Phase transition in stabilizer entropy and efficient purity estimation

Lorenzo Leone,[1,*] Salvatore F. E. Oliviero,[1,†] Gianluca Esposito,[2,‡] and Alioscia Hamma[2,3,§]
[1]*Physics Department, University of Massachusetts Boston, Boston, Massachusetts 02125, USA*
[2]*Dipartimento di Fisica "Ettore Pancini," Università degli Studi di Napoli Federico II, Via Cintia 80126, Napoli, Italy*
[3]*INFN, Sezione di Napoli, 80126 Napoli, Italy*

(Received 23 February 2023; accepted 5 January 2024; published 1 March 2024)

Stabilizer entropy (SE) quantifies the spread of a state in the basis of Pauli operators. It is a computationally tractable measure of nonstabilizerness and thus a useful resource for quantum computation. SE can be moved around a quantum system, effectively purifying a subsystem from its complex features. We show that there is a phase transition in the residual subsystem SE as a function of the density of non-Clifford resources. This phase transition has important operational consequences: it marks the onset of a subsystem purity estimation protocol that requires $\mathrm{poly}(n)\exp(t)$ many queries to a circuit containing t non-Clifford gates that prepares the state from a stabilizer state. Then, for $t = O(\log_2 n)$, it estimates the purity with polynomial resources, and, for highly entangled states, attains an exponential speed-up over the known state-of-the-art algorithms.

DOI: 10.1103/PhysRevA.109.032403

I. INTRODUCTION

Quantum information processing promises an advantage over its classical counterpart [1–7]. Since the inception of this field [8], there has been an extensive theoretical investigation as to what ingredients quantum computation possesses such that it is intrinsically computationally more powerful than classical computation.

The two resources that set quantum computers apart are entanglement [9–12] and nonstabilizerness [8,13]. Without either of them, quantum computers cannot perform any advantageous algorithm over classical devices [1,2,14]. In particular, nonstabilizerness measures how many universal gates one can distill from a given quantum state [15–19] and the cost of simulating a quantum state on a classical computer. Indeed, while stabilizer states (the orbit of the Clifford group [20]) can be simulated classically in polynomial time [8], the cost of the simulation scales exponentially in the number of nonstabilizer resources [21,22], i.e., unitary gates outside the Clifford group [8].

At the same time, many information processing tasks in quantum computing become inefficient exactly because of the conspiracy of entanglement and nonstabilizerness. Examples of this kind are state certification [23], disentangling [24–29], or unscrambling [30–32]. In particular, while purity estimation is a resource-intensive task for universal states, it can be achieved efficiently for stabilizer states.

It seems then that quantum computation is plagued by a so-called catch-22 dilemma: on the one hand, stabilizer information can be efficiently processed, but for the same reason it is useless for a fruitful quantum computation. On the other hand, the combination of entanglement and nonstabilizerness, which makes quantum technology powerful, hinders the efficiency of measurement tasks. Given that, the question posed by this paper is the following: can we leverage the efficiency of information processing offered by the stabilizer formalism for nonstabilizer states?

Recently, a novel measure of nonstabilizerness was introduced as *stabilizer entropy* (SE) [33]. Stabilizer states have zero stabilizer entropy, whereas nonstabilizer states (those that are computationally useful) exhibit a nonvanishing stabilizer entropy. Unlike other measures [15,34,35], it is computable (though expensive) and experimentally measurable [36–38]. SE is also involved in the onset of universal, complex patterns of entanglement [28,29], quantum chaos [39–42], complexity in the wave function of quantum many-body systems [43,44], and the decoding algorithms from the Hawking radiation from old black holes [30–32]. In the context of operator spreading, it is akin to the string entropy [45].

As we said, when states possess SE, measurement tasks tend to become inefficient. However, the intriguing aspect of entropy is that it can be transferred from one subsystem to another without altering the total entropy, and thus the total computational power of the system. This parallels the behavior of a Carnot refrigerator that effectively reduces entropy in a system by transferring it to the environment, all while keeping the entropy of the universe unchanged.

In this paper, drawing inspiration from this thermodynamic analogy, we show a general scheme of how to push SE out of a subsystem with Clifford operations, effectively *cooling* the subsystem down from its complex features while preserving the total SE, and explore the consequences and implications of this approach. The two main results of this paper are the following.

(i) There is a phase transition in SE driven by the competition between the creation and spreading of non-Clifford resources versus their localization and erasure.

*lorenzo.leone001@umb.edu
†s.oliviero001@umb.edu
‡gianluca.esposito32@studenti.unina.it
§alioscia.hamma@unina.it

(ii) The localized phase (when a subsystem is successfully cleansed of its nonstabilizerness quantified by SE) allows for a purity estimation algorithm that, for some cases of interest, obtains an exponential speed-up over all the state-of-the-art known algorithms [46–52]. In practice, this result is obtained by constructing a stabilizer state whose subsystem purity is shown to bound the purity of the desired state.

II. SETUP

Consider a system of n qubits with Hilbert space $\mathcal{H} \simeq \mathbb{C}^{\otimes 2n} \simeq \mathcal{H}_E \otimes \mathcal{H}_F$ with dimensions $d_X = 2^{n_X}$ with $(X = E, F)$ and $n = n_E + n_F$. Let $|\mathbf{0}\rangle \equiv |0\rangle^{\otimes n}$. To every pure density operator ψ on $\mathcal{H}$ one can associate a probability distribution P_ψ through its decomposition $\psi = d^{-1}\sum_{P\in\mathbb{P}} \mathrm{tr}(P\psi)P$ in Pauli operators $P \in \mathbb{P}$ by $P_\psi = d^{-1}\mathrm{tr}^2(P\psi)$. Regardless of the purity of ψ, its stabilizer purity is defined as $\mathrm{SP}(\psi) := \sum_P P_\psi^2$. On the other hand, the purity of the state ψ is given by $\mathrm{Pur}(\psi) = \mathrm{tr}(\psi^2)$. Defining the ratio $w(\psi) := d\,\mathrm{SP}(\psi)/\mathrm{Pur}(\psi)$, the 2-Rényi SE is given by $M(\psi) = -\log_2 w$ [33], while the linear SE is defined as $M_{\mathrm{lin}} = 1 - w$. Throughout the paper, we define *stabilizer states* [53]. as those for which $M(\psi) = 0$.

Through standard techniques one can write these purities as $\mathrm{SP}(\psi) = d\,\mathrm{tr}(Q\psi^{\otimes 4})$ and $\mathrm{Pur}(\psi) = \mathrm{tr}(T\psi^{\otimes 2})$ where T is the SWAP operator and $Q = d^{-2}\sum_P P^{\otimes 4}$ is the projector onto the stabilizer code [20]. Given the bipartition $E|F$ defined above, one can define the SE associated with the subsystem X as $M_X(\psi) \equiv M(\psi_X)$ where the partial states are defined as $\psi_E \equiv \mathrm{tr}_F\psi$ and $\psi_F \equiv \mathrm{tr}_E\psi$. In terms of the Q, T operators the w of the partial state reads $w(\psi_X) = d_X\mathrm{tr}(Q_X\psi^{\otimes 4})/\mathrm{tr}(T_X\psi^{\otimes 2})$ where now Q_X, T_X are the Q and SWAP operator on the subsystem $X = E, F$. Note also that if ψ is a stabilizer state then $M_X(\psi) = 0$ for every subsystem X [33]. However, whether the partial trace is, in general, a SE-non-increasing map is an open question.

In the following, we are interested in the SE of states ψ_t parametrized by a number t of non-Clifford resources. To this end, we consider outputs $\psi_t \equiv C_t|\mathbf{0}\rangle\langle\mathbf{0}|C_t^\dagger$ of t-doped Clifford circuits C_t, that is, Clifford circuits in which t non-Clifford gates have been injected, say T gates [27,41]. Now, since the Clifford group is very efficient in entangling, the states ψ_t are typically highly entangled [20,24,26,54]. As a result if $n_F \ll n_E$, the partial state ψ_{tF} is very close to the maximally mixed state, which is a stabilizer state with SE equal zero. Absent from this picture is the characterization of stabilizer entropy behavior within partitions when $n_F \ll n_E$. To gain insights into this behavior, it is often necessary to conduct averaging over the Clifford orbit. In pursuit of this objective, we introduce the following lemma, which will prove beneficial throughout this paper.

Lemma 1. Let ψ be a pure quantum state, $\psi^C \equiv C\psi C^\dagger$ its Clifford orbit, and $\psi_E^C \equiv \mathrm{tr}_F(C\psi C^\dagger)$. Then, for $n_f = n\mathfrak{f}$ and $0 < \mathfrak{f} < 1/2$,

$$\mathbb{E}_C \frac{\mathrm{SP}\left(\psi_E^C\right)}{\mathrm{Pur}\left(\psi_E^C\right)} = \frac{\mathbb{E}_C \mathrm{SP}\left(\psi_E^C\right)}{\mathbb{E}_C\,\mathrm{Pur}\left(\psi_E^C\right)}\left[1 + O\left(2^{-n\frac{1-2\mathfrak{f}}{2}}\right)\right]. \tag{1}$$

Proof. Expanding $\mathrm{Pur}(\psi_E^C)$ around the average $\mathbb{E}_C\,\mathrm{Pur}(\psi_E^C)$ one has

$$\mathrm{Pur}\left(\psi_E^C\right) = \mathbb{E}_C\,\mathrm{Pur}\left(\psi_E^C\right) + \Delta_C\,\mathrm{Pur}\left(\psi_E^C\right), \tag{2}$$

where $\Delta_C\,\mathrm{Pur}(\psi_E^C) \equiv \sqrt{\mathbb{E}_C\,\mathrm{Pur}^2(\psi_E^C) - [\mathbb{E}_C\,\mathrm{Pur}(\psi_E^C)]^2}$, and thus

$$\mathbb{E}_C \frac{\mathrm{SP}\left(\psi_E^C\right)}{\mathrm{Pur}\left(\psi_E^C\right)} = \frac{\mathbb{E}_C \mathrm{SP}\left(\psi_E^C\right)}{\mathbb{E}_C\,\mathrm{Pur}\left(\psi_E^C\right)}\left(1 - \frac{\Delta_C\,\mathrm{Pur}\left(\psi_E^C\right)}{\mathbb{E}_C\,\mathrm{Pur}\left(\psi_E^C\right)}\right), \tag{3}$$

i.e., one has that the average of a ratio is equal to the ratio of the averages up to a relative error that is the relative error of the average purity, i.e., the ratio of the purity fluctuations to the average purity. We can use the fundamental result from [41] that states that the relative error over the Clifford orbit is small as long as the bipartition is not the exact half bipartition, see Eq. (52) of [41]. For the purpose of this paper consider $n_F = \mathfrak{f}n$ and $\mathfrak{f} < 1/2$, one has

$$\frac{\Delta_C\,\mathrm{Pur}\left(\psi_E^C\right)}{\mathbb{E}_C\,\mathrm{Pur}\left(\psi_E^C\right)} = O\left(2^{-n\frac{1-2\mathfrak{f}}{2}}\right). \tag{4}$$

■

Thanks to this lemma one can then write, up to an exponentially small relative error

$$\mathbb{E}_C\left[w\left(\psi_E^C\right)\right] = \frac{d_E \mathbb{E}_C\,\mathrm{SP}(\psi_E)}{\mathbb{E}_C\ \mathrm{Pur}\left(\psi_E^C\right)} \tag{5}$$

and in the same way, one can write

$$\mathbb{E}_C\left[w\left(\psi_F^C\right)\right] = \frac{d_F \mathbb{E}_C\,\mathrm{SP}(\psi_F)}{\mathbb{E}_C\ \mathrm{Pur}\left(\psi_F^C\right)}. \tag{6}$$

From the above lemma, one can indeed show that the average $\mathbb{E}_C$ (over the Clifford group) SE in F is very small while it is all contained in the subsystem E.

Proposition 1. Consider $n_F/n < 1$. Let ψ a pure state and denote $\psi^C \equiv C\psi C^\dagger$, for C a Clifford circuit. The average over the Clifford orbit of the partial SEs of subsystems E, F is given by

$$\begin{aligned}\mathbb{E}_C\left[M_{\mathrm{lin}}\left(\psi_E^C\right)\right] &= M_{\mathrm{lin}}(\psi) + O\left(\frac{d_F}{d_E}\right),\\ \mathbb{E}_C\left[M_{\mathrm{lin}}\left(\psi_F^C\right)\right] &= O\left(\frac{d_F}{d_E}\right),\end{aligned} \tag{7}$$

for large n. See Appendix B for the proof.

In the paper, we frequently make use of the big-O notation, see Appendix A for a brief review.

The above formulas show how the SE is all in the larger system. As a corollary (by plugging the above formulas), on average over the Clifford orbit of ψ, the partial trace preserves SE in the sense that $M(\psi) \simeq M_E(\psi) + M_F(\psi)$.

III. CLEANSING ALGORITHM

As we have seen, typically all the SE is contained in the greater of the two subsystems, namely, the subsystem E. As such, quantum information contained in the state ψ_E cannot be efficiently processed [21]: it contains almost all of the quantum complexity induced by the circuit C_t. One asks

whether it is possible to *cleanse* E from this complexity. If one could do that, one would be able to manipulate E efficiently by means of the stabilizer formalism [55]. Since the SE is an entropy, one wonders if one could toss it in the other subsystem F by means of a suitable protocol. Here, we set the problem up in the following way: can we find a quantum map $\mathcal{E}(\cdot) := \mathrm{tr}_Y W(\cdot)W^\dagger$ with W a Clifford unitary and Y a subset of n_Y qubits $Y \subseteq F$, such that $M_E[\mathcal{E}(\psi_t)] = 0$?

To this end, we utilize a fundamental result obtained in [31]: given a t-doped circuit C_t the so-called *Clifford Completion algorithm* can learn [by $\mathrm{poly}(n)\exp(t)$ query accesses to a unknown C_t] a Clifford operator D called *diagonalizer* such that $C_t = D^\dagger c_t D V$, where c_t is a t-doped Clifford circuit acting on a subsystem with only t qubits (for $t \leqslant n$) and V is another suitable Clifford unitary operator. This result ensures that $t = O(\log_2 n)$ the operator D will (with negligibly small failure probability) be found in polynomial time.

Now, consider the permutation operator T_π on n qubits, namely $T_\pi|x\rangle = |\pi(x)\rangle$ with $\pi \in S_n$ the symmetric group of n objects. Note that the T_π's also belong to the Clifford group [56]. Then, one can choose a suitable permutation π_Y such that the dressed operator $T_{\pi_Y} c_t T_{\pi_Y}^\dagger \equiv c_t^Y$ acts nontrivially only on any desired subsystem $Y \subset E \cup F$, where $n_Y = t$. In particular, one can choose $Y \subset F$ if $n_Y \leqslant n_F$. As a result, one has $T_{\pi_Y} D|\psi_t\rangle = T_{\pi_Y} c_t T_{\pi_Y}^\dagger |\Phi\rangle = c_t^Y|\Phi\rangle$ with $|\Phi\rangle := T_{\pi_Y} DV|\mathbf{0}\rangle$ being a stabilizer state (with density operator $\Phi \equiv |\Phi\rangle\langle\Phi|$). Then, by picking $W \equiv T_{\pi_Y} D$, we define the Clifford map $\mathcal{E}(\cdot) := \mathrm{tr}_Y W(\cdot)W^\dagger$ that first localizes the non-Clifford resources in Y and then erases them by tracing Y out. We can now prove the following.

Proposition 2. For $n_Y \leqslant n_F$, the map $\mathcal{E}$ moves the t non-Clifford gates in the subsystem $Y \subset F$, and by tracing out Y makes the SE on E zero, i.e., $M_E[\mathcal{E}(\psi_t)] = 0$.

Proof. Start with

$$M_E[\mathcal{E}(\psi_t)] = -\log_2 d_E \frac{\mathrm{tr}[Q_E \mathcal{E}(\psi_t)^{\otimes 4}]}{\mathrm{tr}[T_E \mathcal{E}(\psi_t)^{\otimes 2}]}, \tag{8}$$

and recall that $\mathcal{E}(\psi_t) = \mathrm{tr}_Y(c_t^Y \Phi c_t^{Y\dagger})$. Now, since $[c_t^{Y\otimes 4}, Q_E] = [c_t^{Y\otimes 2}, T_E] = 0$, one obtains $\mathrm{tr}[Q_E \mathcal{E}(\psi_t)^{\otimes 4}] = \mathrm{tr}_E(Q_E \Phi_E^{\otimes 4})$ and $\mathrm{tr}[T_E \mathcal{E}(\psi_t)^{\otimes 2}] = \mathrm{tr}_E(T_E \Phi_E^{\otimes 2})$ from which one gets $M_E[\mathcal{E}(\psi_t)] = M_E(\Phi) = 0$, where the last equality follows from Φ being a stabilizer state [33].

IV. PHASE TRANSITION IN SE

We now show that doped Clifford circuits feature a phase transition in SE due to the competition between a term that creates and spreads SE and a term that localizes and then erases it. The first term is the quantum circuit C_t. As we saw in the previous section, this unitary operator can be written as $C_t = D^\dagger c_t D V$. The unitary c_t does insert a number t of non-Clifford gates in a t-qubits subsystem, and then the term D spreads them around the entire system. In the large n limit, we can define $\mathfrak{t} := t/n$ the density of non-Clifford gates and $\mathcal{C}_\mathfrak{t}$ its adjoint action. The spreading strength of this channel is given by the depth of C_t, while its SE strength is given by the density $\mathfrak{t}$. The map $\mathcal{E}$, on the other hand, first localizes the SE in the subsystem $Y \subset F$ and then erases it by entangling and tracing out. Thanks to Proposition 2, its localizing strength is given by the density of qubits belonging to F, namely, $\mathfrak{f} := n_F/n$. Altogether, we consider the map composing the two terms, namely, $\mathcal{E} \circ \mathcal{C}_\mathfrak{t}$ and study the behavior of the SE induced by such a map. In the limit of $\mathfrak{t}/\mathfrak{f} \ll 1$, we expect the map to leave the subsystem E clear of SE. However, for $\mathfrak{t}/\mathfrak{f} \gg 1$, all the SE should be intact in E.

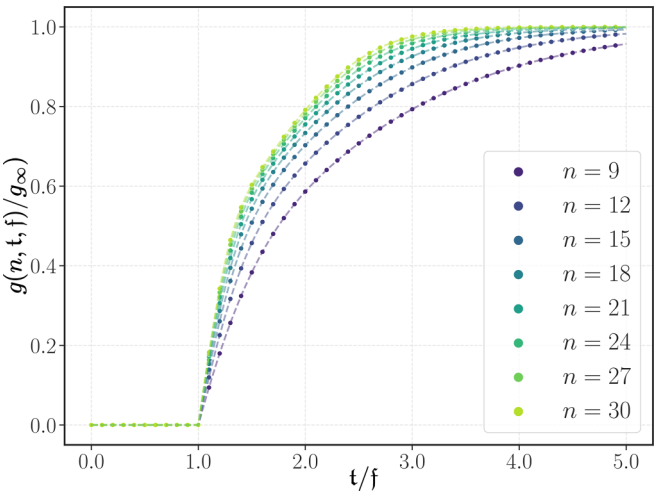

FIG. 1. Plot of the ratio between $g(n,\mathfrak{t},\mathfrak{f})$ and g_∞, which is a lower bound for the localized SE power of the map $\mathcal{E} \circ \mathcal{C}_\mathfrak{t}$, as a function of the ratio of the density $\mathfrak{t}$ of non-Clifford gates to the density $\mathfrak{f}$ of qubits belonging to F. We set $\mathfrak{f} = \frac{1}{3}$. The critical point for the phase transition is $\mathfrak{t}/\mathfrak{f} = 1$, with critical index one.

We compute the localizing SE power of the map $\mathcal{E} \circ \mathcal{C}_\mathfrak{t}$ by averaging the value of $M_E[\mathcal{E} \circ \mathcal{C}_\mathfrak{t}(\omega)]$ over all the maps $\mathcal{E} \circ \mathcal{C}_\mathfrak{t}$ and all the (pure) stabilizer input states ω. We denote such an average $\mathbb{E}$ (see Appendix C). For $\mathfrak{t}/\mathfrak{f} \leqslant 1$ we obtain, in virtue of Proposition 2, $\mathbb{E}\,[M_E(\mathcal{E} \circ \mathcal{C}_\mathfrak{t}[\omega])] = 0$. This is the localized phase, where the localizing power of the map $\mathcal{E}$ prevails. However, for $\mathfrak{t}/\mathfrak{f} \geqslant 1$, direct computation of the average yields, in the delocalized phase,

$$\mathbb{E}[M_E(\mathcal{E} \circ \mathcal{C}_\mathfrak{t}[\omega])] \geqslant g(n,\mathfrak{t},\mathfrak{f}). \tag{9}$$

See Appendix C for details and the explicit expression for $g(n,\mathfrak{t},\mathfrak{f})$. For large n, one has

$$g_\infty \equiv \lim_{t\to\infty} g(n,\mathfrak{t},\mathfrak{f}) \simeq n(1-2\mathfrak{f}). \tag{10}$$

In Fig. 1, we plot $g(n,\mathfrak{t},\mathfrak{f})/g_\infty$ (for $\mathfrak{f} = 1/3$). In the neighborhood of the critical value $\mathfrak{t}/\mathfrak{f} = 1$ this ratio behaves as $g(n,\mathfrak{t},\mathfrak{f})/g_\infty \simeq \frac{\mathfrak{f}}{1-\mathfrak{f}}(\mathfrak{t}/\mathfrak{f} - 1)$, which shows a critical index one, see Fig. 1.

V. EFFICIENT PURITY ESTIMATION

The phase transition described above has relevant operational applications in terms of quantum information processing. We see that in the localized phase $\mathfrak{t}/\mathfrak{f} \leqslant 1$, the SE can be cleansed from the subsystem E, making the subsystem E manipulable by means of the stabilizer formalism. We now show that in this phase it is possible to probe the bipartite entanglement in a way that, for cases of interest, gains an exponential speed-up over the state-of-the-art algorithms in the literature [46,52].

The best-known way to evaluate the purity of a quantum state within an error ϵ is the SWAP test, which requires a number of resources scaling as $O(\epsilon^{-2})$. However, typical states possess a subsystem purity $\Theta(e^{-\tilde{\beta} n_F})$, the so-called *volume law* scenario. In this case, to evaluate the purity, one needs to resolve an exponentially small error and therefore exponential (in n_F) resources, which, since $n_F = \mathfrak{f} n$, is $\Theta(e^{2\beta n})$, with $\beta \equiv \tilde{\beta}\mathfrak{f}$.

If the purity one wants to estimate scales polynomially, that is, $\mathrm{Pur}(\psi_E) = \Omega[\mathrm{poly}^{-1}(n)]$, one will need a polynomial number N_{shot} of measurements to resolve the quantity. Notice that one would not know beforehand what is the number N_{shot} necessary. In practice, one sets a number of experiments N_{shot} to obtain upper bound thresholds $\mathrm{Pur}(\psi_E) \leqslant O(N_{\mathrm{shot}}^{-1/2})$. If $\mathrm{poly}(n)$ is a large polynomial, one would, in practice, be forced to halt the procedure without knowing how tight is the bound. In the worst-case scenario (which is also typical), the purity to evaluate is exponentially small and the estimation will always be exponentially costly.

We want to show that the cleansing algorithm can give an exponential advantage over the known protocols. The intuition is that if one can cleanse the SE from ψ_E, one would obtain a stabilizer state whose purity can be evaluated with polynomial resources [57].

Let us start with some technical preliminaries. Consider a state initialized in $|\mathbf{0}\rangle$ and be $|\psi_t\rangle = C_t|\mathbf{0}\rangle$ the output of a t-doped Clifford circuit. Its marginal state to E will be denoted by $\psi_E = \mathrm{tr}_F(\psi_t)$. This is the state whose purity we want to evaluate. In the localized phase of the cleansing algorithm, we know that the output state $\mathcal{E}(\psi_t)$ is a stabilizer state of $E \cup F \setminus Y$ for $Y \subset F$. Its purity would be easy to evaluate, but it is not directly related to the purity of the original state ψ_E because of the action of W. We now show that we can manipulate the *cleansed state* $\mathcal{E}(\psi_t)$ and construct a stabilizer state ρ whose purity bounds the purity of the desired state ψ_E.

Starting from the cleansed state $\mathcal{E}(\psi_t)$, we first append the maximally mixed state $d_Y^{-1} I_Y$ on Y and then we act with the diagonalizer $W^\dagger$ back, obtaining the state $\rho := W^\dagger[\mathcal{E}(\psi_t) \otimes d_Y^{-1} I_Y] W$. Let ρ_X with $X = E, F$ be its marginals and notice that ρ, ρ_E, ρ_F are stabilizer states. What we have effectively done is to re-entangle the state $\mathcal{E}(\psi_t)$ so that it gives a bound to the purity of ψ_E.

After these preliminaries, we are ready to establish our protocol. Set $t = O(\log_2 n)$. Utilizing the cleansing algorithm, we first prepare the stabilizer state ρ by learning the diagonalizer W, which requires $O[\mathrm{poly}(n)]$ resources. Then, since ρ is a stabilizer state, by means of $O(n^3)$ shot measurements, we evaluate $\mathrm{Pur}(\rho_X)$ with no error [57]. Indeed it is sufficient to first learn the stabilizer group S associated to ρ using the algorithm in [58] and then use the methods developed in [57] to compute entanglement from the knowledge of S in a computationally efficient fashion. We have two scenarios (i) $\mathrm{Pur}(\rho_X) = \Omega[\mathrm{poly}^{-1}(n)]$ or (ii) $\mathrm{Pur}(\rho_X) = \Theta(2^{-\alpha n})$ and the two following propositions.

Proposition 3. The purity of the state ψ_E is lower bounded by

$$\mathrm{Pur}(\rho_X) \leqslant \mathrm{Pur}(\psi_E), \tag{11}$$

while we do not know, in principle, whether $\mathrm{Pur}(\rho_E) \leqslant \mathrm{Pur}(\rho_F)$.

Proof. Consider the state ψ_t and its partial state on E, i.e., ψ_E. We are interested in knowing $\mathrm{Pur}(\psi_E)$. In general, the purity of ψ_E is written as

$$\mathrm{Pur}(\psi_E) = \frac{1}{d_E} \sum_{P \in \mathbb{P}_E} \mathrm{tr}^2(P_E \psi_E) \geqslant \frac{1}{d_E} \sum_{P_E\,:\,\mathrm{tr}(P_E\psi_E)=\pm 1}, \tag{12}$$

i.e., is of course lower bounded by the number of Pauli operators on E that have expectation value 1 in absolute value. Now, let $G \subset \mathbb{P}$ the subset of the Pauli group such that

$$G = \{P \in \mathbb{P} \,|\, P|\psi_t\rangle = \pm|\psi_t\rangle\}, \tag{13}$$

taking the partial trace of the set G, one create the following set

$$G_E := \{P_E \in \mathbb{P}_E \,|\, P_E = \mathrm{tr}_F(P),\ P \in G\}, \tag{14}$$

and trivially one has $\frac{1}{d_E} \sum_{P_E\,:\,\mathrm{tr}(P_E\psi_E)=\pm 1} = \frac{|G_E|}{d_E}$. Now let W be the diagonalizer and let it act on the state $|\psi\rangle$. We know that the diagonalizer takes G to a $WGW^\dagger \equiv G'$. Define $G_{\bar{Y}} \equiv \mathrm{tr}_Y(G') := \{P_{\bar{Y}} \in \mathbb{P}_{\bar{Y}} \,|\, P_{\bar{Y}} = \mathrm{tr}_Y(P), P \in G'\}$. Note that, by construction, $G_{\bar{Y}}$ is the same stabilizer group of $\mathrm{tr}_Y(\Phi) \equiv \Phi_{\bar{Y}}$. This is because $W\psi W^\dagger = c_t^Y \Phi c_t^{Y\dagger}$ and $\mathrm{tr}_Y(c_t^Y \Phi c_t^{Y\dagger}) = \mathrm{tr}_Y(\Phi)$. By applying the diagonalizer back $W^\dagger$, one has that $W^\dagger G'_{\bar{Y}} W \in G$, where $G'_{\bar{Y}} \equiv \{P \in \mathbb{P} \,|\, P = P_{\bar{Y}} \otimes \mathbb{1}_Y, P_{\bar{Y}} \in G_{\bar{Y}}\}$; note that $W^\dagger G'_{\bar{Y}} W$ is now the stabilizer group of ρ; first of all $W^\dagger G'_{\bar{Y}} W \subset G$ and, in general, $|W^\dagger G'_{\bar{Y}} W| \leqslant |G|$; consequently, defining $\tilde{G}_E \equiv \{P_E \in \mathbb{P}_E \,|\, P_E = \mathrm{tr}_F(P), P \in W^\dagger G'_{\bar{Y}} W\}$, we have $|\tilde{G}_E| \leqslant |G_E|$. Noting that $\mathrm{Pur}(\rho_E) = |\tilde{G}_E|/d_E$, we have the following inequality:

$$\mathrm{Pur}(\rho_E) \leqslant \frac{|G_E|}{d_E} = \frac{1}{d_E} \sum_{P_E\,:\,\mathrm{tr}(P_E\psi_E)=\pm 1} \leqslant \mathrm{Pur}(\psi_E). \tag{15}$$

The proof is concluded. The bound $\mathrm{Pur}(\psi_E) \geqslant \mathrm{Pur}(\rho_F)$ is proven in the same way by noting that $\mathrm{Pur}(\psi_E) = \mathrm{Pur}(\psi_F) \geqslant \mathrm{Pur}(\rho_F)$.□

Proposition 4. The purity of the state ψ_E is upper bounded by

$$\mathrm{Pur}(\psi_E) \leqslant d_Y^2\, \mathrm{Pur}(\rho_X), \tag{16}$$

see Appendix D 1 for the proof.

Let us now explain the application of the protocol. Without loss of generality, posit $\mathrm{Pur}(\rho_E) > \mathrm{Pur}(\rho_F)$ and thus $X \equiv E$. After evaluating $\mathrm{Pur}(\rho_E)$, Proposition 3 immediately tells us what is a sufficient number of measurements N_{shot} to resolve the purity of ψ_E by the SWAP test. In case (i) this is a polynomial number, and thus purity can be efficiently estimated with a polynomial algorithm. In case (ii), recalling that $n_Y = t$ and thus $d_Y = O[\mathrm{poly}(n)]$, Proposition 4 implies that one can estimate the purity as [59]

$$\mathrm{Pur}(\psi_E) = 2^{-\alpha n + O(\log_2 n)}, \tag{17}$$

i.e., we estimate the bipartite entanglement in $E|F$ up to a second-order logarithmic correction. This is the second main result of our paper: we can estimate an exponentially small purity by a polynomial number of measurements, thereby

achieving an exponential improvement over the known state-of-the-art algorithms.

VI. CONCLUSION

In this paper, we showed that the stabilizer entropy can be moved around subsystems. Effectively, this results in reducing the complexity of a selected subsystem. The tension between the spreading of nonstabilizerness and its localization is akin to an insulator-superfluid transition. In the localized phase, one can exploit this reduction of complexity in relevant quantum information protocols: we show a way of estimating an exponentially small purity by polynomial resources, thereby improving dramatically on known methods.

In perspective, there are a number of questions raised by this paper that we find of interest. First of all, the scope of the purity estimation algorithm presented here can be extended to an efficient SE estimation. Similarly, the cleansing algorithm can potentially be utilized as a starting point for a whole family of quantum algorithms aimed at exploiting the easiness of handling stabilizer states even in nonstabilizer settings.

Then, more generally, how does the complexity cleansing algorithm generalize to time evolution generated by a Hamiltonian? What is the connection between SE cleansing and quantum error-correcting codes? Finally, being an entropy, can SE be evaluated geometrically in the general context of AdS/CFT [60]?

ACKNOWLEDGMENTS

The authors thank the anonymous referee for prompting us to significantly improve our algorithm. The authors acknowledge support from NSF Award No. 2014000. A.H. acknowledges financial support from PNRR MUR Project No. PE0000023-NQSTI and PNRR MUR Project No. CN 00000013 -ICSC.

L.L. and S.F.E.O. contributed equally to this work.

APPENDIX A: BIG-O NOTATION: A BRIEF SUMMARY

In this section, we briefly review: $O(\cdot)$, $\Omega(\cdot)$, and $\Theta(\cdot)$ notations. Consider, for simplicity, positive functions $f(n)$, $g(n)$ of natural numbers $n \in \mathbb{N}$. We can think n as the number of qubits in a multiqubit quantum system. Then

$$f(n) = O(g(n)), \tag{A1}$$

iff there exists a constant $a \geqslant 0$ and a certain value $n_> \in \mathbb{N}$ such that

$$\forall n \geqslant n_>, \quad f(n) \leqslant a\, g(n). \tag{A2}$$

Conversely,

$$f(n) = \Omega(g(n)), \tag{A3}$$

iff there exists a constant $b \geqslant 0$ and a value $n_< \in \mathbb{N}$ such that

$$\forall n \geqslant n_<, \quad f(n) \geqslant b\, g(n). \tag{A4}$$

Lastly,

$$f(n) = \Theta(g(n)), \tag{A5}$$

iff there exist two constant $c_1, c_2 \geqslant 0$ and a value $n_= \in \mathbb{N}$ such that

$$\forall n \geqslant n_=, \quad c_1\, g(n) \leqslant f(n) \leqslant c_2\, g(n). \tag{A6}$$

Let us give some clarifying examples. Consider $f(n) = 10n^3$. We say that $f(n) = O(2^n)$ because there exists $n_0 = 15$ after which $f(n) \leqslant 2^n$. We can also write $f(n) = \Omega(n^2)$ because for any $n \geqslant 1$ $f(n) \geqslant n^2$. Lastly, we write $f(n) = \Theta(n^3)$ because there exist two constants $c_1 = 9$, $c_2 = 11$ such that for any $n \geqslant 1$, one has $9n^3 \leqslant f(n) \leqslant 11n^3$. After the above trivial warm-up example, let us make another last example. Consider $f(n)$ as a sum of two exponentials $f(n) = e^{-2n} + e^{-n}$. We say that $f(n) = \Theta(e^{-n})$ because there exist two constant $c_1 = 1$ and $c_2 = 1 + e^{-4}$ such that for every $n \geqslant 2$ one has $e^{-n} \geqslant f(n) \leqslant (1 + e^{-4})e^{-n}$.

APPENDIX B: PROOF OF PROPOSITION 1

In this section, we compute the average value of $w(\psi_X^C) := \frac{d_X\ \mathrm{SP}(\psi_X)}{\mathrm{Pur}(\psi_X^C)}$, with $X = \{E, F\}$, over the Clifford group

$$\mathbb{E}_C\big[w\big(\psi_E^C\big)\big] = \mathbb{E}_C \frac{d_E\ \mathrm{SP}(\psi_E)}{\mathrm{Pur}\left(\psi_E^C\right)},$$
$$\mathbb{E}_C\big[w\big(\psi_F^C\big)\big] = \mathbb{E}_C \frac{d_F\ \mathrm{SP}(\psi_F)}{\mathrm{Pur}\left(\psi_F^C\right)}. \tag{B1}$$

We would like to calculate the average of the ratio in (B1) as the ratio of the individual averages. Starting from Lemma 1 one can then proceed with the evaluation of the averages. Let us recall a result from [41,42]: let ψ be a pure quantum state, then its Clifford orbit $\mathbb{E}_C[\psi^{C\otimes 4}]$ where $\psi^C \equiv C\psi C^\dagger$ reads

$$\mathbb{E}_C[\psi^{C\otimes 4}] = \alpha Q \Pi_{\mathrm{sym}} + \beta \Pi_{\mathrm{sym}}, \tag{B2}$$

where

$$\Pi_{\mathrm{sym}} = \frac{1}{4!} \sum_{\pi \in S_4} T_\pi, \tag{B3}$$

and

$$\alpha = \frac{\mathrm{SP}(\psi)}{(d+1)(d+2)/6} - \frac{1 - \mathrm{SP}(\psi)}{(d-1)(d+1)(d+2)(d+4)/24},$$
$$\beta = \frac{1 - \mathrm{SP}(\psi)}{(d-1)(d+1)(d+2)(d+4)/24}. \tag{B4}$$

Remark 1. Notice that, despite the similar notation, the permutation operators introduced in (B3) are very different from the permutation operator used to define W. The first are a representation over $\mathcal{H}^{\otimes 4}$ of the permutation group of 4 objects S_4, regardless of the size of the system, which act by switching the basis element of the single Hilbert space $\mathcal{H}$ among the other three copies, according to $\pi \in S_4$, whereas the operators T_{π_Y} are a representation of the permutation group of n elements over $\mathcal{H}$, that act as simply shuffling the qubits of the system according to the permutation of S_n of choice.

Using this knowledge let us evaluate the numerator of Eq. (5). Since $\mathrm{SP}(\psi)$ is a linear operator of $\psi^{\otimes}4$, we are able to slide the Clifford expectation value into it. One then gets

$$\mathrm{tr}(Q_E \mathbb{E}_C[\psi^{C\otimes 4}])$$
$$= \mathrm{tr}[Q_E \otimes \mathbb{1}_F (\alpha Q \Pi_{\mathrm{sym}} + \beta \Pi_{\mathrm{sym}})]$$

$$= \alpha \mathrm{tr}(Q\Pi_{\mathrm{sym}}) + \beta \sum_{\pi} \mathrm{tr}\big(Q_E T_\pi^{(E)}\big)\mathrm{tr}\big(T_\pi^{(F)}\big)$$

$$= \alpha(d+1)(d+2)/6 + \beta \sum_{\pi} \mathrm{tr}\big(Q_E T_\pi^{(E)}\big)\mathrm{tr}\big(T_\pi^{(F)}\big), \quad \text{(B5)}$$

where $\mathrm{tr}(Q^{(E)}T_\pi^{(E)})$ are displayed in [41]. Conversely, the denominator in Eq. (5) reads [12]

$$\mathbb{E}_C\big[\,\mathrm{Pur}\,\big(\psi_E^C\big)\big] = \frac{(d_E + d_F)}{d_E d_F + 1}. \quad \text{(B6)}$$

Taking the ratio between Eqs. (B5) and (B6), we get

$$\mathbb{E}_C\big[M_{\mathrm{lin}}\big(\psi_E^C\big)\big] = 1 - \mathbb{E}_C\big[w\big(\psi_E^C\big)\big] = \frac{M_{\mathrm{lin}}(\psi)\big(d_E^2 - 1\big)d}{(d-1)\big(d + d_E^2\big)}, \quad \text{(B7)}$$

in the limit of $n_E \gg n_F$:

$$\mathbb{E}_C\big[M_{\mathrm{lin}}\big(\psi_E^C\big)\big] = M_{\mathrm{lin}}(\psi) + O\left(\frac{d_F}{d_E}\right). \quad \text{(B8)}$$

In a similar fashion, one could evaluate the other ratio in Eq. (6). Recall that, thanks to the Schmidt decomposition, one has $\mathbb{E}_C[\mathrm{Pur}(\psi_F^C)] = \mathbb{E}_C[\mathrm{Pur}(\psi_E^C)] = \frac{d_E+d_F}{d_E d_F+1}$. The numerator of Eq. (6) reads

$$\mathrm{tr}(Q_F \mathbb{E}_C[\psi^{C\otimes 4}])$$

$$= \alpha \mathrm{tr}(Q\Pi_{\mathrm{sym}}) + \beta \sum_{\pi} \mathrm{tr}\big(Q_F T_\pi^{(F)}\big)\mathrm{tr}\big(T_\pi^{(E)}\big)$$

$$= \alpha(d+1)(d+2)/6 + \beta \sum_{\pi} \mathrm{tr}\big(Q_F T_\pi^{(F)}\big)\mathrm{tr}\big(T_\pi^{(E)}\big), \quad \text{(B9)}$$

by taking the ratio between Eqs. (B9) and (B6), we get

$$\mathbb{E}_C\big[M_{\mathrm{lin}}\big(\psi_F^C\big)\big] = 1 - \mathbb{E}_C\big[w\big(\psi_F^C\big)\big] = \frac{M_{\mathrm{lin}}(\psi)\big(d^2 - d_E^2\big)}{(d-1)\big(d + d_E^2\big)} \quad \text{(B10)}$$

in the limit of $n_E \gg n_F$, one gets

$$\mathbb{E}_C\big[M_{\mathrm{lin}}\big(\psi_F^C\big)\big] = O\left(\frac{d_F}{d_E}\right), \quad \text{(B11)}$$

which proves Proposition 1. ■

APPENDIX C: PROOF OF EQ. (9)

1. Average over the cleansing map

In this section, we show that averaging over the maps $\mathcal{E}\circ\mathcal{C}_{\mathfrak{t}}(\omega)$ denoted as $\mathbb{E}$ is equivalent to averaging over the doped Clifford circuits C_t. First note that $C_t = D^\dagger c_t D V$; then thanks to the left or right invariance of the Haar measure one has the average over t-doped Clifford circuits $\mathbb{E}_{C_t}$ introduced in [41] equals $\mathbb{E}_{C_t} = \mathbb{E}_V \mathbb{E}_D \mathbb{E}_{c_t}$, where c_t is a t-doped Clifford on t qubits and D, V belong to the Clifford group. Finally, note that $\mathcal{E}\circ\mathcal{C}_{\mathfrak{t}}(\omega) \equiv \mathrm{tr}_Y[WC_t(\omega)C_t^\dagger W^\dagger] = \mathrm{tr}_Y(c_t^Y T_{\pi_Y} DV(\omega)(c_t^Y T_{\pi_Y} DV)^\dagger)$. Consequently, the average over $\mathcal{E}\circ\mathcal{C}_{\mathfrak{t}}(\omega)$, hereby denoted as $\mathbb{E}$, of a function $f[\mathcal{E}\circ\mathcal{C}_{\mathfrak{t}}(\omega)]$ obey $\mathbb{E} f[\mathcal{E}\circ\mathcal{C}_{\mathfrak{t}}(\omega)] \equiv \mathbb{E}_{c_t^Y}\mathbb{E}_D\mathbb{E}_V\mathbb{E}_\omega f[\mathcal{E}\circ\mathcal{C}_{\mathfrak{t}}(\omega)] \equiv \mathbb{E}_{c_t}\mathbb{E}_D\mathbb{E}_V f[\mathcal{E}\circ\mathcal{C}_{\mathfrak{t}}(\omega)] \equiv \mathbb{E}_{C_t} f[\mathcal{E}\circ\mathcal{C}_{\mathfrak{t}}(\omega)]$, where we used the invariance of the Haar measure over the Clifford group. Note that

$$\mathbb{E}\{M_E[\mathcal{E}\circ\mathcal{C}_{\mathfrak{t}}(\omega)]\} \geqslant -\log_2 \mathbb{E}\left[\frac{\mathrm{SP}[\mathrm{tr}_{F\setminus Y}\mathcal{E}\circ\mathcal{C}_{\mathfrak{t}}(\omega)]}{\mathrm{Pur}[\mathrm{tr}_{F\setminus Y}\mathcal{E}\circ\mathcal{C}_{\mathfrak{t}}(\omega)]}\right], \quad \text{(C1)}$$

where $\mathbb{E}$ is the average over $\mathcal{E}\circ\mathcal{C}_{\mathfrak{t}}(\omega)$ discussed in Appendix C 1. The object we intend to calculate is

$$\mathbb{E}_{c_t}\mathbb{E}_D\mathbb{E}_\omega SP[\mathrm{tr}_{F\setminus Y}\mathcal{E}\circ\mathcal{C}_{\mathfrak{t}}(\omega)] = \mathbb{E}_{c_t}\mathbb{E}_D\mathbb{E}_V \mathrm{tr}\big(Q_E\otimes \mathbb{1}_F^{\otimes 4}\big)\big(c_t^{Y^{\otimes 4}}\otimes \mathbb{1}_{\bar{Y}}^{\otimes 4}\big)(T_{\pi_Y}DV)^{\otimes 4}|\mathbf{0}\rangle\langle\mathbf{0}|^{\otimes 4}(T_{\pi_Y}DV)^{\dagger\otimes 4}\big(c_t^{Y^{\otimes 4}}\otimes \mathbb{1}_{\bar{Y}}^{\otimes 4}\big)^\dagger, \quad \text{(C2)}$$

where $\bar{Y} = (E\cup F)\cap Y$. Due to the left and right unitary invariance of the Haar measure over the Clifford group, the average over D gets absorbed into the one over V, and the permutation operator gets absorbed as well. One then gets

$$\mathbb{E}_{c_t}\mathbb{E}_D\mathbb{E}_V SP(\mathrm{tr}_{F\setminus Y}\mathcal{E}\circ\mathcal{C}_{\mathfrak{t}}(\omega)) = \mathbb{E}_{c_t}\mathbb{E}_V SP[\mathrm{tr}_{F\setminus Y}\mathcal{E}\circ\mathcal{C}_{\mathfrak{t}}(\omega)] = \mathrm{tr}\big(Q_E\otimes \mathbb{1}_F^{\otimes 4}\big)\mathbb{E}_{c_t}\big(c_t^{Y^{\otimes 4}}\otimes \mathbb{1}_{\bar{Y}}^{\otimes 4}\big)\mathbb{E}_V V^{\otimes 4}|\mathbf{0}\rangle\langle\mathbf{0}|^{\otimes 4}V^{\dagger\otimes 4}\big(c_t^{Y^{\otimes 4}}\otimes \mathbb{1}_{\bar{Y}}^{\otimes 4}\big)^\dagger. \quad \text{(C3)}$$

The average over V gives the Clifford n-qubit state orbit, as shown in Eq. (B2). By substituting the expression we get

$$\mathbb{E}_{c_t}\mathbb{E}_V \mathrm{SP}[\mathrm{tr}_{F\setminus Y}\mathcal{E}\circ\mathcal{C}_{\mathfrak{t}}(\omega)] = \frac{1}{4!}\sum_{\pi} \mathrm{tr}\big(Q_E\otimes \mathbb{1}_F^{\otimes 4}\big)\mathbb{E}_{c_t}\big(c_t^{Y^{\otimes 4}}\otimes \mathbb{1}_{\bar{Y}}^{\otimes 4}\big)(\alpha Q T_\pi + \beta T_\pi)\big(c_t^{Y^{\otimes 4}}\otimes \mathbb{1}_{\bar{Y}}^{\otimes 4}\big)^\dagger. \quad \text{(C4)}$$

Here we will exploit the fact that the permutation operators T_π and Q in $\mathcal{B}(\mathcal{H}^{\otimes 4})$ can always be factorized as $T_\pi^{(Y)}\otimes T_\pi^{(\bar{Y})}$ and $Q_Y\otimes Q_{\bar{Y}}$ with $T_\pi^{(Y)}, Q_Y \in \mathcal{B}(\mathcal{H}_Y^{\otimes 4})$ and $T_\pi^{(\bar{Y})}, Q_{\bar{Y}} \in \mathcal{B}(\mathcal{H}_{\bar{Y}}^{\otimes 4})$, being $\mathcal{H} = \mathcal{H}_Y\otimes\mathcal{H}_{\bar{Y}}$. In this way we can write

$$\begin{aligned}\mathbb{E}_{c_t}\mathbb{E}_V \mathrm{SP}[\mathrm{tr}_{F\setminus Y}\mathcal{E}\circ\mathcal{C}_t(\omega)] &= \frac{1}{24}\sum_{\pi} \mathrm{tr}\big(Q_E\otimes \mathbb{1}_F^{\otimes 4}\big)\mathbb{E}_{c_t}\big(c_t^{Y^{\otimes 4}}\otimes \mathbb{1}_{\bar{Y}}^{\otimes 4}\big)\big(\alpha Q_Y T_\pi^{(Y)}\otimes Q_{\bar{Y}} T_\pi^{(\bar{Y})} + \beta T_\pi^{(Y)}\otimes T_\pi^{(\bar{Y})}\big)\big(c_t^{Y^{\otimes 4}}\otimes \mathbb{1}_{\bar{Y}}^{\otimes 4}\big)^\dagger \\ &= \frac{1}{24}\sum_{\pi} \mathrm{tr}\big(Q_E\otimes \mathbb{1}_F^{\otimes 4}\big)\big(\alpha\mathbb{E}_{c_t}\big(c_t^{Y^{\otimes 4}} Q_Y T_\pi^{(Y)} c_t^{Y^{\dagger\otimes 4}}\big)\otimes Q_{\bar{Y}} T_\pi^{(\bar{Y})} + \beta\mathbb{E}_{c_t}\big(c_t^{Y^{\otimes 4}} T_\pi^{(Y)} c_t^{Y^{\dagger\otimes 4}}\big)\otimes T_\pi^{(\bar{Y})}\big) \\ &\equiv R_1 + R_2. \end{aligned} \quad \text{(C5)}$$

As before, due to the left and right invariance of the Haar measure over the Clifford group, the average over c_t of c_t^Y is the same of the average of c_t over c_t since the permutation operator which dresses c_t is a Clifford operator. The way to calculate the

averages over c_t is shown in [41] and the result reads

$$\mathbb{E}_{c_t} c_t^{Y\otimes 4} Q_Y T_\pi^{(Y)} c_t^{Y\dagger\otimes 4} = \sum_\sigma \eta_\sigma\big(Q_Y T_\pi^{(Y)}\big) Q_Y T_\sigma^{(Y)} + \mu_\sigma\big(Q_Y T_\pi^{(Y)}\big) T_\sigma^{(Y)},$$
$$\mathbb{E}_{c_t} c_t^{Y\otimes 4} T_\pi^{(Y)} c_t^{Y\dagger\otimes 4} = \sum_\sigma \eta_\sigma\big(T_\pi^{(Y)}\big) Q_Y T_\sigma^{(Y)} + \mu_\sigma\big(T_\pi^{(Y)}\big) T_\sigma^{(Y)}, \tag{C6}$$

with

$$\eta_\sigma(\mathcal{O}) = \sum_\pi \Xi^t_{\sigma\pi} c_\pi(\mathcal{O}),$$
$$\mu_\sigma(\mathcal{O}) = b(\mathcal{O}) + \sum_\pi \Gamma^{(t)}_{\sigma\pi} c_\pi(\mathcal{O}), \tag{C7}$$

and c_π, b_π, $\Xi_{\pi\sigma}$, and $\Gamma^{(t)}_{\pi\sigma}$ as defined in [41]. Since both the averages have nonzero components both on $Q_Y T_\sigma^{(Y)}$ and $T_\sigma^{(Y)}$, the terms R_1, R_2 in Eq. (C5) will have the same structure, with the only difference being the value of the coefficients η_σ and μ_σ, so we will carry on the calculation for just the first one:

$$R_1 = \frac{\alpha}{24}\sum_{\pi\sigma} \mathrm{tr}\big(\eta_\sigma\big(Q_Y T_\pi^{(Y)}\big) Q_Y T_\sigma^{(Y)} + \mu_\sigma\big(Q_Y T_\pi^{(Y)}\big) T_\sigma^{(Y)}\big) \otimes Q_{\bar{Y}} T_\pi^{(\bar{Y})} (Q_E \otimes \mathbb{1}_F). \tag{C8}$$

Since $\mathfrak{t}/\mathfrak{f} > 1$ we can write the partition E as $E = E \cup F \cap Y \cup Y \cap F = \bar{Y} \cup Y \cap F = \bar{Y} \cup Y'$ with $Y' = Y \cap F$. In the same fashion, we can write $Y = F \cup Y'$ and factorize the Q and T operators accordingly as

$$Q_E = Q_{\bar{Y}} \otimes Q_{Y'}$$
$$T_\pi^{(E)} = T_\pi^{(\bar{Y})} \otimes T_\pi^{(Y')},$$
$$Q_Y = Q_F \otimes Q_{Y'},$$
$$T_\pi^{(Y)} = T_\pi^{(F)} \otimes T_\pi^{(Y')}. \tag{C9}$$

Substituting into Eq. (C8) one gets

$$\begin{aligned}
R_1 &= \frac{\alpha}{24}\sum_{\pi\sigma} \mathrm{tr}\big(\eta_\sigma\big(Q_Y T_\pi^{(Y)}\big) Q_Y T_\sigma^{(Y)} + \mu_\sigma\big(Q_Y T_\pi^{(Y)}\big) T_\sigma^{(Y)}\big) \otimes Q_{\bar{Y}} T_\pi^{(\bar{Y})} (\mathbb{1}_F \otimes Q_{Y'} \otimes Q_Y) \\
&= \frac{\alpha}{24}\sum_{\pi\sigma} \mathrm{tr}\big(\eta_\sigma\big(Q_Y T_\pi^{(Y)}\big) Q_Y T_\sigma^{(Y)} + \mu_\sigma\big(Q_Y T_\pi^{(Y)}\big) T_\sigma^{(Y)}\big) (\mathbb{1}_F \otimes Q_{Y'}) \mathrm{tr}\big(Q_{\bar{Y}} T_\pi^{(\bar{Y})}\big) \\
&= \frac{\alpha}{24}\sum_{\pi\sigma} \mathrm{tr}\big(\eta_\sigma\big(Q_Y T_\pi^{(Y)}\big) Q_F T_\sigma^{(F)} \otimes Q_{Y'} T_\sigma^{(Y')} + \mu_\sigma\big(Q_Y T_\pi^{(Y)}\big) T_\sigma^{(F)} \otimes T_\sigma^{(Y')}\big) (\mathbb{1}_F \otimes Q_{Y'}) \mathrm{tr}\big(Q_{\bar{Y}} T_\pi^{(\bar{Y})}\big) \\
&= \frac{\alpha}{24}\sum_{\pi\sigma} \mathrm{tr}\big(\eta_\sigma\big(Q_Y T_\pi^{(Y)}\big) Q_Y T_\sigma^{(Y)} + \mu_\sigma\big(Q_Y Y_\pi^{(Y)}\big) Q_{Y'} T_\sigma^{(Y')} \otimes T_\sigma^{F}\big) \mathrm{tr}\big(Q_{\bar{Y}} T_\pi^{(\bar{Y})}\big) \\
&= \frac{\alpha}{24}\sum_{\pi\sigma} \eta_\sigma\big(Q_Y T_\pi^{(Y)}\big) \mathrm{tr}\big(Q_Y T_\sigma^{(Y)}\big) \mathrm{tr}\big(Q_{\bar{Y}} T_\pi^{(\bar{Y})}\big) + \mu_\sigma\big(Q_Y T_\pi^{(Y)}\big) \mathrm{tr}\big(Q_{Y'} T_\sigma^{(Y')}\big) \mathrm{tr}(T_\sigma^{(F)}) \mathrm{tr}\big(Q_{\bar{Y}} T_\pi^{(\bar{Y})}\big).
\end{aligned} \tag{C10}$$

By plugging the values of the coefficients η and μ computing the traces one gets the expression of the function $g(n, \mathfrak{t}, \mathfrak{f})$ (fully displayed in the following section). ■

2. Explicit formula for the partial SE in E

In this section, we show the full expression for $g(n, \mathfrak{t}, \mathfrak{f})$,

$$\begin{aligned}
g(n,\mathfrak{t},\mathfrak{f}) = {} & (3(d+2)(d+4)(d+2^{2n(1-\mathfrak{f})})(2^{2n\mathfrak{t}}-9))^{-1} \\
& \times (2^{-6n\mathfrak{t}-1}(2^{2n(1-\mathfrak{f})+4n\mathfrak{t}+1}(2^{n\mathfrak{t}}(3\,2^{2n\mathfrak{t}} f_-^{n\mathfrak{t}} + 2^{3n\mathfrak{t}} f_-^{n\mathfrak{t}} - 5\,2^{n\mathfrak{t}+1} f_-^{n\mathfrak{t}} - 24 f_-^{n\mathfrak{t}} + \\
& + (2^{n\mathfrak{t}}-4)(2^{n\mathfrak{t}}-2)(2^{n\mathfrak{t}}+3) f_+^{n\mathfrak{t}} - 9\,2^{n\mathfrak{t}+4} + 9\,2^{3n\mathfrak{t}+1}) + \\
& - 2(-13\,2^{2n\mathfrak{t}} + 2^{4n\mathfrak{t}} + 36) g^{n\mathfrak{t}}) + 2^{2n+2n\mathfrak{t}}(2^{4n\mathfrak{t}}(2^{2n(1-\mathfrak{f})}(f_-^{n\mathfrak{t}} + f_+^{n\mathfrak{t}} + 18) + \\
& - 2(f_-^{n\mathfrak{t}} + f_+^{n\mathfrak{t}} - 2g^{n\mathfrak{t}})) - 2^{2n\mathfrak{t}+1}(2^{2n(1-\mathfrak{f})}(5 f_-^{n\mathfrak{t}} + 5 f_+^{n\mathfrak{t}} + 24 g^{n\mathfrak{t}} + 72) + \\
& - 10 f_-^{n\mathfrak{t}} - 10 f_+^{n\mathfrak{t}} + 26 g^{n\mathfrak{t}} + 228) + 3(2^{2n(1-\mathfrak{f})} - 2) 2^{3n\mathfrak{t}}(f_-^{n\mathfrak{t}} - f_+^{n\mathfrak{t}}) + \\
& - 3(2^{2n(1-\mathfrak{f})} - 2) 2^{n\mathfrak{t}+3}(f_-^{n\mathfrak{t}} - f_+^{n\mathfrak{t}}) + 144(2^{2n(1-\mathfrak{f})} + 1) g^{n\mathfrak{t}} + 9\,2^{6n\mathfrak{t}+2}) +
\end{aligned}$$

$$
\begin{aligned}
&+ 2^{4n}(2^{nt}(24(f_-^{nt} - f_+^{nt}) + 2^{nt}(3\,2^{nt}(f_+^{nt} - f_-^{nt}) - 2^{2nt}(f_-^{nt} + f_+^{nt} + 72)+\\
&+ 2(5f_-^{nt} + 5f_+^{nt} + 72) + 3\,2^{4nt+1})) + 48(2^{2nt} - 3)g^{nt})+\\
&+ 3(24(f_-^{nt} + f_+^{nt}) + 2^{nt}(-3\,2^{nt}f_-^{nt} - 2^{2nt}f_-^{nt} + 10f_-^{nt}+\\
&+ (2^{nt} - 5)(2^{nt} + 2)f_+^{nt} - 15\,2^{nt+2} + 2^{5nt+1}))2^{3n+2nt} + 3\,2^{n+4nt}t\\
&\times (-5(f_-^{nt} - f_+^{nt})2^{2n(1-\mathfrak{f})+nt+1} + (f_-^{nt} - f_+^{nt})2^{2n(1-\mathfrak{f})+3nt}+\\
&- 3\,2^{2n(1-\mathfrak{f})+3}(f_-^{nt} + f_+^{nt}) + 3\,2^{2nt}(2^{2n(1-\mathfrak{f})}(f_-^{nt} + f_+^{nt} - 20) - 48)+\\
&+ (3\,2^{2n(1-\mathfrak{f})} + 8)2^{4nt+1}))),
\end{aligned}
\tag{C11}
$$

with

$$
f_- = \frac{3\times 4^{nt} - 3\times 2^{nt} - 4}{4^{nt} - 1}, \quad f_+ = \frac{3\times 4^{nt} + 3\times 2^{nt} - 4}{4^{nt} - 1}, \quad g = \frac{3\times 4^{nt} - 4}{4^{nt} - 1}. \tag{C12}
$$

APPENDIX D: EFFICIENT PURITY ESTIMATION

Let us first of all establish some useful notation. Recalling the definition for the stabilizer state $|\Phi\rangle := WV|\mathbf{0}\rangle$ and $\Phi \equiv |\Phi\rangle\langle\Phi|$, we obtain the identity $\mathrm{tr}_Y(W\psi_t W^\dagger) = \mathrm{tr}_Y(\Phi) \equiv \Phi_{\bar{Y}}$ and notice that this is a stabilizer state. Notice that $\mathcal{E}(\psi_t) = \Phi_{\bar{Y}}$ and thus the stabilizer state $\rho = W^\dagger(\mathcal{E}(\psi_t)\otimes d_Y^{-1}I_Y)W = W^\dagger(\Phi_{\bar{Y}}\otimes d_Y^{-1}I_Y)W$.

1. Proof of Proposition 4

Following the notations of the main paper, let us denote $|\Phi\rangle = WV|\mathbf{0}\rangle$ and let $|\psi_t\rangle = W^\dagger c_t^Y WV|\mathbf{0}\rangle$ with W being the diagonalizer and c_t^Y being a t-doped Clifford circuit acting on the system Y with $n_Y \leqslant t$. Let us prove the upper bound in Proposition 4, i.e.,

$$
\mathrm{Pur}(\psi_E) \leqslant d_Y^2\,\mathrm{Pur}(\rho_X), \tag{D1}
$$

for $X = E, F$ and $\rho = W^\dagger(\Phi_{\bar{Y}}\otimes d_Y^{-1}I_Y)W$ and $\Phi_{\bar{Y}} = \mathrm{tr}_Y|\Phi\rangle\langle\Phi|$.

Define $S_E = \{P\in\mathbb{P}\,|\,P = WP_EW^\dagger\}$. First note that

$$
\mathrm{Pur}(\psi_E) = \frac{1}{d_E}\sum_{P_E\in\mathbb{P}_E}\mathrm{tr}^2(\psi_E P_E) = \frac{1}{d_E}\sum_{P\in S_E}\mathrm{tr}^2(c_t\Phi c_t^\dagger P). \tag{D2}
$$

Then, thanks to their tensor product structure, Pauli operators $S_E \ni P$ can be decomposed over $Y\cup\bar{Y}$. Therefore, let us define $S_E|_{\bar{Y}} = \{d_Y\sum_{P_Y}\mathrm{tr}_Y(P_Y P)\,|\,P\in S_E\}$ the restriction of S_E to $\bar{Y}$. Define the completion of $S_E|_{\bar{Y}}$. For $P_{\bar{Y}}\in S_E|_{\bar{Y}}$, define the set $T_{P_{\bar{Y}}} = \{d_{\bar{Y}}\mathrm{tr}_{\bar{Y}}(PI_Y\otimes P_{\bar{Y}})\,|\,P\in S_E\}$. Notice that S_E can be written as

$$
S_E = \bigcup_{P_{\bar{Y}}\in S_E|_{\bar{Y}}}\{T_{P_{\bar{Y}}}\otimes P_{\bar{Y}}\}. \tag{D3}
$$

We can thus write

$$
\mathrm{Pur}(\psi_E) = \frac{1}{d_E}\sum_{P_{\bar{Y}}\in S_E|_{\bar{Y}}}\sum_{P_Y\in T_{P_{\bar{Y}}}}\mathrm{tr}^2(c_t\Phi c_t^\dagger P_Y\otimes P_{\bar{Y}}) = \frac{1}{d_E}\sum_{P_{\bar{Y}}\in S_E|_{\bar{Y}}}\sum_{P_Y\in T_{P_{\bar{Y}}}}\mathrm{tr}^2(\Phi c_t^\dagger P_Y c_t\otimes P_{\bar{Y}}). \tag{D4}
$$

Since there is a sum of positive terms, we can upper bound it as

$$
\mathrm{Pur}(\psi_E) \leqslant \frac{1}{d_E}\sum_{P_{\bar{Y}}\in S_E|_{\bar{Y}}}\sum_{P_Y\in\mathbb{P}_Y}\mathrm{tr}^2(\Phi c_t^\dagger P_Y c_t\otimes P_{\bar{Y}}) = \frac{d_Y}{d_E}\sum_{P_{\bar{Y}}\in S_E|_{\bar{Y}}}\mathrm{tr}[\Phi^{\otimes 2}T_Y(I_Y\otimes P_{\bar{Y}})^{\otimes 2}]. \tag{D5}
$$

Now, note that $\mathrm{tr}[\Phi^{\otimes 2}T_Y(I_Y\otimes P_{\bar{Y}})^{\otimes 2}] = \mathrm{tr}_{\bar{Y}}[\Phi^{\otimes 2}T_{\bar{Y}}(I_Y\otimes P_{\bar{Y}})^{\otimes 2}] = \mathrm{tr}_{\bar{Y}}(\Phi_{\bar{Y}}^{\otimes 2}T_{\bar{Y}}P_{\bar{Y}}^{\otimes 2}) = \frac{1}{d_Y}\mathrm{tr}[\Phi_{\bar{Y}}^{\otimes 2}T_{\bar{Y}}(I_Y\otimes P_{\bar{Y}})^{\otimes 2}]$. Where the equality follows from the fact that $T|\Phi^{\otimes 2}\rangle = |\Phi^{\otimes 2}\rangle$. Therefore, we arrived to

$$
\mathrm{Pur}(\psi_E) \leqslant \frac{1}{d_E}\sum_{P_{\bar{Y}}\in S_E|_{\bar{Y}}}\mathrm{tr}\big[\Phi_{\bar{Y}}^{\otimes 2}T_{\bar{Y}}(I_Y\otimes P_{\bar{Y}})^{\otimes 2}\big] = \frac{1}{d_E}\sum_{P_{\bar{Y}}\in S_E|_{\bar{Y}}}\sum_{P_Y\in T_{P_{\bar{Y}}}}\mathrm{tr}\big[\Phi_{\bar{Y}}^{\otimes 2}T_{\bar{Y}}(P_Y\otimes P_{\bar{Y}})^{\otimes 2}\big], \tag{D6}
$$

the second equality follows from the fact that $\mathrm{tr}[\Phi_{\bar{Y}}^{\otimes 2}T_{\bar{Y}}(P_Y\otimes P_{\bar{Y}})^{\otimes 2}] = 0$ for every $P_Y\neq I_Y$. Therefore, we have the bound

$$
\mathrm{Pur}(\psi_E) \leqslant \frac{1}{d_E}\sum_{P\in S_E}\mathrm{tr}\big(\Phi_{\bar{Y}}^{\otimes 2}T_{\bar{Y}}P^{\otimes 2}\big) = \mathrm{tr}\big(W^{\otimes 2}T_E W^{\dagger\otimes 2}\Phi_{\bar{Y}}^{\otimes 2}T_{\bar{Y}}\big) = d_Y\big(d_Y^2 - 1\big)\Lambda_1, \tag{D7}
$$

with

$$\Lambda_1 := \frac{1}{d_Y\left(d_Y^2-1\right)}\mathrm{tr}\left(W^{\dagger\otimes 2}T_E W^{\otimes 2}\Phi_{\bar{Y}}^{\otimes 2}T_{\bar{Y}}\right). \tag{D8}$$

The term Λ_1 is proportional to $\mathrm{Pur}(\rho_F)$. First, by evaluating $\mathrm{Pur}(\rho_F)$, which reads

$$\begin{aligned}\mathrm{Pur}(\rho_F) &= \frac{1}{d_Y^2}\mathrm{tr}\left(W^{\dagger\otimes 2}T_E W^{\otimes 2}\Phi_{\bar{Y}}^{\otimes 2}T\right) = \frac{1}{d_E d_Y^2}\sum_{P_E}\mathrm{tr}[P_E(W)\Phi_{\bar{Y}}P_E(W)\Phi_{\bar{Y}}]\\ &= \frac{1}{d_Y d d_E}\sum_{P_E,Q_{\bar{Y}}}\mathrm{tr}[P_E(W)Q_{\bar{Y}}P_E(W)\Phi_{\bar{Y}}]\mathrm{tr}(Q_{\bar{Y}}\Phi_{\bar{Y}}) = \frac{1}{d^2 d_E}\sum_{P_E,Q_{\bar{Y}}}\mathrm{tr}[P_E(W)Q_{\bar{Y}}P_E Q_{\bar{Y}}],\end{aligned} \tag{D9}$$

we now show the proportionality between $\mathrm{Pur}(\rho_F)$ and Λ_1:

$$\begin{aligned}\Lambda_1 &= \frac{1}{d_Y\left(d_Y^2-1\right)}\mathrm{tr}\left(W^{\dagger\otimes 2}T_E W^{\otimes 2}\Phi_{\bar{Y}}^{\otimes 2}T_{\bar{Y}}\right) = \frac{1}{d_Y d_E\left(d_Y^2-1\right)}\sum_{P_E}\mathrm{tr}\left(W^{\dagger\otimes 2}P_E^{\otimes 2}W^{\otimes 2}\Phi_{\bar{Y}}^{\otimes 2}T_{\bar{Y}}\right)\\ &= \frac{1}{d_Y d_E\left(d_Y^2-1\right)}\sum_{P_E}\mathrm{tr}_{\bar{Y}}\{\mathrm{tr}_Y[P_E(W)]\Phi_{\bar{Y}}\mathrm{tr}_Y[P_E(W)]\Phi_{\bar{Y}}\}\\ &= \frac{1}{d_E d\left(d_Y^2-1\right)}\sum_{P_E,Q_{\bar{Y}}}\mathrm{tr}_{\bar{Y}}\{\mathrm{tr}_Y[P_E(W)]Q_{\bar{Y}}\mathrm{tr}_Y[P_E(W)]\Phi_{\bar{Y}}\}\mathrm{tr}(Q_{\bar{Y}}\Phi_{\bar{Y}})\\ &= \frac{1}{d_{\bar{Y}} d_E d\left(d_Y^2-1\right)}\sum_{P_E,Q_{\bar{Y}}}\mathrm{tr}_{\bar{Y}}\{\mathrm{tr}_Y[P_E(W)]Q_{\bar{Y}}\mathrm{tr}_Y[P_E(W)]Q_{\bar{Y}}\}\\ &= \frac{d_Y}{d_E d^2\left(d_Y^2-1\right)}\sum_{P_E,Q_{\bar{Y}}}\mathrm{tr}[P_E(W)Q_{\bar{Y}}P_E(W)Q_{\bar{Y}}]\\ &= \frac{d_Y}{\left(d_Y^2-1\right)}\mathrm{Pur}(\rho_F).\end{aligned} \tag{D10}$$

Therefore, recalling the inequality in Eq. (D7), we proved that

$$\mathrm{Pur}(\psi_E) \leqslant d_Y^2\,\mathrm{Pur}(\rho_F). \tag{D11}$$

Similarly, one can obtain the bound for

$$\mathrm{Pur}(\psi_F) \leqslant d_Y\left(d_Y^2-1\right)\Lambda_2, \tag{D12}$$

with

$$\Lambda_2 := \frac{1}{d_Y\left(d_Y^2-1\right)}\mathrm{tr}\left(W^{\dagger\otimes 2}T_E W^{\otimes 2}\Phi_{\bar{Y}}^{\otimes 2}T_Y\right). \tag{D13}$$

In a similar fashion to Λ_1, the term Λ_2 is proportional to $\mathrm{Pur}(\rho_E)$. The latter reads

$$\begin{aligned}\mathrm{Pur}(\rho_E) &= \frac{1}{d_{\bar{Y}}^2}\mathrm{tr}\left(W^{\dagger\otimes 2}T_E W^{\otimes 2}\Phi_{\bar{Y}}^{\otimes 2}\right) = \frac{1}{d_{\bar{Y}}^2}\mathrm{tr}\left(TW^{\dagger\otimes 2}T_F W^{\otimes 2}\Phi_{\bar{Y}}^{\otimes 2}\right)\\ &= \frac{1}{d_{\bar{Y}}^2 d_F}\sum_{P_F}\mathrm{tr}[P_F(W)\Phi_{\bar{Y}}P_F(W)\Phi_{\bar{Y}}] = \frac{1}{d_Y d d_F}\sum_{P_F,Q_{\bar{Y}}}\mathrm{tr}[P_F(W)Q_{\bar{Y}}P_F(W)\Phi_{\bar{Y}}]\mathrm{tr}_{\bar{Y}}(Q_{\bar{Y}}\Phi_{\bar{Y}})\\ &= \frac{1}{d^2 d_F}\sum_{P_F,Q_{\bar{Y}}}\mathrm{tr}[P_F(W)Q_{\bar{Y}}P_F(W)Q_{\bar{Y}}],\end{aligned} \tag{D14}$$

and we can show the proportionality between $\mathrm{Pur}(\rho_E)$ and Λ_2 as follows:

$$\begin{aligned}
\Lambda_2 &= \frac{1}{d_Y\left(d_Y^2-1\right)}\mathrm{tr}\left(W^{\dagger\otimes 2}T_E W^{\otimes 2}\Phi_{\bar{Y}}^{\otimes 2}T_Y\right) = \frac{1}{d_Y\left(d_Y^2-1\right)}\mathrm{tr}\left(W^{\dagger\otimes 2}T_F W^{\otimes 2}\Phi_{\bar{Y}}^{\otimes 2}T_{\bar{Y}}\right)\\
&= \frac{1}{d_Y d_F\left(d_Y^2-1\right)}\sum_{P_F}\mathrm{tr}\left(W^{\dagger\otimes 2}P_F W^{\otimes 2}\Phi_{\bar{Y}}^{\otimes 2}T_{\bar{Y}}\right)\\
&= \frac{1}{d_Y d_F\left(d_Y^2-1\right)}\sum_{P_F}\mathrm{tr}_{\bar{Y}}\{\mathrm{tr}_Y[P_F(W)]\Phi_{\bar{Y}}\mathrm{tr}_Y[P_F(W)]\Phi_{\bar{Y}}\}\\
&= \frac{1}{d_F d\left(d_Y^2-1\right)}\sum_{P_F,Q_{\bar{Y}}}\mathrm{tr}_{\bar{Y}}\{\mathrm{tr}_Y[P_F(W)]Q_{\bar{Y}}\mathrm{tr}_Y[P_F(W)]\Phi_{\bar{Y}}\}\mathrm{tr}_{\bar{Y}}(Q_{\bar{Y}}\Phi_{\bar{Y}})\\
&= \frac{1}{d_F d\left(d_Y^2-1\right)d_{\bar{Y}}}\sum_{P_F,Q_{\bar{Y}}}\mathrm{tr}_{\bar{Y}}\{\mathrm{tr}_Y[P_F(W)]Q_{\bar{Y}}\mathrm{tr}_Y[P_F(W)]Q_{\bar{Y}}\}\\
&= \frac{1}{d_F d\left(d_Y^2-1\right)d_{\bar{Y}}}\sum_{P_F,Q_{\bar{Y}}}\mathrm{tr}[P_F(W)Q_{\bar{Y}}P_F(W)Q_{\bar{Y}}]\\
&= \frac{d_Y}{d_F d^2\left(d_Y^2-1\right)}\sum_{P_F,Q_{\bar{Y}}}\mathrm{tr}[P_F(W)Q_{\bar{Y}}P_F(W)Q_{\bar{Y}}]\\
&= \frac{d_Y}{\left(d_Y^2-1\right)}\mathrm{Pur}(\rho_E).
\end{aligned} \tag{D15}$$

Finally, since ψ_t is pure, one has $\mathrm{Pur}(\psi_E)=\mathrm{Pur}(\psi_F)$, hence one obtains $\mathrm{Pur}(\psi_E)\leqslant d_Y^2\,\mathrm{Pur}(\rho_E)$, thus concluding the proof.

[1] P. W. Shor, in *Proceedings 35th Annual Symposium on Foundations of Computer Science* (ACM, New York, 1994), pp. 124–134.
[2] A. Y. Kitaev, Russ. Math. Surv. **52**, 1191 (1997).
[3] E. Farhi and A. W. Harrow, arXiv:1602.07674.
[4] S. Boixo, S. V. Isakov, V. N. Smelyanskiy, R. Babbush, N. Ding, Z. Jiang, M. J. Bremner, J. M. Martinis, and H. Neven, Nat. Phys. **14**, 595 (2018).
[5] A. W. Harrow and A. Montanaro, Nature (London) **549**, 203 (2017).
[6] S. Bravyi, D. Gosset, and R. König, Science **362**, 308 (2018).
[7] F. Arute, K. Arya, R. Babbush, D. Bacon *et al.*, Nature (London) **574**, 505 (2019).
[8] D. Gottesman, The Heisenberg representation of quantum computers, in *Group22: Proceedings of the XXII International Colloquium on Group Theoretical Methods in Physics*, edited by S. P. Corney, R. Delbourgo, and P. D. Jarvis (International, Cambridge, MA, 1999), pp. 32–43.
[9] J. S. Bell, Phys. Phys. Fiz. **1**, 195 (1964).
[10] J. S. Bell, Rev. Mod. Phys. **38**, 447 (1966).
[11] D. M. Greenberger, M. A. Horne, A. Shimony, and A. Zeilinger, Am. J. Phys. **58**, 1131 (1990).
[12] D. N. Page, Phys. Rev. Lett. **71**, 1291 (1993).
[13] S. Bravyi and A. Kitaev, Phys. Rev. A **71**, 022316 (2005).
[14] A. W. Harrow, A. Hassidim, and S. Lloyd, Phys. Rev. Lett. **103**, 150502 (2009).
[15] E. T. Campbell and D. E. Browne, Phys. Rev. Lett. **104**, 030503 (2010).
[16] E. T. Campbell and M. Howard, Phys. Rev. Lett. **118**, 060501 (2017).
[17] E. T. Campbell, Phys. Rev. A **83**, 032317 (2011).
[18] E. T. Campbell and M. Howard, Phys. Rev. A **95**, 022316 (2017).
[19] S. Bravyi and J. Haah, Phys. Rev. A **86**, 052329 (2012).
[20] H. Zhu, R. Kueng, M. Grassl, and D. Gross, arXiv:1609.08172.
[21] S. Bravyi and D. Gosset, Phys. Rev. Lett. **116**, 250501 (2016).
[22] S. Bravyi, D. Browne, P. Calpin, E. Campbell, D. Gosset, and M. Howard, Quantum **3**, 181 (2019).
[23] L. Leone, S. F. E. Oliviero, and A. Hamma, Phys. Rev. A **107**, 022429 (2023).
[24] C. Chamon, A. Hamma, and E. R. Mucciolo, Phys. Rev. Lett. **112**, 240501 (2014).
[25] D. Shaffer, C. Chamon, A. Hamma, and E. R. Mucciolo, J. Stat. Mech.: Theory Exp. (2014) P12007.
[26] Z.-C. Yang, A. Hamma, S. M. Giampaolo, E. R. Mucciolo, and C. Chamon, Phys. Rev. B **96**, 020408(R) (2017).
[27] S. Zhou, Z.-C. Yang, A. Hamma, and C. Chamon, SciPost Physics **9**, 087 (2020).
[28] S. True and A. Hamma, Quantum **6**, 818 (2022).
[29] S. Piemontese, T. Roscilde, and A. Hamma, Phys. Rev. B **107**, 134202 (2023).
[30] L. Leone, S. F. E. Oliviero, S. Piemontese, S. True, and A. Hamma, Phys. Rev. A **106**, 062434 (2022).
[31] L. Leone, S. F. E. Oliviero, S. Lloyd, and A. Hamma, Phys. Rev. A **109**, 022429 (2024).
[32] S. F. E. Oliviero, L. Leone, S. Lloyd, and A. Hamma, Phys. Rev. Lett. **132**, 080402 (2024).

[33] L. Leone, S. F. E. Oliviero, and A. Hamma, Phys. Rev. Lett. **128**, 050402 (2022).
[34] M. Howard and E. Campbell, Phys. Rev. Lett. **118**, 090501 (2017).
[35] Z.-W. Liu and A. Winter, PRX Quantum **3**, 020333 (2022).
[36] S. F. E. Oliviero, L. Leone, A. Hamma, and S. Lloyd, npj Quantum Inf. **8**, 148 (2022).
[37] T. Haug and M. S. Kim, PRX Quantum **4**, 010301 (2023).
[38] J. Odavić, T. Haug, G. Torre, A. Hamma, F. Franchini, and S. M. Giampaolo, SciPost Phys. **15**, 131 (2023).
[39] L. Leone, S. F. E. Oliviero, and A. Hamma, Entropy **23**, 1073 (2021).
[40] S. F. E. Oliviero, L. Leone, F. Caravelli, and A. Hamma, SciPost Physics **10**, 076 (2021).
[41] L. Leone, S. F. E. Oliviero, Y. Zhou, and A. Hamma, Quantum **5**, 453 (2021).
[42] S. F. E. Oliviero, L. Leone, and A. Hamma, Phys. Lett. A **418**, 127721 (2021).
[43] S. F. E. Oliviero, L. Leone, and A. Hamma, Phys. Rev. A **106**, 042426 (2022).
[44] T. Haug and L. Piroli, Phys. Rev. B **107**, 035148 (2023).
[45] C. Chamon, E. R. Mucciolo, and A. E. Ruckenstein, Ann. Phys. **446**, 169086 (2022).
[46] A. K. Ekert, C. M. Alves, D. K. L. Oi, M. Horodecki, P. Horodecki, and L. C. Kwek, Phys. Rev. Lett. **88**, 217901 (2002).
[47] R. Islam, R. Ma, P. M. Preiss, M. Eric Tai, A. Lukin, M. Rispoli, and M. Greiner, Nature (London) **528**, 77 (2015).
[48] A. M. Kaufman, M. E. Tai, A. Lukin, M. Rispoli, R. Schittko, P. M. Preiss, and M. Greiner, Science **353**, 794 (2016).
[49] N. M. Linke, S. Johri, C. Figgatt, K. A. Landsman, A. Y. Matsuura, and C. Monroe, Phys. Rev. A **98**, 052334 (2018).
[50] S. J. van Enk and C. W. J. Beenakker, Phys. Rev. Lett. **108**, 110503 (2012).
[51] A. Elben, B. Vermersch, M. Dalmonte, J. I. Cirac, and P. Zoller, Phys. Rev. Lett. **120**, 050406 (2018).
[52] T. Brydges, A. Elben, P. Jurcevic, B. Vermersch, C. Maier, B. P. Lanyon, P. Zoller, R. Blatt, and C. F. Roos, Science **364**, 260 (2019).
[53] Note that the null set of $M(\cdot)$ does not contain convex combinations of stabilizer states.
[54] H. Zhu, Phys. Rev. A **96**, 062336 (2017).
[55] S. Aaronson and D. Gottesman, Phys. Rev. A **70**, 052328 (2004).
[56] Any permutation of the qubits belongs to the Clifford group because the permutation group S_n is generated by swaps operator and the swap operator S_{ij} between the qubits ith and jth is made out of three CNOTs.
[57] D. Fattal, T. S. Cubitt, Y. Yamamoto, S. Bravyi, and I. L. Chuang, arXiv:quant-ph/0406168.
[58] A. Montanaro, arXiv:1707.04012.
[59] Set $\mathrm{Pur}(\rho_E) = \Theta(2^{-\alpha n})$ and $t = O(\log_2 n)$. Asymptotically (in n) there exist two constants A, B such that $Be^{-\alpha n} \leqslant \mathrm{Pur}(\rho_E) \leqslant Ae^{-\alpha n}$. From Eq. (16), we can thus write $Be^{-\alpha n} \leqslant \mathrm{Pur}(\psi_E) \leqslant A' \,\mathrm{poly}(n) 2^{-\alpha n}$. Therefore, noticing that $\mathrm{poly}(n) = 2^{\log_2 \mathrm{poly}\, n} = 2^{O(\log_2 n)}$, we can write $\mathrm{Pur}(\psi_E) = 2^{-\alpha n + f(n)}$, where $f(n) = O(\log_2 n)$ and $f(n) = \Omega(1)$.
[60] C. D. White, C. J. Cao, and B. Swingle, Phys. Rev. B **103**, 075145 (2021).