Optimality of universal conclusive entanglement concentration protocols

Alexandre C. Orthey Jr.,[1,2,*] Aby Philip,[1] Tulja Varun Kondra,[3] and Alexander Streltsov[1]
[1]*Institute of Fundamental Technological Research, Polish Academy of Sciences, Pawińskiego 5B, 02-106 Warsaw, Poland*
[2]*Faculty of Mathematics, Informatics and Mechanics, University of Warsaw, ulica Banacha 2, 02-097 Warsaw, Poland*
[3]*Institute for Theoretical Physics III, Heinrich Heine University Düsseldorf, Universitätsstraße 1, D-40225 Düsseldorf, Germany*

(Received 12 September 2025; revised 10 April 2026; accepted 13 May 2026; published 1 June 2026)

Entanglement concentration is essential for quantum technologies yet, rigorous bounds on the success probability for universal protocols (those requiring no prior knowledge about the input state) have remained underexplored. We establish such fundamental limits for conclusive protocols distilling a perfect Bell state from pure two-qubit states by deriving the optimal success probability starting with two copies of a state with known Schmidt basis and four copies of a state with unknown Schmidt basis, using concatenated two-qubit operations. We prove that a known protocol achieves these bounds, confirming its optimality. Crucially, universality imposes an inherent efficiency trade-off, yielding an average success probability of just $2/105$ over Haar measure.

DOI: 10.1103/bz3j-9njg

I. INTRODUCTION

Entanglement is a fundamental resource in quantum information processing [1], allowing for applications such as quantum teleportation [2], cryptography [3], and computation [4]. A critical challenge in practical settings is the *concentration* of entanglement: distilling near-perfect Bell states $|\phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$ from multiple copies of partially entangled states shared between distant parties. While protocols exist for this task [5–12] with several experimental implementations [13–18], their *universality*, i.e., effectiveness across a large set of unknown input states, and *optimality*, that is, if they produce $|\phi^+\rangle$ with the highest possible success probability, remain not well explored in quantum information theory. An exception is the result from Ref. [19], a no-go theorem stating that there is no local operations and classical communication (LOCC) protocol capable of generating a state with higher fidelity than its input states for all two-qubit states, including mixed ones.

This work establishes rigorous optimality bounds for universal conclusive entanglement concentration protocols (ECPs), strictly for finite copies of pure input states and two-qubit operations. Unlike faithful protocols [20] that deterministically produce a state with higher fidelity with the target state, *conclusive* ECPs [21] use probabilistic LOCC maps to either produce a perfect target state or fail, with no intermediate outcomes. By considering a smaller set of states and not allowing for intermediate outcomes, we avoid contradiction with the no-go theorem in Ref. [19].

In this work, we are interested in producing $|\phi^+\rangle$ outright with some probability, rather than in just increasing the fidelity of the input states with $|\phi^+\rangle$. The core challenge addressed here is the following question: For $n = 2$ and $n = 4$, what is the optimal (or maximal) success probability for distilling a Bell pair from n copies of an arbitrary pure two-qubit state, when no prior knowledge of the state's entanglement structure is available?

We present a comprehensive optimality analysis for universal conclusive ECPs, considering two scenarios: (i) states with a known Schmidt basis and (ii) completely arbitrary states with an unknown basis. For the first case, we prove the optimal probability of distilling a Bell state from two copies is a simple function of the entanglement of the input state. For the second, more challenging case, we derive a tight upper bound on the probability of transforming two copies into an entangled state within a fixed Schmidt basis. Furthermore, we prove a tight bound for the direct distillation of a Bell state from four copies of an arbitrary pure state, using protocols restricted to concatenated two-qubit operations.

The structural framework of concatenated two-qubit operations is motivated by both fundamental and practical considerations. Physically, since the input states possess a completely unknown Schmidt basis while the target Bell state has a fixed basis, any universal protocol must eventually bridge this gap; our concatenated approach explicitly isolates and optimally solves this necessary intermediate step. Furthermore, from an experimental standpoint, native entangling operations in near-term quantum architectures are predominantly limited to two qubits. This can be seen in several experimental realizations of entanglement distillation [15,17,18,22,23]. Restricting our analysis to two-qubit interactions avoids the prohibitive complexity and noise associated with global four-qubit joint measurements, ensuring our derived limits and protocols are directly relevant for practical implementations.

Our analysis employs a unified framework, established by Definitions 1 and 2 below, that formalizes universal conclusive ECPs as LOCC maps requiring unit fidelity to ϕ^+, while maximizing success probability across all inputs. For unknown Schmidt basis states, we derive tight upper bounds on success probabilities by imposing algebraic constraints on Kraus

*Contact author: alexandre.orthey@gmail.com

operators that enforce universality through the nullification of separable subspaces. We then construct explicit LOCC protocols with operators that saturate these bounds. Our results show that Kálmán *et al.*'s [24] protocol is an optimal universal ECP in both scenarios. While Vidal's formula [21] gives higher conversion probabilities, as one is allowed to tailor the protocol to the specific input, we prove that universal protocols necessarily achieve strictly lower probabilities due to the universality constraint. Finally, we compute the average performance using Dirichlet distributions for Haar-random states via moment formulas [25–27].

II. OPTIMALITY OF A UNIVERSAL CONCLUSIVE ECP

Let ρ_{AB} be a bipartite state shared between Alice and Bob. Let $\mathcal{E}_1$ and $\mathcal{E}_2$ be two completely positive trace nonincreasing LOCC maps, such that $\mathcal{E} = \mathcal{E}_1 + \mathcal{E}_2$ is a completely positive trace-preserving map. Having said that, we define as follows.

Definition 1 (Universal conclusive ECP). Let $\mathcal{S}$ be the subset of states ρ that are shared between Alice and Bob. The n-to-1 map $\mathcal{E}_1$ is called a universal conclusive ECP for a given number of copies n if, for every $\rho \in \mathcal{S}$, it holds that $\langle\phi^+|\mathcal{E}_1(\rho^{\otimes n})|\phi^+\rangle = \mathrm{Tr}[\mathcal{E}_1(\rho^{\otimes n})]$.

Definition 2 (Optimal universal conclusive ECP). Let $\mathcal{S}$ be the subset of states ρ that are shared between Alice and Bob. Let $\mathcal{E}_1$ be any protocol satisfying Definition 1. The map $\mathcal{E}_1^{\text{optimal}}$ is called the optimal universal conclusive ECP for a given number of copies n if, for every $\rho \in \mathcal{S}$ and for any $\mathcal{E}_1$, it holds that $\mathrm{Tr}[\mathcal{E}_1(\rho^{\otimes n})] \leqslant \mathrm{Tr}[\mathcal{E}_1^{\text{optimal}}(\rho^{\otimes n})]$.

First, we would like to note that there can be no universal conclusive ECP for the set of pure states shared between Alice and Bob when $n = 1$. This follows from some simple arguments. Consider, toward a contradiction, that there exists a universal conclusive ECP for $n = 1$ which means $\mathcal{E}_1(|\psi\rangle\langle\psi|) = p|\phi^+\rangle\langle\phi^+|$, where $p > 0$. Let $\rho = (1 - \epsilon)\,|\psi\rangle\langle\psi| + (\epsilon/4)\,\mathbb{I}$, then $\mathcal{E}_1(\rho) = p'|\phi^+\rangle\langle\phi^+|$ where $p' > 0$. (Note, we are allowed to do this because ρ can be expressed as a convex combination of pure states) This means that we will transform a full rank state to a pure entangled state by a LOCC protocol. However, this is not possible since it is impossible to transform a full rank state into a pure resourceful state using free operations [28,29]. Hence, we arrive at a contradiction, and hence there is no universal conclusive ECP for the set of pure states shared between Alice and Bob and $n = 1$.

Similarly, using the aforementioned fact from Refs. [28,29], it is easy to see that there can be no universal conclusive ECP for the set of all states of a given dimension shared between Alice and Bob for any n. Hence, in this work, we shall only be looking at universal conclusive ECPs for sets of pure states.

III. CONCENTRATION PROTOCOL FOR STATES IN A KNOWN SCHMIDT BASIS

We will start by looking into universal conclusive ECPs over a particular set of pure states: Schmidt states. Let $\mathcal{S}_2$ be the set of bipartite pure states that have the same Schmidt basis. Without loss of generality, let the Schmidt basis under consideration be $\{|00\rangle_{AB}, |11\rangle_{AB}\}$. Then, any state in $\mathcal{S}_2$ can be written as $|\psi_S\rangle_{AB} = \alpha|00\rangle_{AB} + \beta|11\rangle_{AB}$, where $\alpha, \beta \in \mathbb{C}$ satisfy $|\alpha|^2 + |\beta|^2 = 1$. We will drop the subindex notation AB whenever suitable. Also, for any pure state $|\psi\rangle$, we define $\psi \equiv |\psi\rangle\langle\psi|$. Our first result is summarized in the following theorem.

Theorem 1. Following Definition 2, the optimal universal conclusive ECP over the set of Schmidt states $|\psi_S\rangle = \alpha|00\rangle + \beta|11\rangle$ transforms two copies of such a state into ϕ^+ with optimal probability $P_{\psi_S^{\otimes 2}\to\phi^+} := 2|\alpha\beta|^2$.

Sketch of the proof. The reader can find the details of the proof in Appendix A. Since both input and output states are pure under the map $\mathcal{E}_{\phi^+}(\cdot) = \sum_i M_i(\cdot)M_i^\dagger$, the action of *each* Kraus operator M_i over the product $|\psi_S\rangle^{\otimes 2}$ must result into a state proportional to $|\phi^+\rangle_{AB}|\mathrm{garb}\rangle_{A'B'}$, where $|\mathrm{garb}\rangle_{A'B'}$ is some garbage state to be discarded. In addition to that, copies of product states such as $|00\rangle|00\rangle$ and $|11\rangle|11\rangle$ should not yield entangled states under LOCCs. Both of these conditions together impose an upper bound on the probability of obtaining ϕ^+ as $2|\alpha\beta|^2$.

Note that there cannot be a universal conclusive ECP for the set of pure states shared between Alice and Bob when $n = 2$. To see this, assume that such a protocol exists and apply it to the convex combination of any two states with different Schmidt bases, such as $|\psi_S\rangle\langle\psi_S|^{\otimes 2}$ and $|++\rangle\langle++|^{\otimes 2}$. Then we will again transform a full rank state to a pure entangled state by a LOCC protocol. However, this is not possible [28,29]. Hence, we arrive at a contradiction, and hence there is no universal conclusive ECP for the set of pure states shared between Alice and Bob and $n = 2$.

IV. CONCENTRATION PROTOCOL FOR STATES IN AN UNKNOWN SCHMIDT BASIS

Now we consider concentration protocols given by the map $\mathcal{E}_1$ satisfying Definition 2 applied to four copies of a state $|\psi\rangle \in \mathcal{S}$, such that $\mathcal{S}$ is the set of all pure two-qubit states. For this case, we are not going to assume a known Schmidt basis, and therefore, we can write $|\psi\rangle$ in the computational basis as

$$|\psi\rangle = c_1|00\rangle + c_2|01\rangle + c_3|10\rangle + c_4|11\rangle, \tag{1}$$

where $c_i \in \mathbb{C}$ for every i and $\sum_{i=1}^4 |c_i|^2 = 1$. By Definition 2, we have that $\mathcal{E}_1(\psi^{\otimes 4}) = P_{\psi^{\otimes 4}\to\phi^+}\,\phi^+$, where $P_{\psi^{\otimes 4}\to\phi^+}$ is the probability of success given by the renormalization required by some measurement included in the protocol, i.e., $P_{\psi^{\otimes 4}\to\phi^+} = \mathrm{Tr}[\mathcal{E}_1(\psi^{\otimes 4})]$.

From Theorem 1, we know the optimal probability of obtaining state $|\phi^+\rangle$ from two copies of $|\psi_S\rangle$. However, that requires that we know the Schmidt basis before the beginning of the protocol. To use arbitrary states in Eq. (1) as input states, we can find a universal protocol that converts two copies of $|\psi\rangle$ into a state with a known Schmidt basis $|\psi_S\rangle$, and then we can apply the protocol that yields Theorem 1 onto two copies of ψ_S to obtain state ϕ^+. As we are going to show, the same protocol that produces Schmidt states is also the protocol that satisfies Theorem 1 (see Fig. 1). Note that, in principle, the coefficients α and β of the output state are not specified; what matters is that the protocol must universally produce a state in a known Schmidt basis.

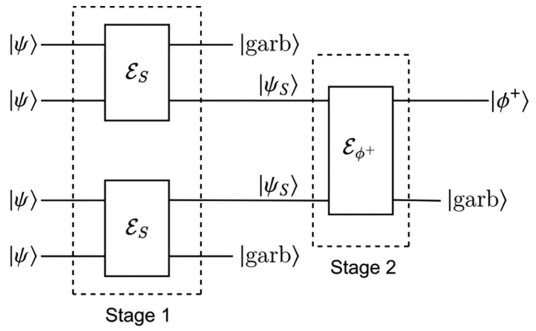

FIG. 1. Optimal universal conclusive ECP applied to four copies of an unknown two-qubit state $|\psi\rangle$. In the first stage of the protocol, four copies of $|\psi\rangle$ are transformed into two copies of $|\psi_S\rangle$, a state with a known Schmidt basis. In the second stage, two copies of $|\psi_S\rangle$ are transformed into one copy of $|\phi^+\rangle$.

V. PRODUCING ENTANGLED STATES IN A KNOWN SCHMIDT BASIS

Let us look for the optimal probability of obtaining an entangled state in a known Schmidt basis from two copies of Eq. (1) by using a universal conclusive protocol. We denote the completely positive trace nonincreasing map as $\mathcal{E}_S$ and the corresponding Kraus operators as M_i, such that

$$\mathcal{E}_S(\cdot) = \sum_i M_i(\cdot)M_i^\dagger, \tag{2}$$

where $\sum_i M_i M_i^\dagger \leqslant \mathbb{I}$. As in the proof of Theorem 1, we know that product states must not produce any entanglement under LOCCs. Hence, we must have that

$$\begin{aligned} M_i|0000\rangle_{ABA'B'} &= 0, \quad M_i|0101\rangle_{ABA'B'} = 0, \\ M_i|1010\rangle_{ABA'B'} &= 0, \quad M_i|1111\rangle_{ABA'B'} = 0, \end{aligned} \tag{3}$$

for every i. From the above, we can see that M_i must act such that

$$\begin{aligned} &M_i|\psi\rangle^{\otimes 2} \\ &= M_i[c_1c_2(|0001\rangle + |0100\rangle) + c_1c_3(|0010\rangle + |1000\rangle) \\ &\quad + c_2c_4(|0111\rangle + |1011\rangle) + c_3c_4(|1011\rangle + |1110\rangle) \\ &\quad + c_1c_4(|0011\rangle + |1100\rangle) + c_2c_3(|0110\rangle + |1001\rangle)], \end{aligned} \tag{4}$$

where $|jklm\rangle \in \mathcal{H}_{ABA'B'}$ whenever $j, k, l, m \in \{0, 1\}$.

We can impose further constraints on the action of M_i. Consider now the case when the input state in Eq. (1) is such that $c_3 = c_4 = 0$, which implies that the input state is a product state $|\psi\rangle = |0\rangle_A \otimes (c_1|0\rangle_B + c_2|1\rangle_B)$. We know that applying map $\mathcal{E}_S$ on copies of such a state must never produce an entangled state ψ_S. Using the same argument when $c_1 = c_2 = 0$, $c_2 = c_4 = 0$, or $c_1 = c_3 = 0$, we infer that, from Eq. (4), M_i must satisfy the following equations:

$$\begin{aligned} M_i(|0001\rangle + |0100\rangle) &= 0, \quad M_i(|0010\rangle + |1000\rangle) = 0, \\ M_i(|0111\rangle + |1011\rangle) &= 0, \quad M_i(|1011\rangle + |1110\rangle) = 0, \end{aligned} \tag{5}$$

for all i. Now, we can use constraints (3) and (5) to help us find the explicit form of the Kraus operators that can do the transformation $\psi^{\otimes 2} \to \psi_S$. In fact, each operator M_i must be a product of two identical Kraus operators acting on spaces $\mathcal{H}_{AA'}$ and $\mathcal{H}_{BB'}$, that is,

$$(M_i)_{AA'BB'} = (K_i)_{AA'} \otimes (K_i)_{BB'}, \quad \forall i. \tag{6}$$

This symmetry is necessary because of the universality requirement of our protocol; since we do not know how entanglement is distributed between Alice and Bob, the protocol should treat each party equally. From Eqs. (3), (5), and (6), we can state our next result.

Theorem 2. To produce arbitrary states in a known Schmidt basis from two copies of arbitrary states $|\psi\rangle_{AB}$ given by Eq. (1), a map $\mathcal{E}_S(\cdot) = \sum_i M_i(\cdot)M_i^\dagger$ must be composed of Kraus operators $M_i = K_i \otimes K_i$ acting on $\mathcal{H}_{AA'} \otimes \mathcal{H}_{BB'}$ such that

$$K_i = a(|00\rangle\langle 01| + |00\rangle\langle 10|) + b(|10\rangle\langle 01| - |10\rangle\langle 10|) \tag{7}$$

for every i, where a and b are any nonzero complex numbers satisfying $2(|a|^4 + |b|^4) \leqslant 1$. The success probability $P_{\psi^{\otimes 2}\to\psi_S}$ of such a protocol is bounded by $\overline{P}_{\psi^{\otimes 2}\to\psi_S} = 2(|c_1c_4| + |c_2c_3|)^2$, i.e., $P_{\psi^{\otimes 2}\to\psi_S} < \overline{P}_{\psi^{\otimes 2}\to\psi_S}$.

Sketch of the proof. The reader can find the details of the proof in Appendix B. From Eqs. (3) and (5), we find that M_i must act on $|\psi\rangle^{\otimes 2}$ such that $M_i|\psi\rangle^{\otimes 2} = \sqrt{2}M_i(c_1c_4|\phi_1\rangle + c_2c_3|\phi_2\rangle)$, where we introduce the notation $|\phi_1\rangle := (|0011\rangle + |1100\rangle)/\sqrt{2}$ and $|\phi_2\rangle := (|0110\rangle + |1001\rangle)/\sqrt{2}$. It follows that M_i applied to $|\phi_j\rangle$ must result in $|\psi_S^{(j)}\rangle = \alpha_j|00\rangle + \beta_j|11\rangle$ with some probability $p_{i,j}$ for $j = 1, 2$. The linear combination of these Schmidt states also configures a Schmidt state. Then, we can write the Kraus operators using a Pauli basis, i.e., $K_i = \sum_{kl} r_{kl}\sigma_k \otimes \sigma_l$, and impose conditions (3) and (5). Normalization conditions then give us the constraint $2(|a|^4 + |b|^4) \leqslant 1$, where $a = r_{14}/4$ and $b = r_{34}/4$. Following this constraint, we can directly calculate the success probability $P_{\psi^{\otimes 2}\to\psi_S}$ and derive the upper bound $\overline{P}_{\psi^{\otimes 2}\to\psi_S} = 2(|c_1c_4| + |c_2c_3|)^2$, where $\overline{P}_{\psi^{\otimes 2}\to\psi_S} - P_{\psi^{\otimes 2}\to\psi_S} > 4(1-f)|c_1c_2c_3c_4|$, $0 < f < 1$, and $f = 2||a|^4 - |b|^4|$.

Note that the success probability $P_{\psi^{\otimes 2}\to\psi_S}$ can reach the bound $\overline{P}_{\psi^{\otimes 2}\to\psi_S}$ if and only if $f = 1$, which, in turn, requires $|a| = 1/\sqrt[4]{2}$ and $b = 0$ (or the opposite). However, the Kraus operators in Eq. (7) with such parameters can only generate product states (see Appendix B), which would be useless in the second stage of the protocol. In Fig. 2, we plot f as a function of $|a|$ and $|b|$. Only in the dark regions does the success probability get closer to the upper bound $\overline{P}_{\psi^{\otimes 2}\to\psi_S}$.

VI. OPTIMAL PROBABILITY OF PRODUCING ϕ^+ FROM FOUR COPIES OF UNKNOWN TWO-QUBIT STATE

Now, we shall combine Theorems 1 and 2 and derive the following result.

Theorem 3. The success probability of transforming four copies of $|\psi\rangle$ into ϕ^+ using a universal conclusive ECP based on concatenated two-qubit operations is bounded by

$$P_{\psi^{\otimes 4}\to\phi^+} \leqslant 2\left|c_2^4c_3^4\right| + 2\left|c_1^4c_4^4\right| - 4\mathrm{Re}\left[c_1^2c_4^2(c_2^*)^2(c_3^*)^2\right]. \tag{8}$$

Proof. First, note that the action of $M_i = K_i \otimes K_i$ from Theorem 2 on two copies of $|\psi\rangle = c_1|00\rangle + c_3|00\rangle + c_3|00\rangle + c_4|11\rangle$ results in

$$M_i|\psi\rangle^{\otimes 2} = (\alpha'|00\rangle + \beta'|11\rangle) \otimes |00\rangle =: |\psi'\rangle \otimes |00\rangle, \tag{9}$$

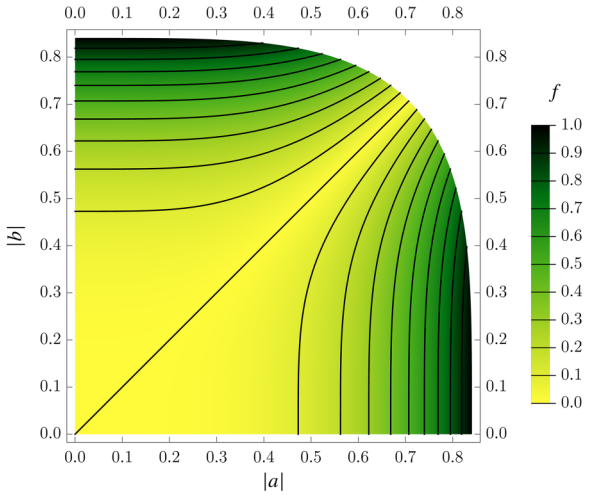

FIG. 2. Domain of the absolute value of coefficients a and b for Kraus operators satisfying Theorem 2: $\{a, b \in \mathbb{C} \mid 2(|a|^4 + |b|^4) \leqslant 1$ and $(a \neq 0$ or $b \neq 0)\}$. Function $f = 2||a|^4 - |b|^4|$ determines how much the success probability can get close to the bound $\overline{P}_{\psi^{\otimes 2} \to \psi_S}$; if $f \sim 1$, the bound is achieved.

where $|\psi'\rangle$ is an unnormalized state, $\alpha' = 2a^2(c_1c_4 + c_2c_3)$, and $\beta' = 2b^2(c_1c_4 - c_2c_3)$. The probability of success is given by the normalization constant $P' = \mathrm{Tr}|\psi'\rangle\langle\psi'|$. From Theorem 1, we know that the optimal universal conclusive ECP applied to two copies of a renormalized $|\psi'\rangle$ produces ϕ^+ with a success probability $P'' = 2|\alpha'\beta'|^2$. The full protocol applied to four copies is only successful if all steps are successful. Therefore, the final success probability is given by

$$\begin{aligned} P_{\psi^{\otimes 4} \to \phi^+} &= (P')^2 P'' \\ &= 2^5 |ab|^4 \left\{ \left| c_2^4 c_3^4 \right| + \left| c_1^4 c_4^4 \right| - 2\mathrm{Re}\left[c_1^2 c_4^2 (c_2^*)^2 (c_3^*)^2 \right] \right\}. \end{aligned} \tag{10}$$

From the domain of f, one can check that $2^5|ab|^4 \leqslant 2$, and this upper bound can be achieved when $a = b = \sqrt{2}/2$, which results in Eq. (8).

Next, we elaborate on why we focused on this concatenated structure.

First, there is a fundamental physical motivation. Because the input states have a completely unknown Schmidt basis, but the target maximally entangled state has a fixed, known basis, any universal protocol must eventually bridge this gap. Our concatenated approach explicitly isolates this requirement: the first stage transforms the unknown states into an intermediate state with a fixed, known Schmidt basis. Because we rigorously found the absolute optimal protocol for this specific basis-fixing task, our concatenation represents the optimal strategy among any protocol that enforces this intermediate step.

Second, there is a strong operational and experimental motivation. In current and near-term quantum architectures, the native entangling operations are strictly limited to two qubits. Implementing a genuine, collective four-qubit joint LOCC measurement is experimentally challenging and requires complex circuit decompositions highly susceptible to noise. Commonly used entanglement distillation protocols, such the Bennet-Brassard-Popescu-Schumacher-Smolin-Wootters (BBPSSW) [5] protocol and the Deutsch-Ekert-Jozsa-Macchiavello-Popescu-Sanpera (DEJMPS) [6] protocol, have a similar structure to the protocol considered in this work: they operate on two copies at a time and the same operations applied repeatedly throughout the protocol. This can also be seen in several experimental realizations of entanglement distillation [15,17,18,22]. Restricting our analysis to protocols based on two-qubit operations ensures that our derived limits and optimal protocols are practically relevant and experimentally feasible.

A direct consequence of the above result is stated below.

Corollary 1. If $a = b = \sqrt{2}/2$, then the map $\mathcal{E}_S$ from Theorem 2 with Kraus operators (7) is an optimal universal conclusive ECP when acting on two copies of $|\psi_S\rangle = \alpha|00\rangle + \beta|11\rangle$. Therefore, $\mathcal{E}_S = \mathcal{E}_{\phi^+}$.

Proof. The direct calculation of $\mathrm{Tr}\sum_i M_i \psi_S^{\otimes 2} M_i^\dagger$ with $a = b = \sqrt{2}/2$ provides the upper bound $2|\alpha\beta|^2$.

Using Corollary 1, we can see that Kálmán *et al.*'s protocol [24] is the optimal protocol that can conclusively produce Bell pairs in this scenario, following the same scheme in Fig. 1. The successful branch of the map has Kraus operators $M = K_{AA'} \otimes K_{BB'}$, such that $K_{AA'} = K_{BB'} =: K$ is given by

$$K = (H \otimes |0\rangle\langle 0|)U_{\mathrm{CNOT}}(\mathbb{1} \otimes \sigma_x), \tag{11}$$

where H is the Hadamard operator and U_{CNOT} is the cnot gate. The success probability of the first stage of Kálmán *et al.*'s protocol when applied to an arbitrary pure two-qubit state (1) is given by $P_{\mathrm{K}} = 2(|c_2c_3|^2 + |c_1c_4|^2)$ [24]. The second stage of the protocol, when applied to two copies of states resulting from the first stage, yields Bell pairs with a success probability given by

$$P'_{\mathrm{K}} = \frac{|(c_1c_4)^2 - (c_2c_3)^2|^2}{2(|c_2c_3|^2 + |c_1c_4|^2)^2}. \tag{12}$$

The entire protocol is only successful if every stage is successful. Therefore, the probability of obtaining one Bell pair from four copies of an arbitrary state (1) is given by $(P_{\mathrm{K}})^2 P'_{\mathrm{K}}$, which reproduces Eq. (8). In fact, if we use $a = b = \sqrt{2}/2$ derived in Theorem 3 in Eq, (7), we reproduce Eq. (11).

VII. COMPARING UNIVERSAL AND NONUNIVERSAL BOUNDS

When successful, a conclusive protocol always transforms the initial state ψ into the target state ϕ using LOCCs. The maximum success probability $P_{\mathrm{Vidal}}(\psi \to \phi)$ achievable by a conclusive protocol is given by Vidal's formula [21]

$$P_{\mathrm{Vidal}}(\psi \to \phi) = \min_{l \in \{1,\ldots,n\}} \frac{E_l(\psi)}{E_l(\phi)}, \tag{13}$$

where $E_l(\varrho) = \sum_{i=l}^n \alpha_i$, for every $l \in \{1, \ldots, n\}$, are entanglement monotones written as functions of the Schmidt coefficients $\{\alpha_i\}_{i=1}^n$ of a state ϱ, that is,

$$|\varphi\rangle = \sum_{i=1}^n \sqrt{\alpha_i}|ii\rangle_{AB}, \quad \alpha_i \geqslant \alpha_{i+1} \geqslant 0, \quad \sum_{i=1}^n \alpha_i = 1. \tag{14}$$

Note that Vidal's formula (13) provides the maximum probability P_{Vidal} for converting a specific known state ψ to ϕ but

does not yield a universal protocol (one that works equally for all inputs without prior knowledge). Rather, it states that for a given pair of states, the optimal conclusive protocol achieves probability P_{Vidal}. In this sense, Vidal's formula does not provide the optimal *universal* probability stated in Definition 2 since it optimizes the success probability on a per-state basis.

Let us consider the input state $|\psi_\lambda\rangle = \sqrt{\lambda}|00\rangle + \sqrt{1-\lambda}|11\rangle$ such that $\lambda \in (1/2, 1)$. The Schmidt coefficients $\{\alpha_i\}_{i=1}^4$ of $\psi_\lambda^{\otimes 2}$ are organized in nonincreasing order as $\{\lambda^2, \lambda(1-\lambda), \lambda(1-\lambda), (1-\lambda)^2\}$. From that, the entanglement monotones $E_l(\psi_\lambda^{\otimes 2})$ are given by

$$\left\{E_l\left(\psi_\lambda^{\otimes 2}\right)\right\}_{l=1}^4 = \{1, 1-\lambda^2, 1-\lambda, (1-\lambda)^2\}. \tag{15}$$

Note that the target state ϕ^+ has only two Schmidt coefficients since it is composed of only two qubits, unlike $\psi_\lambda^{\otimes 2}$. To use Vidal's formula, we can embed ϕ^+ in a bigger space, specifically $\mathcal{H}_{AB} \otimes \mathcal{H}_{A'B'}$, by attaching it to a pure state, e.g., $|00\rangle_{A'B'}$. Thus, the embedded target state is

$$|\phi^+\rangle_{AB} \otimes |00\rangle_{A'B'} = \frac{1}{\sqrt{2}}(|0000\rangle_{ABA'B'} + |1100\rangle_{ABA'B'}), \tag{16}$$

$$= \sum_{k=0}^{3} m_k |k\rangle_{AA'} \otimes |k\rangle_{BB'}, \tag{17}$$

where k is the transformation from binary to decimal basis, $m_0 = m_2 = 1/\sqrt{2}$, and $m_1 = m_3 = 0$. The Schmidt coefficients $\{\beta_i\}_{i=1}^4$ of $|\phi^+\rangle \otimes |00\rangle$ are organized in nonincreasing order as $\{\beta_i\}_{i=1}^4 := \{1/2, 1/2, 0, 0\}$. Similarly, the entanglement monotones $E_l(\phi^+)$ are given by

$$\{E_l(\phi^+)\}_{l=1}^4 = \left\{1, \tfrac{1}{2}, 0, 0\right\}. \tag{18}$$

Applying Eqs. (15) and (18) in Eq. (13) gives us $P_{\text{Vidal}}(\psi_\lambda^{\otimes 2} \to \phi^+) = \min_{i\in\{1,\ldots,4\}}\{1, 2(1-\lambda^2), +\infty, +\infty\}$. Finally, the optimal probability of conclusively transforming two copies of ψ_λ into ϕ^+ is given by

$$P_{\text{Vidal}}\left(\psi_\lambda^{\otimes 2} \to \phi^+\right) = \begin{cases} 1, & \text{if} \lambda \in (1/2, 1/\sqrt{2}), \\ 2(1-\lambda^2), & \text{if} \lambda \in [1/\sqrt{2}, 1). \end{cases} \tag{19}$$

While P_{Vidal} is optimized over all possible protocols given a specific state ψ_λ (i.e., a specific λ), the bound from Theorem 1 is the optimal probability that a single conclusive universal protocol can achieve for any state ψ_λ, i.e., for all λ's. From Theorem 1, we have

$$P_{\psi_\lambda^{\otimes 2} \to \phi^+} = 2\lambda(1-\lambda), \tag{20}$$

for $\lambda \in [1/2, 1]$. In this context, if we do not know which state is being produced by the source, but we know its Schmidt basis, we can always apply the same fixed optimal universal protocol to obtain ϕ^+ with some average probability. This average probability $\langle P_{\psi_S^{\otimes 2} \to \phi^+}\rangle$ can be obtained by integrating $P_{\psi_S^{\otimes 2} \to \phi^+}$ over $\lambda \in (1/2, 1)$, following a Haar measure $f(\lambda) = 6(2\lambda - 1)^2$ (see Eq. (3.7) in Ref. [25]). Explicitly,

$$\left\langle P_{\psi_\lambda^{\otimes 2} \to \phi^+}\right\rangle = \int_{\frac{1}{2}}^{1} P_{\psi_\lambda^{\otimes 2} \to \phi^+} f(\lambda) \mathrm{d}\lambda = \frac{1}{5}. \tag{21}$$

By applying the optimal universal protocol to an unknown source of states ψ_λ, we can obtain perfect Bell pairs with 20% probability. In Fig. 3 we present a comparison between bounds (19) and (20).

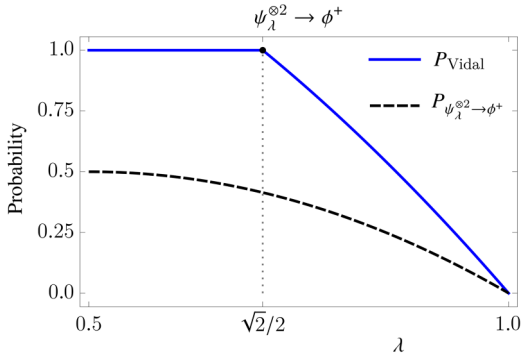

FIG. 3. Probability of conclusively transforming two copies of Schmidt states ψ_λ into a maximally entangled state ϕ^+ as a function of λ: (blue full line) when optimized over all possible protocols given by Vidal's formula (19); and (black dashed line) for the optimal universal protocol given by Eq. (20).

Now, let us assume that we do not know the Schmidt basis of the states produced by the source. If the source produces copies of an unknown pure two-qubit state, we can always apply Kálmán *et al.*'s protocol [24] to four copies of such states to obtain one Bell pair with some probability. The expected final probability can be calculated by using moments of a Dirichlet distribution, resulting in $\langle P_{\psi^{\otimes 4} \to \phi^+}\rangle = 2/105 \approx 1.9\%$ (see Appendix C).

VIII. CONCLUSION

In this work, we established rigorous optimality bounds for universal conclusive entanglement concentration protocols that transformed multiple copies of pure two-qubit states into a maximally entangled Bell state. We focused on protocols that operated without prior knowledge of the input state's entanglement structure, a key requirement for practical applications.

Our main contributions were threefold. First, for states with a known Schmidt basis, we derived the optimal success probability ($2|\alpha\beta|^2$) for converting two copies into a Bell state using LOCC (Theorem 1). Second, for arbitrary two-qubit pure states with unknown Schmidt basis, we determined the optimal probability of converting two copies into a state within a fixed Schmidt basis (Theorem 2), bounded by $2(|c_1c_4| + |c_2c_3|)^2$. Third, by combining these results, we obtained the tight bound in Eq. (8) for converting four copies of an arbitrary two-qubit pure state directly into a Bell state using concatenated protocols based on two-qubit operations (Theorem 3).

We also demonstrated that the protocol by Kalman *et al.* [24] saturated these bounds, proving its optimality as a universal conclusive ECP. Furthermore, we computed the expected success probability over Haar-random states, obtaining an average value of $\approx$1.9%, which served as a benchmark for practical implementations.

A critical insight from our work is the inherent trade-off between universality and efficiency: While Vidal's formula achieves higher conversion probabilities for specific input

states (e.g., $P_{\text{Vidal}} = 1$ for $\lambda < 1/\sqrt{2}$), universal protocols necessarily incur a success probability reduction [e.g., $2\lambda(1-\lambda)$] due to the constraint of operating without prior state knowledge. This quantifies the fundamental cost of universality in entanglement distillation.

While our work establishes tight bounds for the direct distillation of a Bell state from four copies of an arbitrary pure state, it is important to emphasize that these fundamental limits are strictly derived within the class of concatenated LOCC protocols based on two-qubit operations. We showed that this two-stage structure is highly motivated by both the physical necessity of fixing the unknown Schmidt basis and the practical constraints of near-term experimental hardware.

For future work, protocols that process four copies *directly* without intermediate Schmidt-state conversion warrant investigation. We conjecture that such protocols cannot surpass the success probability bound in Eq. (8), although our two-stage approach only rigorously saturates the fundamental limits for concatenated LOCC protocols based on two-qubit operations. This implies that Schmidt-basis alignment remains an optimal strategy for universal operations. Our framework provides a foundation for extending this analysis to mixed states and larger systems, with implications for real-world quantum communication architectures.

ACKNOWLEDGMENTS

The authors acknowledge the support from the National Science Centre Poland (Grant No. 2022/46/E/ST2/00115) and within the QuantERA II Program (Grant No. 2021/03/Y/ST2/00178, acronym ExTRaQT, and Grant No. 2021/03/Y/ST2/00177, acronym PhoMemtor) that received funding from the European Union's Horizon 2020 Research and Innovation Program under Grant Agreement No. 101017733.

DATA AVAILABILITY

There are no publicly available research data or software supporting this manuscript. Requests for further information or data should be sent to the authors.

APPENDIX A: PROOF OF THEOREM 1

Here, we present the detailed proof of Theorem 1 stated in the main part of the text.

Theorem A 1. Following Definition 2, the optimal universal conclusive ECP over the set of Schmidt states $|\psi_S\rangle = \alpha|00\rangle + \beta|11\rangle$ transforms two copies of such a state into ϕ^+ with optimal probability $P_{\psi_S^{\otimes 2}\to\phi^+} := 2|\alpha\beta|^2$.

Proof. Let M_i be the Kraus operators of the completely positive trace nonincreasing map $\mathcal{E}_1$ that is also a universal conclusive ECP. Any universal conclusive ECP $\mathcal{E}_1$ acting on pure states $|\psi_S\rangle_{AB} \otimes |\psi_S\rangle_{A'B'}$ to produce a maximally entangled state $|\phi^+\rangle$ must consist solely of Kraus operators M_i such that

$$M_i(|\psi_S\rangle_{AB} \otimes |\psi_S\rangle_{A'B'}) = \sqrt{p_i}|\phi^+\rangle_{AB} \otimes |\text{garb}\rangle_{A'B'}, \tag{A1}$$

where $|\text{garb}\rangle_{A'B'}$ is some product garbage state and $p_i \in [0, 1]$.

Since we are considering only LOCC protocols, product states cannot produce any entanglement, regardless of the universal conclusive ECPs. Therefore, the Kraus operator K_i must be such that

$$M_i|00\rangle_{AB}|00\rangle_{A'B'} = 0, \tag{A2a}$$

$$M_i|11\rangle_{AB}|11\rangle_{A'B'} = 0. \tag{A2b}$$

From the constraints above, the action of $\mathcal{E}_1$ on $|\psi_S\rangle^{\otimes 2}$ is

$$M_i|\psi_S\rangle^{\otimes 2} = \sqrt{2}\alpha\beta\ M_i\left(\frac{|0011\rangle + |1100\rangle}{\sqrt{2}}\right). \tag{A3}$$

Hence, the probability of success is given by

$$\text{Tr}\left[\mathcal{E}_1\left(\psi_S^{\otimes 2}\right)\right] = \text{Tr}\left[\sum_i p_i(2|\alpha\beta|^2)\phi^+ \otimes \text{garb}\right] \leqslant 2|\alpha\beta|^2, \tag{A4}$$

where p_i is the probability of converting $(|0011\rangle + |1100\rangle)/\sqrt{2}$ into ϕ^+. The inequality above implies that the maximal probability of obtaining ϕ^+ from $|\psi_S\rangle^{\otimes 2}$ using a universal conclusive ECP is bounded by $2|\alpha\beta|^2$. Therefore, following Definition 2, to be considered an optimal universal conclusive ECP from two copies of Schmidt states, a protocol must produce ϕ^+ with optimal universal probability $P_{\psi_S^{\otimes 2}\to\phi^+} := 2|\alpha\beta|^2$.

APPENDIX B: PROOF OF THEOREM 2

Below, we present the detailed proof of Theorem 2 stated in the main part of the text.

Theorem B 1. To produce arbitrary states in a known Schmidt basis from two copies of arbitrary states $|\psi\rangle_{AB}$ given by Eq. (1), a map $\mathcal{E}_S(\cdot) = \sum_i M_i(\cdot)M_i^\dagger$ must be composed of Kraus operators $M_i = K_i \otimes K_i$ acting on $\mathcal{H}_{AA'} \otimes \mathcal{H}_{BB'}$ such that

$$K_i = a(|00\rangle\langle 01| + |00\rangle\langle 10|) + b(|10\rangle\langle 01| - |10\rangle\langle 10|) \tag{B1}$$

for every i, where a and b are any nonzero complex numbers satisfying $2(|a|^4 + |b|^4) \leqslant 1$. The success probability $P_{\psi^{\otimes 2}\to\psi_S}$ of such a protocol is bounded by $\overline{P}_{\psi^{\otimes 2}\to\psi_S} = 2(|c_1c_4| + |c_2c_3|)^2$, i.e., $P_{\psi^{\otimes 2}\to\psi_S} < \overline{P}_{\psi^{\otimes 2}\to\psi_S}$.

Proof. From Eq. (5), Eq. (4) reduces to

$$M_i(|\psi\rangle^{\otimes 2}) = M_i[\sqrt{2}c_1c_4|\phi_1\rangle + \sqrt{2}c_2c_3|\phi_2\rangle], \tag{B2}$$

where we introduce the notation $|\phi_1\rangle := (|0011\rangle + |1100\rangle)/\sqrt{2}$ and $|\phi_2\rangle := (|0110\rangle + |1001\rangle)/\sqrt{2}$. Now, let us consider the situation when $c_2 = c_3 = 0$, which results in $M_i(|\psi\rangle^{\otimes 2}) = \sqrt{2}c_1c_4M_i|\phi_1\rangle$. Therefore, $\sum_i M_i\psi^{\otimes 2}M_i^\dagger = 2|c_1c_4|^2\sum_i M_i\phi_1M_i^\dagger$, which, by definition in Eq. (2), means that

$$2|c_1c_4|^2\sum_i M_i\phi_1M_i^\dagger = N_1\psi_S^{(1)}, \tag{B3}$$

where $|\psi_S^{(1)}\rangle = \alpha_1|00\rangle + \beta_1|11\rangle \in \mathcal{H}_{AB}$ for some pair of coefficients α_1, β_1, and N_1 is a normalization constant. Since both ϕ_1 and $\psi_S^{(1)}$ are pure states, we have that Eq. (B3) implies that

$$M_i|\phi_1\rangle = \sqrt{p_{i,1}}\left|\psi_S^{(1)}\right\rangle, \tag{B4}$$

for every i, where $p_{i,1}$ is a probability distribution satisfying $\sum_i p_{i,1} = N_1/2|c_1c_4|^2$. Note that this last equation implies that $N_1 \leqslant 2|c_1c_4|^2$. A similar argument provides

$$M_i|\phi_2\rangle = \sqrt{p_{i,2}}\left|\psi_S^{(2)}\right\rangle \tag{B5}$$

for every i, where $p_{i,2}$ is a probability distribution satisfying $\sum_i p_{i,2} = N_2/2|c_2c_3|^2$. As before, we also have that $N_2 \leqslant 2|c_2c_3|^2$. Now, using Eqs. (B4) and (B5) in Eq. (B2) gives us

$$\begin{aligned} M_i|\psi\rangle^{\otimes 2} &= c_1c_4\sqrt{2p_{i,1}}\left|\psi_S^{(1)}\right\rangle + c_2c_3\sqrt{2p_{i,2}}\left|\psi_S^{(2)}\right\rangle \\ &= (c_1c_4\sqrt{2p_{i,1}}\alpha_1 + c_2c_3\sqrt{2p_{i,2}}\alpha_2)|00\rangle \\ &\quad + (c_1c_4\sqrt{2p_{i,1}}\beta_1 + c_2c_3\sqrt{2p_{i,2}}\beta_2)|11\rangle. \end{aligned} \tag{B6}$$

Finally, from Eq. (2), the probability $P_{\psi^{\otimes 2}\to\psi_S} = \mathrm{Tr}\sum_i M_i\psi^{\otimes 2}M_i^\dagger$ is expressed as

$$\begin{aligned} P_{\psi^{\otimes 2}\to\psi_S} = 2\sum_i (&|c_1c_4\sqrt{p_{i,1}}\alpha_1 + c_2c_3\sqrt{p_{i,2}}\alpha_2|^2 \\ &+ |c_1c_4\sqrt{p_{i,1}}\beta_1 + c_2c_3\sqrt{p_{i,2}}\beta_2|^2). \end{aligned} \tag{B7}$$

Now, let us write each Kraus operator K_i in Eq. (6) using a Pauli basis as follows:

$$K_i = \sum_{k,l=1}^{4} r_{kl}\sigma_k \otimes \sigma_l, \tag{B8}$$

where $(\sigma_1, \sigma_2, \sigma_3, \sigma_4) = (\sigma_x, \sigma_y, \sigma_z, \mathbb{1})$ and $r_{kl}^i \in \mathbb{R}_+$ for every k, l, i. In principle, r_{kl} could be distinct for every i, but we are going to omit this extra index unless necessary. By using Eq. (B8) in system (3), we obtain, after some algebra,

$$\begin{aligned} &r_{21} = -r_{12}, \quad r_{22} = r_{11}, \quad r_{23} = \mathbb{i}r_{14}, \quad r_{24} = \mathbb{i}r_{13}, \\ &r_{41} = -\mathbb{i}r_{32}, \quad r_{42} = \mathbb{i}r_{31}, \quad r_{43} = -r_{34}, \quad r_{44} = -r_{33}, \end{aligned} \tag{B9}$$

for every i. We can insert Eqs. (B9) into (B8) and solve Eqs. (B4) and (B5). Note that the auxiliary state on the right-hand side of Eqs. (B4) and (B5) is omitted. Without loss of generality and to make the equations more tractable, we can suppose that $|\mathrm{aux}\rangle = |00\rangle_{A'B'}$. Also, both $|\psi_S^{(1,2)}\rangle$ are vectors such that 14 out of 16 entries are null. The resolution of these 14 equations yields eight equivalent solutions, up to a complex phase that can be accommodated by the remaining free coefficients. One of those solutions is provided below:

$$\begin{aligned} &r_{11} = r_{34}, \quad r_{12} = \mathbb{i}r_{34}, \quad r_{13} = r_{14}, \\ &r_{31} = r_{14}, \quad r_{32} = \mathbb{i}r_{14}, \quad r_{33} = r_{34}. \end{aligned} \tag{B10}$$

From the relations above and Eq. (B9), the Kraus operators K_i can be expressed in terms of only two complex parameters, namely, $r_{14} := a/4$ and $r_{34} := b/4$, as

$$K_i = a(|00\rangle\langle 01| + |00\rangle\langle 10|) + b(|10\rangle\langle 01| - |10\rangle\langle 10|). \tag{B11}$$

Remember here that M_i acts on $\mathcal{H}_{AA'} \otimes \mathcal{H}_{BB'}$ such that $M_i = K_i \otimes K_i$, where each K_i acts locally and identically on each laboratory, i.e., AA' and BB'. Finally, by using Eq. (B11) in Eqs. (B4) and (B5), we obtain $\sqrt{p_{i,j}}\alpha_j = \sqrt{2}a^2$ and $\sqrt{p_{i,j}}\beta_j = (-1)^{j-1}\sqrt{2}b^2$ for $j = 1, 2$. By definition, $|\alpha_j|^2 + |\beta_j|^2 = 1$ and $0 \leqslant p_{i,j} \leqslant 1$ for $j = 1, 2$. Therefore, coefficients a and b must satisfy

$$2(|a|^4 + |b|^4) \leqslant 1. \tag{B12}$$

In addition, states $|\psi_S^{(1,2)}\rangle$ must be entangled, which requires α_1, β_1 or α_2, β_2 to be nonzero. That implies that both a and b must be nonzero.

Now, let us obtain the upper bound for $P_{\psi^{\otimes 2}\to\psi_S}$. In Eq. (B7), define $A_i = c_1c_4\sqrt{p_{i,1}}\alpha_1 + c_2c_3\sqrt{p_{i,2}}\alpha_2$ and $B_i = c_1c_4\sqrt{p_{i,1}}\beta_1 + c_2c_3\sqrt{p_{i,2}}\beta_2$. Therefore,

$$\begin{aligned} |A_i|^2 + |B_i|^2 &= |c_1c_4|^2 p_{i,1} + |c_2c_3|^2 p_{i,2} \\ &\quad + 2\mathrm{Re}(c_1c_4c_2^*c_3^*g)\sqrt{p_{i,1}p_{i,2}}, \end{aligned} \tag{B13}$$

where $g := \alpha_1\alpha_2^* + \beta_1\beta_2^*$. Since for any complex number z, it is always true that $\mathrm{Re}(z) \leqslant |z|$, we have

$$\mathrm{Re}(c_1c_4c_2^*c_3^*g) \leqslant |c_1c_4||c_2c_3||g|. \tag{B14}$$

Observe also that $|g| = f/\sqrt{p_{i,1}p_{i,2}}$, where $f = 2||a|^4 - |b|^4|$. From Eq. (B12), we have that $f \in [0, 1]$. In particular, $f = 1$ iff $|a| = \frac{1}{\sqrt[4]{2}}$ and $b = 0$ or the opposite, which results in unentangled Schmidt states. Therefore, we must impose $f < 1$, which makes Eq, (B13) become

$$\begin{aligned} \sum_i |A_i|^2 + |B_i|^2 &< |c_1c_4|^2\left(\sum_i p_{i,1}\right) + |c_2c_3|^2\left(\sum_i p_{i,2}\right) \\ &\quad + 2|c_1c_4||c_2c_3|. \end{aligned} \tag{B15}$$

Since $\sum_i p_{i,j} \leqslant 1$ for $j = 1, 2$, we obtain

$$P_{\psi^{\otimes 2}\to\psi_S} < 2(|c_1c_4| + |c_2c_3|)^2 =: \overline{P}_{\psi^{\otimes 2}\to\psi_S}. \tag{B16}$$

Consequently, the maximum value of $P_{\psi^{\otimes 2}\to\psi_S}$ is $1/2$.

APPENDIX C: EXPECTED PROBABILITY WITH HAAR MEASURE

Here we calculate the expected final probability $\langle P_{\psi^{\otimes 4}\to\phi^+}\rangle$. When the source produces copies of an unknown pure two-qubit state, we can always apply Kálmán *et al.*'s protocol to a set of four states to obtain one Bell pair with some probability. Like before, this average probability can be obtained by integrating Eq. (8) over $\{c_i\}_{i=1}^4$ using a Haar measure. Under Haar measure, the squared amplitudes $x_i = |c_i|^2$ follow a Dirichlet distribution Dir(1, 1, 1, 1) [26], while the phases $\theta_i = \arg(c_i)$ are independent and uniformly distributed in the interval $[0, 2\pi]$. First, let us show that the average of the last term in Eq. (8) is zero. Since $c_1^2c_4^2(c_2^*)^2(c_3^*)^2 = x_1x_2x_3x_4e^{\mathbb{i}2(\theta_1+\theta_4-\theta_2-\theta_3)}$, then

$$\mathrm{Re}\left[c_1^2c_4^2(c_2^*)^2(c_3^*)^2\right] = x_1x_2x_3x_4\cos\eta, \tag{C1}$$

where we define $\eta = 2(\theta_1 + \theta_4 - \theta_2 - \theta_3) \mod 2\pi$, which is uniformly distributed in $[0, 2\pi)$. Let us fix $\mathbf{x} = (x_1, x_2, x_3, x_4)$. A direct integration provides the conditional expectation $\mathbb{E}[\cos\eta|\mathbf{x}] = 0$. Because $|c_i^4| = |c_i|^4 = x_i^2$, the final expectation value is given by

$$\mathbb{E}\left[P_{\mathrm{K}}^{\mathrm{final}}\right] = \mathbb{E}\left[\mathbb{E}\left[P_{\mathrm{K}}^{\mathrm{final}}|\mathbf{x}\right]\right] = 2\mathbb{E}\left[x_1^2x_4^2\right] + 2\mathbb{E}\left[x_2^2x_3^2\right]. \tag{C2}$$

By symmetry of the Dirichlet distribution Dir(1, 1, 1, 1), we have that $\mathbb{E}[x_1^2x_4^2] = \mathbb{E}[x_2^2x_3^2]$. The general formula for the

moments of the distribution Dir($\boldsymbol{\alpha}$) is given by [27]

$$\mathbb{E}\left[\prod_i x_i^{\beta_i}\right] = \frac{\Gamma\left(\sum_i \alpha_i\right)}{\Gamma\left(\sum_i \alpha_i + \sum_i \beta_i\right)} \prod_i \frac{\Gamma(\alpha_i + \beta_i)}{\Gamma(\alpha_i)}, \quad \text{(C3)}$$

where the Gamma function for integers is simply $\Gamma(n) = (n-1)!$. To calculate $\mathbb{E}(x_1^2 x_4^2)$ we must use $(\alpha_1, \alpha_2, \alpha_3, \alpha_4) = (1, 1, 1, 1)$ and $(\beta_1, \beta_2, \beta_3, \beta_4) = (2, 0, 0, 2)$. Having said that, from Eq. (C3), we obtain $\mathbb{E}(x_1^2 x_4^2) = 1/210$. Identical arguments give us $\mathbb{E}(x_2^2 x_3^2) = 1/210$, which results in the expected final probability $\langle P_{\psi^{\otimes 4} \to \phi^+} \rangle = 2/105$. Numerically, we obtain $\langle P_{\psi^{\otimes 4} \to \phi^+} \rangle \sim 0.0190$ for 10^4 initial states (1) where the real and imaginary parts of each c_i follow a Gaussian probability distribution with a mean of zero and standard deviation $1/\sqrt{2}$.

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

[2] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels, Phys. Rev. Lett. **70**, 1895 (1993).

[3] A. K. Ekert, Quantum cryptography based on Bell's theorem, Phys. Rev. Lett. **67**, 661 (1991).

[4] P. W. Shor, Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer, SIAM J. Comput. **26**, 1484 (2006).

[5] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, Purification of noisy entanglement and faithful teleportation via noisy channels, Phys. Rev. Lett. **76**, 722 (1996).

[6] D. Deutsch, A. Ekert, R. Jozsa, C. Macchiavello, S. Popescu, and A. Sanpera, Quantum privacy amplification and the security of quantum cryptography over noisy channels, Phys. Rev. Lett. **77**, 2818 (1996).

[7] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Mixed-state entanglement and quantum error correction, Phys. Rev. A **54**, 3824 (1996).

[8] W Dür and H. J. Briegel, Entanglement purification and quantum error correction, Rep. Prog. Phys. **70**, 1381 (2007).

[9] J. M. Torres and J. Z. Bernád, Conditions for entanglement purification with general two-qubit states, Phys. Rev. A **94**, 052329 (2016).

[10] F. Rozpędek, T. Schiet, L. P. Thinh, D. Elkouss, A. C. Doherty, and S. Wehner, Optimizing practical entanglement distillation, Phys. Rev. A **97**, 062333 (2018).

[11] F. Preti, T. Calarco, J. M. Torres, and J. Z. Bernád, Optimal two-qubit gates in recurrence protocols of entanglement purification, Phys. Rev. A **106**, 022422 (2022).

[12] J. Miguel-Ramiro, A. Pirker, and W. Dür, Improving entanglement purification through coherent superposition of roles, Quantum **9**, 1702 (2025).

[13] J.-W. Pan, C. Simon, Č. Brukner, and A. Zeilinger, Entanglement purification for quantum communication, Nature (London) **410**, 1067 (2001).

[14] J.-W. Pan, S. Gasparoni, R. Ursin, G. Weihs, and A. Zeilinger, Experimental entanglement purification of arbitrary unknown states, Nature (London) **423**, 417 (2003).

[15] X.-M. Hu, C.-X. Huang, Y.-B. Sheng, L. Zhou, B.-H. Liu, Y. Guo, C. Zhang, W.-B. Xing, Y.-F. Huang, C.-F. Li, and G.-C. Guo, Long-distance entanglement purification for quantum communication, Phys. Rev. Lett. **126**, 010503 (2021).

[16] S. Ecker, P. Sohr, L. Bulla, M. Huber, M. Bohmann, and R. Ursin, Experimental single-copy entanglement distillation, Phys. Rev. Lett. **127**, 040506 (2021).

[17] R. Reichle, D. Leibfried, E. Knill, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland, Experimental purification of two-atom entanglement, Nature (London) **443**, 838 (2006).

[18] N. Kalb, A. A. Reiserer, P. C. Humphreys, J. J. W. Bakermans, S. J. Kamerling, N. H. Nickerson, S. C. Benjamin, D. J. Twitchen, M. Markham, and R. Hanson, Entanglement distillation between solid-state quantum network nodes, Science **356**, 928 (2017).

[19] A. Zang, X. Chen, E. Chitambar, M. Suchara, and T. Zhong, No-go theorems for universal entanglement purification, Phys. Rev. Lett. **134**, 190803 (2025).

[20] G. Vidal, D. Jonathan, and M. A. Nielsen, Approximate transformations and robust manipulation of bipartite pure-state entanglement, Phys. Rev. A **62**, 012304 (2000).

[21] G. Vidal, Entanglement of pure states for a single copy, Phys. Rev. Lett. **83**, 1046 (1999).

[22] L.-K. Chen, H.-L. Yong, P. Xu, X.-C. Yao, T. Xiang, Z.-D. Li, C. Liu, H. Lu, N.-L. Liu, L. Li, *et al.*, Experimental nested purification for a linear optical quantum repeater, Nat. Photon. **11**, 695 (2017).

[23] L. Zhou, C.-X. Huang, Y.-B. Sheng, Y. Guo, X.-M. Hu, Y.-F. Huang, C.-F. Li, G.-C. Guo, and B.-H. Liu, Observation of residual entanglement in entanglement purification, Phys. Rev. Lett. **135**, 050801 (2025).

[24] O. Kálmán, A. Gábris, I. Jex, and T. Kiss, Universal, unambiguous concentration and distillation of Bell pairs, Phys. Rev. Lett. **135**, 260202 (2025).

[25] K. Zyczkowski and H.-J. Sommers, Induced measures in the space of mixed quantum states, J. Phys. A: Math. Gen. **34**, 7111 (2001).

[26] I. Bengtsson and K. Życzkowski, Quantum operations, in *Geometry of Quantum States: An Introduction to Quantum Entanglement*, 2nd ed. (Cambridge University Press, Cambridge, 2017), p. 56.

[27] L. D. Schiavo, Characteristic functionals of Dirichlet measures, Electron. J. Probab. **24**, 1 (2019).

[28] K. Fang and Z.-W. Liu, No-go theorems for quantum resource purification, Phys. Rev. Lett. **125**, 060405 (2020).

[29] B. Regula, K. Bu, R. Takagi, and Z.-W. Liu, Benchmarking one-shot distillation in general quantum resource theories, Phys. Rev. A **101**, 062315 (2020).