A_i through (3);

(2) p_i the probability of implementing the optical network corresponding to $\tilde{A}_i$;

(3) $\text{Tr}(\tilde{A}_i \rho \tilde{A}_i^\dagger) = \text{Tr}(\rho \tilde{A}_i^\dagger \tilde{A}_i)$ the probability of realizing, without leaving the encoding, the Kraus operator $\tilde{A}_i$ for each random choice of optical network.

Thus the logical channel Λ will be stochastically implemented with probability p_{succ} (independent of the input ρ).

Given that we want the channel to be realized perfectly, the figure of merit we care about is the probability of success p_{succ}, which we want to be maximal. One possible choice for the distribution $\{p_i\}$ and the operators $\tilde{A}_i$ is trivially $p_i = 1/n$ and $\tilde{A}_i = A_i$; this choice leads to a probability of success $p_{\text{succ}} = 1/n$. This strategy is independent of the properties of the Kraus operator $\{A_i\}$ for the particular channel Λ, and depends only on the number of Kraus operators. As such, one can expect it to be nonoptimal, and it certainly is in the case of a random-unitary channel

$$\Lambda[\rho] = \sum_i q_i U_i \rho U_i^\dagger,$$

with $\{U_i\}$ unitaries and $\{q_i\}$ a probability distribution. Indeed, in this case an obvious better choice—and actually optimal—is $p_i = q_i$, $\tilde{A}_i = U_i$, for all i, so that $p_{\text{succ}} = 1$.

The following theorem provides the optimal choice of the probability distribution $\{p_i\}$ and of the operators $\tilde{A}_i$ to maximize p_{succ}, for any fixed Kraus decomposition $\{A_i\}$.

Theorem 1. Given the Kraus decomposition $\{A_i\}$ for the channel Λ, the optimal probability of success for its realization is

$$p_{\text{succ}}(\{A_i\}) = \frac{1}{\sum_i \|A_i\|_\infty^2}. \tag{4}$$

This can be achieved by the choice $p_i = \frac{\|A_i\|_\infty^2}{\sum_j \|A_j\|_\infty^2}$ and $\tilde{A}_i = \frac{1}{\|A_i\|_\infty} A_i$, for all i.

Proof. From the condition $\sqrt{p_i}\tilde{A}_i = \sqrt{p_{\text{succ}}}A_i$, for all i, one finds $p_i \geqslant p_i \|\tilde{A}_i\|_\infty^2 = p_{\text{succ}}\|A_i\|_\infty^2$, where we used the fact that $\|\tilde{A}_i\|_\infty \leqslant 1$, because each $\tilde{A}_i$ must be a proper Kraus operator. Summing over i and using $\sum_i p_i = 1$, one arrives at $p_{\text{succ}} \leqslant 1/\sum_i \|A_i\|_\infty^2$. The probability distribution and the Kraus operators in the statement of the theorem saturate the inequality. ∎

Thus, the maximal probability of simulating the channel adopting the Kraus decomposition $\{A_i\}$ in our scheme is the inverse of $\sum_i \|A_i\|_\infty^2$. This quantity will in general depend on the specific Kraus decomposition. By optimizing over all Kraus decompositions we have the following.

Corollary 1 (optimal probability of success). In our scheme, the optimal probability of success in the implementation of Λ is

$$p_{\text{succ}}(\Lambda) = \max_{\{A_i\}} \frac{1}{\sum_i \|A_i\|_\infty^2}, \tag{5}$$

where the maximization is over all Kraus decompositions $\{A_i\}$ of the channel Λ.

For convenience in the analysis to follow, we define the *stochasticity* of a channel as

$$\sigma(\Lambda) = \min_{\{A_i\}} \sum_i \|A_i\|_\infty^2, \tag{6}$$

where the minimization is over all Kraus decompositions $\{A_i\}$ of the channel Λ, so that

$$p_{\text{succ}}(\Lambda) = \frac{1}{\sigma(\Lambda)}.$$

The name "stochasticity" is justified by the fact that the larger $\sigma(\Lambda)$, the lower the probability of a successful realization of the channel. We remark that any specific Kraus decomposition will give an upper (lower) bound on the stochasticity (optimal probability of success).

As mentioned, the channel implementation we propose requires switching between different linear-optical networks, each corresponding to a single Kraus operator. This may appear hard to implement, especially given that in the realization of a single Kraus operator proposed in Sec. IV A we used beam splitters whose transmittivity depended on the Kraus operator. Fortunately, it follows immediately from the work of Reck *et al.* [19] that the switching can always be realized by adjusting only optical phases in a multiport linear network.

V. RELATION WITH ENTANGLEMENT MEASURES

The optimal success probability $p_{\text{succ}}(\Lambda)$ of implementing a channel is clearly just a property of the channel itself. Therefore it appears natural to look for a representation of the channel that is independent of any specific Kraus

decomposition. This can be done by considering the Choi-Jamiołkowski isomorphism.

The latter is a one-to-one mapping between maps and operators [16,17]. The isomorphism—explicitly in the direction from maps to operators—is defined as

$$J(\Lambda) = (\Lambda \otimes I)[\psi_d^+] = \frac{1}{d} \sum_{i,j=1}^{d} \Lambda(|i\rangle\langle j|) \otimes |i\rangle\langle j|, \quad (7)$$

for some fixed choice of a maximally entangled state

$$|\psi_d^+\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^{d} |i\rangle|i\rangle. \qquad (8)$$

For our purpose, the interesting observation is that pure ensemble decompositions $\{p_i, \psi_i\}$ of the Choi-Jamiołkowski state J_Λ isomorphic to a channel Λ are in one-to-one correspondence with Kraus decompositions of Λ. This follows from the the fact that for any vector $|\bar{\psi}\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$ there is an operator $A^{\bar{\psi}}$ such that

$$|\bar{\psi}\rangle = (A^{\bar{\psi}} \otimes I)|\psi_d^+\rangle. \qquad (9)$$

Here, the bar in $|\bar{\psi}\rangle$ denotes that the vector need not be normalized. In general, unless it is specified to the contrary with use of the bar notation, all states $|\psi\rangle$ are assumed to be normalized. Thus,

$$\begin{aligned} J(\Lambda) &= \sum_i p_i |\psi_i\rangle\langle\psi_i| = \sum_i |\bar{\psi}_i\rangle\langle\bar{\psi}_i| \\ &= \sum_i (A^{\bar{\psi}_i} \otimes I)|\psi_d^+\rangle\langle\psi_d^+|(A^{\bar{\psi}_i\dagger} \otimes I), \end{aligned} \qquad (10)$$

for $|\bar{\psi}_i\rangle = \sqrt{p_i}|\psi_i\rangle = (A^{\bar{\psi}_i} \otimes I)|\psi_d^+\rangle$.

As we have seen, without the use of randomness the only channels that can be realized deterministically are unitaries. Using randomness and switching among optical networks we can extend this result only to random unitaries. Thus, we have that the only channels that can be realized deterministically in our framework are those whose Choi-Jamiołkowski state admits an ensemble consisting only of maximally entangled states.

One then expects that channels whose probability of realization is high admit Kraus decompositions that are close to random unitary. In turn this would mean that their Choi-Jamiołkowski states admit ensemble decompositions that are highly entangled. We will show that this intuition is correct.

The relation (9) implies

$$\|A^{\bar{\psi}}\|_\infty^2 = d\lambda_{\max}(\bar{\psi}), \qquad (11)$$

if we consider the Schmidt decomposition $|\bar{\psi}\rangle = \sum_i \sqrt{\lambda_i}|i\rangle|i\rangle$, with $\lambda_i \geqslant 0$, $\sum_i \lambda_i = \langle\bar{\psi}|\bar{\psi}\rangle$, and $\lambda_{\max} = \max_i\{\lambda_i\}$. Thus, we find for the stochasticity

$$\sigma(\Lambda) = \min_{\{A_i\}} \sum_i \|A_i\|_\infty^2 \qquad (12a)$$

$$= d \min_{\{p_i,\psi_i\}} \sum_i p_i \lambda_{\max}(\psi_i) \qquad (12b)$$

$$= d\left(1 - \max_{\{p_i,\psi_i\}} \sum_i p_i[1 - \lambda_{\max}(\psi_i)]\right) \qquad (12c)$$

$$= d\left(1 - \max_{\{p_i,\psi_i\}} \sum_i p_i E_G(\psi_i)\right), \qquad (12d)$$

where we used (11) to move from the minimization over Kraus decompositions for Λ to the minimization over ensemble decompositions for $J(\Lambda)$. The quantity

$$E_G(\psi) = 1 - \lambda_{\max}(\psi) = 1 - \max_{\alpha,\beta} |\langle\alpha,\beta|\psi\rangle|^2,$$

where the maximum is taken with respect to factorized pure states $|\alpha,\beta\rangle = |\alpha\rangle|\beta\rangle$, is the *geometric measure of entanglement* for a bipartite pure state [22]. More generally, for a multipartite pure state, the geometric measure of entanglement is defined as $E_G(\psi) = 1 - \max_{\phi_{\mathrm{sep}}} |\langle\phi_{\mathrm{sep}}|\psi\rangle|^2$, with ϕ_{sep} a fully separable state. In the bipartite case, it coincides with the entanglement measure E_2, which was defined in [23] as one of a whole family of entanglement measures. The geometric measure of entanglement has received a good deal of attention [24,25] because of its intuitive—even in the multipartite case—geometric interpretation as maximal overlap of the state of interest with a fully separable state, and because of its connections to other well-known entanglement measures, such as relative entropy of entanglement [26,27]. In the bipartite qudit case we are interested in here, one sees immediately that

$$0 \leqslant E_G(\psi) \leqslant 1 - \frac{1}{d}. \qquad (13)$$

The lower bound is achieved for a factorized pure state, while the upper bound corresponds to a maximally entangled state such as the one in Eq. (8).

The geometric measure of entanglement is extended to the mixed-state case by the usual convex-roof construction [28]:

$$E_G^\cup(\rho) = \min_{\{p_i,\psi_i\}} \sum_i p_i E_G(\psi_i), \qquad (14)$$

where we use $\cup$ to stress that the resulting quantity is convex on the set of mixed states.

The standard convex-roof is defined in terms of the ensemble containing, on average, the *minimum* amount of entanglement as quantified, in this case, by the geometric measure of entanglement for pure states. Equation (12) involves instead the ensemble containing on average the *maximum* amount of entanglement. This corresponds to the *concave*-roof construction

$$E_G^\cap(\rho) = \max_{\{p_i,\psi_i\}} \sum_i p_i E_G(\psi_i), \qquad (15)$$

where we use $\cap$ to stress that in this way we are defining a concave function on the set of mixed states.

For the sake of comparison with quantities better known in the literature, let us mention that in the same way in which the *entanglement of formation* [29] $E_F(\rho^{AB}) = \min_{\{p_i,\psi_i^{AB}\}} \sum_i p_i S(\rho_i^A)$, with $\rho^A = \mathrm{Tr}_B(|\psi^{AB}\rangle\langle\psi^{AB}|)$ and $S(\sigma) = -\mathrm{Tr}(\sigma \log_2 \sigma)$ the von Neumann entropy of a state σ, is the paradigmatic example of convex roof construction, the *entanglement of assistance* [30]

$$E_a = \max_{\{p_i,\psi_i^{AB}\}} \sum_i p_i S(\rho_i^A) \qquad (16)$$

is the paradigmatic example of concave-roof construction.

From (12) it follows that the stochasticity is given by

$$\sigma(\Lambda) = d\big(1 - E_G^\cap(J(\Lambda))\big), \tag{17}$$

and, as a result, the relation between the probability of success $p_{\text{succ}}(\Lambda)$ for our scheme to realize a channel Λ and the entanglement properties of the related Choi-Jamiołkowski state $J(\Lambda)$ can be expressed as

$$p_{\text{succ}}(\Lambda) = \frac{1}{d\big(1 - E_G^\cap(J(\Lambda))\big)}. \tag{18}$$

We remark that, because $E_G^\cap$ is a concave function on states, the probability of success p_{succ} is a concave function on channels, i.e.,

$$p_{\text{succ}}((1-q)\Lambda_1 + q\Lambda_2) \geqslant (1-q)p_{\text{succ}}(\Lambda_1) + q p_{\text{succ}}(\Lambda_2),$$

for $0 \leqslant q \leqslant 1$. Of course, this could be concluded directly from (5).

VI. BOUNDS

The evaluation of the stochasticity (6) for a given channel is in general a nontrivial computational problem. The connection with entanglement that was developed in Sec. V, more precisely Eq. (17), shows that calculating the stochasticity is equivalent to evaluating $E_G^\cap(J(\Lambda))$. In principle, this requires one to check for all possible ensemble decompositions of $J(\Lambda)$, although one can use convexity arguments to restrict the search to ensembles of r^2 pure states for a Choi-Jamiołkowski state of rank r, similarly to the case of entanglement of formation [28]. In this section we will be able to provide analytic upper and lower bounds that do not require any search.

Entanglement of assistance and other concave-roof constructions have not been studied as well as convex-roof constructions. This is due to the fact that they are not entanglement measures [25,31]. Nonetheless, they are of interest because, e.g., they capture some properties of multipartite entanglement. For example, the entanglement of assistance quantifies the average amount of entanglement that two parties—Alice and Bob—can share thanks to a measurement of a third party who holds the purification of the state. Thus, we will be able to make use of some results already derived in the literature, in particular in [30] and [32], to provide upper and lower bounds for the stochasticity σ and the probability of success p_{succ}.

We first start by illustrating the range over which p_{succ} can vary, focusing on the best and worst cases. We then identify a simple bound based uniquely on the mathematical properties of the operator norm. As we will see, such a bound will turn out to be pretty useful in investigating the examples of Sec. VII. We then proceed to consider bounds based on the entanglement properties of the Choi-Jamiołkowski state isomorphic to the channel of interest.

A. Best and worst cases

Given that E_G—and therefore $E_G^\cap$—satisfies (13), it follows from (18) that

$$\frac{1}{d} \leqslant p_{\text{succ}}(\Lambda) \leqslant 1. \tag{19}$$

As pointed out earlier, the upper bound in (19) can only be achieved by random-unitary channels, whose Choi-Jamiołkowski states can be written as convex combinations of maximally entangled states. The lower bound corresponds to $E_G^\cap(J(\Lambda)) = 0$, i.e., to the case where no ensemble for $J(\Lambda)$ contains any entangled state. Such an occurrence was considered in the context of the study of the entanglement of assistance in [30], where it was proved that any state ρ^{AB} with vanishing entanglement of assistance must be of the form $\rho^{AB} = |\alpha\rangle\langle\alpha| \otimes \rho^B$ or $\rho^{AB} = \rho^A \otimes |\beta\rangle\langle\beta|$. Given that we are not considering general bipartite states, but states that are isomorphic to channels via the isomorphism (7), for the first inequality in (19) to be saturated it must be $J(\Lambda) = |\alpha\rangle\langle\alpha| \otimes I/d$. The latter condition implies that the output of the channel is a pure state independent of the input; i.e., $\Lambda[\rho] = \text{Tr}(\rho)|\alpha\rangle\langle\alpha|$.

It may seem strange that the channel that is almost the most trivial theoretically is the one that is the most difficult to implement under our constraints. One can provide the following intuitive explanation. The output state must be independent of the input, but at the same time still be in the encoding. Thus, the output state must include the photon of the input encoding, because the ancillary modes are initially in the vacuum state. This can be accomplished in the following way. In the scheme proposed in Fig. 1, a random rotation is first applied to the input. Subsequently, $d - 1$ of the encoding modes are measured while the remaining one is transmitted—that is, the transitivity of $d - 1$ of the d beam splitters is set to 0, while the remaining one is set to 1. Upon finding the vacuum in the measured modes, we know that the photon is in the only unmeasured mode, i.e., in some known logical basis state of the encoding. Then we can rotate such a state to the desired output state. Given the random rotation of the input, the probability that this procedure succeeds is exactly $1/d$, independently of the input.

B. Triangle-inequality bound

By using the triangle inequality, it is straightforward to derive an upper limit on the success probability.

Observation 1 (triangle-inequality bound). For any quantum channel Λ,

$$p_{\text{succ}}(\Lambda) \leqslant \frac{1}{\|\Lambda(I)\|_\infty}. \tag{20}$$

Proof. If $\{A_i\}$ is any Kraus decomposition for the channel Λ, then we have for the stochasticity

$$\sigma(\Lambda) = \min_{\{A_i\}} \sum_i \|A_i\|_\infty^2 = \min_{\{A_i\}} \sum_i \big\|A_i A_i^\dagger\big\|_\infty$$

$$\geqslant \min_{\{A_i\}} \Big\| \sum_i A_i A_i^\dagger \Big\|_\infty = \|\Lambda(I)\|_\infty,$$

where the inequality is due to the triangle inequality, and the dependence on the choice of the Kraus decomposition is lost because $\sum_i A_i A_i^\dagger = \Lambda(I)$, for any Kraus decomposition of Λ. ∎

This bound proves that it is necessary for a channel to be unital in order for us to implement it deterministically using our scheme, because only for a unital channel $\|\Lambda(I)\|_\infty = 1$.

This is consistent with the already argued fact that under our scheme only random-unitary channels can be deterministically implemented. The bound is easily evaluated, being independent of any particular Kraus decomposition.

We remark that any choice of a specific Kraus decomposition provides a lower bound on the probability of success. If such a lower bound matches the upper bound in (20), then the given decomposition is proven to be optimal.

C. Bounds based on entanglement properties of the Choi-Jamiołkowski state

Now we will move to bounds that exploit the connection we observed between the success probability of our scheme and the entanglement properties of the Choi-Jamiołkowski state $J(\Lambda)$.

1. Qubit channels

We will first focus on the qubit case. Not surprisingly, this is the case where we can employ a large number of results from entanglement theory. In particular, we will be concerned with one of the most adopted entanglement measures, known as concurrence [33,34]. The concurrence of a pure two-qubit state can be expressed as

$$C(\psi) = |\langle \tilde{\psi} | \psi \rangle|,$$

where

$$|\tilde{\psi}\rangle = (\sigma_y \otimes \sigma_y)|\psi^*\rangle,$$

with the complex conjugation taken in the computational basis and $\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$. The definition extends to density matrices via the standard convex-roof construction:

$$C^{\cup}(\rho) = \min_{\{p_i, \psi_i\}} \sum_i p_i C(\psi_i).$$

It is straightforward to check that for a pure state the relation

$$E_G(\psi) = \tfrac{1}{2}(1 - \sqrt{1 - C(\psi)^2}) \qquad (21)$$

holds. In [22] it was argued that $C^{\cup}$ and $E_G^{\cup}$ are related by $E_G^{\cup} = \frac{1}{2}(1 - \sqrt{1 - C^{\cup}(\rho)^2})$. We will instead be interested in the connection between $E_G^{\cap}$ and the concave-roof of the concurrence,

$$C^{\cap}(\rho) = \max_{\{p_i, \psi_i\}} \sum_i p_i c(\psi_i). \qquad (22)$$

The examples of Sec. VII will prove that the relation (21) does not hold for the concave-roof version of the two quantities. Nonetheless, in order to obtain easily computable bounds for p_{succ}, we will exploit the remarkable fact that there is a closed expression for both $C^{\cup}(\rho)$ and $C^{\cap}(\rho)$. For the former it reads [34]

$$C^{\cup}(\rho) = \max\{0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4\},$$

where the λ_i's are the eigenvalues of $\sqrt{\sqrt{\rho}\tilde{\rho}\sqrt{\rho}}$ in decreasing order, with the state $\tilde{\rho} = \sigma_y \otimes \sigma_y \rho^* \sigma_y \otimes \sigma_y$. For $C^{\cap}(\rho)$ instead it holds [32]

$$C^{\cap}(\rho) = F(\rho, \tilde{\rho}), \qquad (23)$$

with $F(\sigma, \tau) = \text{Tr}(\sqrt{\sqrt{\sigma}\tau\sqrt{\sigma}})$ the fidelity between two states σ and τ. We start by providing the following lemma that relates $C^{\cap}(\rho)$ and $E_G^{\cap}(\rho)$.

Lemma 1. Given any state ρ of two qubits, the following inequalities hold:

$$\frac{1}{2}\left(1 - \sqrt{1 - C^{\cap}(\rho)^2}\right) \leqslant E_G^{\cap}(\rho) \leqslant \frac{C^{\cap}(\rho)}{2}. \qquad (24)$$

Proof. See Appendix A. ∎

By using the lemma together with the relation (18) for $d = 2$, and (23), we immediately obtain the following result:

Theorem 2 (concurrence bounds). If $J(\Lambda)$ is the Choi-Jamiołkowski state isomorphic to the qubit channel Λ, then

$$\frac{1}{2 - F(J(\Lambda), \tilde{J}(\Lambda))} \geqslant p_{\text{succ}}(\Lambda)$$

$$\geqslant \frac{1}{1 + \sqrt{1 - F(J(\Lambda), \tilde{J}(\Lambda))^2}}. \qquad (25)$$

2. Qudit channels and entanglement of assistance

In the previous section we focused on the concurrence because its concave-roof version $C^{\cap}$ can be easily evaluated. Concurrence was generalized to higher dimensional systems in a number of different ways [35–37], and even high-dimensional "assisted" versions—i.e., concave-roof constructions—were considered [38]. As we mentioned, the most studied example of concave-roof construction is the entanglement of assistance (16). For this reason, we will provide bounds for the probability of success in terms of the entanglement of assistance.

We will use the following generalization of the binary entropy that depends only on the number of possible outcomes d and one probability parameter p:

$$h_d(p) := -p \log_2 p - (1 - p) \log_2\left(\frac{1 - p}{d - 1}\right) \qquad (26)$$

That is, $h_d(p)$ is the Shannon entropy of the probability distribution of d symbols $(p, \frac{1-p}{d-1}, \ldots, \frac{1-p}{d-1})$, with one symbol having probability p and the remaining $d - 1$ symbols being equally likely. It coincides with the binary entropy for $d = 2$. We remark that $h_d(p)$ is a concave function of p and is monotonically decreasing for $p \geqslant 1/d$. This means that the inverse function $h_d^{-1} : [0, \log_2 d] \to [1/d, 1]$ is well defined.

We are now ready to state the theorem that links entanglement of assistance and probability of success.

Theorem 3 (entanglement-of-assistance bounds). For a given qudit channel Λ, the following inequalities hold:

$$\frac{2^{E_a(J_\Lambda)}}{d} \geqslant p_{\text{succ}}(\Lambda) \geqslant \frac{1}{dh_d^{-1}(E_a(J_\Lambda))}, \qquad (27)$$

where E_a is the entanglement of assistance, $J(\Lambda)$ is the Choi-Jamiołkowski state isomorphic to the channel, and $\sigma(\Lambda)$ is the stochasticity of the channel.

Proof. See Appendix B. ∎

VII. EXAMPLES

In this section, we consider two examples for the qubit case: (i) the amplitude-damping channel and (ii) the probabilistic

constant-output map. For these examples we are able to find analytic results for the probability of success, and we compare these exact results with the bounds we obtained in Sec. VI C. The analytic results are obtained by using the triangle-inequality bound of Observation 1 and the already remarked fact that any specific Kraus decomposition provides a upper (lower) bound on the stochasticity (optimal probability of success).

A. Amplitude-damping channel

The qubit amplitude-damping channel is used to model the decay of an excited state $|1\rangle$ into the ground state $|0\rangle$. With probability ϵ the channel causes the de-excitation of the input state. This de-excitation process is described by the Kraus operator $A_1 = \sqrt{\epsilon}|0\rangle\langle 1|$. A second Kraus operator guarantees that the process preserves probability, i.e., that the channel is trace preserving: $A_2 = |0\rangle\langle 0| + \sqrt{1-\epsilon}|1\rangle\langle 1|$.

Using this specific decomposition and Observation 1 we find

$$\|\Lambda(I)\|_\infty \leqslant \sigma(\Lambda) \leqslant \|A_1\|_\infty^2 + \|A_2\|_\infty^2.$$

For the lower bound, one finds

$$\|\Lambda(I)\|_\infty = \|A_1 A_1^\dagger + A_2 A_2^\dagger\|_\infty$$
$$= \left\| \begin{pmatrix} 1 & 0 \\ 0 & 1-\epsilon \end{pmatrix} + \begin{pmatrix} \epsilon & 0 \\ 0 & 0 \end{pmatrix} \right\|_\infty = 1+\epsilon,$$

and for the upper bound we get

$$\|A_1\|_\infty^2 + \|A_2\|_\infty^2 = 1 + \epsilon.$$

Because these two bounds coincide, the Kraus decomposition is optimal, with a stochasticity of $\sigma(\Lambda) = 1 + \epsilon$ and the optimal success probability $p_{\text{succ}}(\Lambda) = 1/(1 + \epsilon)$. When $\epsilon = 0$ the channel is trivially the identity channel and can be performed deterministically. However at the other extreme, $\epsilon = 1$, the channel becomes the constant map $\Lambda(\rho) = \text{Tr}(\rho)|0\rangle\langle 0|$, and it can only be realized with probability 50%. As we found in Sec. VI A, the constant map is the map that has the lowest success probability in our scheme. Therefore the parameter ϵ that describes the probability of de-excitation lets us move from one extreme to the other of the stochasticity (or probability of success). For the amplitude-damping channel we find that the concave-roof of concurrence (22) satisfies $C^\cap(J(\Lambda)) = F(J(\Lambda), \tilde{J}(\Lambda)) = \sqrt{1-\epsilon}$. We can then compare the analytic bounds from (25) with the exact result we just found (see Fig. 3):

$$\frac{1}{2 - \sqrt{1-\epsilon}} \geqslant p_{\text{succ}}(\Lambda) = \frac{1}{1+\epsilon} \geqslant \frac{1}{1+\sqrt{\epsilon}}. \quad (28)$$

We remark that our scheme achieves a probability of success that depends on ϵ and is close to 100% for ϵ small. On the contrary, the scheme of [14] has a 50% probability of success, independently of ϵ. In our scheme, such a low success probability is just the worst case ($\epsilon = 1$).

B. Probabilistic constant-output channel

The second channel that we choose to analyze is a convex combination of the constant output channel and the identity
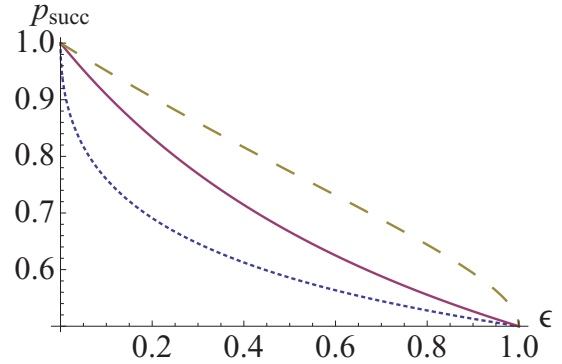

FIG. 3. (Color online) Optimal probability of success for the amplitude-damping channel: exact (solid line) and upper and lower bounds in Eq. (28) (dashed and fine-dashed lines, respectively).

map. Such a channel returns the input state with a probability of $1 - p$ or a fixed output state τ with probability p. The map is then

$$\Lambda : \rho \mapsto (1-p)\rho + p\,\text{Tr}(\rho)\tau$$

and its Choi-Jamiołkowski isomorphic state is simply

$$J(\Lambda) = (1-p)|\psi_2^+\rangle\langle\psi_2^+| + p\,\tau \otimes \frac{I}{2}.$$

We find the stochasticity of this channel by checking that the upper and lower bounds for the stochasticity that are generated from Observation 1 and a particular decomposition of the state match. Without loss of generality we can consider τ to be diagonal in the computational basis; i.e.,

$$\tau = s|0\rangle\langle 0| + (1-s)|1\rangle\langle 1|. \quad (29)$$

In fact, only the degree of mixedness of τ and not the specific basis influences the probability of success p_{succ}. This can be understood at the formal level by considering that a different choice of basis for τ can be taken into account via a rotation U, which does not influence the entanglement properties of $J(\Lambda)$ (see Appendix C).

Using this state, the bound stated in Observation 1 becomes

$$\sigma(\Lambda) \geqslant \|\Lambda(I)\|_\infty = \|(1-p)I + 2p\tau\|_\infty = 1 - p + 2ps, \quad (30)$$

where we have assumed, without loss of generality, that $s \geqslant 1/2$. From this and by using Eq. (17), we find $E_G^\cap(J(\Lambda)) \leqslant 1/2 - p(s - 1/2)$. One can find an ensemble decomposition of $J(\Lambda)$ that saturates the latter inequality (see Appendix D); therefore $p_{\text{succ}} = (1 - p + 2ps)^{-1}$.

This means that also for this channel we find that as the probability parameter p varies from 0 to 1 we move from the identity map to a constant map. However, we can see from Fig. 4 that the success probability of the constant map depends on how mixed the output state is. As expected from the discussion of Sec. VI A, the lowest value for the success probability, $p_{\text{succ}}(\Lambda) = 1/2$, is only attained when the constant output state is pure.
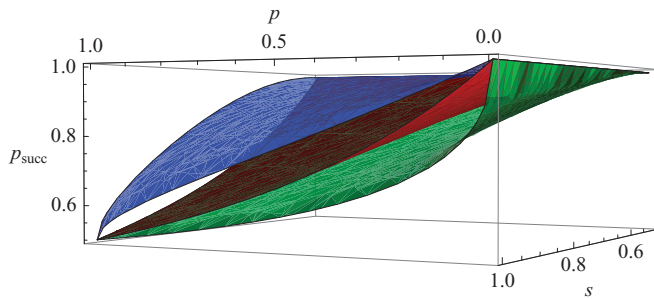
FIG. 4. (Color online) Optimal probability of success for the probabilistic constant-output channel, exact result, and bounds. Upper and lower surfaces (blue and green, respectively) are the upper and lower concurrence bounds from Eq. (25), evaluated for the Choi-Jamiołkowski state isomorphic to the channel; the in-between surface (red) is the exact success probability p_{succ} for realizing the channel. Because p_{succ} is symmetric about the $s = 1/2$ plane, we only plot this function for $s \in [1/2, 1]$.

VIII. DISCUSSION

We have provided a scheme to realize an arbitrary channel on a d-rail-encoded optical qudit, taking into account practical restrictions. In particular: (i) we only allow for operations that are realizable using linear optics; (ii) we only allow ancillary modes that are initially in the vacuum state, thus limiting the need for sources of single photons that are, as of now, still difficult to produce on demand; (iii) we do not allow feed forward (i.e., adaptive schemes), which significantly reduces the cost of the necessary equipment and the loss that is inevitably involved in such schemes due to the need of long fibers for optical delays; and (iv) we consider only photon-number measurements (actually, readily available threshold detectors suffice). The conventional linear-optics toolbox (phase shifters and beam splitters), as well as the possibility for randomly switching between different optical networks, are the only elements needed for the realization of our method. These restrictions render our technique of immediate interest to linear-optical implementations that can be realized using state-of-the-art experimental techniques. Within this framework, it turns out that any channel can, in principle, be realized perfectly, albeit only stochastically. The only channels that can be realized deterministically are random-unitary channels. Given that postselection is a commonly used technique in linear-optics experiments, this restriction effectively only slightly reduces the success probability of an experimental realization, and we are able to provide an expression for the optimal probability of success. This probability turns out to be related to the entanglement properties of the Choi-Jamiołkowski state isomorphic to the channel of interest. More precisely, we were led to evaluate the "assisted version" of the geometric measure of entanglement, i.e., the concave-roof extension of the measure to mixed states, for the Choi-Jamiołkowski state. While we are not aware of a closed formula for it, not even for two-qubit states, we were able to provide upper and lower bounds in terms of the concave roof of concurrence (for qubits) and of entanglement of assistance (for general qudits).

Besides tackling the problem of evaluating, in general, the concave roof of the geometric measure of entanglement, i.e., the probability of successful realization of our scheme with the restrictions considered in this paper, future research will focus on the relaxation of said restrictions, that is, on the analysis of more general schemes for the realization of channels.

For example, the use of ancillary states that are not initially in the vacuum certainly improves the realization of some channels. Indeed, we saw that the worst-case scenario is that of a channel with a fixed—i.e., independent of the input—pure output. We argued that the difficulty—that is, the low probability of success—in the realization of such a channel is essentially due to the necessity of using for the output the same single photon by which the input logical state is encoded in the d modes. This is exactly because no photons are available in the ancillary ports. Obviously, if such a fixed pure output is readily available as an ancillary state, the realization of the pure-fixed-output channel becomes trivial: the ancillary input state becomes the output. It is therefore evident that introducing nonvacuum ancillas would strongly affect the performance of our scheme. Another addition that we plan to consider is feed-forward, which is becoming a powerful and reliable tool in linear-optics quantum-information processing [7,20].

Another possible line of research is that of focusing on channels that are linked to interesting effects in quantum information processing. Indeed, our results can be thought of as a toolbox to be used in any optical experiment where some specific channel has to be applied, be it for the sake of simulating noise or for implementing a specific protocol.

We expect our findings to trigger further theoretical studies on channel realization. In particular, we linked channel realization with more abstract notions of entanglement theory, and we hope that the study of less explored entanglement properties of states will consequently be stimulated. From a more practical point of view, our results provide a simple method for realizing arbitrary quantum channels using linear optics and standard experimental techniques. Our results are ideal for experimental implementation relying on linear optics in combination with postselection. While quantum channels have been a widely discussed topic in theoretical quantum information, we expect our work to trigger an increased interest in the experimental study of this intriguing topic.

APPENDIX A: PROOF OF LEMMA 1

Proof. By substituting (21) in the definition (15) of $E_G^\cap(\rho)$ we obtain $E_G^\cap(\rho) = \max_{\{p_i, \psi_i\}} \sum p_i \frac{1}{2}(1 - \sqrt{1 - C(\psi_i)^2})$. Then, in order to obtain the lower bound it is sufficient to observe that $\sqrt{1 - x^2}$ is a concave function in x that is monotonically decreasing:

$$E_G^\cap(\rho) = \max_{\{p_i, \psi_i\}} \sum p_i \frac{1}{2}(1 - \sqrt{1 - C(\psi_i)^2})$$

$$\geqslant \max_{\{p_i,\psi_i\}} \frac{1}{2}\left(1 - \sqrt{1 - \left[\sum_i p_i C(\psi_i)\right]^2}\right)$$

$$= \frac{1}{2}\left(1 - \sqrt{1 - \left[\max_{\{p_i,\psi_i\}} \sum_i p_i C(\psi_i)\right]^2}\right)$$

$$= \frac{1}{2}(1 - \sqrt{1 - C^\cap(\rho)^2}). \tag{A1}$$

The upper bound can be derived from the relation $\sqrt{1-x^2} \geqslant 1-x$. ∎

Proof. We will use properties of the Shannon entropy $H(\{r_i\}) = -\sum_i r_i \log_2 r_i$, defined for a probability distribution $\{r_i\}$. The von Neumann entropy of a quantum state ρ is equal to the Shannon entropy of its eigenvalues r_i. In particular, for a pure bipartite state with Schmidt decomposition $|\psi\rangle^{AB} = \sum_i \sqrt{\lambda_i}|i\rangle_A|i\rangle_B$, the entropy of the reduced one-party states ρ^A and ρ^B is $H(\{\lambda_i\})$. For any pure ensemble $\{p_a, \psi_a^{AB}\}$ we will denote by $\{\lambda_i^a\}_i$ the set of the squares of the Schmidt coefficients of ψ_a^{AB}, and define $\lambda_{a,\max} = \max_i\{\lambda_i^a\}_i$. For the entanglement of assistance it then holds

$$E_a(J_\Lambda) = \max_{\{p_a,\psi_a\}} \sum_a p_a S(\rho_a^A) = \max_{\{p_a,\psi_a\}} \sum_a p_a H(\{\lambda_i^a\}_i) \leqslant \max_{\{p_a,\psi_a\}} \sum_a p_a H\left(\lambda_{a,\max}, \frac{1-\lambda_{a,\max}}{d-1}, \ldots, \frac{1-\lambda_{a,\max}}{d-1}\right)$$

$$\leqslant \max_{\{p_a,\psi_a\}} h_d\left(\sum_a p_a \lambda_{a,\max}\right) = h_d\left(\min_{\{p_a,\psi_a\}} \sum_a p_a \lambda_{a,\max}\right) = h_d\left(\frac{\sigma(\Lambda)}{d}\right). \tag{B1}$$

The first inequality is due to the fact that substituting any subset of probabilities of some distribution with equally weighted probabilities can only increase the total Shannon entropy. This is easily checked by knowing that the flat probability distribution is the one with highest Shannon entropy, and that for any grouping of probabilities $\{r_i\}$ into two subsets $\{r_i^{(1)}\}$ and $\{r_i^{(2)}\}$ of weight q and $1-q$, respectively, we have $H(\{r_i\}) = h_2(q) + qH(\{r_i^{(1)}/q\}) + (1-q)H(\{r_i^{(2)}/(1-q)\})$. The second inequality is due to the concavity of entropy. The second-to-last equality is due to the the monotonicity of h_d in the interval $[1/d,1]$. Indeed, $\sum_a p_a \lambda_{a,\max} \geqslant 1/d$ because $\lambda_{a,\max} \geqslant 1/d$ for all a. Finally, the last equality comes from the relation (12b). Thus, using the fact that h_d is invertible and monotonically decreasing in the range of interest, we obtain $\sigma(\Lambda) \leqslant dh_d^{-1}(E_a(J_\Lambda))$, i.e., $p_{\text{succ}} \geqslant 1/[dh_d^{-1}(E_a(J_\Lambda))]$.

For the upper bound we have

$$E_a(\rho_\Lambda) = \max_{\{p_a,\psi_a\}} \sum_a p_a H(\{\lambda_i^a\}_i)$$

$$\geqslant \max_{\{p_a,\psi_a\}} \sum_a [-p_a \log_2(\lambda_{a,\max})] \tag{B2}$$

$$\geqslant \max_{\{p_a,\psi_a\}} \left[-\log_2\left(\sum p_a \lambda_{a,\max}\right)\right]$$

$$= -\log_2\left(\min_{\{p_a,\psi_a\}} \sum p_a \lambda_{a,\max}\right)$$

$$= -\log_2\left(\frac{\sigma(\Lambda)}{d}\right). \tag{B3}$$

The first inequality comes from the fact that the min-entropy $H_{\min}(\{r_i\}) = -\log_2 r_{\max}$ of a probability distribution $\{r_i\}$, with $r_{\max} = \max\{r_i\}$, satisfies $H_{\min}(\{r_i\}) \leqslant H(\{r_i\})$. The second inequality is due to the concavity of the logarithm. The second-to-last equality is due to the monotonicity of the logarithm. We finally arrive at the desired relation by exponentiation. ∎

Suppose $\tau' = U\tau U^\dagger$; then

$$J(\Lambda) = (1-p)|\psi_d^+\rangle\langle\psi_d^+| + p\,\tau' \otimes \frac{I}{2}$$

$$= (1-p)|\psi_d^+\rangle\langle\psi_d^+| + p\,U\tau U^\dagger \otimes \frac{I}{2}$$

$$= (U \otimes U^*)\left[(1-p)|\psi_d^+\rangle\langle\psi_d^+| + p\,\tau \otimes \frac{I}{2}\right](U \otimes U^*)^\dagger,$$

where we have used the invariance of the maximally entangled state $|\psi_d^+\rangle = (U \otimes U^*)|\psi_d^+\rangle$, valid for all unitaries U.

One can write the Choi-Jamiołkowski state as the convex combination

$$J(\Lambda) = (1-p)|\psi_2^+\rangle\langle\psi_2^+| + 2p(1-s)\frac{I}{2} \otimes \frac{I}{2}$$

$$+ p(s-(1-s))|0\rangle\langle0| \otimes \frac{I}{2}, \tag{D1}$$

such that for the concave-roof of the geometric measure we find

$$E_G^\cap(J(\Lambda)) \geqslant (1-p)E_G^\cap(|\psi_2^+\rangle\langle\psi_2^+|)$$

$$+ 2p(1-s)E_G^\cap\left(\frac{I}{2} \otimes \frac{I}{2}\right)$$

$$+ p(s-(1-s))E_G^\cap\left(|0\rangle\langle0| \otimes \frac{I}{2}\right)$$

$$= (1-p)\frac{1}{2} + 2p(1-s)\frac{1}{2} = 1/2 - p(s-1/2).$$

Here we used the fact that $E_G^\cap(|0\rangle\langle 0| \otimes I/2) = 0$—see the discussion just after Eq. (19)—and that $E_G^\cap(I/2 \otimes I/2) = 1/2$, because the maximally mixed state of two qubits can be seen as the convex combination of pure maximally entangled states.

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

[2] G. Smith and J. Yard, Science **321**, 1812 (2008).

[3] E. Knill, R. Laflamme, and G. Milburn, Nature (London) **409**, 46 (2001).

[4] K. Mattle, H. Weinfurter, P. G. Kwiat, and A. Zeilinger, Phys. Rev. Lett. **76**, 4656 (1996).

[5] D. Bouwmeester, J. W. Pan, K. Mattle, M. Eibl, H. Weinfurter, and A. Zeilinger, Nature (London) **390**, 575 (1997).

[6] J.-W. Pan, S. Gasparoni, R. Ursin, G. Weihs, and A. Zeilinger, Nature (London) **423**, 417 (2003).

[7] R. Prevedel, P. Walther, F. Tiefenbacher, P. Bohi, R. Kaltenbaek, T. Jennewein, and A. Zeilinger, Nature (London) **445**, 65 (2007).

[8] B. He, J. A. Bergou, and Z. Wang, Phys. Rev. A **76**, 042326 (2007).

[9] R. Prevedel, M. S. Tame, A. Stefanov, M. Paternostro, M. S. Kim, and A. Zeilinger, Phys. Rev. Lett. **99**, 250503 (2007).

[10] K. M. Schreiter, A. Pasieka, R. Kaltenbaek, K. J. Resch, and D. W. Kribs, Phys. Rev. A **80**, 022311 (2009).

[11] E. Amselem and M. Bourennane, Nature Phys. **5**, 748 (2009).

[12] J. Lavoie, R. Kaltenbaek, M. Piani, and K. J. Resch, Phys. Rev. Lett. **105**, 130501 (2010).

[13] M. Almeida, F. D. Melo, M. Hor-Meyll, A. Salles, S. Walborn, P. Ribeiro, and L. Davidovich, Science **316**, 579 (2007).

[14] L. Qing, L. Jian, and G. Guang-Can, Chin. Phys. Lett. **24**, 1809 (2007).

[15] J.-C. Lee, Y.-C. Jeong, Y.-S. Kim, and Y.-H. Kim, Optics Express **19**, 16309 (2011).

[16] A. Jamiołkowski, Rep. Math. Phys. **3**, 275 (1972).

[17] M.-D. Choi, Linear Algebr. Appl. **10**, 285 (1975).

[18] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Rev. Mod. Phys. **81**, 865 (2009).

[19] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, Phys. Rev. Lett. **73**, 58 (1994).

[20] D. N. Biggerstaff, R. Kaltenbaek, D. R. Hamel, G. Weihs, T. Rudolph, and K. J. Resch, Phys. Rev. Lett. **103**, 240504 (2009).

[21] K. Kieling, Ph.D. thesis, Imperial College London, 2008.

[22] T.-C. Wei and P. M. Goldbart, Phys. Rev. A **68**, 042307 (2003).

[23] G. Vidal, Phys. Rev. Lett. **83**, 1046 (1999).

[24] T.-C. Wei, M. Ericsson, P. Goldbart, and W. Munro, Quantum Inf. Comput. **4**, 252 (2004).

[25] M. Plenio and S. Virmani, Quant. Inf. Comput. **7**, 1 (2007).

[26] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, Phys. Rev. Lett. **78**, 2275 (1997).

[27] V. Vedral and M. B. Plenio, Phys. Rev. A **57**, 1619 (1998).

[28] A. Uhlmann, Open Systems and Information Dynamics **5**, 209 (1998).

[29] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Phys. Rev. A **54**, 3824 (1996).

[30] D. P. DiVincenzo, C. A. Fuchs, H. Mabuchi, J. A. Smolin, A. Thapliyal, and A. Uhlmann, Lect. Notes Comput. Sci. **1509**, 247 (1999).

[31] G. Gour and R. W. Spekkens, Phys. Rev. A **73**, 062331 (2006).

[32] T. Laustsen, F. Verstraete, and S. J. van Enk, Quantum Inf. Comput. **3**, 64 (2003).

[33] S. Hill and W. K. Wootters, Phys. Rev. Lett. **78**, 5022 (1997).

[34] W. K. Wootters, Phys. Rev. Lett. **80**, 2245 (1998).

[35] K. Audenaert, F. Verstraete, and B. De Moor, Phys. Rev. A **64**, 052304 (2001).

[36] P. Rungta, V. Bužek, C. M. Caves, M. Hillery, and G. J. Milburn, Phys. Rev. A **64**, 042315 (2001).

[37] G. Gour, Phys. Rev. A **71**, 012318 (2005).

[38] G. Gour, Phys. Rev. A **72**, 042318 (2005).