Complete Characterization of Entanglement Embezzlement

Elia Zanoni[1,2], Thomas Theurer[1,2], and Gilad Gour[1,2]

[1]Department of Mathematics and Statistics, University of Calgary, Calgary, AB T2N 1N4, Canada

[2]Institute for Quantum Science and Technology, University of Calgary, Calgary, AB T2N 1N4, Canada

Using local operations and classical communication (LOCC), entanglement can be manipulated but not created. However, entanglement can be *embezzled*. In this work, we completely characterize universal embezzling families and demonstrate how this singles out the original family introduced by van Dam and Hayden. To achieve this, we first give a full characterization of pure to mixed state LOCC-conversions. Then, we introduce a new conversion distance and derive a closed-form expression for it. These results might be of independent interest.

1 Introduction

Quantum entanglement [1–6] describes correlations between different particles with no classical counterpart and has both deep foundational implications [7] and numerous applications in quantum information science [8–18]. If two or more parties are far apart, in practice, they are restricted to local operations and classical communication (LOCC) because during transmission over long distances, physical systems will unavoidably interact with an environment and eventually lose the quantum information they carry. In contrast, it is much simpler to exchange classical information that can easily be amplified and protected. The states that can be prepared with LOCC, i.e., the separable ones, are therefore considered free and the ones that cannot, which are exactly the entangled ones, are considered costly or resourceful.

According to this reasoning, entanglement is a resource and thus studied within the framework of quantum resource theories (QRTs) [19, 20]. In a QRT, a physically motivated restriction such as the one discussed above divides both states and operations into free or resourceful in a consistent manner: Free operations map free states into free states. Once the free operations and states have been fixed, a QRT studies which quantum advantages depend on the resource under consideration and, closely related, how the consumption of resourceful states can help to overcome the restriction. An example is quantum teleportation [11]: By consuming entangled states, LOCC allows to teleport quantum systems and thereby simulate arbitrary operations outside LOCC.

In this sense, consuming entangled states or more generally resourceful states can lead to operational advantages, e.g., in communication scenarios [8–15]. Importantly, if a state can be converted with a free operation into another one, then the former is at least as valuable as the latter in any application that only allows for free operations. Answering the question which states can be converted into each other is thus a central question in

Elia Zanoni: elia.zanoni@ucalgary.ca

any QRT. In entanglement theory, Nielsen's Theorem [3] provides a characterization of the deterministic conversion between pure states that was later extended to probabilistic conversions [21]. Here, we generalize these results and characterize the conversion of a pure state into a mixed state.

The results on state conversions mentioned so far require that an initial state is exactly converted into a target state. However, every physical setup is affected by noise and finitely many measurements will only lead to finite accuracy. It is therefore practically impossible to distinguish the desired target state from a close approximation. This is the physical motivation to investigate whether a state can be *approximately* converted to the target using free operations [22]. There are several equivalent distances that formalize a notion of closeness between quantum states [23–27], which can be used to define conversion distances [28–31]. These conversion distances measure the error in approximate conversions, i.e., given two states, they are the smallest distance between the second state and the result of any free operation on the first one. In this article, we introduce a new conversion distance under LOCC defined on pure states. Using our result on exact conversions, we prove that it is topologically equivalent to the other conversion distances and derive a closed formula for it.

This in turn will allow us to completely characterize entanglement embezzlement [32, 33]: Whilst it is impossible to create entanglement with LOCC [2, 5, 19, 34], it is possible to embezzle it in the sense that one converts a given entangled state approximately to itself *and* a copy of another entangled target state. A family of pure states such that one can do this for any target state with arbitrary accuracy using LOCC and one of its members is called a universal embezzling family. Embezzlement is thus a generalization of quantum catalysis, a phenomenon discovered in the early years of the resource theory of entanglement [35]: Whilst in catalysis, the catalyst must be preserved exactly, embezzlement allows to change it by an arbitrarily small amount. Very recently, we have witnessed a renewed interest in the catalysis of various quantum resources [36–52], with applications far beyond state conversion. In Ref. [46], it was for example shown that a catalytic quantum teleportation protocol outperforms the standard teleportation protocol [11, 16, 17] in terms of teleportation fidelity. For further details and examples, see the recent review articles Refs. [53, 54].

Universal embezzling families are valuable resources in various applications. They are, for example, necessary resources for the efficient simulation of noisy quantum channels with noiseless channels [15, 18], a result known in quantum information theory as 'Quantum Reverse Shannon Theorem'. They are also an important component in the elementary proofs of Grothendiek theorems [55, 56], which are of fundamental importance in the theories of Banach spaces and C^*- algebras (see, e.g., Ref [57]). Furthermore, embezzling families are necessary to win various quantum guessing games with certainty [58–61]. Lastly, due to the close relation to catalysis, we expect that they will prove useful in applications such as teleportation, where catalysis provides advantages.

The first universal embezzling family was introduced in Ref. [32] and we call it van Dam and Hayden embezzling family. More recently, additional families have been proposed in Ref. [33]. In these works, embezzlement is considered using only local operations. Here, we extend the framework and additionally allow for classical communication, provide a complete characterization of universal embezzling families under LOCC, and discuss in what sense the van Dam and Hayden family is unique.

2 Notation and Preliminaries

In this article, we restrict ourselves to finite-dimensional Hilbert spaces and denote them with capital Latin letters such as A, B. The dimension of a Hilbert space C is denoted by $|C|$, and the set of density matrices acting on it by $\mathfrak{D}(C)$. For the set of pure states in $\mathfrak{D}(C)$, we write PURE (C). Small Greek letters such as ρ and σ denote density matrices, with ψ, φ, and χ reserved for pure states. For $\rho, \sigma \in \mathfrak{D}(C)$, the trace distance between ρ and σ is defined as $\frac{1}{2} \|\rho - \sigma\|_1$, where $\|\cdot\|_1$ is the trace norm.

Probability vectors are represented by bold small Latin letters, e.g., $\mathbf{p}$, with p_x the x-th component of $\mathbf{p}$. The set of probability vectors of length d is denoted by $\mathrm{Prob}(d)$, which contains the subset $\mathrm{Prob}^{\downarrow}(d)$ consisting of all d-dimensional probability vectors with non-increasing entries. The k-th Ky Fan norm of $\mathbf{p} \in \mathrm{Prob}^{\downarrow}(d)$ is defined as

$$\|\mathbf{p}\|_{(k)} = \sum_{x=1}^{k} p_x \tag{1}$$

for $k \in [d]$, where $[d]$ is a shorthand notation for $\{1, \ldots, d\}$. We write $\mathbf{p} \succ \mathbf{q}$ if $\mathbf{p}$ majorizes $\mathbf{q}$ [62–64] and $\|\mathbf{p} - \mathbf{q}\|_1$ for $\sum_x |p_x - q_x|$.

Since this article is concerned with two spatially separated parties, call them Alice and Bob, it is important to make clear which system is under the control of whom: Systems belonging to Alice will always be denoted by A or A', and systems belonging to Bob by B or B'. Quantum channels are represented by calligraphic large Latin letters such as $\mathcal{M}$, $\mathcal{N}$, and the set of quantum channels from a bipartite system AB to $A'B'$ that Alice and Bob can implement if they are restricted to local operations and classical communication is denoted by $\mathrm{LOCC}(AB \to A'B')$. For bipartite systems AB, we assume w.l.o.g. that $|A| = |B| = d$ (since with LOCC, Alice and Bob can always attach and remove local auxiliary systems). Fixing an orthonormal basis $\{|x\rangle_A\}_{x \in [d]}$ for A and $\{|x\rangle_B\}_{x \in [d]}$ for B, under LOCC, every $\psi \in \mathrm{PURE}(AB)$ is then equivalent to its standard form $\sum_x \sqrt{p_x} |xx\rangle_{AB}$, where $\mathbf{p} \in \mathrm{Prob}^{\downarrow}(d)$ are the Schmidt coefficients of ψ (which, w.l.o.g., we will always assume to be ordered non-increasingly from here on). We denote with $\mathrm{SR}(\psi)$ the Schmidt rank of a pure bipartite state $\psi \in \mathrm{PURE}(AB)$, i.e., the number of non-zero Schmidt coefficients of ψ, and use the symbol $\Phi_m = \sum_{x=1}^{m} \frac{1}{\sqrt{m}} |xx\rangle_{AB}$ for the maximally entangled state on AB, where $|A| = |B| = m$. Finally, for $\rho \in \mathfrak{D}(AB)$ and $\sigma \in \mathfrak{D}(A'B')$, we write $\rho \xrightarrow{\mathrm{LOCC}} \sigma$ whenever there exists an $\mathcal{N} \in \mathrm{LOCC}(AB \to A'B')$ such that $\sigma = \mathcal{N}(\rho)$ and $\rho \xrightarrow{\mathrm{LOCC}} \{t_z, \tau_z\}$ whenever there exists a probabilistic LOCC protocol that converts ρ to τ_z with probability t_z.

3 State conversions with LOCC

As motivated in the introduction, the question of how entangled states can be interconverted is at the core of the resource theory of entanglement. The exact deterministic LOCC-conversion problem between pure states is solved by Nielsen's Theorem [3]: For ψ, $\varphi \in \mathrm{PURE}(AB)$ and $\mathbf{p}$, $\mathbf{q} \in \mathrm{Prob}^{\downarrow}(|A|)$ their associated Schmidt vectors, $\psi \xrightarrow{\mathrm{LOCC}} \varphi$ if and only if [65]

$$E_k(\psi) \geq E_k(\varphi) \quad \forall k \in [|A|] , \tag{2}$$

where $E_k(\psi) := 1 - \|\mathbf{p}\|_{(k)}$. In Ref. [21], this result was generalized to the case where the target state φ is replaced with an ensemble of states: For ψ, φ_1, $\ldots$, $\varphi_n \in \mathrm{PURE}(AB)$,

$\psi \xrightarrow{\text{LOCC}} \{ t_z, \varphi_z \}_{z \in [n]}$ if and only if

$$E_k(\psi) \geq \sum_{z=1}^{n} t_z E_k(\varphi_z) \quad \forall k \in [|A|], \tag{3}$$

which can be rewritten as

$$\min_{k \in [|A|]} \left\{ E_k(\psi) - \sum_z t_z E_k(\varphi_z) \right\} \geq 0. \tag{4}$$

We extend this result considering a generic mixed state as target.

Proposition 1. *For $\psi \in \text{PURE}(AB)$ and $\sigma \in \mathfrak{D}(AB)$, $\psi \xrightarrow{\text{LOCC}} \sigma$ if and only if there exists a pure state decomposition $\{ p_z, \chi_z \}$ of σ (i.e., $\sigma = \sum_z p_z \chi_z$) such that*

$$\min_{k \in [|A|]} \left\{ E_k(\psi) - \sum_z p_z E_k(\chi_z) \right\} \geq 0. \tag{5}$$

That this condition is sufficient follows directly from Eq. (4). That it is also necessary follows from the Lo-Popescu Theorem [66] which implies that if $\rho \xrightarrow{\text{LOCC}} \sigma$, then there exists a pure state decomposition $\{ p_z, \chi_z \}$ of σ that satisfies Eq. (4). The details of this proof can be found in Appendix A.

For certain choices of $\rho \in \mathfrak{D}(AB)$ and $\sigma \in \mathfrak{D}(A'B')$, ρ cannot be converted to σ by LOCC. It is then interesting to determine how well we can *approximate* σ with ρ and LOCC. To this end, one can associate with every distance D defined on quantum states the conversion distance

$$D(\rho \to \sigma) = \inf_{\mathcal{N} \in \text{LOCC}(AB \to A'B')} D(\mathcal{N}(\rho), \sigma). \tag{6}$$

This conversion distance determines how close to σ, with respect to the distance D, one can convert ρ using only LOCC. A commonly used distance in quantum information is the trace distance $T(\sigma, \tau) = \frac{1}{2} \|\sigma - \tau\|_1$. The conversion distance associated with it is

$$T(\rho \to \sigma) = \inf_{\mathcal{N} \in \text{LOCC}(AB \to A'B')} \frac{1}{2} \|\mathcal{N}(\rho) - \sigma\|_1. \tag{7}$$

This conversion distance has an operational interpretation in terms of a result in state discrimination known as Holevo-Helstrom Theorem [67, 68] (see Ref. [69, Theorem 3.4] for a review). Indeed, if a single copy of either $\sigma \in \mathfrak{D}(AB)$ or $\mathcal{N}(\rho) \in \mathfrak{D}(AB)$, with $\mathcal{N} \in \text{LOCC}$, is given with equal probability, then the maximum probability $p_{\max}$ of correctly identifying the given state is bounded by

$$p_{\max} \geq \frac{1}{2}(1 + T(\rho \to \sigma)). \tag{8}$$

Moreover, for every fixed ρ and σ, there always exists an $\mathcal{N} \in \text{LOCC}$ such that $p_{\max}$ is arbitrarily close to this lower bound. In this sense, $T(\rho \to \sigma)$ describes how well we can approximate σ given access to ρ and LOCC.

Another distance used often in quantum information is the purified distance $P(\sigma, \tau) = \inf_{\psi_\sigma, \psi_\tau} T(\psi_\sigma, \psi_\tau)$ [26, 27], where the infimum runs over all purifications ψ_σ and ψ_τ of σ and τ, respectively. In this case, the conversion distance is analogously defined as

$$P(\rho \to \sigma) = \inf_{\mathcal{N} \in \text{LOCC}(AB \to A'B')} P(\mathcal{N}(\rho), \sigma). \tag{9}$$

On pure states ψ, $\varphi \in \text{PURE}(AB)$, we define what we call star conversion distances via

$$D_\star(\psi \to \varphi) = \min_{\mathbf{r} \succ \mathbf{p}} D(\mathbf{r}, \mathbf{q}), \tag{10}$$

where $\mathbf{p}$, $\mathbf{q} \in \text{Prob}^\downarrow(|A|)$ are the Schmidt coefficients of ψ and φ, respectively, and $D(\mathbf{r}, \mathbf{q}) = D(\text{diag}(\mathbf{r}), \text{diag}(\mathbf{q}))$. In general $D_\star(\psi \to \varphi) \neq D(\psi \to \varphi)$. However, we show in Appendix B that in the case of the purified distance, the two conversion distances coincide on pure states.

Theorem 2. *Let ψ, $\varphi \in \text{PURE}(AB)$. Then $P(\psi \to \varphi) = P_\star(\psi \to \varphi)$.*

This, in turn, is useful to show that $T(\psi \to \varphi)$ and $T_\star(\psi \to \varphi)$ are topologically equivalent on pure states, which means that if one can approximate φ arbitrarily well with a sequence of states $\{\psi_n\}$ in the sense that $\lim_{n \to \infty} T(\psi_n \to \varphi) = 0$, then the same is true for $T_\star$, and vice-versa. This is guaranteed by the following result (Appendix C).

Lemma 3. *Let ψ, $\varphi \in \text{PURE}(AB)$. Then*

$$\frac{1}{2} \left[T_\star(\psi \to \varphi) \right]^2 \leq T(\psi \to \varphi) \leq \sqrt{2 T_\star(\psi \to \varphi)}. \tag{11}$$

We are particularly interested in $T_\star(\psi \to \varphi)$ because using tools from approximate majorization [62–64, 70–73], one can derive a closed-form expression for it (Appendix C).

Theorem 4. *Let ψ, $\varphi \in \text{PURE}(AB)$ and let $\mathbf{p}$, $\mathbf{q} \in \text{Prob}^\downarrow(|A|)$ be their corresponding Schmidt coefficients. Then,*

$$T_\star(\psi \to \varphi) = \max_{k \in [\text{SR}(\psi)]} \left\{ \|\mathbf{p}\|_{(k)} - \|\mathbf{q}\|_{(k)} \right\}. \tag{12}$$

If $\xi \in \text{PURE}(AB)$ is separable and ψ, φ are not,

$$T_\star(\psi \to \varphi) < T_\star(\xi \to \varphi). \tag{13}$$

This theorem provides, to the best of our knowledge, the first algorithm to compute a conversion distance in LOCC with a finite number of steps. Indeed, if ψ, φ, and their Schmidt coefficients are known, then one can compute $T_\star(\psi \to \varphi)$ using a finite memory, and a finite (perhaps very large) number of operations (additions or subtractions). This is not the case for the other conversion distances because they require a minimization over LOCC, which is in general unfeasible. The second part of the theorem shows that any entangled state is more useful in the approximation of all other entangled states than any separable state.

Whilst Theorem 4 is of independent interest, for example in the characterization of entanglement distillation and dilution [74], we discuss in the following how it yields new results concerning the embezzlement of entanglement. When referring to the conversion distance and star conversion distance, we will thus refer to the versions based on the trace distance from here on.

4 Entanglement embezzlement

As mentioned in the introduction, it is impossible to create additional entanglement with LOCC alone [2, 5, 19, 34]. If $\rho \in \mathfrak{D}(A'B')$ is an entangled state, this implies that there cannot exist a $\psi \in \text{PURE}(AB)$ and a channel $\mathcal{N} \in \text{LOCC}(AB \to ABA'B')$ such that

$\mathcal{N}(\psi) = \psi \otimes \rho$, because this would increase the total amount of entanglement between systems AA' and BB' with respect to any additive entanglement measure. It might however be possible to *approximate* $\psi \otimes \rho$ in the sense that $T(\psi \to \psi \otimes \rho) \le \varepsilon$ for a small ε. In this case, it is hard to distinguish $\psi \otimes \rho$ from the approximation. By keeping the systems $A'B'$, one would thus embezzle entanglement from the owner of ψ - and if one would be able to make ε arbitrarily small, it would be impossible to detect.

In fact, in Ref. [32], van Dam and Hayden showed that it is possible to embezzle *any* bipartite state $\sigma \in \mathfrak{D}(A'B')$ arbitrarily well from a family of pure states $\left\{ \chi_n^{AB} \right\}_{n \in \mathbb{N}}$ in the sense that $\lim_{n \to \infty} T(\chi_n \to \chi_n \otimes \sigma) = 0$. This implies that an arbitrarily good approximation of any σ can be embezzled from χ_n whilst changing χ_n arbitrarily little, as long as n is large enough. This motivates the following definition.

Definition 5. *A family of pure bipartite states $\left\{ \chi_n \right\}_{n \in \mathbb{N}}$ is called a* universal embezzling family *if $\lim_{n \to \infty} T(\chi_n \to \chi_n \otimes \sigma) = 0$ for every bipartite finite dimensional state σ.*

This definition is very similar to the one provided in Refs. [32, 33]. The main difference is that these works only consider protocols using local operations (LO), while in this work, we allow for classical communication too. The set of operations that we consider is therefore larger, and as a result, if a family of states is not an embezzling family according to our definition, then it is not an embezzling family in the sense of Refs. [32, 33]. Surprisingly, the original embezzling family proposed in Ref. [32] is rather unique in a sense that we will specify later, even when we allow for classical communication. Another difference with Refs. [32, 33] is that in those works the fidelity was used to quantify the conversion error. Since the fidelity and trace distance are topologically equivalent, this is irrelevant.

The term 'universal' in Definition 5 underlines the property that any bipartite state can be embezzled. Since for every bipartite state $\sigma \in \mathfrak{D}(AB)$, where $A = B = m$, it is possible to LOOC-convert the maximally entangled state Φ_m into σ (see, e.g., Proposition 1), it is enough to check whether $\lim_{n \to \infty} T(\chi_n \to \chi_n \otimes \Phi_m) = 0$ for all $m \in \mathbb{N}$ to determine if $\left\{ \chi_n \right\}_{n \in \mathbb{N}}$ is a universal embezzling family: This follows for example by combining Lemma 3 and the triangular inequality for the star conversion distance proven in Appendix C. Even simpler, it is in fact equivalent to only require $\lim_{n \to \infty} T(\chi_n \to \chi_n \otimes \Phi_2) = 0$ [33, Lemma 2 for the case of embezzlement with LO], because if one can embezzle enough copies of Φ_2, then one can convert them into Φ_m with LOCC (see Appendix D for more details). It is also important to note that technically, one could have required that $\liminf_{n \to \infty} T(\chi_n \to \chi_n \otimes \sigma) = 0$, since this would also allow to embezzle any state arbitrarily well. However, since one can always choose a subfamily, we decided to keep the definition in line with Ref. [33]. Lastly, due to Eq. (11), we can replace $T(\chi_n \to \chi_n \otimes \sigma)$ in Definition 5 with $T_\star(\chi_n \to \chi_n \otimes \Phi_2)$. Thanks to the closed formula in Theorem 4, we obtain the following complete characterization of universal embezzling families (Appendix D).

Theorem 6. *A family of pure bipartite states $\left\{ \chi_n \right\}_{n \in \mathbb{N}}$ with corresponding Schmidt coefficients $\left\{ \mathbf{p}^{(n)} \right\}_{n \in \mathbb{N}}$ is a universal embezzling family if and only if*

$$\lim_{n \to \infty} \max_{l \in [A_n]} \left\{ \left\| \mathbf{p}^{(n)} \right\|_{(2l-1)} - \left\| \mathbf{p}^{(n)} \right\|_{(l-1)} \right\} = 0, \tag{14}$$

where $A_n = \lceil \mathrm{SR}(\chi_n)/2 \rceil$ is the ceiling of half the Schmidt rank of χ_n.

The problem of determining if a family of states is a universal embezzling family has therefore been restated as a rather simple optimization problem, which in many cases can be solved numerically or even analytically. By choosing $l = 1$ in Eq. (14) one obtains the following necessary condition for a universal embezzling family.

Corollary 7 (cf. Ref. [33, Lemma 3 (LO)]). *If a family of pure bipartite states $\{\,\chi_n\,\}_{n\in\mathbb{N}}$ is a universal embezzling family, then $\lim_{n\to\infty} p_1^{(n)} = 0$, where $\mathbf{p}^{(n)} \in \mathrm{Prob}^{\downarrow}(d_n)$ are the Schmidt coefficients of χ_n.*

This implies that if a family of pure bipartite states $\{\,\chi_n\,\}_{n\in\mathbb{N}}$ is a universal embezzling family, then $\lim_{n\to\infty} \mathrm{SR}(\chi_n) = +\infty$, where $\mathrm{SR}(\chi_n)$ is the Schmidt rank of χ_n. Indeed, considering that the entries of any $\mathbf{p}^{(n)}$ sum to one, if the largest entry converges to zero, then the number of non-zero entries must diverge.

When looking for candidates for universally embezzling families, it is common [32, 33] to consider families of bipartite states

$$|\chi_n\rangle = \frac{1}{\sqrt{F_n}} \sum_{x=1}^{n} \sqrt{f(x)}\, |xx\rangle_{AB}\,, \tag{15}$$

defined by a function $f\colon \mathbb{N} \to \mathbb{R}^+$, where $F_n = \sum_{x=1}^{n} f(x)$ is a normalization constant. With this choice, one ensures that the families of states have a common structure, i.e., $\chi_{m>n}$ is obtained from χ_n by appending additional coefficients and renormalizing. By choosing $f(x) = x^{-1}$, one recovers the universal embezzling family introduced by van Dam and Hayden [32]. This family is rather unique: If we assume that f has a reasonable asymptotic behavior, specifically that it is asymptotically non-increasing and $f(x)/x^{\alpha}$ is asymptotically monotonic for all $\alpha \in \mathbb{R}$, then the family of states $\{\,\chi_n\,\}$ is an embezzling family if and only if f is asymptotically close to x^{-1} in the sense that for every $\varepsilon > 0$,

$$\lim_{x\to\infty} \frac{f(x)}{x^{-1-\varepsilon}} = \infty\,, \quad \lim_{x\to\infty} \frac{f(x)}{x^{-1+\varepsilon}} = 0. \tag{16}$$

As we will discuss now, the assumptions on the asymptotic behavior of f, which we require for technical reasons, are not particularly restrictive. In entanglement theory, it is possible to consider, w.l.o.g., only states with non-increasing Schmidt coefficients. With the first assumption, we require that the family of states, at least asymptotically, has this behavior. The second assumption rules out functions that are asymptotically monotonic but have oscillating components, for an example of such a function see Appendix G, Eq. (G23).

In Ref. [33], it was shown that certain functions that differ from x^{-1} by logarithmic factors lead to universal embezzling families too. Since these functions satisfy the constraints in Eq. (16), this is in accordance with our findings. In addition, in Appendix G, we also show that many non-decreasing functions do not lead to universal embezzling families (see Propositions G.3 and G.4 for details).

A special case of the functions discussed so far are the functions $f_{\alpha}(x) = x^{\alpha}$ with $\alpha \in \mathbb{R}$. For the families of states $\{\,\chi_n^{\alpha}\,\}$ generated by such functions, we analytically compute the exact value of $\lim_{n\to\infty} T_{\star}(\chi_n^{\alpha} \to \chi_n^{\alpha} \otimes \Phi_m)$ for $\alpha \geq -1$ and lower and upper bound it for $\alpha < -1$. The details of the computations can be found in Appendix F and the results for $m = 2$ are shown in Figure 1. Clearly, the limit of the star conversion distance is zero only for $\alpha = -1$, showing again the uniqueness of the choice made by van Dam and Hayden amongst the functions $f(x) = x^{\alpha}$.

5 Conclusions

In Proposition 1 we provided necessary and sufficient conditions for a deterministic LOCC-conversion from a pure bipartite state to a mixed bipartite state. This extends the results already known for pure to pure state LOCC-conversions, whether deterministic or probabilistic [3, 21]. We then exploited this result to prove the topological equivalence of
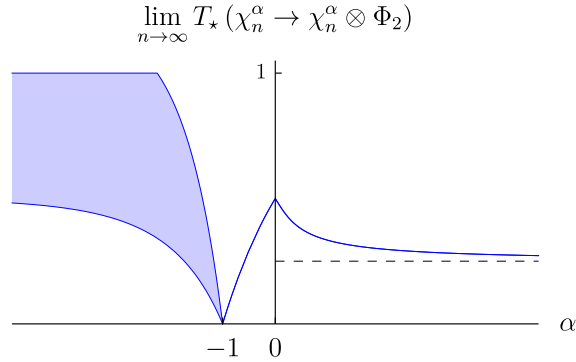
$$\lim_{n \to \infty} T_\star (\chi_n^\alpha \to \chi_n^\alpha \otimes \Phi_2)$$

Figure 1: *Uniqueness of the van Dam and Hayden embezzling family* — The family of states $\{\chi_n^\alpha\}_{n\in\mathbb{N}}$ introduced in the main text is a universal embezzling family if and only if $\alpha = -1$. This can be seen in the above plot showing the analytically derived exact value of $\lim_{n\to\infty} T_\star(\chi_n^\alpha \to \chi_n^\alpha \otimes \Phi_2)$ for $\alpha \geq -1$ and lower and upper bounds for $\alpha < -1$.

the newly defined star conversion distance between pure states and the trace conversion distance commonly used in literature. The star conversion distance exhibits a closed formula (Theorem 4). This is remarkable, since the mixed state LOCC-conversion problem is NP-hard [75].

The closed formula in turn allowed us to completely characterize universal embezzling families (see Definition 5 and Refs. [32, 33] for an analogous definition for LO) in terms of a simple optimization problem stated in Theorem 6. With this characterization at hand, we discussed the uniqueness of the van Dam and Hayden family [32]. For specific families of states generalizing the van Dam and Hayden family, we explicitly evaluated their star conversion distance to maximally entangled states and showed that they are only universally embezzling if they are exactly the van Dam and Hayden family, see Figure 1. Therefore, the van Dam and Hayden embezzling family shows unique properties even for protocols that involve classical communication, as already noticed in Ref. [32]. This suggests a direction for future work, namely, to determine whether LOCC embezzlement implies LO embezzlement (the other direction is trivial), and therefore to investigate if classical communication is relevant in entanglement embezzlement or not. It is worth noting that so far, the research on entanglement embezzlement focuses solely on families composed of pure states. A more comprehensive theory of embezzlement that includes families of mixed states is yet to be developed.

Originally introduced in the resource theory of entanglement as a generalization of catalysis [32], embezzlement has recently also been studied in other resource theories including non-uniformity [76, 77], coherence [78–81], and athermality [76, 82–85]. Moreover, fundamental limits for embezzlement have been proved in Ref. [86] and applied to the resource theories mentioned above. The aforementioned resource theories are related to the resource theory of entanglement via majorization, which is the tool that we used to derive the closed formula for the star conversion distance. A natural next step is to investigate whether it is possible to derive a similar formula in these other majorization-based resource theories. However, this task is not trivial: In the resource theory of non-uniformity, the majorization relation is inverted, and the free state is the maximally mixed state, which is fundamentally different from the free states in the resource theory of entanglement; in the resource theory of athermality, state conversion is described by relative majorization, which is a generalization of majorization. It is worth mentioning that in Ref. [84] it has been shown that embezzlement of athermality allows to violate the second law of thermodynamics. However, by using the work distance [84] as conversion distance, or by imposing

physical constraints on the catalyst [76], e.g., finite dimension or finite energy expectation value, athermality embezzlement is no longer possible, thus restoring the validity of the second law of thermodynamics.

As described in the introduction, universal embezzling families have been useful in many applications [15, 18, 55–61]. Our results provide an easy way to check whether a family is universally embezzling. Moreover, we ruled out large classes of potential candidates that are fundamentally different from the van Dam and Hayden family. The complete characterization of universally embezzling families therefore contributes to a more efficient usage of entanglement in technology.

Acknowledgments

References

[1] Charles H. Bennett, Herbert J. Bernstein, Sandu Popescu, and Benjamin Schumacher. "Concentrating partial entanglement by local operations". Phys. Rev. A **53**, 2046–2052 (1996).

[2] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight. "Quantifying entanglement". Phys. Rev. Lett. **78**, 2275–2279 (1997).

[3] M. A. Nielsen. "Conditions for a class of entanglement transformations". Phys. Rev. Lett. **83**, 436–439 (1999).

[4] Charles H. Bennett, Sandu Popescu, Daniel Rohrlich, John A. Smolin, and Ashish V. Thapliyal. "Exact and asymptotic measures of multipartite pure-state entanglement". Phys. Rev. A **63**, 012307 (2000).

[5] Martin B. Plenio and Shashank Virmani. "An introduction to entanglement measures". Quantum Inf. Comput. **7**, 1–51 (2007).

[6] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki. "Quantum entanglement". Rev. Mod. Phys. **81**, 865–942 (2009).

[7] A. Einstein, B. Podolsky, and N. Rosen. "Can quantum-mechanical description of physical reality be considered complete?". Phys. Rev. **47**, 777–780 (1935).

[8] C. H. Bennet and G Brassard. "Quantum cryptography: Public key distribution and coin tossing". In Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing. Volume 1, pages 175–179. IEEE Bangalore (1984). url: https://web.archive.org/web/20200130165639/http://researcher.watson.ibm.com/researcher/files/us-bennetc/BB84highest.pdf.

[9] Charles H. Bennett and Gilles Brassard. "Quantum cryptography: Public key distribution and coin tossing". Theor. Comput. Sci. **560**, 7–11 (2014).

[10] Charles H. Bennett and Stephen J. Wiesner. "Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states". Phys. Rev. Lett. **69**, 2881–2884 (1992).

[11] Charles H. Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K. Wootters. "Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels". Phys. Rev. Lett **70**, 1895–1899 (1993).

[12] Charles H. Bennett, David P. DiVincenzo, John A. Smolin, and William K. Wootters. "Mixed-state entanglement and quantum error correction". Phys. Rev. A **54**, 3824–3851 (1996).

[13] Charles H. Bennett, Peter W. Shor, John A. Smolin, and Ashish V. Thapliyal. "Entanglement-assisted classical capacity of noisy quantum channels". Phys. Rev. Lett. **83**, 3081–3084 (1999).

[14] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal. "Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem". IEEE Trans. Inf. Theory **48**, 2637–2655 (2006).

[15] Charles H. Bennett, Igor Devetak, Aram W. Harrow, Peter W. Shor, and Andreas Winter. "The quantum reverse Shannon theorem and resource tradeoffs for simulating quantum channels". IEEE Trans. Inf. Theory **60**, 2926–2959 (2014).

[16] Sandu Popescu. "Bell's inequalities versus teleportation: What is nonlocality?". Phys. Rev. Lett. **72**, 797–799 (1994).

[17] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. "General teleportation channel, singlet fraction, and quasidistillation". Phys. Rev. A **60**, 1888–1898 (1999).

[18] Mario Berta, Matthias Christandl, and Renato Renner. "The quantum reverse shannon theorem based on one-shot information theory". Commun. Math. Phys. **306**, 579–615 (2011).

[19] Eric Chitambar and Gilad Gour. "Quantum resource theories". Rev. Mod. Phys. **91**, 025001 (2019).

[20] Bob Coecke, Tobias Fritz, and Robert W. Spekkens. "A mathematical theory of resources". Inf. Comput. **250**, 59–86 (2016).

[21] Daniel Jonathan and Martin B. Plenio. "Minimal conditions for local pure-state entanglement manipulation". Phys. Rev. Lett. **83**, 1455–1458 (1999).

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

[23] A. Uhlmann. "The "transition probability" in the state space of a *-algebra". Rep. Math. Phys. **9**, 273–279 (1976).

[24] Richard Jozsa. "Fidelity for mixed quantum states". J. Mod. Opt **41**, 2315–2323 (1994).

[25] C.A. Fuchs and J. van de Graaf. "Cryptographic distinguishability measures for quantum-mechanical states". IEEE Trans. Inf. Theory **45**, 1216–1227 (1999).

[26] Marco Tomamichel, Roger Colbeck, and Renato Renner. "Duality between smooth min- and max-entropies". IEEE Trans. Inf. Theory **56**, 4674–4681 (2010).

[27] Gilad Gour and Marco Tomamichel. "Optimal extensions of resource measures and their applications". Phys. Rev. A **102**, 062401 (2020).

[28] Xin Wang and Mark M. Wilde. "Resource theory of asymmetric distinguishability". Phys. Rev. Res. **1**, 033170 (2019).

[29] Gaurav Saxena, Eric Chitambar, and Gilad Gour. "Dynamical resource theory of quantum coherence". Phys. Rev. Res. **2**, 023298 (2020).

[30] Gaurav Saxena and Gilad Gour. "Quantifying multiqubit magic channels with completely stabilizer-preserving operations". Phys. Rev. A **106**, 042422 (2022).

[31] Gilad Gour. "Role of quantum coherence in thermodynamics". PRX Quantum **3**, 040323 (2022).

[32] Wim van Dam and Patrick Hayden. "Universal entanglement transformations without communication". Phys. Rev. A **67**, 060302 (2003).

[33] Debbie Leung and Bingjie Wang. "Characteristics of universal embezzling families". Phys. Rev. A **90**, 042331 (2014).

[34] Richard Cleve, Li Liu, and Vern I. Paulsen. "Perfect embezzlement of entanglement". J. Math. Phys. **58**, 012204 (2017).

[35] Daniel Jonathan and Martin B. Plenio. "Entanglement-assisted local manipulation of pure quantum states". Phys. Rev. Lett. **83**, 3566–3569 (1999).

[36] Ludovico Lami, Bartosz Regula, and Alexander Streltsov. "Catalysis cannot overcome bound entanglement" (2023). arXiv:2305.03489.

[37] Ray Ganardi, Tulja Varun Kondra, and Alexander Streltsov. "Catalytic and asymptotic equivalence for quantum entanglement" (2023). arXiv:2305.03488.

[38] Jeongrak Son and Nelly H. Y. Ng. "A hierarchy of thermal processes collapses under catalysis" (2023). arXiv:2303.13020.

[39] Seok Hyung Lie and Nelly H. Y. Ng. "Catalysis always degrades external quantum correlations" (2023). arXiv:2303.02376.

[40] Lauritz van Luijk, Reinhard F. Werner, and Henrik Wilming. "Covariant catalysis requires correlations and good quantum reference frames degrade little" (2023). arXiv:2301.09877.

[41] Chandan Datta, Ray Ganardi, Tulja Varun Kondra, and Alexander Streltsov. "Is there a finite complete set of monotones in any quantum resource theory?". Phys. Rev. Lett. **130**, 240204 (2023).

[42] Jeongrak Son and Nelly H. Y. Ng. "Catalysis in action via elementary thermal operations" (2022). arXiv:2209.15213.

[43] Patryk Lipka-Bartosik, Martí Perarnau-Llobet, and Nicolas Brunner. "Operational definition of the temperature of a quantum state". Phys. Rev. Lett. **130**, 040401 (2023).

[44] Kamil Korzekwa and Matteo Lostaglio. "Optimizing thermalization". Phys. Rev. Lett. **129**, 040602 (2022).

[45] Chandan Datta, Tulja Varun Kondra, Marek Miller, and Alexander Streltsov. "Entanglement catalysis for quantum states and noisy channels" (2022). arXiv:2202.05228.

[46] Patryk Lipka-Bartosik and Paul Skrzypczyk. "Catalytic quantum teleportation". Phys. Rev. Lett. **127**, 080502 (2021).

[47] Tulja Varun Kondra, Chandan Datta, and Alexander Streltsov. "Catalytic transformations of pure entangled states". Phys. Rev. Lett. **127**, 150503 (2021).

[48] Martti Karvonen. "Neither contextuality nor nonlocality admits catalysts". Phys. Rev. Lett. **127**, 160402 (2021).

[49] H. Wilming. "Entropy and reversible catalysis". Phys. Rev. Lett. **127**, 260402 (2021).

[50] Seok Hyung Lie and Hyunseok Jeong. "Randomness for quantum channels: Genericity of catalysis and quantum advantage of uniformness". Phys. Rev. Res. **3**, 013218 (2021).

[51] Patryk Lipka-Bartosik and Paul Skrzypczyk. "All states are universal catalysts in quantum thermodynamics". Phys. Rev. X **11**, 011061 (2021).

[52] Naoto Shiraishi and Takahiro Sagawa. "Quantum thermodynamics of correlated-catalytic state conversion at small scale". Phys. Rev. Lett. **126**, 150502 (2021).

[53] Patryk Lipka-Bartosik, Henrik Wilming, and Nelly H. Y. Ng. "Catalysis in quantum information theory" (2023). arXiv:2306.00798.

[54] Chandan Datta, Tulja Varun Kondra, Marek Miller, and Alexander Streltsov. "Catalysis of entanglement and other quantum resources" (2022). arXiv:2207.05694.

[55] Oded Regev and Thomas Vidick. "Elementary proofs of Grothendieck theorems for completely bounded norms". J. Operat. Theor. **71**, 491–505 (2014).

[56] Alexander Grothendieck. "Résumé de la théorie métrique des produits tensoriels topologiques". Boll. Soc. Mat. São-Paulo **8**, 1–79 (1953).

[57] Gilles Pisier. "Grothendieck's theorem, past and present". Bull. Amer. Math. Soc. **49**, 237–323 (2012).

[58] Debbie Leung, Ben Toner, and John Watrous. "Coherent state exchange in multi-prover quantum interactive proof systems". Chic. J. Theoret. Comput. Sci. **2013**, 11 (2013).

[59] Irit Dinur, David Steurer, and Thomas Vidick. "A parallel repetition theorem for entangled projection games". Comput. Complex. **24**, 201–254 (2015).

[60] Oded Regev and Thomas Vidick. "Quantum xor games". ACM Trans. Comput. Theory **7**, 15 (2015).

[61] Zhengfeng Ji, Debbie Leung, and Thomas Vidick. "A three-player coherent state embezzlement game". Quantum **4**, 349 (2020).

[62] Albert W. Marshall, Ingram Olkin, and Barry C. Arnold. "Inequalities: Theory of majorization and its applications". Springer Series in Statistics. Springer. (2011).

[63] Michael A. Nielsen and Guifré Vidal. "Majorization and the interconversion of bipartite states". Quantum Inf. Comput. **1**, 76–93 (2001).

[64] Michael A Nielsen. "An introduction to majorization and its applications to quantum mechanics". Lecture Notes, Department of Physics, University of Queensland, AustraliaPage 53 (2002). url: `https://michaelnielsen.org/papers/maj-book-notes.pdf`.

[65] Guifré Vidal. "Entanglement of pure states for a single copy". Phys. Rev. Lett. **83**, 1046–1049 (1999).

[66] Hoi-Kwong Lo and Sandu Popescu. "Concentrating entanglement by local actions: Beyond mean values". Phys. Rev. A **63**, 022301 (2001).

[67] Carl W Helstrom. "Quantum detection and estimation theory". J. Stat. Phys. **1**, 231–252 (1969).

[68] Alexander S Holevo. "Statistical decision theory for quantum systems". J. Multivariate Anal. **3**, 337–394 (1973).

[69] John Watrous. "Similarity and distance among states and channels". Pages 124–200. Cambridge University Press. (2018).

[70] Erik Nikolai Torgersen. "Comparison of experiments when the parameter space is finite". Probab. Theory Relat. Fields **16**, 219–249 (1970).

[71] Erik Torgersen. "Comparison of statistical experiments". Encyclopedia of Mathematics and its Applications. Cambridge University Press. Cambridge (1991).

[72] Joseph M. Renes. "Relative submajorization and its use in quantum resource theories". J. Math. Phys **57**, 122202 (2016).

[73] Michał Horodecki, Jonathan Oppenheim, and Carlo Sparaciari. "Extremal distributions under approximate majorization". J. Phys. A **51**, 305301 (2018).

[74] Thomas Theurer, Kun Fang, and Gilad Gour. "Single-shot entanglement manipulation of states and channels revisited" (2023). arXiv:2312.17088.

[75] Leonid Gurvits. "Classical deterministic complexity of Edmonds' problem and quantum entanglement". In Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing. Pages 10–19. STOC '03New York, NY, USA (2003). Association for Computing Machinery.

[76] N H Y Ng, L Mančinska, C Cirstoiu, J Eisert, and S Wehner. "Limits to catalysis in quantum thermodynamics". New J. Phys. **17**, 085004 (2015).

Accepted in 〈 〉uantum 2024-04-08, click title to verify. Published under CC-BY 4.0.

12

[77] Gilad Gour, Markus P. Müller, Varun Narasimhachar, Robert W. Spekkens, and Nicole Yunger Halpern. "The resource theory of informational nonequilibrium in thermodynamics". Phys. Rep. **583**, 1–58 (2015).

[78] T. Baumgratz, M. Cramer, and M. B. Plenio. "Quantifying Coherence". Phys. Rev. Lett. **113**, 140401 (2014).

[79] Shuanping Du, Zhaofang Bai, and Yu Guo. "Conditions for coherence transformations under incoherent operations". Phys. Rev. A **91**, 052120 (2015).

[80] Senrui Chen, Xingjian Zhang, You Zhou, and Qi Zhao. "One-shot coherence distillation with catalysts". Phys. Rev. A **100**, 042323 (2019).

[81] Ryuji Takagi and Naoto Shiraishi. "Correlation in catalysts enables arbitrary manipulation of quantum coherence". Phys. Rev. Lett. **128**, 240501 (2022).

[82] Fernando G. S. L. Brandão, Michał Horodecki, Jonathan Oppenheim, Joseph M. Renes, and Robert W. Spekkens. "Resource Theory of Quantum States Out of Thermal Equilibrium". Phys. Rev. Lett. **111**, 250404 (2013).

[83] Michał Horodecki and Jonathan Oppenheim. "Fundamental limitations for quantum and nanoscale thermodynamics". Nat. Commun. **4**, 2059 (2013).

[84] Fernando Brandão, Michał Horodecki, Nelly Ng, Jonathan Oppenheim, and Stephanie Wehner. "The second laws of quantum thermodynamics". Proc. Natl. Acad. Sci. U.S.A. **112**, 3275–3279 (2015).

[85] Patryk Lipka-Bartosik and Paul Skrzypczyk. "All states are universal catalysts in quantum thermodynamics". Phys. Rev. X **11**, 011061 (2021).

[86] Roberto Rubboli and Marco Tomamichel. "Fundamental limits on correlated catalytic state transformations". Phys. Rev. Lett. **129**, 120506 (2022).

[87] Giulio Chiribella and Carlo Maria Scandolo. "Entanglement and thermodynamics in general probabilistic theories". New J. Phys **17**, 103027 (2015).

[88] A. Wayne Roberts and Dale E. Varberg. "Convex functions". Volume 57 of Pure and Applied Mathemathics; A Series of Monographs and Textbooks. ACADEMIC PRESS New York and London. (1973). url: https://www.sciencedirect.com/bookseries /pure-and-applied-mathematics/vol/57.

[89] John von Neumann. "Some matrix-inequalities and metrization of matrix-space". Tomsk. Univ. Rev. **1**, 286–300 (1937).

[90] L. Mirsky. "A trace inequality of John von Neumann". Monatsh. Math **79**, 303–306 (1975).

[91] Thomas S Ferguson. "Mathematical statistics: A decision theoretic approach". Probability and Mathematical Statistics: A Series of Monographs and Textbooks. Academic Press. (1967).

[92] Thomas H Cormen, Charles E Leiserson, Ronald L Rivest, and Clifford Stein. "Introduction to algorithms". MIT press. (2022). url: https://mitpress.mit.edu/97802 62046305/introduction-to-algorithms/.

[93] Robert A. Wijsman. "A useful inequality on ratios of integrals, with application to maximum likelihood estimation". J. Am. Stat. Assoc **80**, 472–475 (1985).

[94] W. Rudin. "Real and complex analysis". Higher Mathematics Series. McGraw-Hill Education. (1987). url: https://www.mheducation.com/highered/product/M9780 070542341.html?exactIsbn=true.

A Pure to Mixed State Conversions with LOCC

In this section, we prove Proposition 1 of the main text, which we restate for readability.

Proposition 1. *For $\psi \in \mathrm{PURE}(AB)$ and $\sigma \in \mathfrak{D}(AB)$, $\psi \xrightarrow{\mathrm{LOCC}} \sigma$ if and only if there exists a pure state decomposition $\{\,p_z, \chi_z\,\}$ of σ (i.e., $\sigma = \sum_z p_z \chi_z$) such that*

$$\min_{k \in [|A|]} \left\{ E_k(\psi) - \sum_z p_z E_k(\chi_z) \right\} \geq 0. \tag{A1}$$

Proof. We first assume that ψ can be converted into σ with LOCC operations. Then, according to Refs. [66, 87], there exists a protocol in which Alice performs a generalized measurement $\{\,M_z\,\}$ and Bob performs a unitary transformation U_z conditioned on the measurement's outcome such that $\sigma = \sum_z (M_z \otimes U_z)\psi(M_z \otimes U_z)^\dagger$. If Alice and Bob record the outcome of the measurement in a classical system X, the output of the protocol is $\sum_z p_z |z\rangle\langle z|^X \otimes \chi_z^{AB} := \{\,p_z, \chi_z\,\}$, where

$$|\chi_z\rangle = \frac{(M_z \otimes U_z)\,|\psi\rangle}{\|(M_z \otimes U_z)\,|\psi\rangle\|}, \quad p_z = \|(M_z \otimes U_z)\,|\psi\rangle\|^2. \tag{A2}$$

This shows that Alice and Bob can convert ψ into the ensemble $\{\,p_z, \chi_z\,\}$ with LOCC operations, which is equivalent to the condition [21, 65]

$$\min_{k \in [d]} \left(E_k(\psi) - \sum_z p_z E_k(\chi_z) \right) \geq 0, \tag{A3}$$

where E_k was introduced in the main text and $d = |A| = |B|$ as per our convention. This proves the necessary condition.

For the reverse, we assume that $\{\,p_z, \chi_z\,\}$ is a pure state decomposition of σ that satisfies

$$\min_{k \in [d]} \left(E_k(\psi) - \sum_z p_z E_k(\chi_z) \right) \geq 0. \tag{A4}$$

This is equivalent to $\psi \xrightarrow{\mathrm{LOCC}} \{\,p_z, \chi_z\,\} = \sum_z p_z |z\rangle\langle z|^X \otimes \chi_z^{AB}$ [21, 65]. Alice and Bob can trace out the classical system and they obtain $\sum_z p_z \chi_z = \sigma$, thus ψ can be converted into σ with LOCC operations. $\qquad\square$

B Purified Conversion Distance

In this section, we present some results concerning the purified conversion distance [26, 27], which are useful to prove the theorems about the star conversion distance presented in the main text. The purified distance between two states $\rho, \sigma \in \mathfrak{D}(AB)$ is defined as

$$P(\rho, \sigma) = \sqrt{1 - F^2(\rho, \sigma)}, \tag{B1}$$

where $F(\rho, \sigma) = \|\sqrt{\rho}\sqrt{\sigma}\|_1$ is the fidelity. The purified distance is a metric and according to Uhlmann's Theorem [23],

$$P(\rho, \sigma) = \min_{\psi, \varphi} T(\psi, \varphi), \tag{B2}$$

where $T(\rho, \sigma) = \frac{1}{2}\|\rho - \sigma\|_1$ is the trace distance and the minimization runs over all purifications ψ and φ of ρ and σ, respectively. Importantly, the purified distance is topologically equivalent to the trace distance [26],

$$T(\rho, \sigma) \leq P(\rho, \sigma) \leq \sqrt{2T(\rho, \sigma)}. \tag{B3}$$

We now recall the definition of the purified conversion distance and purified star conversion distance introduced in the main text,

$$P\left(\rho \to \sigma\right) = \inf_{\mathcal{N} \in \text{LOCC}} P\left(\mathcal{N}\left(\rho\right), \sigma\right), \quad P_{\star}\left(\psi \to \varphi\right) = \min_{\mathbf{r} \succ \mathbf{p}} P\left(\mathbf{r}, \mathbf{q}\right), \quad \text{(B4)}$$

where $\mathbf{p}, \mathbf{q} \in \text{Prob}^{\downarrow}(d)$ are the Schmidt coefficients of ψ and φ, respectively, and $P\left(\mathbf{r}, \mathbf{q}\right) = \sqrt{1 - F^2(\mathbf{r}, \mathbf{q})}$ is the classical version of the purified distance with

$$F(\mathbf{r}, \mathbf{q}) \coloneqq F(\text{diag}(\mathbf{r}), \text{diag}(\mathbf{q})) = \sum_{x} \sqrt{r_x q_x}. \quad \text{(B5)}$$

We next prove that the two purified conversion distances recalled above are equal on pure states. To this end, we need the following Lemma.

Lemma B.1. *Let* $\mathbf{q} \in \text{Prob}(d)$. *The function* $f \colon \text{Prob}(d) \to [0,1]$ *defined as* $f(\mathbf{v}) = \left(\sum_{x=1}^{d} \sqrt{q_x v_x}\right)^2$ *is concave.*

Proof. For details about concave functions, see Ref. [88]. Here, we are going to use that a twice differentiable function is concave if and only if its Hessian matrix is negative semi-definite. The functions $f_{x,y}(v_x, v_y) \coloneqq \sqrt{q_x v_x}\sqrt{q_y v_y}$ are twice differentiable in v_x, v_y for $v_x, v_y \in (0,1]$ and their Hessian matrices are given by

$$H\left(v_x, v_y\right) = \frac{1}{4} \begin{bmatrix} -\sqrt{\frac{q_x q_y v_y}{v_x^3}} & \sqrt{\frac{q_x q_y}{v_x v_y}} \\ \sqrt{\frac{q_x q_y}{v_x v_y}} & -\sqrt{\frac{q_x q_y v_x}{v_y^3}} \end{bmatrix}. \quad \text{(B6)}$$

Since $\det H(v_x, v_y) = 0$, one of the eigenvalues of $H(v_x, v_y)$ is 0 and the other is equal to the trace of $H(v_x, v_y)$, which is smaller than or equal to zero. This implies that the functions $f_{x,y}(v_x, v_y)$ are concave for $v_x, v_y \in (0,1]$. Moreover, since $0 = f_{x,y}(0, v_y) = f_{x,y}(v_x, 0) = f_{x,y}(0,0)$, it is easy to see that $f_{x,y}(v_x, v_y)$ is in fact concave for $v_x, v_y \in [0,1]$. From this follows that f is concave, since

$$f\left(\mathbf{v}\right) = \sum_{x} v_x q_x + \sum_{y} \sum_{x \neq y} \sqrt{q_x v_x}\sqrt{q_y v_y} \quad \text{(B7)}$$

is the sum of concave functions. $\qquad \square$

We are now ready to prove the promised theorem.

Theorem 2. *Let* $\psi, \varphi \in \text{PURE}(AB)$, *then* $P\left(\psi \to \varphi\right) = P_{\star}\left(\psi \to \varphi\right)$.

Proof. We assume w.l.o.g. that all pure states are in standard form, that is, $|\psi\rangle = \sum_{x} \sqrt{p_x} |xx\rangle_{AB}$, where $\mathbf{p} \in \text{Prob}^{\downarrow}(d)$ and $\{|x\rangle_A\}$ and $\{|x\rangle_B\}$ are fixed bases for A and B, respectively (see main text for more details). Let $\mathbf{p}, \mathbf{q} \in \text{Prob}^{\downarrow}(d)$ be the Schmidt coefficients or ψ and φ, respectively. With $\mathbf{r} \in \text{Prob}^{\downarrow}(d)$, as a consequence of Nielsen's Theorem [3], $\mathbf{r} \succ \mathbf{p}$ if and only if there exists an $\mathcal{N} \in \text{LOCC}$ such that $\mathcal{N}(\psi) \in \text{PURE}(AB)$ has Schmidt coefficients $\mathbf{r}$. We notice that

$$P(\mathcal{N}(\psi), \varphi) = \sqrt{1 - F^2(\mathcal{N}(\psi), \varphi)} = \sqrt{1 - |\langle \mathcal{N}(\psi) | \varphi \rangle|^2}$$

$$= \sqrt{1 - \left(\sum_{x} \sqrt{r_x q_x}\right)^2} = \sqrt{1 - F^2(\mathbf{r}, \mathbf{q})} \quad \text{(B8)}$$

$$= P(\mathbf{r}, \mathbf{q}).$$

This implies that

$$
\begin{aligned}
P_\star (\psi \to \varphi) &= \min_{\mathbf{r} \succ \mathbf{p}} P(\mathbf{r}, \mathbf{q}) \\
&= \min_{\substack{\mathcal{N} \in \mathrm{LOCC} \\ \mathcal{N}(\psi) \in \mathrm{PURE}(AB)}} P(\mathcal{N}(\psi), \varphi) \\
&\geq \inf_{\mathcal{N} \in \mathrm{LOCC}} P(\mathcal{N}(\psi), \varphi) \\
&= P(\psi \to \varphi).
\end{aligned}
\tag{B9}
$$

This inequality follows from the definition of the star purified conversion distance because the minimization in the star conversion distance is done over a smaller set. The non-trivial part is to show that the opposite inequality holds as well. To this end, we want to show that for every mixed state σ such that $\psi \xrightarrow{\mathrm{LOCC}} \sigma$, there exists a pure state χ such that $\psi \xrightarrow{\mathrm{LOCC}} \chi$ and $P(\sigma, \varphi) \geq P(\chi, \varphi)$. It is then sufficient to consider only pure output states for the computation of the purified conversion distance, which implies the reverse inequality.

Let $\sigma = \mathcal{M}(\psi) \in \mathfrak{D}(AB)$, where $\mathcal{M} \in \mathrm{LOCC}$. Also, let $\{\, t_z, \chi_z \,\}$ be a pure state decomposition of σ that satisfies Eq. (A1) (where the χ_z are not necessarily in standard form), and $\mathbf{s}^{(z)}$ be the Schmidt coefficient of χ_z for every z. Furthermore, for every χ_z, let $\tilde{\chi}_z$ be the pure state in standard form that is equal to χ_z up to local unitaries, let $\tilde{\sigma} = \sum_z t_z \tilde{\chi}_z$, and define $\chi \in \mathrm{PURE}(AB)$ as the pure bipartite state (in standard form) with Schmidt coefficients $\mathbf{s} = \sum_z t_z \mathbf{s}^{(z)}$. We notice that for all $k \in [d]$

$$
\begin{aligned}
E_k(\chi) &= 1 - \sum_{x=1}^{k} s_x = 1 - \sum_{x=1}^{k} \sum_z t_z s_x^{(z)} \\
&= 1 - \sum_z t_z (1 - E_k(\chi_z)) = \sum_z t_z E_k(\chi_z) \\
&\leq E_k(\psi),
\end{aligned}
\tag{B10}
$$

where the last inequality follows from the fact that $\{\, t_z, \chi_z \,\}$ satisfies Eq. (A1). This implies that $\mathbf{s} \succ \mathbf{p}$.

The next step is to show that $P(\sigma, \varphi) \geq P(\mathbf{s}, \mathbf{q})$. First, we note that due to the von Neumann trace inequality [89, 90],

$$
F^2(\tilde{\sigma}, \varphi) = \sum_z t_z \left(\sum_x \sqrt{q_x s_x^{(z)}} \right)^2 \geq F^2(\sigma, \varphi)
\tag{B11}
$$

and

$$
F^2(\chi, \varphi) = F^2(\mathbf{s}, \mathbf{q}) = \left(\sum_x \sqrt{q_x \sum_z t_z s_x^{(z)}} \right)^2.
\tag{B12}
$$

Second, we introduce the concave function $f(\mathbf{v}) = \left(\sum_x \sqrt{q_x v_x} \right)^2$, for $\mathbf{v} \in \mathrm{Prob}(d)$ (see Lemma B.1), and rewrite Eq. (B11) and Eq. (B12) as

$$
F^2(\sigma, \varphi) \leq F^2(\tilde{\sigma}, \varphi) = \sum_z t_z f\left(\mathbf{s}^{(z)} \right), \quad F^2(\mathbf{s}, \mathbf{q}) = f\left(\sum_z t_z \mathbf{s}^{(z)} \right).
\tag{B13}
$$

Finally, the concavity of f and the multidimensional Jensen's inequality [91] imply

$$
F^2(\sigma, \varphi) \leq F^2(\tilde{\sigma}, \varphi) = \sum_z t_z f\left(\mathbf{s}^{(z)} \right) \leq f\left(\sum_z t_z \mathbf{s}^{(z)} \right) = F^2(\mathbf{s}, \mathbf{q}),
\tag{B14}
$$

which is equivalent to $P(\sigma, \varphi) \geq P(\mathbf{s}, \mathbf{q})$. So far, we have shown that for every $\sigma = \mathcal{M}(\psi)$, with $\mathcal{M} \in \text{LOCC}$, there exists an $\mathbf{s}^{(\mathcal{M}(\psi))} \in \text{Prob}^{\downarrow}(d)$ such that $\mathbf{s}^{(\mathcal{M}(\psi))} \succ \mathbf{p}$ and $P(\sigma, \varphi) \geq P(\mathbf{s}^{(\mathcal{M}(\psi))}, \mathbf{q})$, thus

$$
\begin{aligned}
P(\psi \to \varphi) &= \inf_{\mathcal{M} \in \text{LOCC}} P(\mathcal{M}(\psi), \varphi) \\
&\geq \inf_{\mathcal{M} \in \text{LOCC}} P\left(\mathbf{s}^{(\mathcal{M}(\psi))}, \mathbf{q}\right) \\
&\geq \min_{\mathbf{r} \succ \mathbf{p}} P(\mathbf{r}, \mathbf{q}) \\
&= P_\star(\psi \to \varphi).
\end{aligned}
\tag{B15}
$$

Eq. (B9) and Eq. (B15) imply that $P(\psi \to \varphi) = P_\star(\psi \to \varphi)$, and this concludes the proof. $\square$

C The Star Conversion Distance

In this section, we discuss properties of the star conversion distance based on the trace distance and provide proofs omitted in the main text. We begin by proving that the trace star conversion distance is topologically equivalent to the standard conversion distance defined via the trace distance.

Lemma 3. *Let ψ, $\varphi \in \text{PURE}(AB)$. Then*

$$
\frac{1}{2}\left[T_\star(\psi \to \varphi)\right]^2 \leq T(\psi \to \varphi) \leq \sqrt{2 T_\star(\psi \to \varphi)}.
\tag{C1}
$$

Proof. Using Eq. (B3) and Theorem 2, we obtain

$$
\begin{aligned}
T(\psi \to \varphi) &= \inf_{\mathcal{N} \in \text{LOCC}} T(\mathcal{N}(\psi), \varphi) \\
&\leq \inf_{\mathcal{N} \in \text{LOCC}} P(\mathcal{N}(\psi), \varphi) = P(\psi \to \varphi) = P_\star(\psi \to \varphi) = \min_{\mathbf{r} \succ \mathbf{p}} P(\mathbf{r}, \mathbf{q}) \\
&\leq \min_{\mathbf{r} \succ \mathbf{p}} \sqrt{2 T(\mathbf{r}, \mathbf{q})} = \sqrt{2 T_\star(\psi \to \varphi)},
\end{aligned}
\tag{C2}
$$

where $T(\mathbf{p}, \mathbf{q}) := T(\text{diag}(\mathbf{p}), \text{diag}(\mathbf{q}))$. Analogously,

$$
\begin{aligned}
T_\star(\psi \to \varphi) &= \min_{\mathbf{r} \succ \mathbf{p}} T(\mathbf{r}, \mathbf{q}) \\
&\leq \min_{\mathbf{r} \succ \mathbf{p}} P(\mathbf{r}, \mathbf{q}) = P_\star(\psi \to \varphi) = P(\psi \to \varphi) = \inf_{\mathcal{N} \in \text{LOCC}} P(\mathcal{N}(\psi), \varphi) \\
&\leq \inf_{\mathcal{N} \in \text{LOCC}} \sqrt{2 T(\mathcal{N}(\psi), \varphi)} = \sqrt{2 T(\psi \to \varphi)}.
\end{aligned}
\tag{C3}
$$

$\square$

This proves that $T_\star(\psi \to \varphi)$ is topologically equivalent to $T(\psi \to \varphi)$.

In the following, we provide the proof of Theorem 4 of the main text, which we restate below for readability. In other words, we derive a closed formula for $T_\star(\psi \to \varphi)$ based on the Schmidt coefficients $\mathbf{p}, \mathbf{q} \in \text{Prob}^{\downarrow}(d)$ of ψ and φ, respectively.

Theorem 4. *Let ψ, $\varphi \in \text{PURE}(AB)$ and let $\mathbf{p}$, $\mathbf{q} \in \text{Prob}^{\downarrow}(|A|)$ be their corresponding Schmidt coefficients. Then,*

$$
T_\star(\psi \to \varphi) = \max_{k \in [\text{SR}(\psi)]} \left\{ \|\mathbf{p}\|_{(k)} - \|\mathbf{q}\|_{(k)} \right\}.
\tag{C4}
$$

If $\xi \in \text{PURE}(AB)$ is separable and ψ, φ are not, then

$$
T_\star(\psi \to \varphi) < T_\star(\xi \to \varphi).
\tag{C5}
$$

Proof. Let $\varepsilon \in [0, 1]$ and $\mathfrak{B}_{\mathbf{q}}^{\varepsilon} = \left\{ \mathbf{q}' : \frac{1}{2} \|\mathbf{q} - \mathbf{q}'\|_1 \leq \varepsilon \right\}$. As shown in Ref. [73], there exist probability vectors $\bar{\mathbf{q}}^{(\varepsilon)} \in \mathfrak{B}_{\mathbf{q}}^{\varepsilon}$ called steepest ε-approximations of $\mathbf{q}$ such that $\bar{\mathbf{q}}^{(\varepsilon)} \succ \mathbf{q}'$ for all $\mathbf{q}' \in \mathfrak{B}_{\mathbf{q}}^{\varepsilon}$. Moreover, these steepest ε-approximations can be constructed explicitly: If $\frac{1}{2} \|\mathbf{q} - \mathbf{e}_1\|_1 \leq \varepsilon$, then $\bar{\mathbf{q}}^{(\varepsilon)} = \mathbf{e}_1$, otherwise let $k_\varepsilon \in [d]$ be the index satisfying

$$\sum_{x=1}^{k_\varepsilon} q_x \leq 1 - \varepsilon \quad \text{and} \quad \sum_{x=1}^{k_\varepsilon+1} q_x > 1 - \varepsilon. \tag{C6}$$

The components of $\bar{\mathbf{q}}^{(\varepsilon)}$ are then given by

$$\bar{q}_x^{(\varepsilon)} = \begin{cases} q_1 + \varepsilon & \text{if } x = 1, \\ q_x & \text{if } x \in \{ 2, \ldots, k_\varepsilon \}, \\ 1 - \varepsilon - \sum_{x=1}^{k_\varepsilon} q_x & \text{if } x = k_\varepsilon + 1, \\ 0 & \text{otherwise.} \end{cases} \tag{C7}$$

We now show that

$$\min_{\mathbf{r} \succ \mathbf{p}} \left\{ \frac{1}{2} \|\mathbf{q} - \mathbf{r}\|_1 \right\} = \min \left\{ \varepsilon \in [0, 1] : \bar{\mathbf{q}}^{(\varepsilon)} \succ \mathbf{p} \right\}. \tag{C8}$$

First, we notice that if $\mathbf{r}_\star$ is an optimizer of $\min_{\mathbf{r} \succ \mathbf{p}} \left\{ \frac{1}{2} \|\mathbf{q} - \mathbf{r}\|_1 \right\}$ and $\tilde{\varepsilon} = \frac{1}{2} \|\mathbf{q} - \mathbf{r}_\star\|_1$, then $\mathbf{r}_\star \in \mathfrak{B}_{\mathbf{q}}^{\tilde{\varepsilon}}$ and therefore $\bar{\mathbf{q}}^{(\tilde{\varepsilon})} \succ \mathbf{r}_\star$. By transitivity, we also have $\bar{\mathbf{q}}^{(\tilde{\varepsilon})} \succ \mathbf{p}$, which implies that

$$\min \left\{ \varepsilon \in [0, 1] : \bar{\mathbf{q}}^{(\varepsilon)} \succ \mathbf{p} \right\} \leq \tilde{\varepsilon} = \min_{\mathbf{r} \succ \mathbf{p}} \left\{ \frac{1}{2} \|\mathbf{q} - \mathbf{r}\|_1 \right\}. \tag{C9}$$

For the reverse inequality, let $\varepsilon_\star$ be the optimizer of $\min \left\{ \varepsilon \in [0, 1] : \bar{\mathbf{q}}^{(\varepsilon)} \succ \mathbf{p} \right\}$. By definition, $\bar{\mathbf{q}}^{(\varepsilon_\star)} \succ \mathbf{p}$ and $\frac{1}{2} \left\| \mathbf{q} - \bar{\mathbf{q}}^{(\varepsilon_\star)} \right\|_1 \leq \varepsilon_\star$, thus

$$\min_{\mathbf{r} \succ \mathbf{p}} \left\{ \frac{1}{2} \|\mathbf{q} - \mathbf{r}\|_1 \right\} \leq \varepsilon_\star = \min \left\{ \varepsilon \in [0, 1] : \bar{\mathbf{q}}^{(\varepsilon)} \succ \mathbf{p} \right\}. \tag{C10}$$

This shows that

$$T_\star(\psi \to \varphi) = \min_{\mathbf{r} \in \mathrm{Prob}^{\downarrow}(d)} \left\{ \frac{1}{2} \|\mathbf{q} - \mathbf{r}\|_1 : \mathbf{r} \succ \mathbf{p} \right\} = \min \left\{ \varepsilon \in [0, 1] : \bar{\mathbf{q}}^{(\varepsilon)} \succ \mathbf{p} \right\}. \tag{C11}$$

We observe that $\left\| \bar{\mathbf{q}}^{(\varepsilon)} \right\|_{(k)} = \min \left\{ 1, \|\mathbf{q}\|_{(k)} + \varepsilon \right\}$, thus the condition $\bar{\mathbf{q}}^{\varepsilon} \succ \mathbf{p}$ is equivalent to

$$\|\mathbf{p}\|_{(k)} \leq \min \left\{ 1, \|\mathbf{q}\|_{(k)} + \varepsilon \right\}, \quad \forall k \in [d]. \tag{C12}$$

This expression is further simplified by noticing that $\mathbf{p}$ is a probability vector, and therefore $\|\mathbf{p}\|_{(k)} \leq 1$ for all $k \in [d]$. Consequently,

$$\bar{\mathbf{q}}^{\varepsilon} \succ \mathbf{p} \quad \Leftrightarrow \quad \|\mathbf{p}\|_{(k)} \leq \|\mathbf{q}\|_{(k)} + \varepsilon, \forall k \in [d] \quad \Leftrightarrow \quad \max_{k \in [d]} \left\{ \|\mathbf{p}\|_{(k)} - \|\mathbf{q}\|_{(k)} \right\} \leq \varepsilon. \tag{C13}$$

In combination with the minimization in Eq. (C11) follows that

$$T_\star(\psi \to \varphi) = \max_{k \in [d]} \left\{ \|\mathbf{p}\|_{(k)} - \|\mathbf{q}\|_{(k)} \right\}. \tag{C14}$$

To conclude the proof of the first part of the theorem, we observe that for $k \geq \mathrm{SR}(\psi^{AB})$, $\|\mathbf{p}\|_{(k)} = 1$ and $\|\mathbf{q}\|_{(k)}$ is non-decreasing with k, thus we can restrict the maximization to $k \leq \mathrm{SR}(\psi^{AB})$.

For the second part, we notice that according to Eq. (C11),

$$T_\star(\xi \to \varphi) = \min_{\mathbf{r} \in \mathrm{Prob}^\downarrow(d)} \left\{ \frac{1}{2} \|\mathbf{q} - \mathbf{r}\|_1 : \mathbf{r} \succ \mathbf{e}_1 \right\} = \frac{1}{2} \|\mathbf{q} - \mathbf{e}_1\|_1 = 1 - q_1. \qquad \text{(C15)}$$

Furthermore, from $\mathbf{e}_1 \succ \mathbf{p}$ and the transitivity of the majorization-relation, it follows that

$$\begin{aligned} T_\star(\psi \to \varphi) &= \min_{\mathbf{r} \in \mathrm{Prob}^\downarrow(d)} \left\{ \frac{1}{2} \|\mathbf{q} - \mathbf{r}\|_1 : \mathbf{r} \succ \mathbf{p} \right\} \\ &\leq \min_{\mathbf{r} \in \mathrm{Prob}^\downarrow(d)} \left\{ \frac{1}{2} \|\mathbf{q} - \mathbf{r}\|_1 : \mathbf{r} \succ \mathbf{e}_1 \right\} \\ &= T_\star(\xi \to \varphi). \end{aligned} \qquad \text{(C16)}$$

To rule out equality, suppose that $T_\star(\psi \to \varphi) = \max_{k \in [d]} \left\{ \|\mathbf{p}\|_{(k)} - \|\mathbf{q}\|_{(k)} \right\} = 1 - q_1$, and denote with $k_\star$ an index that achieves this maximum. Then

$$\|\mathbf{p}\|_{(k_\star)} = 1 + \|\mathbf{q}\|_{(k_\star)} - q_1. \qquad \text{(C17)}$$

From this expression follows that $\|\mathbf{p}\|_{(k_\star)} \leq 1$ only if either $\mathbf{q} = \mathbf{e}_1$ or if $k_\star = 1$, and therefore $\mathbf{p} = \mathbf{e}_1$. These conditions are in contrast with the assumption that ψ and φ are not separable. As a consequence, equality in Eq. (C16) is unachievable, which proves the second part of the theorem. $\qquad \square$

Next, we show that the star conversion distance satisfies a triangle inequality. Let ψ, φ, and $\chi \in \mathrm{PURE}(AB)$ and let $\mathbf{p}$, $\mathbf{q}$, and $\mathbf{r}$ be their Schmidt coefficients. With the help of Eq. (C14), this implies that

$$\begin{aligned} T_\star(\psi \to \chi) &= \max_{n \in [d]}(\|\mathbf{p}\|_{(n)} - \|\mathbf{r}\|_{(n)}) \\ &= \max_{n \in [d]}(\|\mathbf{p}\|_{(n)} - \|\mathbf{q}\|_{(n)} + \|\mathbf{q}\|_{(n)} - \|\mathbf{r}\|_{(n)}) \\ &\leq \max_{n \in [d]}(\|\mathbf{p}\|_{(n)} - \|\mathbf{q}\|_{(n)}) + \max_{n \in [d]}(\|\mathbf{q}\|_{(n)} - \|\mathbf{r}\|_{(n)}) \\ &= T_\star(\psi \to \varphi) + T_\star(\varphi \to \chi). \end{aligned} \qquad \text{(C18)}$$

Remark. The star conversion distance $T_\star(\psi \to \varphi)$ was so far only defined for pure bipartite states belonging to the same Hilbert space. This restriction is easily lifted by noting that one can always add separable auxiliary states such that the Hilbert spaces (or dimensions) match. This can be done in multiple ways: Let $\psi \in \mathrm{PURE}(AB)$, $\varphi \in \mathrm{PURE}(A'B')$, and let $d = |A| = |B|$, $d' = |A'| = |B'|$. Then

$$T_\star(\psi_{AB} \to \varphi_{A'B'}) \equiv T_\star(\psi_{AB} \otimes |11\rangle\langle 11|_{A'B'} \to |11\rangle\langle 11|_{AB} \otimes \varphi_{A'B'}). \qquad \text{(C19)}$$

Since in fact we are only interested in the dimension of systems and their spatial separation, a more compact equivalent notation that we will use later can be defined as follows: Denote with m, $m' > 0$ the smallest integers such that $md = m'd'$, and with A_m, B_m, $A_{m'}$, $B_{m'}$ Hilbert spaces of corresponding dimensions. The conversion distance $T_\star(\psi_{AB} \to \varphi_{A'B'})$ is then

$$T_\star(\psi_{AB} \to \varphi_{A'B'}) \equiv T_\star(\psi_{AB} \otimes |11\rangle\langle 11|_{A_m B_m} \to \varphi_{A'B'} \otimes |11\rangle\langle 11|_{A_{m'} B_{m'}}). \qquad \text{(C20)}$$

D Universal Embezzling Families

In the main text, we provided the following definition of universal embezzling families, which we repeat here for readability.

Definition 5. *A family of pure bipartite states* $\{\chi_n\}_{n\in\mathbb{N}}$ *is called a* universal embezzling family *if* $\lim_{n\to\infty} T(\chi_n \to \chi_n \otimes \sigma) = 0$ *for every bipartite finite dimensional state* σ.

Thanks to the triangle inequality for the star conversion distance proven in Eq. (C18), a family of states $\{\chi_n\}_{n\in\mathbb{N}}$ is an embezzling family if and only if it can embezzle the state Φ_2, as already shown in Ref [33, Lemma 2] for the case of embezzling with LO.

Lemma D.1 (cf. Ref. [33, Lemma 2 (LO)]). *Let* $\{\chi_n\}_{n\in\mathbb{N}}$ *be a family of pure bipartite states. The following three statements are equivalent*

1. $\{\chi_n\}_{n\in\mathbb{N}}$ *is a universal embezzling family.*

2. $\lim_{n\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_2) = 0.$

3. $\lim_{n\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) = 0$ *for every* $m \geq 2$.

Proof. Clearly 2. and 3. follow from 1. due to the definition of universal embezzling families and the topological equivalence of T and $T_\star$. Moreover, 3. follows from 2. because $\Phi_2^{\otimes\lceil \log_2 m\rceil} \xrightarrow{\text{LOCC}} \Phi_m$, and therefore

$$
\begin{aligned}
T_\star(\chi_n &\to \chi_n \otimes \Phi_m) \\
&\leq T_\star(\chi_n \to \chi_n \otimes \Phi_2^{\otimes\lceil \log_2 m\rceil}) \\
&\leq T_\star(\chi_n \to \chi_n \otimes \Phi_2^{\otimes\lceil \log_2 m\rceil-1}) + T_\star(\chi_n \otimes \Phi_2^{\otimes\lceil \log_2 m\rceil-1} \to \chi_n \otimes \Phi_2^{\otimes\lceil \log_2 m\rceil}) \\
&\leq T_\star(\chi_n \to \chi_n \otimes \Phi_2) + T_\star(\chi_n \otimes \Phi_2 \to \chi_n \otimes \Phi_2^{\otimes 2}) + \dots \\
&\quad + T_\star(\chi_n \otimes \Phi_2^{\otimes\lceil \log_2 m\rceil-1} \to \chi_n \otimes \Phi_2^{\otimes\lceil \log_2 m\rceil}) \\
&\leq \lceil \log_2 m\rceil T_\star(\chi_n \to \chi_n \otimes \Phi_2).
\end{aligned}
\tag{D1}
$$

By taking the limit $n \to \infty$ on both sides we obtain the desired result. To conclude, we note that 3. implies 1., since for all σ^{AB}, $\Phi_{|A|} \xrightarrow{\text{LOCC}} \sigma^{AB}$. $\qquad\square$

Next, we derive a formula for the conversion distance $T_\star(\chi \to \chi \otimes \Phi_m)$ using Eq. (C14) and Eq. (C20).

Lemma D.2. *Let* $\mathbf{p} \in \text{Prob}^\downarrow(d)$ *be the Schmidt coefficients of* χ. *Then*

$$
T_\star(\chi \to \chi \otimes \Phi_m) = \max_{k\in[\text{SR}(\chi)]} \left\{ \|\mathbf{p}\|_{(k)} - \|\mathbf{p}\|_{(a_k)} - \frac{b_k}{m} p_{a_k+1} \right\},
\tag{D2}
$$

where $a_k = \lfloor k/m\rfloor$, $\lfloor\cdot\rfloor$ *denotes the floor, and* $b_k = k - a_k m$.

Proof. The Schmidt coefficients of the input and target state are

$$
\begin{aligned}
\mathbf{p} \otimes \mathbf{e}_1 &= (p_1, \dots, p_d, \underbrace{0, \dots, 0}_{d\cdot(m-1)\,\text{times}}), \\
\mathbf{p} \otimes \mathbf{u}^{(m)} &= \frac{1}{m}(\underbrace{p_1, \dots, p_1}_{m\,\text{times}}, \dots, \underbrace{p_d, \dots, p_d}_{m\,\text{times}}),
\end{aligned}
\tag{D3}
$$

where $\mathbf{u}^{(m)} = (1/m, \ldots, 1/m)$. It is straightforward to see that

$$\|\mathbf{p} \otimes \mathbf{e}_1\|_{(k)} = \begin{cases} \|\mathbf{p}\|_{(k)} & \text{if } k \in [d], \\ 1 & \text{if } d < k \leq d \cdot m, \end{cases} \tag{D4}$$

and by writing $k = a_k m + b_k$, where $a_k = \lfloor k/m \rfloor$,

$$\left\|\mathbf{p} \otimes \mathbf{u}^{(m)}\right\|_{(k)} = \sum_{x=1}^{a_k} p_x + b_k \frac{p_{a_k+1}}{m} = \|\mathbf{p}\|_{(a_k)} + b_k \frac{p_{a_k+1}}{m}. \tag{D5}$$

Using the closed formula for the star conversion distance given in Theorem 4, we obtain

$$T_\star(\chi \to \chi \otimes \Phi_m) = \max_{k \in [\mathrm{SR}(\chi)]} \left\{ \|\mathbf{p}\|_{(k)} - \|\mathbf{p}\|_{(a_k)} - \frac{b_k}{m} p_{a_k+1}^{(n)} \right\}. \tag{D6}$$

$\square$

A simplified version of this expression can be used to characterize embezzling families.

Theorem 6. *A family of pure bipartite states $\{\chi_n\}_{n \in \mathbb{N}}$ with corresponding Schmidt coefficients $\left\{\mathbf{p}^{(n)}\right\}_{n \in \mathbb{N}}$ is a universal embezzling family if and only if*

$$\lim_{n \to \infty} \max_{l \in [A_n]} \left\{ \left\|\mathbf{p}^{(n)}\right\|_{(2l-1)} - \left\|\mathbf{p}^{(n)}\right\|_{(l-1)} \right\} = 0, \tag{D7}$$

where $A_n = \lceil \mathrm{SR}(\chi_n)/2 \rceil$.

Proof. Due to Lemma D.1 and Eq. (D2), $\{\chi_n\}_{n \in \mathbb{N}}$ is a universal embezzling family if and only if

$$0 = \lim_{n \to \infty} T_\star(\chi_n \to \chi_n \otimes \Phi_2) = \lim_{n \to \infty} \max_{k \in [\mathrm{SR}(\chi_n)]} \left\{ \left\|\mathbf{p}^{(n)}\right\|_{(k)} - \left\|\mathbf{p}^{(n)}\right\|_{(a_k)} - \frac{b_k}{2} p_{a_k+1}^{(n)} \right\}, \tag{D8}$$

where $a_k = \lfloor k/2 \rfloor$ and $k = 2a_k + b_k$. First, we prove the necessary condition. Let $\{\chi_n\}_{n \in \mathbb{N}}$ be a universal embezzling family. We observe that

$$\max_{k \in [\mathrm{SR}(\chi_n)]} \left\{ \left\|\mathbf{p}^{(n)}\right\|_{(k)} - \left\|\mathbf{p}^{(n)}\right\|_{(a_k)} - \frac{b_k}{2} p_{a_k+1}^{(n)} \right\} \geq \left\|\mathbf{p}^{(n)}\right\|_{(1)} - \left\|\mathbf{p}^{(n)}\right\|_{(a_1)} - \frac{b_1}{2} p_{a_1+1}^{(n)}$$
$$= p_1^{(n)} - 0 - \frac{1}{2} p_1^{(n)} > 0. \tag{D9}$$

Taking the limit for $n \to \infty$ in the expression above, we obtain $\lim_{n \to \infty} p_1^{(n)} = 0$. Since $p_1^{(n)}$ is the largest Schmidt coefficient, $p_{a_k+1}^{(n)}$ converges to zero too. Since $0 \leq \frac{b_k}{2} \leq \frac{1}{2}$,

$$\lim_{n \to \infty} \frac{b_k}{2} p_{a_k+1}^{(n)} = 0 \tag{D10}$$

and thus

$$0 = \lim_{n \to \infty} \max_{k \in [\mathrm{SR}(\chi_n)]} \left\{ \left\|\mathbf{p}^{(n)}\right\|_{(k)} - \left\|\mathbf{p}^{(n)}\right\|_{(a_k)} - \frac{b_k}{2} p_{a_k+1}^{(n)} \right\}$$
$$= \lim_{n \to \infty} \max_{k \in [\mathrm{SR}(\chi_n)]} \left\{ \left\|\mathbf{p}^{(n)}\right\|_{(k)} - \left\|\mathbf{p}^{(n)}\right\|_{(a_k)} \right\} \tag{D11}$$
$$\geq \lim_{n \to \infty} \max_{l \in [\lceil \mathrm{SR}(\chi_n)/2 \rceil]} \left\{ \left\|\mathbf{p}^{(n)}\right\|_{(2l-1)} - \left\|\mathbf{p}^{(n)}\right\|_{(l-1)} \right\} \geq 0.$$

For the sufficient condition, we observe that

$$0 = \lim_{n\to\infty} \max_{l\in[A_n]} \left\{ \left\|\mathbf{p}^{(n)}\right\|_{(2l-1)} - \left\|\mathbf{p}^{(n)}\right\|_{(l-1)} \right\} \geq \lim_{n\to\infty} p_1^{(n)}, \tag{D12}$$

which implies that $\lim_{n\to\infty} p_1^{(n)} = 0$. Furthermore,

$$\begin{aligned}
\lim_{n\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_2) &= \lim_{n\to\infty} \max_{k\in[\mathrm{SR}(\chi_n)]} \left\{ \left\|\mathbf{p}^{(n)}\right\|_{(k)} - \left\|\mathbf{p}^{(n)}\right\|_{(a_k)} - \frac{b_k}{2} p_{a_k+1}^{(n)} \right\} \\
&\leq \lim_{n\to\infty} \max_{k\in[\mathrm{SR}(\chi_n)]} \left\{ \left\|\mathbf{p}^{(n)}\right\|_{(k)} - \left\|\mathbf{p}^{(n)}\right\|_{(a_k)} \right\}
\end{aligned} \tag{D13}$$

At this point, we have a closer look at

$$\max_{k\in[\mathrm{SR}(\chi_n)]} \left\{ \left\|\mathbf{p}^{(n)}\right\|_{(k)} - \left\|\mathbf{p}^{(n)}\right\|_{(a_k)} \right\}. \tag{D14}$$

If k is even, then $a_k = k/2 = a_{k+1}$, and thus

$$\left\|\mathbf{p}^{(n)}\right\|_{(k)} - \left\|\mathbf{p}^{(n)}\right\|_{(a_k)} \leq \left\|\mathbf{p}^{(n)}\right\|_{(k+1)} - \left\|\mathbf{p}^{(n)}\right\|_{(a_{k+1})} \tag{D15}$$

If $\mathrm{SR}(\chi_n)$ is odd, we can thus without loss of generality restrict the maximization to run over odd $k \in [\mathrm{SR}(\chi_n)]$. If $\mathrm{SR}(\chi_n)$ is even, we must additionally consider $k = \mathrm{SR}(\chi_n)$. Assume that this is the case: It then holds that

$$\begin{aligned}
&\left| \left\|\mathbf{p}^{(n)}\right\|_{(\mathrm{SR}(\chi_n))} - \left\|\mathbf{p}^{(n)}\right\|_{(\mathrm{SR}(\chi_n)/2)} - \left(\left\|\mathbf{p}^{(n)}\right\|_{(\mathrm{SR}(\chi_n)-1)} - \left\|\mathbf{p}^{(n)}\right\|_{\left(\frac{\mathrm{SR}(\chi_n)}{2}-1\right)} \right) \right| \\
&= \left| p_{\mathrm{SR}(\chi_n)}^{(n)} - p_{\mathrm{SR}(\chi_n)/2}^{(n)} \right| \\
&\leq p_1^{(n)},
\end{aligned} \tag{D16}$$

which vanishes in the limit $n \to \infty$. As a consequence,

$$\begin{aligned}
\lim_{n\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_2) &\leq \lim_{n\to\infty} \max_{k\in[\mathrm{SR}(\chi_n)]} \left\{ \left\|\mathbf{p}^{(n)}\right\|_{(k)} - \left\|\mathbf{p}^{(n)}\right\|_{(a_k)} \right\} \\
&= \lim_{n\to\infty} \max_{k\in[\mathrm{SR}(\chi_n)],\, k\ \mathrm{odd}} \left\{ \left\|\mathbf{p}^{(n)}\right\|_{(k)} - \left\|\mathbf{p}^{(n)}\right\|_{(a_k)} \right\} \\
&= \lim_{n\to\infty} \max_{l\in\lceil\mathrm{SR}(\chi_n)/2\rceil} \left\{ \left\|\mathbf{p}^{(n)}\right\|_{(2l-1)} - \left\|\mathbf{p}^{(n)}\right\|_{(l-1)} \right\} = 0.
\end{aligned} \tag{D17}$$

Eq. (D17) shows that $\{\chi_n\}_{n\in\mathbb{N}}$ is an embezzling family and concludes the proof. $\qquad\square$

An important and easy to check necessary condition for a universal embezzling family is given in the following Corollary.

Corollary 7 (cf. Ref. [33, Lemma 3 (LO)])**.** *If a family of pure bipartite states $\{\chi_n\}_{n\in\mathbb{N}}$ is a universal embezzling family, then $\lim_{n\to\infty} p_1^{(n)} = 0$, where $\mathbf{p}^{(n)} \in \mathrm{Prob}^{\downarrow}(d_n)$ are the Schmidt coefficients of χ_n.*

E Regular Embezzling Families of States

Often, families of states are defined in terms of a positive, monotonic, and continuous function. This motivates the following definition (compare to Ref. [33]).

Definition E.1. *A family of states $\{\,\chi_n\,\}_{n\in\mathbb{N}}$ is a regular family if there exists a monotonic function $f\colon\mathbb{N}\to(0,\infty)$ such that*

$$|\chi_n\rangle = \frac{1}{\sqrt{F_n}}\sum_{x=1}^{n}\sqrt{f(x)}\,|xx\rangle \quad \textit{for all} \quad n\in\mathbb{N}, \tag{E1}$$

where $F_n = \sum_{x=1}^{n}f(x)$.

Remark. If $\{\,\chi_n\,\}_{n\in\mathbb{N}}$ is a regular family, then there exists a sequence $\{\,n_j\,\}_{j\in\mathbb{N}}$ such that the family $\left\{\,\chi_{n_j}\,\right\}_{j\in\mathbb{N}}$ is a universal embezzling family if and only if $\liminf_{n\to\infty}T_\star(\chi_n\to\chi_n\otimes\Phi_m)=0$. However, according to our definition, the family $\left\{\,\chi_{n_j}\,\right\}_{j\in\mathbb{N}}$ is then no longer regular, unless $n_j=j$ for all $j\in\mathbb{N}$.

For the following proofs, it is beneficial to extend f to a monotonic continuous function on $[1,\infty)$. This can often be done trivially by simply extending the domain of f. An example is the van Dam and Hayden family where $f(x)=x^{-1}$. Otherwise, we can extend f by connecting two consecutive points with straight lines. If the function f is multiplied by a constant factor, this does not change the corresponding family of states. Therefore, from now on we assume for simplicity that $f(1)=1$.

Let $\{\,\chi_n\,\}_{n\in\mathbb{N}}$ be a regular family of states and let f be the function associated to it. We define the non-increasing functions f_n, $n\in\mathbb{N}$, by

$$f_n(x) = \begin{cases} f(x) & \text{if } f \text{ is non-increasing,} \\ \frac{f(n+1-x)}{f(n)} & \text{if } f \text{ is non-decreasing.} \end{cases} \tag{E2}$$

By our extension of f, f_n is also naturally extended to a continuous function $g(x,y)$ such that $g(x,n)=f_n(x)$ via the definition

$$g(x,y) = \begin{cases} f(x) & \text{if } f \text{ is non-increasing,} \\ \frac{f(y+1-x)}{f(y)} & \text{if } f \text{ is non-decreasing.} \end{cases} \tag{E3}$$

The (by definition non-increasing) Schmidt coefficients of χ_n are therefore given by

$$p_x^{(n)} = \frac{f_n(x)}{F_n}. \tag{E4}$$

Since $p_1^{(n)}=\frac{1}{F_n}$, we can restate Corollary 7 for regular families.

Corollary E.2 (cf. Ref. [33, Lemma 3 (LO)]). *If $\{\,\chi_n\,\}_{n\in\mathbb{N}}$ is a regular universal embezzling family, then $\lim_{n\to\infty}F_n=+\infty$, where F_n is defined in Definition E.1.*

In the following proposition, we present bounds on the limit of the conversion distance $T_\star(\chi_n\to\chi_n\otimes\Phi_m)$ in terms of the function $g(x,y)$ introduced in Eq. (E3), which will be of use later.

Proposition E.3. *Let $\{\,\chi_n\,\}_{n\in\mathbb{N}}$ be a regular family of states, f be the function associated to it, and g be defined as in Eq. (E3). If $\lim_{n\to\infty} F_n = \infty$, then*

$$\liminf_{n\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) \geq \liminf_{y\to\infty} \frac{\max_{1\leq a\leq y/m}\left\{\int_a^{am} g(x,y)\,\mathrm{d}x\right\}}{\int_1^y g(x,y)\,\mathrm{d}x}, \tag{E5}$$

$$\limsup_{n\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) \leq \limsup_{y\to\infty} \frac{\max_{1\leq a\leq y/m}\left\{\int_a^{am} g(x,y)\,\mathrm{d}x\right\}}{\int_1^y g(x,y)\,\mathrm{d}x}. \tag{E6}$$

If $\lim_{y\to\infty} \frac{\max_{1\leq a\leq y/m}\left\{\int_a^{am} g(x,y)\,\mathrm{d}x\right\}}{\int_1^y g(x,y)\,\mathrm{d}x}$ exists, this implies that

$$\lim_{n\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) = \lim_{y\to\infty} \frac{\max_{1\leq a\leq y/m}\left\{\int_a^{am} g(x,y)\,\mathrm{d}x\right\}}{\int_1^y g(x,y)\,\mathrm{d}x}. \tag{E7}$$

Proof. For better readability, we divide the proof into steps and use the notation $G(x,y) = \int_1^x g(t,y)\,\mathrm{d}t$.

Step 1: Starting from Eq. (D2), we find

$$\liminf_{n\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m)$$

$$= \liminf_{n\to\infty} \max_{k\in[n]} \left\{ \left\|\mathbf{p}^{(n)}\right\|_{(k)} - \left\|\mathbf{p}^{(n)}\right\|_{(a_k)} - \frac{b_k}{m} p_{a_k+1}^{(n)} \right\}$$

$$= \liminf_{n\to\infty} \max_{k\in[n]} \left\{ \left\|\mathbf{p}^{(n)}\right\|_{(a_km)} + \sum_{x=a_km+1}^{k} p_x^{(n)} - \left\|\mathbf{p}^{(n)}\right\|_{(a_k)} - \frac{b_k}{m} p_{a_k+1}^{(n)} \right\} \tag{E8}$$

$$= \liminf_{n\to\infty} \max_{k\in[n]} \left\{ \left\|\mathbf{p}^{(n)}\right\|_{(a_km)} - \left\|\mathbf{p}^{(n)}\right\|_{(a_k)} + \frac{\sum_{x=a_km+1}^{a_km+b_k} f_n(x) - \frac{b_k}{m} f_n(a_k+1)}{F_n} \right\}$$

Now we observe that the last contribution in the expression above vanishes because by construction $f_n(x) \leq 1$ and

$$0 \leq \frac{\sum_{x=a_km+1}^{a_km+b_k} f_n(x)}{F_n} \leq \frac{m}{F_n},$$

$$0 \leq \frac{\frac{b_k}{m} f_n(a_k+1)}{F_n} \leq \frac{1}{F_n}. \tag{E9}$$

Moreover, by assumption, the right-hand sides converge to zero in the limit $n \to \infty$, and this ensures that the limit inferior is additive. As a result,

$$\liminf_{n\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) = \liminf_{n\to\infty} \max_{a\in\{\,0,\dots,\lfloor n/m\rfloor\,\}} \left\{ \left\|\mathbf{p}^{(n)}\right\|_{(am)} - \left\|\mathbf{p}^{(n)}\right\|_{(a)} \right\}$$

$$= \liminf_{n\to\infty} \max_{a\in[\lfloor n/m\rfloor]} \left\{ \left\|\mathbf{p}^{(n)}\right\|_{(am)} - \left\|\mathbf{p}^{(n)}\right\|_{(a)} \right\}, \tag{E10}$$

where we used that $\left\|\mathbf{p}^{(n)}\right\|_{(0)} = 0$.

Step 2: We start with finding bounds for $\sum_{x=1}^{n} f_n(x) - \int_1^n f_n(x)\,\mathrm{d}x$. To this end we observe that since $f_n(x)$ is by construction non-increasing,

$$\int_1^n f_n(x)\,\mathrm{d}x \leq \int_1^n f_n(x)\,\mathrm{d}x + f_n(n) \leq \sum_{x=1}^n f_n(x) \leq f_n(1) + \int_1^n f_n(x)\,\mathrm{d}x. \tag{E11}$$

By subtracting $\int_1^n f_n(x)\,\mathrm{d}x$ from all terms, the desired bounds follow:

$$0 \le \sum_{x=1}^n f_n(x) - \int_1^n f_n(x)\,\mathrm{d}x \le f_n(1) = 1. \tag{E12}$$

Dividing by $F_n = \sum_{x=1}^n f_n(x)$ and taking the limit $n \to \infty$, we obtain $\lim_{n\to\infty} 1 - \frac{\int_1^n f_n(x)\,\mathrm{d}x}{\sum_{x=1}^n f_n(x)} = 0$. Since $f_n(x) = g(x,n)$, the limit can be rewritten as $\lim_{n\to\infty} \frac{G(n,n)}{F_n} = 1$.

Step 3: In this step we want to show that

$$\liminf_{n\to\infty} \max_{a\in[\lfloor n/m\rfloor]} \left\{ \left\|\mathbf{p}^{(n)}\right\|_{(am)} - \left\|\mathbf{p}^{(n)}\right\|_{(a)} \right\}$$
$$= \liminf_{n\to\infty} \frac{\max_{a\in[\lfloor n/m\rfloor]}\left\{ \sum_{x=a+1}^{am} f_n(x) \right\}}{F_n} \tag{E13}$$
$$= \liminf_{n\to\infty} \frac{\max_{a\in[\lfloor n/m\rfloor]}\left\{ \int_{a+1}^{am} f_n(x)\,\mathrm{d}x \right\}}{\int_1^n f_n(x)\,\mathrm{d}x}.$$

Analogously to the previous step, one can obtain the following bounds for $\frac{\sum_{x=a+1}^{am} f_n(x)}{F_n}$:

$$\frac{\int_{a+1}^{am} f_n(x)\,\mathrm{d}x}{F_n} \le \frac{\sum_{x=a+1}^{am} f_n(x)}{F_n} \le \frac{1 + \int_{a+1}^{am} f_n(x)\,\mathrm{d}x}{F_n}, \tag{E14}$$

and after taking the maximum over $a \in [\lfloor n/m\rfloor]$ and the limit inferior the bounds become

$$\liminf_{n\to\infty} \frac{\max_{a\in[\lfloor n/m\rfloor]}\left\{ \int_{a+1}^{am} f_n(x)\,\mathrm{d}x \right\}}{F_n}$$
$$\le \liminf_{n\to\infty} \frac{\max_{a\in[\lfloor n/m\rfloor]}\left\{ \sum_{x=a+1}^{am} f_n(x) \right\}}{F_n} \tag{E15}$$
$$\le \liminf_{n\to\infty} \frac{1 + \max_{a\in[\lfloor n/m\rfloor]}\left\{ \int_{a+1}^{am} f_n(x)\,\mathrm{d}x \right\}}{F_n}.$$

We can rewrite the last expression as

$$\liminf_{n\to\infty} \left(\frac{1}{F_n} + \frac{\max_{a\in[\lfloor n/m\rfloor]}\left\{ \int_{a+1}^{am} f_n(x)\,\mathrm{d}x \right\}}{F_n} \right)$$
$$= \liminf_{n\to\infty} \left(\frac{1}{F_n} \right) + \liminf_{n\to\infty} \left(\frac{\max_{a\in[\lfloor n/m\rfloor]}\left\{ \int_{a+1}^{am} f_n(x)\,\mathrm{d}x \right\}}{F_n} \right) \tag{E16}$$
$$= \liminf_{n\to\infty} \frac{\max_{a\in[\lfloor n/m\rfloor]}\left\{ \int_{a+1}^{am} f_n(x)\,\mathrm{d}x \right\}}{F_n}.$$

By combining Eq. (E15) with Eq. (E16), one obtains

$$\liminf_{n\to\infty} \frac{\max_{a\in[\lfloor n/m\rfloor]}\left\{ \sum_{x=a+1}^{am} f_n(x) \right\}}{F_n} = \liminf_{n\to\infty} \frac{\max_{a\in[\lfloor n/m\rfloor]}\left\{ \int_{a+1}^{am} f_n(x)\,\mathrm{d}x \right\}}{F_n}. \tag{E17}$$

This is equivalent to

$$\liminf_{n\to\infty} \max_{a\in[\lfloor n/m\rfloor]} \left\{ \left\|\mathbf{p}^{(n)}\right\|_{(am)} - \left\|\mathbf{p}^{(n)}\right\|_{(a)} \right\}$$
$$= \liminf_{n\to\infty} \frac{\max_{a\in[\lfloor n/m\rfloor]}\left\{ G(am,n) - G(a+1,n) \right\}}{G(n,n)}, \tag{E18}$$

where we replaced F_n with $G(n, n)$, which we can do according to step 2.

Step 4: Let c_n be the value that maximizes $\max_{1 \leq a \leq n/m} \{ G(am, n) - G(a + 1, n) \}$ and let c'_n be the natural number satisfying $c'_n \leq c_n < c'_n + 1$. This implies

$$
\begin{aligned}
G(c_n m, n) - G(c_n + 1, n) &= \int_{c_n+1}^{c_n m} g(t, n) \, dt \\
&\leq \int_{c'_n+1}^{c'_n m} g(t, n) \, dt + \int_{c'_n m}^{(c'_n+1)m} g(t, n) \, dt \qquad \text{(E19)} \\
&\leq \int_{c'_n+1}^{c'_n m} g(t, n) \, dt + m.
\end{aligned}
$$

Dividing by $G(n, n)$, taking the limit inferior, and arguing as in the previous step, we obtain

$$
\begin{aligned}
\liminf_{n \to \infty} &\frac{\max_{1 \leq a \leq n/m} \{ G(am, n) - G(a + 1, n) \}}{G(n, n)} \\
&= \liminf_{n \to \infty} \frac{G(c_n m, n) - G(c_n + 1, n)}{G(n, n)} \\
&\leq \liminf_{n \to \infty} \frac{G(c'_n m, n) - G(c'_n + 1, n) + m}{G(n, n)} \qquad \text{(E20)} \\
&= \liminf_{n \to \infty} \frac{G(c'_n m, n) - G(c'_n + 1, n)}{G(n, n)} \\
&\leq \liminf_{n \to \infty} \frac{\max_{a \in [\lfloor n/m \rfloor]} \{ G(am, n) - G(a + 1, n) \}}{G(n, n)}.
\end{aligned}
$$

The reverse inequality follows from the fact that we maximize over a larger set. This proves that

$$
\begin{aligned}
\liminf_{n \to \infty} &\frac{\max_{a \in [\lfloor n/m \rfloor]} \{ G(am, n) - G(a + 1, n) \}}{G(n, n)} \\
&= \liminf_{n \to \infty} \frac{\max_{1 \leq a \leq n/m} \{ G(am, n) - G(a + 1, n) \}}{G(n, n)}.
\end{aligned}
\qquad \text{(E21)}
$$

Step 5: Combining the results of the previous steps, and in particular Eq. (E10), Eq. (E18), and Eq. (E21), we have proven that

$$
\liminf_{n \to \infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) = \liminf_{n \to \infty} \frac{\max_{1 \leq a \leq n/m} \left\{ \int_{a+1}^{am} g(x, n) \, dx \right\}}{\int_1^n g(x, n) \, dx}. \qquad \text{(E22)}
$$

In the above equation, we can replace $\int_{a+1}^{am} g(x, y) \, dx$ with $\int_a^{am} g(x, y) \, dx$ because the difference of the two integrals is finite and divided by a diverging term. Thus

$$
\liminf_{n \to \infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) = \liminf_{n \to \infty} \frac{\max_{1 \leq a \leq n/m} \left\{ \int_a^{am} g(x, y) \, dx \right\}}{\int_1^n g(x, y) \, dx}. \qquad \text{(E23)}
$$

The expression inside the $\liminf$ on the right-hand side of Eq. (E23) is a function of $n \in \mathbb{N} \setminus \{ 1 \}$ and we call it $M(n)$,

$$
M(n) = \frac{\max_{1 \leq a \leq n/m} \left\{ \int_a^{am} g(x, n) \, dx \right\}}{\int_1^n g(x, n) \, dx}. \qquad \text{(E24)}
$$

We extend this function to real numbers $y > 1$ as follows:

$$
M(y) = \frac{\max_{1 \leq a \leq y/m} \left\{ \int_a^{am} g(x, y) \, dx \right\}}{\int_1^y g(x, y) \, dx}. \qquad \text{(E25)}
$$

Since the natural numbers are a subset of the real numbers,

$$\liminf_{n\to\infty} M(n) \geq \liminf_{y\to\infty} M(y), \tag{E26}$$

which proves Eq. (E5). Furthermore, Steps 1-4 can be repeated with exactly the same arguments for the lim sup, and they imply

$$\limsup_{n\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) = \limsup_{n\to\infty} M(n) \leq \limsup_{y\to\infty} M(y). \tag{E27}$$

This proves Eq. (E6).

In the following, we assume that $\lim_{y\to\infty} M(y)$ exists. This implies

$$\lim_{y\to\infty} M(y) = \liminf_{y\to\infty} M(y) = \limsup_{y\to\infty} M(y). \tag{E28}$$

From Eq. (E26), Eq. (E27), and Eq. (E28) we obtain

$$\lim_{y\to\infty} M(y) \leq \liminf_{n\to\infty} M(n) \leq \limsup_{n\to\infty} M(n) \leq \lim_{y\to\infty} M(y). \tag{E29}$$

Therefore, $\lim_{n\to\infty} M(n)$ exists and is equal to $\lim_{y\to\infty} M(y)$. From Eq. (E23) and Eq. (E27) we derive

$$\lim_{n\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) = \lim_{n\to\infty} M(n) = \lim_{y\to\infty} M(y). \tag{E30}$$

This proves Eq. (E7) and concludes the proof.

$\square$

Corollary E.2 combined with Proposition E.3 leads to the following characterization of embezzling families.

Corollary E.4. *Let* $\{\chi_n\}_{n\in\mathbb{N}}$ *be a regular family of states,* f *be the function associated to it, and* g *be defined as in Eq.* (E3)*. If*

$$\lim_{y\to\infty} \frac{\max_{1\leq a\leq y/m} \left\{ \int_a^{am} g(x,y)\,\mathrm{d}x \right\}}{\int_1^y g(x,y)\,\mathrm{d}x} \tag{E31}$$

exists, then the family of states $\{\chi_n\}_{n\in\mathbb{N}}$ *is an universal embezzling family if and only if* $\lim_{n\to\infty} F_n = \infty$ *and*

$$\lim_{y\to\infty} \frac{\max_{1\leq a\leq y/m} \left\{ \int_a^{am} g(x,y)\,\mathrm{d}x \right\}}{\int_1^y g(x,y)\,\mathrm{d}x} = 0. \tag{E32}$$

In addition, thanks to Proposition E.3, the task of determining if a family of states is a universal embezzling family is converted into an optimization problem. It is enough to find the maximum of the differentiable function $G_y(a) = \int_a^{am} g(x,y)\,\mathrm{d}x$ on $1 \leq a \leq y/m$, which is easily done with the help of its derivative

$$mg(am,y) - g(a,y), \quad 1 \leq a \leq y/m. \tag{E33}$$

F Generalization of the van Dam and Hayden Family

The universal embezzling family $\{\chi_n\}_{n\in\mathbb{N}}$ introduced by van Dam and Hayden consists of the states

$$|\chi_n\rangle = \frac{1}{\sqrt{H_n}} \sum_{x=1}^{n} \sqrt{x^{-1}} |xx\rangle, \tag{F1}$$

where $H_n = \sum_{x=1}^{n} x^{-1}$ is the n-th harmonic number. We generalize this family of states as follows: For every $\alpha \in \mathbb{R}$ we introduce the family $\left\{ \chi_n^{(\alpha)} \right\}_{n\in\mathbb{N}}$, where

$$\left|\chi_n^{(\alpha)}\right\rangle = \frac{1}{\sqrt{H_n^{(-\alpha)}}} \sum_{x=1}^{n} \sqrt{x^{\alpha}} |xx\rangle \tag{F2}$$

and $H_n^{(-\alpha)} = \sum_{x=1}^{n} x^{\alpha}$ is the n-th generalized harmonic number. Note that the families of states $\left\{ \chi_n^{(\alpha)} \right\}_{n\in\mathbb{N}}$ are regular families of states, and the van Dam and Hayden family is recovered for $\alpha = -1$. The corresponding Schmidt coefficients, as per our convention arranged in non-increasing order, are

$$\mathbf{p}^{(n|\alpha)} = \begin{cases} \frac{1}{H_n^{(-\alpha)}}(1, 2^{\alpha}, \dots, n^{\alpha}) & \text{if } \alpha \le 0, \\ \frac{1}{H_n^{(-\alpha)}}(n^{\alpha}, \dots, 2^{\alpha}, 1) & \text{if } \alpha > 0. \end{cases} \tag{F3}$$

In the remaining part of this section, we show that the family of state $\left\{ \chi_n^{(\alpha)} \right\}_{n\in\mathbb{N}}$ is a universal embezzling family if and only if $\alpha = -1$ and derive bounds on the star conversion distance.

Case $\alpha < -1$. For $\alpha < -1$, $\lim_{n\to\infty} H_n^{(-\alpha)}$ is finite, thus the regular family of states $\{\chi_n^{\alpha}\}_{n\in\mathbb{N}}$ is not a universal embezzling family (see Corollary E.2). The largest Schmidt coefficient provides a lower bound on the star conversion distance: Using Eq. (C14), we obtain

$$\begin{aligned} T_{\star}(\chi_n^{(\alpha)} \to \chi_n^{(\alpha)} \otimes \Phi_m) &= \max_{k\in[n]} \left\{ \left\| \mathbf{p}^{(n|\alpha)} \right\|_{(k)} - \left\| \mathbf{p}^{(n|\alpha)} \right\|_{(a_k)} - \frac{b_k}{m} p_{a_k+1}^{(n|\alpha)} \right\} \\ &\ge p_1^{(n|\alpha)} \left(1 - \frac{1}{m}\right) \\ &= \frac{1}{H_n^{(-\alpha)}} \left(1 - \frac{1}{m}\right), \end{aligned} \tag{F4}$$

where the inequality followed from choosing $k = 1$. By taking the limit, we obtain

$$\liminf_{n\to\infty} T_{\star}(\chi_n^{(\alpha)} \to \chi_n^{(\alpha)} \otimes \Phi_m) \ge \liminf_{n\to\infty} \frac{1}{H_n^{(-\alpha)}} \left(1 - \frac{1}{m}\right) = \frac{1}{\zeta(-\alpha)} \left(1 - \frac{1}{m}\right), \tag{F5}$$

where $\zeta(-\alpha) = \sum_{x=1}^{\infty} x^{\alpha}$ is the Riemann Zeta function. To obtain an upper bound, we observe that, since $\alpha < -1$,

$$\sum_{x=a+1}^{(a+1)m} x^{\alpha} \le (a+1)^{\alpha} + \int_{a+1}^{(a+1)m} x^{\alpha}\, \mathrm{d}x = (a+1)^{\alpha} + \frac{(a+1)^{\alpha+1}}{\alpha+1} \left(m^{\alpha+1} - 1\right). \tag{F6}$$

The right-hand side is decreasing in a, and thus

$$
\max_{k\in[n]}\left\{\left\|\mathbf{p}^{(n|\alpha)}\right\|_{(k)}-\left\|\mathbf{p}^{(n|\alpha)}\right\|_{(a_k)}-\frac{b_k}{m}p_{a_k+1}^{(n|\alpha)}\right\}
$$
$$
\leq\frac{1}{H_n^{(-\alpha)}}\max_{k\in[n]}\left\{\sum_{x=a_k+1}^{k}x^\alpha\right\}
$$
$$
\leq\frac{1}{H_n^{(-\alpha)}}\max_{a\in\{0,\ldots,\lfloor n/m\rfloor\}}\left\{\sum_{x=a+1}^{(a+1)m}x^\alpha\right\} \tag{F7}
$$
$$
\leq\frac{1}{H_n^{(-\alpha)}}\max_{a\in\{0,\ldots,\lfloor n/m\rfloor\}}\left\{(a+1)^\alpha+\frac{(a+1)^{\alpha+1}}{\alpha+1}\left(m^{\alpha+1}-1\right)\right\}
$$
$$
=\frac{1}{H_n^{(-\alpha)}}\left(1+\frac{m^{\alpha+1}-1}{\alpha+1}\right).
$$

Taking the limit $n\to\infty$, and considering that the conversion distance is by definition smaller than one, we obtain

$$
\limsup_{n\to\infty}T_\star(\chi_n^{(\alpha)}\to\chi_n^{(\alpha)}\otimes\Phi_m)\leq\min\left\{1,\frac{1}{\zeta(-\alpha)}\left(1+\frac{m^{\alpha+1}-1}{\alpha+1}\right)\right\}. \tag{F8}
$$

Case $\alpha=-1$. This is the van Dam and Hayden family. Analogously to the previous case, we observe that

$$
\sum_{x=a+1}^{(a+1)m}x^{-1}\leq(a+1)^{-1}+\int_{a+1}^{(a+1)m}x^{-1}\,\mathrm{d}x=(a+1)^{-1}+\ln m\leq1+\ln m. \tag{F9}
$$

Since $H_n^{(1)}=H_n$ diverges, by following the same steps, we have

$$
\lim_{n\to\infty}T_\star(\chi_n^{(-1)}\to\chi_n^{(-1)}\otimes\Phi_m)\leq\lim_{n\to\infty}\frac{1+\ln m}{H_n}=0. \tag{F10}
$$

Thus, the limit for the conversion distance is zero as expected for the van Dam and Hayden family.

Case $-1<\alpha<0$. In this scenario $H_n^{(-\alpha)}$ diverges when $n\to\infty$, and we can use the results of Section E. Here $g(x,y)=x^\alpha$, and the derivative with respect to a of the function $\int_a^{am}g(x,y)\,\mathrm{d}x$ is

$$
mg(am,y)-g(a,y)=a^\alpha(m^{\alpha+1}-1), \tag{F11}
$$

which is positive in the domain $1\leq a\leq y/m$. This implies that the function $\int_a^{am}g(x,y)\,\mathrm{d}x$ is non-decreasing and the maximum

$$
\max_{1\leq a\leq y/m}\left\{\int_a^{am}g(x,y)\,\mathrm{d}x\right\} \tag{F12}
$$

is obtained for $a=y/m$. Due to Proposition E.3,

$$
\lim_{n\to\infty}T_\star(\chi_n^{(\alpha)}\to\chi_n^{(\alpha)}\otimes\Phi_m)=\lim_{y\to\infty}\frac{\int_{y/m}^{y}x^\alpha\,\mathrm{d}x}{\int_1^{y}x^\alpha\,\mathrm{d}x}=1-\frac{1}{m^{\alpha+1}}. \tag{F13}
$$

Case $\alpha = 0$. Here we have

$$p_x^{(n)} = \frac{1}{n}, \quad (\mathbf{p}^{(n)} \otimes \mathbf{u}^{(m)})_x = \frac{1}{nm}. \tag{F14}$$

We can compute the star conversion distance directly using Eq. (C14) and obtain

$$T_\star(\chi_n^{(0)} \to \chi_n^{(0)} \otimes \Phi_m) = \max_{k \in [n]} \left\{ \frac{k}{n} - \frac{k}{nm} \right\} = 1 - \frac{1}{m}. \tag{F15}$$

Case $\alpha > 0$. Also here $H_n^{(-\alpha)}$ diverges for $n \to \infty$ and we can use the results of Section E. The function $g(x,y)$ is defined as $g(x,y) = (y+1-x)^\alpha / y^\alpha$, because x^α is increasing. From this follows

$$mg(am, y) - g(a, y) \geq 0 \Leftrightarrow m(y + 1 - am)^\alpha \geq (y + 1 - a)^\alpha$$

$$\Leftrightarrow a \leq \frac{(y+1)\left(m^{\frac{1}{\alpha}} - 1\right)}{m^{1 + \frac{1}{\alpha}} - 1} := a_{\max}. \tag{F16}$$

Since $a_{\max}$ increases linearly with y and $a_{\max} \leq \frac{y}{m}$ is equivalent to $y \geq \frac{m(m^{1/\alpha} - 1)}{m - 1}$, for large enough y, the value $a_{\max}$ belongs to the interval $[1, y/m]$ and is a global maximum. Plugging $a_{\max}$ into Proposition E.3, we obtain

$$\lim_{n \to \infty} T_\star(\chi_n^{(\alpha)} \to \chi_n^{(\alpha)} \otimes \Phi_m) = \lim_{y \to \infty} \frac{\int_{a_{\max}}^{a_{\max} m} \frac{(y+1-x)^\alpha}{y^\alpha} \, \mathrm{d}x}{\int_1^y \frac{(y+1-x)^\alpha}{y^\alpha} \, \mathrm{d}x} = (m - 1) \left(\frac{m - 1}{m^{1 + \frac{1}{\alpha}} - 1} \right)^\alpha. \tag{F17}$$

This completes the study of regular families of states defined by $f(x) = x^\alpha$. The star conversion distance vanishes only for $\alpha = -1$, thus the only universal embezzling family of this form is the one introduced by van Dam and Hayden. Furthermore, exact values for the limit of the star conversion distance for $\alpha \geq -1$ and lower and upper bounds for $\alpha < -1$ were provided.

G   Uniqueness of the van Dam and Hayden Embezzling Family

In this section, we provide further results on the uniqueness of the van Dam and Hayden embezzling family. Given a regular family of states $\{\chi_n\}_{n \in \mathbb{N}}$, the asymptotic behavior of the function f associated to it is relevant to determine whether $\{\chi_n\}_{n \in \mathbb{N}}$ is a universal embezzling family or not (e.g., see Corollary E.2). We thus use the notations little-ω and little-o (see, e.g., Ref. [92]) to describe asymptotic relations between two functions $g, h \colon [1, \infty) \to (0, \infty)$,

$$h \in \omega(g) \Leftrightarrow \lim_{x \to \infty} \frac{h(x)}{g(x)} = +\infty, \quad h \in o(g) \Leftrightarrow \lim_{x \to \infty} \frac{h(x)}{g(x)} = 0. \tag{G1}$$

Before we state the results about the uniqueness of the van Dam and Hayden embezzling family, we prove the following Lemma, which is a direct consequence of Theorem 2 in Ref. [93], which we restate here to improve readability: Let μ be a measure on the real line $\mathbb{R}$, and let f_i, g_i $(i = 1, 2)$ be four Borel-measurable functions: $\mathbb{R} \to \mathbb{R}$ such that $f_2 \geq 0$ and $g_2 \geq 0$, and $\int |f_i g_j| \, \mathrm{d}\mu < \infty$ $(i, j = 1, 2)$. If f_1/f_2 and g_1/g_2 are monotonic in the same direction, then

$$\int f_1 g_1 \, \mathrm{d}\mu \int f_2 g_2 \, \mathrm{d}\mu \geq \int f_1 g_2 \, \mathrm{d}\mu \int f_2 g_1 \, \mathrm{d}\mu. \tag{G2}$$

Lemma G.1. *Let f, g be continuous, positive functions on $[a,b]$ such that $f(x)/g(x)$ is a non-decreasing function on $[a,b]$. Then, for every x_1 such that $a < x_1 < b$,*

$$\int_a^b f(x)\,\mathrm{d}x \int_a^{x_1} g(x)\,\mathrm{d}x \geq \int_a^{x_1} f(x)\,\mathrm{d}x \int_a^b g(x)\,\mathrm{d}x. \tag{G3}$$

Proof. Let $\tilde{h}(x) = 1$ and $h(x) = \chi_{[a,x_1]}(x)$, where $\chi_S(x)$ is the characteristic function of the set S. Furthermore, let

$$h_k(x) = \begin{cases} 1 & \text{if } a \leq x \leq x_1, \\ e^{-k(x-x_1)} & \text{if } x_1 < x \leq b. \end{cases} \tag{G4}$$

The sequence of functions $\{\, h_k \,\}_{k \in \mathbb{N}}$ converges to $h(x)$ and satisfies $0 \leq h_k(x) \leq 1$. Due to the dominated convergence theorem (see, e.g., Ref. [94]), we have

$$\lim_{k \to \infty} \int_a^b f(x) h_k(x) = \int_a^{x_1} f(x). \tag{G5}$$

The same result holds for g.

Since both $f(x)/g(x)$ and $\tilde{h}(x)/h_k(x)$ are non-decreasing, we can apply Theorem 2 of Ref. [93] to find

$$\int_a^b f(x)\tilde{h}(x)\,\mathrm{d}x \int_a^b g(x)h_k(x)\,\mathrm{d}x \geq \int_a^b f(x)h_k(x)\,\mathrm{d}x \int_a^b g(x)\tilde{h}(x)\,\mathrm{d}x. \tag{G6}$$

After taking the limit $k \to \infty$ on both sides, we obtain

$$\int_a^b f(x)\,\mathrm{d}x \int_a^{x_1} g(x)\,\mathrm{d}x \geq \int_a^{x_1} f(x)\,\mathrm{d}x \int_a^b g(x)\,\mathrm{d}x, \tag{G7}$$

which finishes the proof. $\qquad\square$

With this lemma at hand, we are ready to present the promised results concerning the uniqueness of the van Dam and Hayden embezzling family.

Theorem G.2. *Let f be a positive non-increasing function such that $f(x)/x^\alpha$ is asymptotically monotonic for all $\alpha \in \mathbb{R}$ and let $\{\, \chi_n \,\}_{n \in \mathbb{N}}$ be the regular family of states associated to it (see Definition E.1). Then $\{\, \chi_n \,\}_{n \in \mathbb{N}}$ is a universal embezzling family if and only if $f \in \omega(x^{-1-\varepsilon}) \cap o(x^{-1+\varepsilon})\ \forall \varepsilon > 0$ and $\sum_{x=1}^\infty f(x) = \infty$. Furthermore, if $f \notin \omega(x^{-1-\varepsilon}) \cap o(x^{-1+\varepsilon})$ for at least one $\varepsilon > 0$, then $\left\{\, \chi_{n_j} \,\right\}_{j \in \mathbb{N}}$, where $\{\, n_j \,\}_{j \in \mathbb{N}}$ is any sequence of natural numbers, is not a universal embezzling family.*

Proof. Before we start, we notice that if $\lim_{j \to \infty} n_j = J < \infty$, then $\lim_{j \to \infty} F_{n_j} = \sum_{x=1}^J f(x) < \infty$, which implies that the family of states $\left\{\, \chi_{n_j} \,\right\}_{j \in \mathbb{N}}$ is not an embezzling family (see Corollary E.2). In this proof we will thus assume, w.l.o.g., that $\lim_{j \to \infty} n_j = \infty$.

Necessary condition — Let $\{\, \chi_n \,\}_{n \in \mathbb{N}}$ be a universal embezzling family with corresponding function f satisfying the assumptions above. We already proved in Corollary E.2 that if $\{\, \chi_n \,\}_{n \in \mathbb{N}}$ is a universal embezzling family, then $\sum_{x=1}^\infty f(x) = +\infty$. To show the remainder, let

$$R = \left\{\, \alpha \in \mathbb{R} \,\middle|\, \lim_{x \to \infty} \frac{f(x)}{x^\alpha} = 0 \,\right\}, \quad L = \left\{\, \alpha \in \mathbb{R} \,\middle|\, \lim_{x \to \infty} \frac{f(x)}{x^\alpha} = +\infty \,\right\}. \tag{G8}$$

Since $\lim_{x\to\infty} f(x)/x^\alpha = 0$ for $\alpha > 0$, we have that $R \neq \emptyset$. Now we prove by contradiction that also $L \neq \emptyset$. Let us assume that L is empty. This implies that $\lim_{x\to\infty} f(x)/x^\alpha = l_\alpha < \infty$ for all α. If there exists an α such that $l_\alpha \neq 0$, then $\lim_{x\to\infty} f(x)/x^{\alpha-1} = l_\alpha(\lim_{x\to\infty} x) = +\infty$. Thus, $\alpha - 1 \in L$, and $L \neq \emptyset$, leading to the desired contradiction. If instead $\lim_{x\to\infty} f(x)/x^\alpha = 0$ for all α, then $\lim_{x\to\infty} f(x)/x^{-2} = 0$. This implies that $f(x)/x^{-2}$ converges monotonically to zero for large x, i.e., there exists an N such that $f(x) < x^{-2}$ for $x > N$. Thus, $\sum_{x=N}^\infty f(x) < \sum_{x=N}^\infty x^{-2} < \infty$. According to Corollary E.2, this contradicts the hypothesis that the family under consideration is universally embezzling. We have therefore shown that $L \neq \emptyset \neq R$.

The next step is to prove that $\inf R = \sup L$. From the definition of R and L follows that $\inf R \geq \sup L$. Let us assume that $\inf R > \sup L$, i.e., that there exists an $\alpha \in \mathbb{R}$ such that $\sup L < \alpha < \inf R$. Since $\alpha \notin L \cup R$, there exists a positive real number l such that $\lim_{x\to\infty} f(x)/x^\alpha = l$. Now pick α_1 such that $\sup L < \alpha_1 < \alpha < \inf R$. We observe that $\lim_{x\to\infty} f(x)/x^{\alpha_1} = l(\lim_{x\to\infty} x^\alpha/x^{\alpha_1}) = +\infty$, thus $\alpha_1 \in L$. This is in contradiction to the choice $\alpha_1 > \sup L$. We therefore showed that $\sup L = \inf R$.

So far, we have shown that if $\{\chi_n\}$ is a universal embezzling family satisfying our assumptions, then there exists a unique α such that $f(x) \in o(x^{\alpha+\varepsilon}) \cap \omega(x^{\alpha-\varepsilon})$ for every $\varepsilon > 0$. What is left to show is that if $\alpha \neq -1$, then the family of states corresponding to f cannot be universally embezzling. From the above discussion, we know that $\{\alpha > 0\} \subseteq R$, thus we can focus on $\alpha \leq 0$ and split our discussion into two scenarios, $\alpha < -1$ and $-1 < \alpha \leq 0$.

If $\alpha < -1$, then there exists an $\varepsilon > 0$ such that $\alpha + \varepsilon < -1$. Since $f \in o(x^{\alpha+\varepsilon})$ and $f(x)/x^{\alpha+\varepsilon}$ converges monotonically to zero for large x, there exists an N such that $f(x) < x^{\alpha+\varepsilon}$ for $x > N$. This implies that $\sum_{x=N}^\infty f(x) < \sum_{x=N}^\infty x^{\alpha+\varepsilon} < \infty$ and therefore $\sum_{x=1}^\infty f(x) < \infty$, which, according to Corollary E.2, contradicts the hypothesis that $\{\chi_n\}$ is a universal embezzling family. Let $\{n_j\}_{j\in\mathbb{N}}$ be any sequence of natural number such that $\lim_{j\to\infty} n_j = \infty$. Then $\lim_{j\to\infty} \sum_{x=1}^{n_j} f(x) = \sum_{x=1}^\infty f(x) < \infty$. This implies, again due to Corollary E.2, that $\left\{\chi_{n_j}\right\}_{j\in\mathbb{N}}$ is not a universal embezzling family.

If $-1 < \alpha \leq 0$, then there exists an $\varepsilon > 0$ such that $\alpha - \varepsilon > -1$. We notice that $f(x)/x^{\alpha-\varepsilon}$ diverges to infinity for large x, because $f \in \omega(x^{\alpha-\varepsilon})$. This implies that there exists an N such that $f(x) \geq x^{\alpha-\varepsilon}$ for $x > N$ and therefore $\sum_{x=N}^\infty f(x) \geq \sum_{x=N}^\infty x^{\alpha-\varepsilon} = \infty$, which allows us to use the results of Section E. Since $f(x)/x^{\alpha-\varepsilon}$ is non-decreasing for $x > N$, for all y such that $y/m > N$, we can use Lemma G.1 to obtain

$$\int_N^y f(x)\,\mathrm{d}x \int_N^{y/m} x^{\alpha-\varepsilon}\,\mathrm{d}x \geq \int_N^{y/m} f(x)\,\mathrm{d}x \int_N^y x^{\alpha-\varepsilon}\,\mathrm{d}x. \tag{G9}$$

This implies that

$$\frac{(y/m)^{\alpha-\varepsilon+1} - N^{\alpha-\varepsilon+1}}{(y)^{\alpha-\varepsilon+1} - N^{\alpha-\varepsilon+1}} \geq \frac{\int_N^{y/m} f(x)\,\mathrm{d}x}{\int_N^y f(x)\,\mathrm{d}x}. \tag{G10}$$

Next we introduce $\delta := \alpha - \varepsilon + 1 > 0$ and take on both sides the limit inferior $y \to \infty$, leading to

$$\begin{aligned}
\frac{1}{m^\delta} &= \liminf_{y\to\infty} \frac{(y/m)^{\alpha-\varepsilon+1} - N^{\alpha-\varepsilon+1}}{(y)^{\alpha-\varepsilon+1} - N^{\alpha-\varepsilon+1}} \\
&\geq \liminf_{y\to\infty} \frac{\int_N^{y/m} f(x)\,\mathrm{d}x}{\int_N^y f(x)\,\mathrm{d}x} \\
&= \liminf_{y\to\infty} \frac{\int_1^{y/m} f(x)\,\mathrm{d}x}{\int_1^y f(x)\,\mathrm{d}x},
\end{aligned} \tag{G11}$$

where in the last equality, we used that $\int_1^N f(x)\,\mathrm{d}x$ is finite, whilst $\int_N^\infty f(x)\,\mathrm{d}x$ diverges. As last step, we observe that according to Proposition E.3,

$$
\begin{aligned}
\liminf_{n\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) &\geq \liminf_{y\to\infty} \frac{\max_{1\leq a\leq y/m}\left\{ \int_a^{am} f(x)\,\mathrm{d}x \right\}}{\int_1^y f(x)\,\mathrm{d}x} \\
&\geq \liminf_{y\to\infty} \frac{\int_{y/m}^y f(x)\,\mathrm{d}x}{\int_1^y f(x)\,\mathrm{d}x} \\
&\geq 1 - \frac{1}{m^\delta} \\
&> 0.
\end{aligned}
\tag{G12}
$$

Using the remark after Definition E.1, we obtain that for $-1 < \alpha \leq 0$, there are no sequences $\{\, n_j \,\}_{j\in\mathbb{N}}$ such that $\left\{\, \chi_{n_j} \,\right\}_{j\in\mathbb{N}}$ is a universal embezzling family, contradicting the hypothesis. Once we combine the results for $\alpha < -1$ and $-1 < \alpha \leq 0$, we have that if $\{\, \chi_n \,\}$ is a universal embezzling family satisfying our assumptions, $f(x) \in o(x^{-1+\varepsilon}) \cap \omega(x^{-1-\varepsilon})$ for all $\varepsilon > 0$. We have also shown that if $f \notin o(x^{-1+\varepsilon}) \cap \omega(x^{-1-\varepsilon})$ for at least one $\varepsilon > 0$, then there are no sequences $\{\, n_j \,\}_{j\in\mathbb{N}}$ such that the family $\left\{\, \chi_{n_j} \,\right\}_{j\in\mathbb{N}}$ is a universal embezzling family. This concludes the first part of the proof.

Sufficient condition — Let f be a function satysfying our assumptions. Since $f(x)/x^{-1}$ is asymptotically monotonic by assumption, $\lim_{x\to\infty} f(x)/x^{-1}$ exists in $[0, +\infty]$. A priori, it can be either 0, $0 < l \in \mathbb{R}$, or $+\infty$. If $\lim_{x\to\infty} f(x)/x^{-1} = l > 0$, then by definition of the limit, for any $l > \tilde{\varepsilon} > 0$ there exists an N such that $\left| \frac{f(x)}{x^{-1}} - l \right| < \tilde{\varepsilon}$ for all $x > N$. This is equivalent to $(l - \tilde{\varepsilon})x^{-1} < f(x) < (l + \tilde{\varepsilon})x^{-1}$ for all $x > N$. Since $\sum_{x=N}^\infty f(x) \geq (l - \tilde{\varepsilon}) \sum_{x=N}^\infty x^{-1} = +\infty$, we can compute the star conversion distance using Proposition E.3, and find

$$
\begin{aligned}
&\lim_{n\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) \\
&= \lim_{y\to\infty} \frac{\max_{1\leq a\leq y/m}\left\{ \int_a^{am} f(x)\,\mathrm{d}x \right\}}{\int_1^y f(x)\,\mathrm{d}x} \\
&\leq \lim_{y\to\infty} \frac{\max_{1\leq a < N}\left\{ \int_a^{am} f(x)\,\mathrm{d}x \right\} + \max_{N\leq a\leq y/m}\left\{ \int_a^{am} f(x)\,\mathrm{d}x \right\}}{\int_1^y f(x)\,\mathrm{d}x} \\
&= \lim_{y\to\infty} \frac{\max_{N\leq a\leq y/m}\left\{ \int_a^{am} f(x)\,\mathrm{d}x \right\}}{\int_1^y f(x)\,\mathrm{d}x} \\
&\leq \lim_{y\to\infty} \frac{\max_{N\leq a\leq y/m}\left\{ \int_a^{am} f(x)\,\mathrm{d}x \right\}}{\int_N^y f(x)\,\mathrm{d}x} \\
&\leq \frac{l + \tilde{\varepsilon}}{l - \tilde{\varepsilon}} \lim_{y\to\infty} \frac{\max_{N\leq a\leq y/m}\left\{ \int_a^{am} x^{-1}\,\mathrm{d}x \right\}}{\int_N^y x^{-1}\,\mathrm{d}x} \\
&= \frac{l + \tilde{\varepsilon}}{l - \tilde{\varepsilon}} \lim_{y\to\infty} \frac{\log m}{\log y - \log N} \\
&= 0.
\end{aligned}
\tag{G13}
$$

Thus if $\lim_{x\to\infty} f(x)/x^{-1} = l \neq 0$, then $\{\, \chi_n \,\}$ is a universal embezzling family.

Suppose now $\lim_{x\to\infty} f(x)/x^{-1} = 0$. Since $f(x)/x^{-1}$ is non-increasing for large x, there exists an $\tilde{N}$ such that $f(x) < x^{-1}$ for all $x > \tilde{N}$. For any fixed $\varepsilon > 0$, by hypothesis, $\lim_{x\to\infty} f(x)/x^{-1-\varepsilon} = +\infty$. Furthermore, since by assumption $\int_1^\infty f(x)\,\mathrm{d}x = \infty$, while $\int_1^\infty x^{-1-\varepsilon}\,\mathrm{d}x < \infty$, there exists an $\tilde{N}_\varepsilon$, such that $\int_1^y f(x)\,\mathrm{d}x \geq \int_1^y x^{-1-\varepsilon}\,\mathrm{d}x$ for $y > \tilde{N}_\varepsilon$.

Let $N_\varepsilon = \max\left\{ \tilde{N}, \tilde{N}_\varepsilon \right\}$. Using again Proposition E.3 and performing the same steps as in Eq. (G13), we obtain

$$
\begin{aligned}
\limsup_{n\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) &\le \limsup_{y\to\infty} \frac{\max_{N_\varepsilon \le a \le y/m}\left\{ \int_a^{am} f(x)\,\mathrm{d}x \right\}}{\int_1^y f(x)\,\mathrm{d}x} \\
&\le \limsup_{y\to\infty} \frac{\max_{N_\varepsilon \le a \le y/m}\left\{ \int_a^{am} x^{-1}\,\mathrm{d}x \right\}}{\int_1^y x^{-1-\varepsilon}\,\mathrm{d}x} \\
&= \log m \lim_{y\to\infty} \frac{\varepsilon}{1 - y^{-\varepsilon}} \\
&= \varepsilon \log m.
\end{aligned} \tag{G14}
$$

Since this is true for all $\varepsilon > 0$,

$$
\limsup_{n\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) \le \lim_{\varepsilon \to 0} \varepsilon \log m = 0. \tag{G15}
$$

This implies that $\lim_{n\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) = 0$, i.e., $\{\chi_n\}$ is a universal embezzling family.

Let us now consider the case $\lim_{x\to\infty} f(x)/x^{-1} = \infty$. From this immediately follows that $\sum_{x=1}^\infty f(x) = +\infty$, so we can use the results of Section E. Furthermore, there exists an N such that $f(x)/x^{-1}$ is non-decreasing for $x \ge N$. We write $f(x) = x^{-1}h(x)$ and thus $h(x)$ is non-decreasing for $x \ge N$. Computing the derivative of $\int_a^{am} f(x)\,\mathrm{d}x$ for $a \ge N$, we find

$$
\frac{\mathrm{d}\int_a^{am} f(x)\,\mathrm{d}x}{\mathrm{d}a} = mf(am) - f(a) = \frac{h(am) - h(a)}{a} \ge 0. \tag{G16}
$$

Thus, $\max_{N \le a \le y/m} \int_a^{am} f(x)\,\mathrm{d}x = \int_{y/m}^y f(x)\,\mathrm{d}x$ and, by following the same steps as in Eq. (G13), we obtain

$$
\begin{aligned}
\limsup_{n\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) &\le \limsup_{y\to\infty} \frac{\max_{N \le a \le y/m}\left\{ \int_a^{am} f(x)\,\mathrm{d}x \right\}}{\int_1^y f(x)\,\mathrm{d}x} \\
&= \limsup_{y\to\infty} \frac{\int_{y/m}^y f(x)\,\mathrm{d}x}{\int_1^y f(x)\,\mathrm{d}x}.
\end{aligned} \tag{G17}
$$

Let us fix $\varepsilon > 0$. Since $f(x)/x^{-1+\varepsilon}$ is non-increasing for $x > N_\varepsilon$ by hypothesis, $x^{-1+\varepsilon}/f(x)$ is non-decreasing for $x > N_\varepsilon$. By applying Lemma G.1 (for y large enough), we obtain

$$
\int_{N_\varepsilon}^y x^{-1+\varepsilon}\,\mathrm{d}x \int_{N_\varepsilon}^{y/m} f(x)\,\mathrm{d}x \ge \int_{N_\varepsilon}^{y/m} x^{-1+\varepsilon}\,\mathrm{d}x \int_{N_\varepsilon}^y f(x)\,\mathrm{d}x. \tag{G18}
$$

From this follows that

$$
\frac{\int_{N_\varepsilon}^{y/m} f(x)\,\mathrm{d}x}{\int_{N_\varepsilon}^y f(x)\,\mathrm{d}x} \ge \frac{(y/m)^\varepsilon - N_\varepsilon^\varepsilon}{y^\varepsilon - N_\varepsilon^\varepsilon}. \tag{G19}
$$

Taking the $\limsup$ on both sides, and adding the finite contributions $\int_1^{N_\varepsilon} f(x)\,\mathrm{d}x$ to the diverging integrals $\int_{N_\varepsilon}^{y/m} f(x)\,\mathrm{d}x$ and $\int_{N_\varepsilon}^y f(x)\,\mathrm{d}x$, we obtain

$$
\limsup_{y\to\infty} \frac{\int_1^{y/m} f(x)\,\mathrm{d}x}{\int_1^y f(x)\,\mathrm{d}x} \ge \frac{1}{m^\varepsilon}. \tag{G20}
$$

Combining Eq. (G17) and Eq. (G20) we get

$$
\limsup_{n\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) \le 1 - \limsup_{y\to\infty} \frac{\int_1^{y/m} f(x)\,\mathrm{d}x}{\int_1^y f(x)\,\mathrm{d}x} \le 1 - \frac{1}{m^\varepsilon}. \tag{G21}
$$

This is true for all $\varepsilon > 0$, thus taking the limit $\varepsilon \to 0$ we obtain the desired result

$$\limsup_{n \to \infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) \leq \lim_{\varepsilon \to 0} 1 - \frac{1}{m^\varepsilon} = 0. \tag{G22}$$

This concludes the proof of the sufficient condition. $\qquad\square$

The above result on the uniqueness of the van Dam and Hayden family can be expressed as follows: Any regular family of states $\{\chi_n\}_{n \in \mathbb{N}}$ satisfying the conditions of Theorem G.2 is a universal embezzling family if and only if f, the function associated to $\{\chi_n\}_{n \in \mathbb{N}}$, is asymptotically close to x^{-1}, the function associated to the van Dam and Hayden family, where asymptotically close means $f \in \omega(x^{-1-\varepsilon}) \cap o(x^{-1+\varepsilon})$ for all $\varepsilon > 0$.

The assumption that $f(x)/x^\alpha$ is asymptotically monotonic for every $\alpha \in \mathbb{R}$ is crucial for our proof and does not follow from the monotonicity of f. There are functions that are non-increasing, but oscillate asymptotically when multiplied by powers of x. An example is the function

$$f(x) = \frac{1 + (1 + \sin \ln \ln x) \ln x}{x}, \tag{G23}$$

which is non-increasing, but $x f(x)$ oscillates between 1 and ∞. Theorem G.2 does not provide any information about families of states associated to such functions, and it cannot be used to determine whether such families are universally embezzling or not.

Next, we prove two related propositions concerning regular families of states associated to non-decreasing functions.

Proposition G.3. *Let f be a positive non-decreasing function such that $f(x)/x^\alpha$ is asymptotically non-increasing for at least one $\alpha > 0$. Then the regular family of states $\{\chi_n\}_{n \in \mathbb{N}}$ associated to f (see Definition E.1) is not a universal embezzling family. Furthermore, there are no sequences $\{n_j\}_{j \in \mathbb{N}}$ such that $\{\chi_{n_j}\}_{j \in \mathbb{N}}$ is a universal embezzling family.*

Proof. Since $f(x)$ is non-decreasing and $\sum_{x=1}^\infty f(x)$ diverges, we can use Proposition E.3 and Lemma G.1, to obtain

$$\begin{aligned}
\liminf_{n \to \infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) &\geq \liminf_{y \to \infty} \frac{\max_{1 \leq a \leq y/m} \left\{ \int_a^{am} g(x,y)\, \mathrm{d}x \right\}}{\int_1^y g(x,y)\, \mathrm{d}x} \\
&= \liminf_{y \to \infty} \frac{\max_{1 \leq a \leq y/m} \left\{ \int_{y-am+1}^{y-a+1} f(x)\, \mathrm{d}x \right\}}{\int_1^y f(x)\, \mathrm{d}x} \\
&\geq \liminf_{y \to \infty} \frac{\left\{ \int_1^{y(1-1/m)+1} f(x)\, \mathrm{d}x \right\}}{\int_1^y f(x)\, \mathrm{d}x} \\
&\geq \liminf_{y \to \infty} \frac{\left\{ \int_1^{y(1-1/m)+1} x^\alpha\, \mathrm{d}x \right\}}{\int_1^y x^\alpha\, \mathrm{d}x} \\
&= \left(1 - \frac{1}{m}\right)^{\alpha+1}.
\end{aligned} \tag{G24}$$

The last inequality is based on Lemma G.1 and is derived as in the previous cases. Using the remark after Definition E.1, we obtain that $\{\chi_n\}_{n \in \mathbb{N}}$ is not a universal embezzling family. Furthermore, there are no sequences $\{n_j\}_{j \in \mathbb{N}}$ such that $\{\chi_{n_j}\}_{j \in \mathbb{N}}$ is a universal embezzling family. $\qquad\square$

Proposition G.4. *Let f be a positive non-decreasing function such that $f(x)/e^{kx}$ is asymptotically non-decreasing for at least one $k > 0$. Then the regular family of states $\{ \chi_n \}_{n \in \mathbb{N}}$ associated to f is not a universal embezzling family. Furthermore, there are no sequences $\{ n_j \}_{j \in \mathbb{N}}$ such that $\{ \chi_{n_j} \}_{j \in \mathbb{N}}$ is a universal embezzling family.*

Proof. From Eq. (D2) follows that

$$
\begin{aligned}
\liminf_{n \to \infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) &\geq \left(1 - \frac{1}{m}\right) \liminf_{n \to \infty} p_1^{(n)} \\
&= \left(1 - \frac{1}{m}\right) \liminf_{n \to \infty} \frac{f(n)}{\sum_{x=1}^n f(x)} \\
&= \left(1 - \frac{1}{m}\right) \liminf_{y \to \infty} \frac{f(y)}{\int_1^y f(x) \, dx} \\
&\geq \left(1 - \frac{1}{m}\right) \liminf_{y \to \infty} \frac{\int_{y-1}^y f(x) \, dx}{\int_1^y f(x) \, dx} \\
&\geq \left(1 - \frac{1}{m}\right) \liminf_{y \to \infty} \frac{\int_{y-1}^y e^{kx} \, dx}{\int_1^y e^{kx} \, dx} \\
&= \left(1 - \frac{1}{m}\right) (1 - e^{-k}).
\end{aligned}
\tag{G25}
$$

Also here the last inequality is due to Lemma G.1 and the family of states is not a universal embezzling family. Furthermore, there are no sequences $\{ n_j \}_{j \in \mathbb{N}}$ such that $\{ \chi_{n_j} \}_{j \in \mathbb{N}}$ is a universal embezzling family. $\qquad\square$

H Asymptotically Regular Families

In the definition of universal embezzling families, Definition 5, and in all the results about embezzlement, only the asymptotic behaviour of a family of states $\{ \chi_n \}_{n \in \mathbb{N}}$ is relevant. This motivates the following definition, which is a generalization of regular families.

Definition H.1. *A family of states $\{ \chi_n \}_{n \in \mathbb{N}}$ is called asymptotically regular if there exists an asymptotically monotonic function $f \colon \mathbb{N} \to (0, \infty)$ such that*

$$
|\chi_n\rangle = \frac{1}{\sqrt{F_n}} \sum_{x=1}^n \sqrt{f(x)} \, |xx\rangle \quad \text{for all} \quad n \in \mathbb{N},
\tag{H1}
$$

where $F_n = \sum_{x=1}^n f(x)$.

We next show that our results hold for asymptotically regular families too. To this end, we start with the following theorem.

Theorem H.2. *Let $\{ \chi_n \}_{n \in \mathbb{N}}$ be an asymptotically regular family and let f be the function associated to it (see Definition H.1). Then one can construct a function $\tilde{f}$ that satisfies*

1. *$\tilde{f}$ corresponds to a regular family $\{ \tilde{\chi}_n \}_{n \in \mathbb{N}}$ (see Definition E.1),*

2. *$\lim_{x \to \infty} \tilde{f}(x)/f(x) = 1$,*

3. *$\{ \tilde{\chi}_n \}_{n \in \mathbb{N}}$ is a universal embezzling family if and only if $\{ \chi_n \}_{n \in \mathbb{N}}$ is a universal embezzling family,*

4. $\{\,\tilde{\chi}_n\,\}_{n\in\mathbb{N}}$ *contains a universal embezzling subfamily if and only if* $\{\,\chi_n\,\}_{n\in\mathbb{N}}$ *contains a universal embezzling subfamily.*

Proof. Since f is by assumption asymptotically monotonic, $\lim_{x\to\infty} f(x)$ exists in the extended domain $[0,\infty]$. We first study the case $\lim_{x\to\infty} f(x) = l$, with $0 < l \in \mathbb{R}$. In this case, we choose $\tilde{f}(x) = l$, which corresponds to a regular family. Moreover, from the definition of $\tilde{f}$ follows that $\lim_{x\to\infty} f(x)/\tilde{f}(x) = 1$. We also notice that for every $0 < \varepsilon < l$, there exists an N such that $l - \varepsilon < f(x) < l + \varepsilon$ for all $x \geq N$. This implies that $\sum_{x=1}^{\infty} f(x) \geq \sum_{x=N}^{\infty} f(x) \geq \sum_{x=N}^{\infty} (l - \varepsilon) = \infty$. Similarly, $\sum_{x=1}^{\infty} \tilde{f}(x) = \sum_{x=1}^{\infty} l = \infty$. Since $\lim_{n\to\infty} F_n^{-1} = \lim_{n\to\infty} \tilde{F}_n^{-1} = 0$, we obtain that $\lim_{n\to\infty} p_1^{(n)} = 0$ (where $\mathbf{p}^{(n)}$ are the Schmidt coefficients of χ_n), and therefore, according to Eq. (D2),

$$\liminf_{n\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) = \liminf_{n\to\infty} \max_{k\in[n]} \left\{ \left\| \mathbf{p}^{(n)} \right\|_{(k)} - \left\| \mathbf{p}^{(n)} \right\|_{(a_k)} - \frac{b_k}{m} p_{a_k+1}^{(n)} \right\}$$

$$= \liminf_{n\to\infty} \max_{k\in\{\,N,\dots,n\,\}} \left\{ \sum_{x=a_k+1}^{k} p_x^{(n)} - \frac{b_k}{m} p_{a_k+1}^{(n)} \right\} \qquad \text{(H2)}$$

$$= \liminf_{n\to\infty} \max_{k\in\{\,N,\dots,n\,\}} \left\{ \sum_{x=a_k+1}^{k} p_x^{(n)} \right\},$$

where $a_k = \lfloor k/m \rfloor$ and $b_k = k - ma_k$. The Schmidt coefficients are by definition non-increasing, i.e., obtained by reordering $\{\, f(x)/F_n \,\}_{x\in\mathbb{N}}$. Since there are at most N natural numbers x that do not satisfy the condition $l - \varepsilon < f(x) < l + \varepsilon$, there are at most N Schmidt coefficients that do not satisfy $\frac{l-\varepsilon}{F_n} < p_x^{(n)} < \frac{l+\varepsilon}{F_n}$. We call the set of indices corresponding to these Schmidt coefficients A (thus $|A| \leq N$) and observe that for any $a, b \in \{\, N, \dots, n \,\}$ such that $a \leq b$,

$$\sum_{x=a}^{b} p_x^{(n)} = \sum_{x\in\{\,a,\dots,b\,\}\setminus A} p_x^{(n)} + \sum_{x\in\{\,a,\dots,b\,\}\cap A} p_x^{(n)}$$

$$= \sum_{x\in\{\,a,\dots,b\,\}\setminus A} p_x^{(n)} + \sum_{x\in\{\,a,\dots,b\,\}\cap A} \frac{l}{F_n} + \sum_{x\in\{\,a,\dots,b\,\}\cap A} \left(p_x^{(n)} - \frac{l}{F_n} \right). \qquad \text{(H3)}$$

This implies that

$$\sum_{x=a}^{b} \frac{(l-\varepsilon)}{F_n} + \sum_{x\in\{\,a,\dots,b\,\}\cap A} \left(p_x^{(n)} - \frac{l}{F_n} \right) < \sum_{x=a}^{b} p_x^{(n)}$$

$$< \sum_{x=a}^{b} \frac{(l+\varepsilon)}{F_n} + \sum_{x\in\{\,a,\dots,b\,\}\cap A} \left(p_x^{(n)} - \frac{l}{F_n} \right), \qquad \text{(H4)}$$

and therefore

$$\liminf_{n\to\infty} \max_{k\in\{\,N,\dots,n\,\}} \left\{ \sum_{x=a_k+1}^{k} \frac{l-\varepsilon}{F_n} \right\} \leq \liminf_{n\to\infty} \max_{k\in\{\,N,\dots,n\,\}} \left\{ \sum_{x=a_k+1}^{k} p_x^{(n)} \right\}$$

$$\leq \liminf_{n\to\infty} \max_{k\in\{\,N,\dots,n\,\}} \left\{ \sum_{x=a_k+1}^{k} \frac{l+\varepsilon}{F_n} \right\}, \qquad \text{(H5)}$$

which implies

$$(l - \varepsilon) \liminf_{n \to \infty} \frac{n - \lfloor n/m \rfloor}{F_n} \le \liminf_{n \to \infty} \max_{k \in \{ N, \dots, n \}} \left\{ \sum_{x=a_k+1}^{k} p_x^{(n)} \right\} \tag{H6}$$
$$\le (l + \varepsilon) \liminf_{n \to \infty} \frac{n - \lfloor n/m \rfloor}{F_n}.$$

Observing that $\sum_{x=1}^{N} f(x) + (l - \varepsilon)(n - N) < F_n < \sum_{x=1}^{N} f(x) + (l + \varepsilon)(n - N)$, we obtain

$$\frac{l - \varepsilon}{l + \varepsilon} \left(1 - \frac{1}{m} \right) \le \liminf_{n \to \infty} \max_{k \in \{ N, \dots, n \}} \left\{ \sum_{x=a_k+1}^{k} p_x^{(n)} \right\} \le \frac{l + \varepsilon}{l - \varepsilon} \left(1 - \frac{1}{m} \right). \tag{H7}$$

Since Eq. (H7) holds for every $\varepsilon > 0$, we conclude that

$$\liminf_{n \to \infty} \max_{k \in \{ N, \dots, n \}} \left\{ \sum_{x=a_k+1}^{k} p_x^{(n)} \right\} = 1 - \frac{1}{m}. \tag{H8}$$

Inserting this result into Eq. (H2), we have

$$\liminf_{x \to \infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) = 1 - \frac{1}{m}. \tag{H9}$$

The function $\tilde{f}$ is a rescaling of $f(x) = x^0$, which we already studied in Section F. This implies that the family $\{ \tilde{\chi}_n \}_{n \in \mathbb{N}}$ is equal to the family $\left\{ \chi_n^{(0)} \right\}_{n \in \mathbb{N}}$ and

$$\liminf_{x \to \infty} T_\star(\tilde{\chi}_n \to \tilde{\chi}_n \otimes \Phi_m) = \liminf_{x \to \infty} T_\star(\chi_n^{(0)} \to \chi_n^{(0)} \otimes \Phi_m)$$
$$= 1 - \frac{1}{m} \tag{H10}$$
$$= \liminf_{x \to \infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m).$$

This proves that neither f nor $\tilde{f}$ corresponds to families with universally embezzling subfamilies (and are therefore also not universally embezzling themselves).

We consider now the case $\lim_{x \to \infty} f(x) = 0$, thus f is asymptotically non-increasing. Let N be such that f is non-increasing on (N, ∞) and let $a = \min_{x \in \{ 1, \dots, N \}} f(x)$. Since f is positive, $a > 0$. Let $M > N$ be such that $f(x) < a$ for every $x \ge M$ (such M exists because $\lim_{x \to \infty} f(x) = 0$). In this case, we define $\tilde{f}$ as

$$\tilde{f}(x) = \begin{cases} f(M) & \text{if } x \le M, \\ f(x) & \text{if } x > M. \end{cases} \tag{H11}$$

Clearly, the family $\{ \tilde{\chi}_n \}_{n \in \mathbb{N}}$ associated to it is regular and $\lim_{x \to \infty} f(x)/\tilde{f}(x) = 1$. Furthermore, the ordered Schmidt coefficients satisfy for all $x \ge M$

$$p_x^{(n)} = \frac{f(x)}{F_n} = \frac{\tilde{f}(x)}{F_n} = \tilde{p}_x^{(n)} \frac{\tilde{F}_n}{F_n}. \tag{H12}$$

If F_n converges, $\tilde{F}_n$ converges too and by Corollary 7 neither $\{ \chi_n \}_{n \in \mathbb{N}}$ nor $\{ \tilde{\chi}_n \}_{n \in \mathbb{N}}$ are universal embezzling families (and do not contain universal embezzling subfamilies). If instead F_n diverges, so does $\tilde{F}_n$ and

$$\lim_{n \to \infty} \frac{\tilde{F}_n}{F_n} = \lim_{n \to \infty} \frac{\sum_{x=1}^{n} \tilde{f}(x)}{\sum_{x=1}^{n} f(x)} = \lim_{n \to \infty} \frac{\sum_{x=M}^{n} \tilde{f}(x)}{\sum_{x=M}^{n} f(x)} = 1. \tag{H13}$$

Using Eq. (D2) and Eq. (H12), we obtain

$$
\begin{aligned}
&\liminf_{n\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) \\
&= \liminf_{n\to\infty} \max_{k\in[n]} \left\{ \left\| \mathbf{p}^{(n)} \right\|_{(k)} - \left\| \mathbf{p}^{(n)} \right\|_{(a_k)} - \frac{b_k}{m} p^{(n)}_{a_k+1} \right\} \\
&= \liminf_{n\to\infty} \max_{k\in\{\,mM,\ldots,n\,\}} \left\{ \left\| \mathbf{p}^{(n)} \right\|_{(k)} - \left\| \mathbf{p}^{(n)} \right\|_{(a_k)} - \frac{b_k}{m} p^{(n)}_{a_k+1} \right\} \\
&= \liminf_{n\to\infty} \frac{\tilde{F}_n}{F_n} \max_{k\in\{\,mM,\ldots,n\,\}} \left\{ \left\| \tilde{\mathbf{p}}^{(n)} \right\|_{(k)} - \left\| \tilde{\mathbf{p}}^{(n)} \right\|_{(a_k)} - \frac{b_k}{m} \tilde{p}^{(n)}_{a_k+1} \right\} \\
&= \liminf_{n\to\infty} \max_{k\in[n]} \left\{ \left\| \tilde{\mathbf{p}}^{(n)} \right\|_{(k)} - \left\| \tilde{\mathbf{p}}^{(n)} \right\|_{(a_k)} - \frac{b_k}{m} \tilde{p}^{(n)}_{a_k+1} \right\} \\
&= \liminf_{n\to\infty} T_\star(\tilde{\chi}_n \to \tilde{\chi}_n \otimes \Phi_m),
\end{aligned}
\tag{H14}
$$

where again $a_k = \lfloor k/m \rfloor$ and $b_k = k - ma_k$. Also in this case, we therefore proved that $\{\,\chi_n\,\}_{n\in\mathbb{N}}$ is a universal embezzling family if and only if $\{\,\tilde{\chi}_n\,\}_{n\in\mathbb{N}}$ is a universal embezzling family (and the same holds for subfamilies).

Lastly, we consider the case when $\lim_{x\to\infty} f(x) = \infty$, and f is asymptotically non-decreasing. Analogously to the previous case, let N be such that f is non-decreasing for $x \in (N,\infty)$. Let a be the maximum of $f(x)$ for $x \in \{\,1,\ldots,N\,\}$, and let $M > N$ be such that $f(x) > a$ for $x \geq M$. Also here, we define $\tilde{f}$ via

$$
\tilde{f} = \begin{cases} f(M) & \text{if } x \leq M, \\ f(x) & \text{if } x > M. \end{cases}
\tag{H15}
$$

The family of states $\{\,\tilde{\chi}_n\,\}_{n\in\mathbb{N}}$ is regular, $\lim_{x\to\infty} \tilde{f}(x)/f(x) = 1$, and $\lim_{x\to\infty} \tilde{F}_n/F_n = 1$. The Schmidt coefficients associated to χ_n and $\tilde{\chi}_n$ are related by

$$
p^{(n)}_x = \frac{f(n+1-x)}{F_n} = \frac{\tilde{f}(n+1-x)}{F_n} = \tilde{p}^{(n)}_x \frac{\tilde{F}_n}{F_n} \quad \forall x \leq n - M.
\tag{H16}
$$

Using Eq. (D2) again and the relation between Schmidt coefficients derived in Eq. (H16) we obtain

$$
\begin{aligned}
&\liminf_{n\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) \\
&= \liminf_{n\to\infty} \max_{k\in[n]} \left\{ \left\| \mathbf{p}^{(n)} \right\|_{(k)} - \left\| \mathbf{p}^{(n)} \right\|_{(a_k)} - \frac{b_k}{m} p^{(n)}_{a_k+1} \right\} \\
&= \liminf_{n\to\infty} \max_{k\in[n-M]} \left\{ \left\| \mathbf{p}^{(n)} \right\|_{(k)} - \left\| \mathbf{p}^{(n)} \right\|_{(a_k)} - \frac{b_k}{m} p^{(n)}_{a_k+1} \right\} \\
&= \liminf_{n\to\infty} \frac{\tilde{F}_n}{F_n} \max_{k\in[n-M]} \left\{ \left\| \tilde{\mathbf{p}}^{(n)} \right\|_{(k)} - \left\| \tilde{\mathbf{p}}^{(n)} \right\|_{(a_k)} - \frac{b_k}{m} \tilde{p}^{(n)}_{a_k+1} \right\} \\
&= \liminf_{n\to\infty} \max_{k\in[n]} \left\{ \left\| \tilde{\mathbf{p}}^{(n)} \right\|_{(k)} - \left\| \tilde{\mathbf{p}}^{(n)} \right\|_{(a_k)} - \frac{b_k}{m} \tilde{p}^{(n)}_{a_k+1} \right\} \\
&= \liminf_{n\to\infty} T_\star(\tilde{\chi}_n \to \tilde{\chi}_n \otimes \Phi_m).
\end{aligned}
\tag{H17}
$$

This proves the theorem. $\qquad\square$

Thanks to Theorem H.2, Theorem G.2 also holds for asymptotically regular families.

Corollary H.3. *Let f be a positive asymptotically non-increasing function such that $f(x)/x^\alpha$ is asymptotically monotonic for all $\alpha \in \mathbb{R}$ and let $\{\chi_n\}_{n\in\mathbb{N}}$ be the asymptotically regular family of states associated to f (see Definition H.1). Then $\{\chi_n\}_{n\in\mathbb{N}}$ is a universal embezzling family if and only if $f \in \omega(x^{-1-\varepsilon}) \cap o(x^{-1+\varepsilon})$ for all $\varepsilon > 0$ and $\sum_{x=1}^{\infty} f(x) = \infty$. Furthermore, if $f \notin \omega(x^{-1-\varepsilon}) \cap o(x^{-1+\varepsilon})$ for at least one $\varepsilon > 0$, then $\left\{\chi_{n_j}\right\}_{j\in\mathbb{N}}$, where $\{n_j\}_{j\in\mathbb{N}}$ is any sequence of natural numbers, is not a universal embezzling family.*

Proof. Combine Theorem H.2 and Theorem G.2. $\qquad\square$

For the same reasons, Proposition G.3 and Proposition G.4 also hold for asymptotically regular families.

Corollary H.4. *Let f be a positive asymptotically non-decreasing function such that $f(x)/x^\alpha$ is asymptotically non-increasing for at least one $\alpha > 0$. Then the asymptotically regular family of states $\{\chi_n\}_{n\in\mathbb{N}}$ associated to f is not a universal embezzling family. Furthermore, there are no sequences $\{n_j\}_{j\in\mathbb{N}}$ such that $\left\{\chi_{n_j}\right\}_{j\in\mathbb{N}}$ is a universal embezzling family.*

Corollary H.5. *Let f be a positive asymptotically non-decreasing function such that $f(x)/e^{kx}$ is asymptotically non-decreasing for at least one $k > 0$. Then the asymptotically regular family of states $\{\chi_n\}_{n\in\mathbb{N}}$ associated with f is not a universal embezzling family. Furthermore, there are no sequences $\{n_j\}_{j\in\mathbb{N}}$ such that $\left\{\chi_{n_j}\right\}_{j\in\mathbb{N}}$ is a universal embezzling family.*