Non-Clifford Cost of Random Unitaries

Lorenzo Leone,[1,2,*] Salvatore F.E. Oliviero,[3,2,†] Alioscia Hamma,[4,5,‡] Jens Eisert,[2,§] and Lennart Bittel[2,‖]

[1] *Dipartimento di Ingegneria Industriale, Università degli Studi di Salerno, Via Giovanni Paolo II, 132, 84084 Fisciano (SA), Italy*

[2] *Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany*

[3] *NEST, Scuola Normale Superiore and Istituto Nanoscienze, Consiglio Nazionale delle Ricerche, Piazza dei Cavalieri 7, IT-56126 Pisa, Italy*

[4] *Dipartimento di Fisica 'Ettore Pancini', Università degli Studi di Napoli Federico II, Via Cintia, 80126 Napoli, Italy*

[5] *INFN Sezione di Napoli, Napoli, Italy*

(Received 12 June 2025; revised 15 October 2025; accepted 3 February 2026; published 4 May 2026)

Recent years have enjoyed a strong interest in exploring properties and applications of random quantum circuits. In this work, we explore the ensemble of t-doped Clifford circuits on n qubits, consisting of Clifford circuits interspersed with t single-qubit non-Clifford gates. We establish rigorous convergence bounds toward unitary k-designs, revealing the intrinsic cost in terms of non-Clifford resources in various flavors. First, we analyze the kth order frame potential, which quantifies how well the ensemble of doped Clifford circuits is spread within the unitary group. We prove that a quadratic doping level, $t = \tilde{\Theta}(k^2)$, is both necessary and sufficient to approximate the frame potential of the full unitary group. As a consequence, we refine existing upper bounds on the convergence of the ensemble toward state k-designs. Second, we derive tight bounds on the convergence of t-doped Clifford circuits toward relative-error k-designs, showing that $t = \tilde{\Theta}(nk)$ is both necessary and sufficient for the ensemble to form a relative ε-approximate k-design. Similarly, $t = \tilde{\Theta}(n)$ is required to generate pseudorandom unitaries. All these results highlight that generating random unitaries is extremely costly in terms of non-Clifford resources, and that such ensembles fundamentally lie beyond the classical simulability barrier. Additionally, we analyze high-order doped-Clifford-Weingarten functions to derive analytic expressions for the twirling operator over the ensemble of random doped Clifford circuits, and we establish their asymptotic behavior in relevant regimes.

DOI: 10.1103/25v1-my1x

I. INTRODUCTION

Randomness is a central and powerful tool in quantum physics, with a wealth of applications spanning quantum cryptography [1–5], quantum algorithms [6–10], quantum sensing [11–14], benchmarking and certification [15–18], and quantum communication [19–28]. The use of random unitary operators is also prevalent in providing a coarse-grained description of the evolution of complex quantum systems, yielding insights into the spectra of complex Hamiltonians and fundamental processes and features such as quantum thermalization [29–34], the holographic principle [35–40], notions of quantum scrambling [41–47], and quantum thermodynamics [48]. Given the ubiquity of quantum randomness in both applications and rather fundamental topics, understanding its precise origin constitutes a fundamental question.

In its simplest form, quantum randomness can be achieved by drawing unitary matrices uniformly at random according to the Haar measure over the unitary group $\mathcal{U}_n = U(2^n)$ on n qubits. However, for most applications in many-body physics (for large n), constructing Haar-random unitaries is practically infeasible, for the circuit complexity becoming prohibitively large. A simple counting argument reveals that an exponential number (in n) of elementary gates is needed to approximate a Haar-random unitary with constant precision. However,

*Contact author: loleone@unisa.it
†Contact author: s.oliviero@fu-berlin.de
‡Contact author: alioscia.hamma@unina.it
§Contact author: jense@fu-berlin.de
‖Contact author: l.bittel@fu-berlin.de

for many applications, sampling Haar-random unitaries is too much to ask for anyway. For most practical purposes, one can be content with constructing an ensemble of unitaries $\mathcal{E}$ reproducing up to the first k moments of the distribution, leading to the concept of *unitary k-designs* [44,49–55]. An even weaker—and conceptually somewhat differently motivated—request is that of computational pseudorandomness, that is, indistinguishably from true randomness given computational constraints. Both concepts have proven to be highly fruitful in a wealth of applications, both practically and conceptually [56–60]. Recently, there has been a breakthrough in Refs. [55,58] with the introduction of sets generating unitary k-designs and pseudorandom unitaries with the minimal possible depths, i.e., scaling logarithmically with the number of qubits n. This represents a significant reduction compared to the exponential scaling of gate complexity for Haar-random unitaries.

Despite the existence of extremely shallow circuits capable of generating pseudorandom unitaries, in many quantum hardware architectures involving early quantum computers, resource costs stem not only from the *number* of gates but also from the *types* of gates used. One of the main bottlenecks in quantum error correction is a consequence of the *Eastin-Knill theorem* [61], which asserts that no universal gate set can be implemented transversely—that is, without incurring spatial overhead. As a result, the best that a fault-tolerant, gate-based quantum computation scheme can achieve is to use only a single non-transversal gate which, when combined with a transversal gate set, enables universal quantum computation. A common approach is to use transversal Clifford operations, which are available in many stabilizer *quantum error-correcting codes* [62], and to invest resources in implementing a single non-Clifford gate—making use of what is called *magic*—that is sufficient to promote the Clifford group to a universal gate set.

In this context, the challenge becomes generating unitary designs using the fewest possible non-Clifford gates, since Clifford operations, being transversal, can be effectively considered *free*. Notably, the *Clifford group* alone is sufficient to generate exact unitary 3-design [63,64], a highly useful property for learning algorithms [65] and benchmarking [66–68]. However, it falls short in producing higher-order designs [69]. Consequently, non-Clifford operations become the critical resource required for achieving higher-order unitary designs, while their usage must be minimized to reduce the associated overhead.

In this work, we explore the generation of unitary designs using *t-doped Clifford circuits*—deep Clifford circuits, which we regard as free resources, augmented with t single-qubit resource-intensive non-Clifford gates (as illustrated in Fig. 1)—with a primary focus on determining the minimal use of non-Clifford resources required to generate unitary k-designs. The main question we seek to answer in this work is the following.

> *How well "spread" is the set of doped Clifford circuits and what relative error design does it achieve as a function of the doping t?*

By developing a general framework for computing averages over doped Clifford circuits—an approach of independent interest—this work provides detailed answers to both of the aforementioned questions.

We remark that addressing this question, concerning the quantum randomness attained by doped Clifford circuits, also has implications for scrambling and for quantum advantage experiments [44,70], where Clifford+T architectures play a crucial role.

First, we establish tight bounds on the kth frame potential—a scalar quantity that measures how uniformly an ensemble of unitaries is spread over the unitary group—as a function of the doping parameter t. Specifically, we show that a quadratic doping level, $t = \Theta(k^2 + \log \varepsilon^{-1})$, is both necessary and sufficient to approximate the frame potential within an additive error ε relative to the frame potential of Haar-random unitaries. While one might have expected a linear scaling of the doping with respect to the degree k of the frame potential, this result reveals that generating

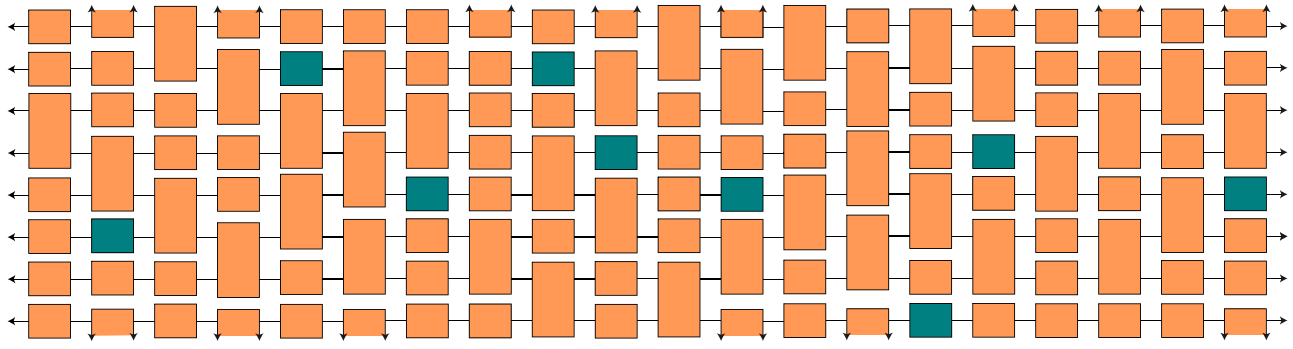

FIG. 1. Pictorial representation of t-doped Clifford circuits. The orange gates are Clifford gates, while the teal gates are single-qubit non-Clifford gates. Hence, non-Clifford gates are interleaved within a circuit that predominantly consists of feasible Clifford operations.

quantum randomness using doped Clifford circuits incurs an intrinsic cost in terms of non-Clifford resources. This convergence bound has a direct application in quantifying the scrambling capabilities of doped Clifford circuits; notably, only a polylogarithmic amount of doping is both necessary and sufficient to achieve maximal scrambling.

Second, we establish tight bounds on the convergence of the t-doped Clifford ensemble toward unitary k-designs. We show that a doping level scaling linearly in both the number of qubits n and the degree k, i.e., $t = \tilde{\Theta}(kn)$ (up to logarithmic factors), is *necessary and sufficient* for the ensemble to form a k-design up to a relative approximation error ε, even for values of ε exponentially large in nk. While the main result of Ref. [51] shows that a doping level independent of the number of qubits n is sufficient to produce a k-design with additive approximation error, our result demonstrates that generating a k-design with relative approximation error is necessarily costly in terms of non-Clifford resources. Similarly, we show that $t = \tilde{\Theta}(n)$ are necessary and sufficient to construct pseudorandom unitaries. The stark contrast between these two results lies in the notion of approximation: additive-error designs guarantee indistinguishability from Haar-random unitaries (up to precision ε) only under a restricted class of protocols, whereas relative-error designs, as well as pseudorandom unitaries, ensure indistinguishability under *any* quantum protocols. In the following sections, we carefully explain our results and place them in context with previously known bounds.

A. Overview of the results

Let us provide an informal overview of our results, also summarized in Table I. Our model of t-doped Clifford circuits consists of randomly drawn Clifford circuits interspersed with single-qubit random gates. As we discuss later, the specific choice of "doping gates" is not crucial to our analysis. In this paper, we mainly examine two commonly used choices: random diagonal single-qubit gates and Haar-random single-qubit gates. Both yield qualitatively similar results. However, the techniques we develop are independent of the particular choice of doping.

1. The spreading of doped Clifford circuits via frame potential

One of the question we address in this paper is how spread is the ensemble of doped Clifford circuits with respect to the amount of doping, quantified by the number t of non-Clifford gates inserted. The frame potential measures whether $\mathcal{E}$ is "spread" or "concentrated" within the full unitary group $\mathcal{U}_n$. Specifically, the frame potential of order k measures how well the finite collection of unitaries $\mathcal{E}^{\otimes k} := \{U^{\otimes k} : U \in \mathcal{E}\}$ is uniformly distributed on $\mathcal{U}_n^{\otimes k}$, and it is defined as

$$\mathcal{F}_{\mathcal{E}}^{(k)} := \mathbb{E}_{U,V\in\mathcal{E}} |\mathrm{tr}(UV^\dagger)|^{2k}. \tag{1}$$

As the name suggests, the frame potential can also be thought of as a "repulsive force": we can think of the unitaries in $\mathcal{E}$ as particles on the unitary group subject to a repulsive force proportional to $|\mathrm{tr}(UV^\dagger)|^k$, and Eq. (1) gives the potential of the configuration [50]. In the extreme case where the ensemble consists of only one element with probability one, the forward evolution perfectly cancels the backward evolution, leading to the maximal possible value of the frame potential. In contrast, if the ensemble $\mathcal{E}$ is sufficiently spread, then two independently drawn samples U and V are likely to be very different, meaning that $|\mathrm{tr}(UV^\dagger)|$ will be very small. As expected, the most spreading ensemble is the ensemble equipped with the uniform measure over the full unitary group, the Haar measure: in fact, this is the equilibrium configuration exactly minimizing the frame potential [50]. It holds that for any ensemble of unitaries, we have $\mathcal{F}_{\mathcal{E}}^{(k)} \geq \mathcal{F}_{\mathrm{Haar}}^{(k)}$. We quantify the spreading of an ensemble $\mathcal{E}$ with the difference $\mathcal{F}_{\mathcal{E}}^{(k)} - \mathcal{F}_{\mathrm{Haar}}^{(k)}$, which is positive, and zero if and only if $\mathcal{E}$ is an (exact) k-design.

Let us first consider the ensemble of unitaries constituted by Clifford operators, belonging to the Clifford group. Such ensemble is discrete and it features a minimum distance between its elements. As such, it is not sufficiently spread across the unitary group,

$$\mathcal{F}_{\mathrm{Cl}}^{(k)} - \mathcal{F}_{\mathrm{Haar}} \geq \exp\big(\Omega(k^2)\big), \tag{2}$$

i.e., the Clifford group gets further away, increasing k, in a quadratic-exponential fashion. The ultimate reason

TABLE I. Summary of results on unitary design constructions and their associated non-Clifford costs. Results (or portions thereof) proven in this work are marked with [⋆].

	non-Clifford cost	Operational meaning
Additive unitary k-design [51,71]	$O(k^4 + k\log\varepsilon^{-1})$	Secure against non-adaptive k-query experiments
Relative unitary k-designs [51][⋆]	$\Theta(nk)$	Secure against any k-query quantum experiment
Pseudorandom unitaries [58][⋆]	$\tilde{\Theta}(n)$	Secure against any poly-time quantum experiment
Unitary frame potential [⋆]	$\tilde{\Theta}(k^2 + \log\varepsilon^{-1})$	Spreading on the unitary group and OTOCs
State k-designs [⋆]	$O(k^2 + \log\varepsilon^{-1})$	Secure against any k-query quantum experiment

for this behavior is that the frame potential of subgroups of the unitary groups is given by the dimension of the commutant, which, for the Clifford group, increases quadratic-exponentially in k [72]. The relevant question then becomes how much doping is necessary for a sufficiently uniform spreading of the doped Clifford ensemble across the unitary group.

The first central result of this work resolves around a tight bound on the frame potential of the ensemble of doped Clifford circuits.

Theorem (Frame potential convergence. Informal of Theorem 3). Let $n \geq \Omega(k^2 + t\log k)$. There exists two constants $A, B = \Theta(1)$ such that the difference between the frame of t-doped Clifford circuits $\mathcal{F}_t^{(k)}$ and the Haar frame potential $\mathcal{F}_{\text{Haar}}^{(k)}$ is upper and lower bounded as

$$2^{\Omega(k^2 - At\log k)} \leq \mathcal{F}_t^{(k)} - \mathcal{F}_{\text{Haar}}^{(k)} \leq 2^{O(k^2 - Bt)}. \tag{3}$$

There are two worth-noticing aspects around the result summarized in Eq. (3). First, the upper bound is completely independent from the number of qubits n of the quantum circuit. However, the result of generating system-size independent additive-error designs with t-doped Clifford circuits was established in the seminal paper [51], and it is well-understood in the literature. The second, and novel, aspect of Eq. (3) is that the bounds showcase a quadratic dependence of t upon the design order k. As the order of the frame potential grows, the convergence requires a quadratic amount of doping. Indeed, to have $\mathcal{F}_t^{(k)} - \mathcal{F}_{\text{Haar}}^{(k)} = \varepsilon$ then a necessary and sufficient amount of doping (up to log factors) is $t = \Theta(k^2 + \log \varepsilon^{-1})$. The result can be understood by the following hand-wavy argument: the ultimate responsible for this convergence bound is the quadratic-exponential growth of the commutant of the Clifford group with respect to k against the factorial growing of the unitary group commutant [72]. Indeed, the doping can effectively "kill" an exponential fraction (in t) of operators within the commutant of the Clifford group, resulting in $\exp(k^2)\exp(-t)$ scaling. The above result already highlights an inherent substantial cost—in terms of non-Clifford resources—associated with generating higher-order designs through doping, a point that will be further explored in the following sections.

2. Implications for approximate state k-designs

Our result in Eq. (3) has implications in the context of unitary k-designs, particularly to *state k-designs*. Given an ensemble of unitaries, it is natural to consider the ensemble of states $\mathcal{S} = \{U|0\rangle \mid U \in \mathcal{E}\}$ induced by $\mathcal{E}$ when acting on the trivial state vector $|0\rangle$. The ensemble of unitaries $\mathcal{E}$ generates an ε-approximate state k-design if and only if $\|\Psi_{\mathcal{E}}^{(k)} - \Psi_{\text{Haar}}^{(k)}\|_1 \leq \varepsilon$, where we define

$$\Psi_{\mathcal{E}}^{(k)} := \mathbb{E}_{U\sim\text{Haar}}(U|0\rangle\langle 0|U^\dagger)^{\otimes k} \tag{4}$$

as the average state of the given ensemble. A key application of our main result in Eq. (3) is in establishing convergence bounds on the ability of doped Clifford circuits to generate approximate state k-designs.

Theorem (Doped Clifford circuits producing state k-designs. See Lemma 18 and Corollary 3). Let $n \geq \Omega(k^2 + t\log k)$. Let $\Psi_t^{(k)}$ the average k-copy state produced by the ensemble of t-doped Clifford circuits, and let $\Psi_{\text{Haar}}^{(k)}$ the Haar k-copy state. Then

$$\|\Psi_{\text{Haar}}^{(k)} - \Psi_t^{(k)}\|_1 \leq \sqrt{\frac{3}{2}\frac{\mathcal{F}_t^{(k)} - \mathcal{F}_{\text{Haar}}^{(k)}}{\mathcal{F}_{\text{Haar}}}}, \tag{5}$$

which, combined with Eq. (3), implies that a quadratic doping level $t = \Omega(k^2 + \log \varepsilon^{-1})$ suffices for doped Clifford circuits to produce an ε-approximate state k-design.

This result constitutes an improvement over previously established bounds. In particular, the best known bound prior to this work follows from the stronger results of Ref. [73], which show that a doping level of $t = \Omega(\log^2 k(k^4 + k\log \varepsilon^{-1}))$ is sufficient to achieve a unitary k-design up to an additive error ε.

Moreover, Eq. (5) provides an operational interpretation of the frame potential. While frame potentials have become a central tool for analyzing quantum randomness and have inspired connections across a broad range of fields—such as circuit complexity lower bounds [74,75] and information scrambling [44]—Eq. (5) directly links the frame potential to the rigorous notion of approximate state k-designs.

3. Applications to quantum scrambling

Scrambling is the phenomenon by which local information—i.e., information encoded in a local subsystem—is dispersed, as a result of quantum evolution, into nonlocal degrees of freedom. Consequently, information is said to be scrambled if no local quantum measurement can recover the initial information. Quantum scrambling is thus a hallmark of chaotic and complex quantum dynamics, making it one of the most intriguing features of quantum evolutions.

Since scrambling is a property of the quantum evolution itself, diagnosing scrambling behavior comes with an intrinsic challenge: to decide whether a unitary U scrambles information, one would in principle need to verify that its action on all possible localized quantum states renders the information unrecoverable. A more practical bottom-up approach is instead to check whether quantum information can be locally recovered via local quantum measurements, which, without loss of generality, can be

taken to be Pauli measurements (as these form a complete operator basis). At the operator level, diagnosing quantum scrambling can then be achieved through exotic correlation functions between local operators. In particular, given a unitary U, one can define k-order Out-of-Time-Ordered Correlation (OTOC) functions [76]. Consider two k-tuples of Pauli operators $A_1,\ldots,A_k$ and $B_1,\ldots,B_k$ and define $\tilde{B}_i := U^\dagger B_i U$. The k-OTOC is then defined as

$$\mathrm{OTOC}_k(U,\{A\},\{B\}) := \frac{1}{d}\mathrm{tr}(A_1\tilde{B}_1\cdots A_k\tilde{B}_k). \tag{6}$$

The careful reader will note that the behavior of $\mathrm{OTOC}_k(U,\{A\},\{B\})$ can strongly depend on the choice of the tuples of observables $\{A\},\{B\}$. To eliminate this spurious dependence, a common approach is to average over all Pauli observables: $\mathrm{OTOC}_k(U) := \mathbb{E}_{\{A\},\{B\}}\,\mathrm{OTOC}_k(U,\{A\},\{B\})$.

A fundamental benchmark of scrambling behavior is given by Haar-random unitary dynamics, which maximally disperse local information into nonlocal degrees of freedom. A natural figure of merit is, therefore, $\delta\,\mathrm{OTOC}_k(U)$, defined as the relative difference between $\mathrm{OTOC}_k(U)$ and the Haar-random average $\mathbb{E}_{U\sim\mathcal{U}_n}\mathrm{OTOC}_k(U)$.

Combining the convergence bound on the frame potential discussed above with the findings of Ref. [44], which relate the frame potential of unitary ensembles to OTOCs, we can estimate the average scrambling behavior of doped Clifford circuits.

Corollary (Scrambling in t-doped Clifford circuits. Informal of Corollary 4). Let $n \geq \Omega(k^2 + t\log k)$ and $c > 0$. The relative difference between the average $\mathrm{OTOC}_k(U)$ of a t-doped Clifford circuit and that of Haar-random unitaries is negligibly small if and only if $t = \tilde{\Theta}(k^2 + \log^{1+c} n)$.

These results establish both upper and lower bounds, revealing the ultimate capabilities and limitations of scrambling in doped Clifford circuits. The conclusion is noteworthy: only a polylogarithmic number of additional non-Clifford gates is sufficient to reproduce scrambling behavior that is indistinguishable—up to negligible relative error—from Haar-like scrambling. In sharp contrast, as we will later demonstrate, relative k-designs (and pseudorandom unitaries) require an extensive $\Omega(n)$ number of non-Clifford gates. This underscores a fundamental distinction: while relative designs entail a linear overhead, reproducing complex scrambling dynamics demands only a polylogarithmic number of non-Clifford resources. These findings are akin to those of Ref. [18], which showed that reproducing the Porter-Thomas distribution of measurement statistics for Haar-random states in the computational basis requires only a polylogarithmic number of non-Clifford gates. This suggests that the two properties, namely *anticoncentration* and *scrambling*, are closely related.

Before concluding this section, one might argue that $\mathrm{OTOC}_k(U)$ inherently washes out the locality aspect of scrambling, since averaging over the full Pauli group includes highly nonlocal measurements. However, because we are concerned with comparing deep unitaries—such as doped Clifford circuits and Haar-random unitaries—locality becomes irrelevant in this regime.

4. The twirling operator over doped Clifford circuits

On a more technical note, proving the tight convergence bounds for t-doped Clifford circuits relies on a formal understanding of the twirling operator over this ensemble. In particular, given an ensemble of unitaries $\mathcal{E}$, the *twirling superoperator* (also commonly referred to as the kth *moment superoperator*) is defined as

$$\Phi_{\mathcal{E}}^{(k)}(O) := \mathbb{E}_{U\in\mathcal{E}}[U^{\otimes k}OU^{\dagger\otimes k}]. \tag{7}$$

If the ensemble $\mathcal{E}$ forms a unitary k-design, then the twirling superoperator coincides with the one over the entire unitary group, i.e., $\Phi_{\mathrm{Haar}}^{(k)}(O) := \mathbb{E}_{U\in\mathcal{U}_n}[U^{\otimes k}OU^{\dagger\otimes k}]$. In Sec. III, we characterize the twirling operator over the doped Clifford circuits, regardless of the particular choice of doping gates, in terms of the elements Ω of the Clifford commutant,

$$\Phi_t^{(k)}(O) = \Phi_{\mathrm{Haar}}^{(k)}(O) + \sum_{\Omega,\Omega'}[\Delta^t\mathcal{W}^{-1}]_{\Omega,\Omega'}\mathrm{tr}(O\Omega)\Omega', \tag{8}$$

the so-called *Pauli monomials*, introduced in Ref. [72]. Here, the components of the matrix $\Delta^t\mathcal{W}^{-1}$ are referred to as *doped-Clifford-Weingarten functions*: the matrix $\mathcal{W}^{-1}$ encodes the Clifford-Weingarten functions, introduced and studied in Refs. [72,77,78], while the matrix Δ captures the influence of doping in the circuit. In Sec. III B, we derive the asymptotic scalings of these functions for $k^2 = O(n)$, which allows for simpler analytic expressions. For the doping architectures analyzed in this paper, one has $\|\Delta\|_\infty \leq C < 1$, with $C = \Theta(1)$ being a constant. Equation (8) illustrates the competition between the Clifford contribution to the twirling, which gets exponentially suppressed with increasing doping t, and the gradual emergence of the twirling operator $\Phi_{\mathrm{Haar}}^{(k)}(\cdot)$ over the full unitary group. From Eq. (8), it immediately follows that whenever $\|\Delta\|_\infty < 1$,

$$\lim_{t\to\infty}\Phi_t^{(k)}(O) = \Phi_{\mathrm{Haar}}^{(k)}(O), \tag{9}$$

the only remaining question being the scaling competition between t, k, and n. The convergence toward Haar random depends on the matrix Δ. In particular, the crux of the lower bound in Eq. (3) resolves around a relation

between the trace of Δ and the *spectral form factor* of the ensemble of non-Clifford gates used to dope the Clifford group and promote it to universality (see Lemma 15). The relationship between the convergence rate of doped Clifford circuits and spectral form factors is intriguing due to their widespread presence across various areas of quantum information. Originally introduced in the context of the Random Matrix Theory [79] as a tool to measure correlations between the energy eigenvalues of quantum systems, spectral form factors have found numerous applications in quantum information theory due to their ability to model complex systems in a simple and effective way. In particular, they have become fundamental tools to characterize chaotic dynamics [80–84], thermalization and ergodicity [85], and also found applications in number theory [86] and quantum transport [87].

5. Convergence of doped Clifford circuits to relative unitary designs

The second question we address in this paper concerns the convergence of t-doped Clifford circuits to unitary k-designs. While Eq. (9) guarantees that, with an appropriately chosen measure, t-doped Clifford circuits always converge to exact unitary designs in the limit of infinite doping, the fundamental question that remains is how much doping t is necessary and sufficient for the ensemble to be effectively considered a unitary k-design on n qubits.

There are essentially two common notions of approximate unitary k-designs for an ensemble $\mathcal{E}$: ε-approximate unitary designs in terms of *additive* error and in terms of *relative* error. The first notion ensures that any quantum algorithm, with resolution ε, making k-fold queries to $U^{\otimes k}$ cannot distinguish $\mathcal{E}$ from unitaries drawn uniformly from the Haar measure. The second, stronger notion guarantees that even quantum algorithms making k (possibly adaptive) queries cannot distinguish $\mathcal{E}$ from the Haar-random unitaries. We refer the reader to Sec. II B for a formal introduction; in the following, we focus on the stronger notion of approximate k-designs in relative error. In particular, an ensemble $\mathcal{E}$ of unitaries is a relative ε-approximate unitary k-design if and only if,

$$(1-\varepsilon)\Phi^{(k)}_{\mathrm{Haar}} \leq \Phi^{(k)}_{t} \leq (1+\varepsilon)\Phi^{(k)}_{\mathrm{Haar}}, \tag{10}$$

where $\Phi - \Phi' \geq 0$ iff $\Phi - \Phi'$ is a completely positive map. Thanks to the characterization of the twirling operator of doped Clifford circuits in Eq. (8), we establish the following convergence bounds to relative approximate k-designs.

Theorem (Optimal convergence to relative unitary designs). Let $\delta > 0$, $0 \leq \varepsilon \leq 2^{n\frac{k}{2}(1-\delta)}$, and $k \leq 2^{\frac{n}{2}}$. Then, for the ensemble of t-doped Clifford circuits $\mathcal{C}_t$, the following results hold.

(1) For $k \leq \delta n/2$, the ensemble $\mathcal{C}_t$ cannot form a relative ε-approximate unitary k-design unless

$$t = \tilde{\Omega}(nk) \text{ (Theorem 4).} \tag{11}$$

(2) For $k = O(\sqrt{n})$, the ensemble $\mathcal{C}_t$ does form an ε-approximate k-design with $t = \tilde{\Omega}(nk)$ ([51], see Proposition 3).
(3) For $k > 8n\delta^{-1}$, the ensemble $\mathcal{C}_t$ cannot form a relative ε-approximate unitary k-design unless

$$t = \Omega(k) \text{ (Theorem 5).} \tag{12}$$

In particular, item (1) and (2) together implies an optimal convergence bound toward relative ε-approximate k-designs with $t = \tilde{\Theta}(nk)$, in the regime $k = O(\sqrt{n})$.

The theorem above implies that constructing relative approximate designs in the context of doped Clifford circuits is highly resource-intensive—even when a relative error exponentially large in nk is allowed. In particular, when $k = O(n)$, the number of non-Clifford gates required scales linearly with both the tensor power k and the number of qubits n (and it is also saturated for $k = O(\sqrt{n})$). Similarly, when $k = \Omega(n)$, the number of non-Clifford gates must grow at least linearly with k. In the intermediate regime $\Omega(n) \leq k \leq O(n^2)$, where the first bound provides a tighter constraint than the second, it follows that at least $t = \Omega(n^2)$ non-Clifford gates are necessary, since any $(k+1)$-design is also at least a k-design. This result strictly generalizes the finding of Ref. [77], where a linear number (in n) of non-Clifford gates was shown to be necessary and sufficient to reproduce an ε-approximate 4-design in relative error.

At this point, the reader may wonder how this result reconciles with the observed convergence of the frame potential that appears independent of n, as well as with the "homeopathic" convergence in additive-error designs reported in Ref. [51]. The key lies in the notion of approximation: while a doping level independent of t is sufficient for the convergence of the frame potential, and for achieving ε-approximate state designs and unitary designs in *additive* error, it does not preclude the existence of a quantum algorithm that, with k queries to U and resolution $\sim\varepsilon$, can distinguish the ensemble from the Haar measure. This leads to the stronger notion of unitary designs in *relative* error. The theorem above shows that such distinguishing algorithms do exist unless the doping is extremely high—specifically, linear in nk. This highlights the fact that an ensemble of unitaries which is indistinguishable from Haar-random unitaries up to the k moment must necessarily lie beyond the classical simulability threshold, which is set by $t = O(\log n)$ [88,89].

B. Remarks on *t*-doped Clifford circuits generating pseudorandom unitaries

As anticipated, a weaker alternative to generating relative ε-approximate unitary k-designs is the construction of pseudorandom unitaries [90]. A set of unitaries $\mathcal{U}$ is called pseudorandom if no polynomial-time quantum algorithm can distinguish whether a unitary was sampled uniformly from $\mathcal{U}$ or from the Haar measure on the full unitary group. In contrast with unitary k-design, pseudorandom unitaries demand indistinguishability only against polynomial-time adversaries.

Although not the primary focus of this work, we briefly examine the construction of pseudorandom unitaries in the framework of t-doped Clifford circuits. We consolidate existing results from the literature to determine the optimal level of doping that is both necessary and sufficient for generating pseudorandom unitaries.

Proposition 1 (Pseudorandom t-doped Clifford circuits). Any pseudorandom set of unitaries must involve at least $\Omega(n)$ non-Clifford gates. Conversely, there exists a pseudorandom set constructed using $\tilde{O}(n)$ (up to log factors) non-Clifford gates.

The proof of the above proposition is entirely based on the results of Refs. [58,91] and is provided in Sec. IV C for completeness. While we have shown that generating relative-error unitary k-designs requires doping with $t = \tilde{\Theta}(nk)$ non-Clifford gates, the proposition above demonstrates that a significantly lower doping level of $t = \tilde{\Theta}(n)$ is both necessary and sufficient for generating pseudorandom unitaries. Although the weaker requirement of indistinguishability from Haar-random unitaries by polynomial-time quantum adversaries allows for more favorable scaling of t, the fact that it still grows linearly with system size underscores the inherent cost of randomness generation in terms of non-Clifford resources, a recurring theme throughout this work.

C. Discussion, outlook, and open questions

In this work, we study how well doped Clifford circuits spread over the unitary group and establish tight bounds on both their frame potential and the approximation of state k-designs. Additionally, we carefully analyze the precise conditions under which the ensemble of doped Clifford circuits constitutes a relative-error unitary k-design, a natural and, at the same time, strong sense of approximation. The recurring theme in our results is that generating randomness is extremely demanding in terms of non-Clifford resources, and, therefore, substantially challenge the fault-tolerant implementation of unitary k-designs. On the other hand, a positive interpretation of the present results is to see them as an encouragement for efforts aimed at producing quantum circuits outside the realm of classical simulability.

The first main result of this work is that a quadratic amount of doping, in terms of the degree k, is both necessary and sufficient for the ensemble to approximate the Haar frame potential. To prove this, we introduce new technical tools for averaging over t-doped Clifford circuits from a mathematical perspective, including the construction of doped-Clifford-Weingarten functions and an analysis of their asymptotic scaling in n. These tools are of independent theoretical interest and are aimed at increasing the portfolio of rigorous methods to study random processes in quantum physics.

Our second main result is a linear lower bound in both n and k for doped Clifford circuits to form a relative ε-approximate k-design. Specifically, we show that the ensemble cannot constitute a relative design unless $t = \Omega(nk)$. This lower bound is optimal in the regime $k = O(\sqrt{n})$: the seminal work of Haferkamp *et al.* [73] establishes a matching upper bound, demonstrating that doped Clifford circuits do yield relative designs when $t = O(nk)$. Seen in this light, our results establish the optimality of the construction from Ref. [73], showing that a linear amount of doping, $t = \tilde{\Theta}(nk)$, is both necessary and sufficient for achieving relative-error unitary designs. Beyond this regime, we show that $\Omega(k)$ many non-Clifford gates are necessary for constructing relative unitary k-designs.

While our results provide a conclusive analysis of the convergence of the frame potential and unitary k-designs in the relative error setting in the regime $k = O(\sqrt{n})$, they leave important open questions regarding other fundamental aspects of unitary k-designs, which we explicitly invite further investigation into.

First, convergence guarantees for t-doped Clifford circuits forming relative ε-approximate k-designs—potentially matching our lower bounds—are currently lacking in the regime $k = \Omega(\sqrt{n})$. Second, we explicitly study Clifford circuits doped with non-Clifford unitaries. However, the same question is also relevant—and, in fact, some of our results carry over directly—to scenarios where the doping is implemented via suitable non-Clifford mid-circuit measurements rather than unitaries, as in what are known as *dynamical quantum circuits* [92]. While the results of Ref. [78] imply convergence to a relative ε-approximate 4-design using $\tilde{\Theta}(n)$ mid-circuit measurements, an analysis of higher-order designs remains lacking.

Third, finally, the frame potential captures the difference between the moment operators associated with an ensemble of unitaries in the 2-norm. Similarly, approximation in relative error for the twirling channel provides a strong indistinguishability notion, even against adversaries capable of adaptive queries to copies of the unitaries. However, alternative approximation measures may be more appropriate and physically plausible in certain contexts. In particular, the diamond distance between the twirling operator defined in Eq. (7) and the Haar twirling

operator has operational significance analogous to the trace distance used in state designs, as formalized through a generalization of the Holevo-Helstrom theorem. Our results do not yield strong implications for this notion of approximation. For this reason, they are also neither contradicting nor incompatible in spirit to the findings of Ref. [51]. Although relative-error designs constitute a robust and practically useful framework—often outperforming additive-error designs—an intriguing open question remains: does the quadratic scaling observed in the convergence of the frame potential also arise in the convergence of additive-error k-designs? Or is this scaling merely a consequence of the spreading properties captured by the frame potential?

While these questions are primarily mathematical in nature, they also invite conceptual investigations into which precise sense of approximation is most natural in what context, and they will be the subject of fruitful future research.

D. Structure of this work

The remainder of this work is organized as follows. We begin in Sec. II by setting the notation and introducing key concepts that are instrumental for proving our main results. In Sec. II D, we briefly review the main findings concerning the Haar measure over the Clifford group—a critical component of our approach—while referring the reader to Ref. [72] for a comprehensive introduction to the topic. In Sec. III, we define the ensemble of t-doped Clifford circuits, introduce an appropriate measure over this ensemble, and present a systematic method for computing averages using doped-Clifford-Weingarten functions—a generalization of the Clifford-Weingarten functions from Ref. [72]. We also derive the asymptotic behavior of these functions in the regime $k = O(\sqrt{n})$. Finally, in Sec. IV, we leverage the framework developed in the preceding sections to prove the main results discussed in Sec. I A.

II. PRELIMINARIES

In this section, we briefly introduce the tools that are necessary to introduce and prove the main results in this work. Let $\mathcal{H}$ be the Hilbert space of n qubits and $d = 2^n$ its dimension. We consider k copies of it $\mathcal{H}^{\otimes k}$. We denote as $\mathcal{B}(\mathcal{H})$ the set of operators acting on $\mathcal{H}$, and $\mathcal{U}_n$ denotes the unitary group on n qubits. We use the notation $[n]$ to denote the set $[n] := \{1, \dots, n\}$, where $n \in \mathbb{N}$. The finite field $\mathbb{F}_2$ consists of the elements $\{0, 1\}$ with addition and multiplication defined modulo 2. For $x \in \mathbb{F}_2^k$, we denote $|x|$ the Hamming weight of x. The space of $k \times m$ binary matrices over $\mathbb{F}_2$ is denoted by $\mathbb{F}_2^{k\times m}$. The set $\mathrm{Sym}(\mathbb{F}_2^{m\times m})$ denotes the set of all symmetric $m \times m$ matrices over the finite field $\mathbb{F}_2$, having a null diagonal. That is,

$$\mathrm{Sym}(\mathbb{F}_2^{m\times m}) := \left\{M \in \mathbb{F}_2^{m\times m} : M^T = M,\, M_{j,j} = 0\ \forall j \in [m]\right\}.$$

Similarly, we define set $\mathrm{Even}(\mathbb{F}_2^{k\times m})$ as the set of binary matrices $V \in \mathbb{F}_2^{k\times m}$ with $m \in [k]$ with column vectors V_i^T of even Hamming weight

$$\mathrm{Even}(\mathbb{F}_2^{k\times m}) := \{V \in \mathbb{F}_2^{k\times m} \ :\ |V_i^T| = 0 \mod 2\ \forall i \in [m]\}. \tag{13}$$

A. Pauli operators

The entire discussion of this work is focusing on capturing properties of quantum circuits consisting on n qubits. To this purpose, let us introduce the qubit Pauli matrices $\{I, X, Y, Z\}$ as

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{14}$$

The Pauli group on n-qubits is defined as the n-fold tensor product of the set of Pauli matrices, multiplied by a phase factor from $\{\pm 1, \pm i\}$. We denote the Pauli group modulo phases as $\mathbb{P}_n := \{I, X, Y, Z\}^{\otimes n}$. Let us define the function

$$\chi(A, B) := \frac{1}{d}\mathrm{tr}(ABA^\dagger B^\dagger), \tag{15}$$

where A and B are operators. The function $\chi(A, B)$ satisfies the symmetry properties,

$$\chi(A, B) = \chi(\phi_A A, \phi_B B), \quad \chi(A, B) = \chi(B^\dagger, A) = \chi(A^\dagger, B^\dagger), \tag{16}$$

where $\phi_A, \phi_B \in \{-1, 1, -i, +i\}$ and the latter follows from the cyclicity of the trace. Note that, the following property holds: $\chi(P, Q)$ can be written as

$$\chi(P, Q) = \begin{cases} 1, & \text{if } [P, Q] = 0, \\ -1, & \text{if } \{P, Q\} = 0, \end{cases} \tag{17}$$

and we have that $PQP = \chi(P, Q)Q$.

A Clifford operator is a unitary operator that maps any Pauli operator to another Pauli operator under conjugation. This family of unitaries is extremely important in quantum information theory, many-body physics [93], and it plays a dominating role in quantum error correction.

Definition 1 (Clifford group). The set of Clifford operators forms a group, called the Clifford group $\mathcal{C}_n$, which is a

subgroup of the unitary group $\mathcal{U}_n$, defined as the normalizer of the Pauli group $\mathbb{P}_n$. Concretely, this means that

$$\mathcal{C}_n = \{U \in \mathcal{U}(2^n) \mid UPU^\dagger \in \mathbb{P}_n, \forall P \in \mathbb{P}_n\}.$$

It is straightforward to verify that this leads to the group property of $\mathcal{C}_n$.

Definition 2 (Stabilizer states). Stabilizer states are pure quantum states obtained by $|0\rangle^{\otimes n}$ via the action of a Clifford unitary operator $C \in \mathbb{C}_n$. The set of stabilizer states on n qubits will be denoted as STAB_n.

B. Unitary k-designs

Unitary designs are sets of unitaries that approximate Haar-random unitaries by reproducing the expectation values of low-order polynomials. In this section, we briefly introduce the relevant concepts and refer the reader to Ref. [52] for a more in-depth overview.

We denote as $\mathrm{d}U$ the Haar measure over the unitary group, which is the only left/right invariant measure on $\mathcal{U}_n$.

Definition 3 (k-fold channel). Let $O \in \mathcal{B}(\mathcal{H})$ be a operator. The k-fold channel (or twirling) is defined as

$$\Phi^{(k)}_{\mathrm{Haar}}(O) := \int \mathrm{d}U\, U^{\otimes k} O U^{\dagger\otimes k}. \tag{18}$$

A general recipe to compute the k-fold channel associated with the unitary group is via the Weingarten functions [94].

Lemma 1 (Weingarten calculus). Let $\mathbf{\Lambda}$ the $k! \times k!$ matrix with components $\mathbf{\Lambda}_{\pi,\sigma} = \mathrm{tr}(T_\pi T_\sigma)$, then for any $O \in \mathcal{B}(\mathcal{H})$ the k-fold twirling in Eq. (18) reads

$$\Phi^{(k)}_{\mathrm{Haar}} = \sum_{\pi,\sigma} (\mathbf{\Lambda}^{-1})_{\pi,\sigma} \mathrm{tr}(T_\pi O) T_\sigma, \tag{19}$$

where the components $(\mathbf{\Lambda}^{-1})_{\pi,\sigma}$ are called Weingarten functions.

Lemma 2 (Asymptotics of Weingarten functions [94]). Let $\mathbf{\Lambda}$ be the $k! \times k!$ matrix with components $\mathbf{\Lambda}_{\pi,\sigma} = \mathrm{tr}(T_\pi T_\sigma)$. Then, in the limit of large d, the components of the inverse matrix $\mathbf{\Lambda}^{-1}$ behave as

$$\begin{aligned} (\mathbf{\Lambda}^{-1})_{\pi,\pi} &= d^{-k} + O(d^{-(k+2)}), \\ (\mathbf{\Lambda}^{-1})_{\pi,\sigma} &= O(d^{-(k+|\pi\sigma|)}), \quad \pi \neq \sigma. \end{aligned} \tag{20}$$

The above asymptotics are extremely useful for various applications, see also Ref. [95]. Let us introduce the concept of unitary k-designs. Denoting $\mathcal{E} \subseteq \mathcal{U}_n$ as an ensemble of unitaries, either discrete or continuous, equipped with a measure $\mathrm{d}\mu(U)$, one can define the k-fold channel of $\mathcal{E}$ on $O \in \mathcal{B}(\mathcal{H})$ simply as

$$\Phi^{(k)}_{\mathcal{E}}(O) := \int_{\mathcal{E}} \mathrm{d}\mu(U)\, U^{\otimes k} O U^{\dagger\otimes k}. \tag{21}$$

This operator is also commonly referred to as the kth moment superoperator. Whenever $\Phi^{(k)}_{\mathcal{E}}(O)$ agrees with the k-fold channel on the full unitary group, we say that $\mathcal{E}$ is (an exact) unitary k-design.

Definition 4 (Unitary k-design). Let $\mathcal{E} \subseteq \mathcal{U}_n$ be an ensemble of unitaries equipped with a measure $\mathrm{d}\mu(U)$. $\mathcal{E}$ is a unitary k-design iff

$$\Phi^{(k)}_{\mathcal{E}}(O) = \Phi^{(k)}_{\mathrm{Haar}}(O), \tag{22}$$

for any $O \in \mathcal{B}(\mathcal{H})$.

Let us now introduce one of the central objects of this work in the context of unitary k-design: the *frame potential* [50]. The frame potential is a scalar function, which can determine whether an ensemble $\mathcal{E}$ is a k-design or not. The frame potential can be seen as a "repulsive force:" The global minimum of the (unitary) frame potential corresponds to an exact unitary design. The frame potential is defined as

$$\mathcal{F}^{(k)}_{\mathcal{E}} := \int_{\mathcal{E}} \mathrm{d}\mu(U)\, \mathrm{d}\mu(V) |\mathrm{tr}(UV^\dagger)|^{2k}. \tag{23}$$

Lemma 3 (Unitary k-design via the frame potential). Let $\mathcal{E}$ be an ensemble of unitaries equipped with the measure $\mathrm{d}\mu(U)$. Then

$$\mathcal{F}^{(k)}_{\mathcal{E}} \geq \mathcal{F}^{(k)}_{\mathrm{Haar}}, \tag{24}$$

with equality if and only if $\mathcal{E}$ is a k-design. Moreover, $\mathcal{F}^{(k)}_{\mathrm{Haar}} = k!$.

Let us now define the concept of approximate unitary k-designs. There are essentially two distinctly different but, at the same time, common definitions of approximate unitary k-designs.

Definition 5 (ε-approximate unitary design in additive error). Let $\varepsilon > 0$. Let $\mathcal{E}$ be an ensemble of unitaries equipped with the measure $\mathrm{d}\mu(U)$. $\mathcal{E}$ is an ε-approximate unitary design in additive error iff

$$\|\Phi_{\mathcal{E}} - \Phi_{\mathrm{Haar}}\|_\diamond \leq \varepsilon, \tag{25}$$

where $\Phi_{\mathcal{E}}$ (resp. Φ_{Haar}) is the k-fold channel defined in Definition 3 and $\|\cdot\|_\diamond$ denotes the diamond norm.

The notion of ε-approximate unitary designs in additive error has the operational meaning that the state produced by any quantum algorithm that queries $U^{\otimes k}$ when is sampled from $\mathcal{E}$ or according to the Haar measure is at most ε close in trace distance. However, recently, the concept of relative error is becoming increasingly prominent in the context of the study of unitary k-designs [58,73,96].

Definition 6 (ε-approximate unitary design in relative error). Let $\varepsilon > 0$. Let $\mathcal{E}$ be an ensemble of unitaries equipped with the measure $\mathrm{d}\mu(U)$. $\mathcal{E}$ is an ε-approximate unitary design in relative error iff

$$(1-\varepsilon)\Phi_{\mathrm{Haar}} \leq \Phi_{\mathcal{E}} \leq (1+\varepsilon)\Phi_{\mathrm{Haar}}, \tag{26}$$

where $\Phi \leq \Phi'$ denotes that $\Phi' - \Phi$ is a completely positive map.

The operational meaning of relative error design in Definition 6 is much stronger than the additive error in Definition 5. This is indeed referred to as the "gold standard" in Ref. [58], making reference to Ref. [52]: if $\mathcal{E}$ is a unitary design in relative error then the state produced by any quantum algorithm that makes k queries to U (which can be adaptive queries even) when U is sampled from $\mathcal{E}$ or from the Haar measure is at most 2ε-close in trace distance (for in-depth discussions, see, e.g., Refs. [52,96] and, for a modern perspective, Ref. [58]). As a matter of fact, Definition 6 implies Definition 5 (up to a factor of 2), while the converse is not true unless exponential factors are in the error:

Lemma 4 (Additive vs. relative error. Lemma 3 in Ref. [52]). Let $\varepsilon > 0$. Let $\mathcal{E}$ be an ensemble of unitaries equipped with the measure $\mathrm{d}\mu(U)$. If $\mathcal{E}$ is ε-approximate k design in additive error, then it is a $d^{2k}\varepsilon$-approximate k-design in relative error. Conversely, if $\mathcal{E}$ is an ε-approximate k-design in relative error, then $\mathcal{E}$ is an 2ε-approximate k-design in additive error.

Lastly, let us introduce the concept of k-expander, tightly related to the relative approximate k-design. We refer to Ref. [52] for more details. Let us define the kth moment operator,

$$M_{\mathcal{E}}^{(k)} := \mathbb{E}_{U\in\mathcal{E}}(U^{\otimes k} \otimes \bar{U}^{\otimes k}), \tag{27}$$

where $\bar{U}$ denotes the complex conjugate of U. Let $M_{\mathrm{Haar}}^{(k)}$ be the equivalent for the Haar measure.

Then a δ-approximate k-expander $\mathcal{E}$ is defined as [7,52,73,97]

$$\|M_{\mathcal{E}}^{(k)} - M_{\mathrm{Haar}}^{(k)}\|_\infty \leq \delta. \tag{28}$$

As anticipated, there is a intimate relation between δ-approximate k-expander and relative ε-approximate k-design, as the following lemma, due to Ref. [52], summarizes.

Lemma 5 (k-expander and relative k-designs, Lemma 4 in Ref. [52]). Denote with $G_{\mathcal{E}}$ the smallest ε for which $\mathcal{E}$ forms an ε-approximate k-design in relative error, so that Eq. (26) holds true. Then

$$2^{-nk/2-1}\|M^{(k)} - M_{\mathrm{Haar}}^{(k)}\|_\infty \leq G_{\mathcal{E}} \leq 2^{2kn}\|M^{(k)} - M_{\mathrm{Haar}}^{(k)}\|_\infty. \tag{29}$$

C. State k-designs and state frame potential

In this section, we introduce the concept of *state k-design*, also called *projective k-design*. This concept is closely related to the one of unitary design but restricted on pure states (i.e., rank one projectors). For the remainder of this work, we indicate as $|\psi\rangle$ state vectors and ψ their density matrices, i.e., quantum states.

Definition 7 (Projective k-design). Consider a set of state vectors $\mathcal{S} = \{|\psi\rangle\}$, equipped with a measure $\mathrm{d}\mu(\psi)$. Defining

$$\Psi_{\mathcal{S}}^{(k)} := \int \mathrm{d}\mu(\psi)\, \psi^{\otimes k}, \tag{30}$$

then $\mathcal{S}$ is a projective k-design iff $\Psi_{\mathcal{S}}^{(k)} = \Psi_{\mathrm{Haar}}^{(k)}$, where $\Psi_{\mathrm{Haar}}^{(k)} := \Phi_{\mathrm{Haar}}^{(k)}(\psi^{\otimes k})$.

Proposition 2 (Haar twirling). For any pure state ψ, the Haar twirling over $\psi^{\otimes k}$ reads

$$\Phi_{\mathrm{Haar}}^{(k)}(\psi^{\otimes k}) = \frac{\Pi_{\mathrm{sym}}}{\mathrm{tr}(\Pi_{\mathrm{sym}})}, \tag{31}$$

where $\Pi_{\mathrm{sym}} = (1/k!)\sum_{\pi\in S_k} T_\pi$ is the projector onto the symmetric subspace of $\mathcal{H}^{\otimes k}$, S_k is the symmetric group of order k and T_π are representation of permutation operators $\pi \in S_k$.

Given an ensemble $\mathcal{E}$ of unitary operators, $\mathcal{E}$ induces an ensemble of states when applied on a reference state vector $|0\rangle$, i.e., $\mathcal{S}_{\mathcal{E}} := \{U|0\rangle, U \in \mathcal{E}\}$ such that $|\mathcal{S}_{\mathcal{E}}| \leq |\mathcal{E}|$. We have that if $\mathcal{E}$ is unitary k-design, then $\mathcal{S}_{\mathcal{E}}$ is a state k-design. The converse, naturally, does not hold. However, if $\mathcal{E}$ is not a unitary k-design, it can happen that, with a accurate choice of the reference state vector $|0\rangle$ in the definition

of $\mathcal{S}_\mathcal{E}$, then $\mathcal{S}_\mathcal{E}$ is state k-design. One relevant example is given by the Clifford orbit of quantum state. In particular, Ref. [69] has shown the existence of fiducial quantum state vectors $|\psi\rangle$ promoting the Clifford orbit of $|\psi\rangle$ to a state 4-design. Similarly, Ref. [98] has shown that ensembles of matrix product states provide good fiducial vectors for state 4-designs as their stabilizer entropies attain universal values.

Similarly to the case of unitary design, there is a notion of a *state frame potential* for projective k-design, which corresponds to the *purity*

$$\mathrm{Pur}(\rho) := \mathrm{tr}(\rho^2) \tag{32}$$

of the state vector $\Psi_S^{(k)}$ in Eq. (30). Similarly, to the frame potential in Lemma 3 [50], also the state frame potential provides a necessary and sufficient condition for state k-designs.

Lemma 6 (State design via frame potential). Let $\mathcal{S}$ be an ensemble of states equipped with the measure $\mathrm{d}\mu(\psi)$. Then

$$\mathrm{Pur}(\Psi_\mathcal{E}^{(k)}) \geq \mathrm{Pur}(\Psi_{\mathrm{Haar}}^{(k)}), \tag{33}$$

with equality if and only if $\mathcal{S}$ is a projective k-design. Moreover,

$$\mathrm{Pur}(\Psi_{\mathrm{Haar}}^{(k)}) = \frac{1}{\mathrm{tr}(\Pi_{\mathrm{sym}})}, \tag{34}$$

where $\mathrm{tr}(\Pi_{\mathrm{sym}}) = \binom{d+k-1}{k}$.

We can define a notion of ε-approximate projective design in trace distance.

Definition 8 (ε-approximate state design in trace distance). Let $\varepsilon > 0$. Let $\mathcal{S}$ be an ensemble of pure states equipped with a measure $\mathrm{d}\mu(\psi)$. $\mathcal{S}$ is an ε-approximate state k-design iff

$$\|\Psi_\mathcal{E}^{(k)} - \Psi_{\mathrm{Haar}}^{(k)}\|_1 \leq \varepsilon. \tag{35}$$

The notion of ε-approximate state k-design in trace distance has a clear operational meaning. Through the Holevo-Helstrom theorem, k-copies of a state drawn uniformly at random from the set $\mathcal{S}$ cannot be dististinguished from a Haar-random state with probability greater than $\frac{1}{2} + \varepsilon$ if $\mathcal{S}$ is an ε-approximate k-design.

The state frame potential, being related to the purity of the state vector $\Psi_\mathcal{E}$, is amenable for practical computation and also provides a useful bounds to the less computable notions of approximate state designs Definition 8. For this purpose, it is useful to define the *relative frame potential* of an ensemble $\mathcal{S}$ of states, and relate it to Definition 8.

Definition 9 (Relative frame potential). Let $\mathcal{S}$ be an ensemble of states equipped with the measure $\mathrm{d}\mu(\psi)$. The relative frame potential is defined as

$$\mathcal{R}_\mathcal{S}^{(k)} := \frac{\mathrm{Pur}(\Psi_\mathcal{S}^{(k)}) - \mathrm{Pur}(\Psi_{\mathrm{Haar}}^{(k)})}{\mathrm{Pur}(\Psi_{\mathrm{Haar}}^{(k)})}. \tag{36}$$

Lemma 7 (Bounds via state relative frame potential). Let $\mathcal{S}$ be an ensemble of pure states equipped with a measure $\mathrm{d}\mu(\psi)$, and $\Psi_\mathcal{S}^{(k)}$ be the mixed state associated with $\mathcal{S}$. Let $\mathcal{R}_\mathcal{S}^{(k)}$ be the relative frame potential of the ensemble $\mathcal{S}$. Then the general bounds

$$\|\Psi_\mathcal{S}^{(k)} - \Psi_{\mathrm{Haar}}^{(k)}\|_1 \leq \sqrt{\mathcal{R}_\mathcal{S}^{(k)}} \tag{37}$$

hold. In particular, Eq. (37) implies that if the relative state frame potential $\mathcal{R}_\mathcal{S}^{(k)} \leq \varepsilon^2$ then $\mathcal{S}$ is an ε-approximate state design in trace distance.

Proof. The bound in Eq. (37) descends from a useful property of Schatten p-norms, i.e., given $A \in \mathcal{B}(\mathcal{H})$, then $\|A\|_1 \leq \sqrt{\mathrm{rank}(A)}\|A\|_2$. Hence, we can readily apply it to Eq. (35), to get

$$\|\Psi_\mathcal{S}^{(k)} - \Psi_{\mathrm{Haar}}^{(k)}\|_1 \leq \sqrt{\mathrm{rank}(\Psi_\mathcal{S}^{(k)} - \Psi_{\mathrm{Haar}}^{(k)})}\|\Psi_\mathcal{S}^{(k)} - \Psi_{\mathrm{Haar}}^{(k)}\|_2. \tag{38}$$

From Lemma 6, we know that $\|\Psi_\mathcal{S}^{(k)} - \Psi_{\mathrm{Haar}}^{(k)}\|_2^2 = \mathrm{Pur}(\Psi_{\mathrm{Haar}}^{(k)}) - \mathrm{Pur}(\Psi_\mathcal{S}^{(k)})$. Let us now show that

$$\mathrm{rank}(\Psi_\mathcal{S}^{(k)} - \Psi_{\mathrm{Haar}}^{(k)}) \leq \mathrm{rank}(\Psi_{\mathrm{Haar}}^{(k)}). \tag{39}$$

First, we have that $\mathrm{rank}(\Psi_{\mathrm{Haar}}^{(k)}) = \mathrm{rank}(\Pi_{\mathrm{sym}})$. Then, since $T_\pi |\psi^{\otimes k}\rangle = |\psi^{\otimes k}\rangle$ for any ψ, we also have that $\Pi_{\mathrm{sym}}\Psi_\mathcal{S}^{(k)} = \Psi_\mathcal{S}^{(k)}$, i.e., Π_{sym} acts identically on $\Psi_\mathcal{S}^{(k)}$, from which it immediately follows that $\mathrm{rank}(\Psi_\mathcal{S}^{(k)} - \Psi_{\mathrm{Haar}}^{(k)}) \leq \mathrm{rank}(\Psi_{\mathrm{Haar}}^{(k)})$. Recalling that $\Psi_{\mathrm{Haar}}^{(k)}$ is proportional to a projector, i.e., Π_{sym}, the statement follows. The validity of Eq. (37) then follows. ■

D. Haar average over the Clifford group

In this section, we summarize the findings of Ref. [72] regarding the average over the Clifford group, as it will be instrumental for the proof of our main result. Let us first define the set of *reduced Pauli monomials*, which will be referred to as the set of Pauli monomials for brevity.

Definition 10 (Pauli monomials). Let $k, m \in \mathbb{N}$, $V \in \mathrm{Even}(\mathbb{F}_2^{k\times m})$ with independent column vectors and $M \in$

$\mathrm{Sym}(\mathbb{F}_2^{m\times m})$. A Pauli monomial, denoted as $\Omega(V,M)\in\mathcal{B}(\mathcal{H}^{\otimes k})$, is defined as

$$\Omega(V,M):=\frac{1}{d^m}\sum_{\boldsymbol{P}\in\mathbb{P}_n^m}P_1^{\otimes v_1}P_2^{\otimes v_2}\cdots P_m^{\otimes v_m}\times\left(\prod_{\substack{i,j\in[m]\\ i<j}}\chi(P_i,P_j)^{M_{i,j}}\right), \tag{40}$$

where $\chi(P_i,P_j)$ is defined in Eq. (15), and we denote a string of Pauli operators in $\mathbb{P}_n$ as $\boldsymbol{P}:=P_1,\ldots,P_k$. We define the set of reduced Pauli monomials as

$$\mathcal{P}:=\{\Omega(V,M)\,|\,V\in\mathrm{Even}(\mathbb{F}_2^{k\times m})\ :\ \mathrm{rank}(V)=m,\ M\in\mathrm{Sym}(\mathbb{F}_2^{m\times m}),\ m\in[k-1]\}. \tag{41}$$

Furthermore, we define $\mathcal{P}_U:=\mathcal{P}\cap\mathcal{U}_{nk}$ the set of reduced Pauli monomial which are also unitary and $\mathcal{P}_P$ the set of reduced Pauli monomials proportional to projectors, formally defined as

$$\mathcal{P}_P:=\{\Omega(V,0)\,|\,V\in\mathrm{Even}(\mathbb{F}_2^{k\times m})\ :\ |V_i^T|=0 \mod 4,\ V_i^T\cdot V_j^T=0 \mod 2,\ i,j\in[m]\}. \tag{42}$$

Lemma 8 (Relevant properties of Pauli monomials). The following facts hold [72]:

(1) The commutant of the Clifford group is spanned by $\mathcal{P}$.
(2) For $n\le k-1$, $\mathcal{P}$ contains linearly independent operators.
(3) $|\mathcal{P}|=\prod_{i=0}^{k-2}(2^i+1)$, and the following bounds holds $2^{\frac{k^2-3k-1}{2}}\le|\mathcal{P}|\le 2^{\frac{k^2-3k+12}{2}}$.
(4) For $\Omega\in\mathcal{P}$, then $\Omega=\omega^{\otimes n}$ (i.e., factorizes on qubits), where

$$\omega=\frac{1}{2^m}\sum_{\boldsymbol{P}\in\{I,X,Y,Z\}^m}P_1^{\otimes v_1}P_2^{\otimes v_2}\cdots P_m^{\otimes v_m}\times\left(\prod_{\substack{i,j\in[m]\\ i<j}}\chi(P_i,P_j)^{M_{i,j}}\right). \tag{43}$$

(5) $\Omega^\dagger=\Omega^T$ for all $\Omega\in\mathcal{P}$.
(6) $\Omega=\Omega_U\Omega_P$ where $\Omega_U\in\mathcal{P}_U$, $\Omega_P\in\mathcal{P}_P$ for all $\Omega\in\mathcal{P}$.
(7) $\|\Omega\|_1=\mathrm{tr}(\Omega_P(V,0))=d^{k-\mathrm{rank}(V)}$ where $\Omega_P(V,0)\in\mathcal{P}_P$, for any $\Omega\in\mathcal{P}$.
(8) $\mathrm{tr}(\Omega\rho^{\otimes k})\le 1$ for any state ρ and any $\Omega\in\mathcal{P}$.

Lemma 9 (Twirling over the Clifford group [72]). Consider the Clifford group $\mathcal{C}_n$. The k-fold channel of $\mathcal{C}_n$, denoted as $\Phi_{\mathrm{Cl}}(\cdot)$, on an operator O reads

$$\Phi_{\mathrm{Cl}}(O)=\sum_{\Omega,\Omega'\in\mathcal{P}}(\mathcal{W}^{-1})_{\Omega,\Omega'}\mathrm{tr}(\Omega^\dagger O)\Omega', \tag{44}$$

where $\mathcal{P}$ is the set of Pauli monomials in Definition 10, and $\mathcal{W}^{-1}$ are the Clifford-Weingarten functions introduced in [72], obtained by inverting the Gram-Schmidt matrix $\mathcal{W}_{\Omega,\Omega'}:=\mathrm{tr}(\Omega^\dagger\Omega')$.

Lemma 10 (Average stabilizer state [99]). Let $\sigma\in\mathrm{STAB}_n$, then

$$\Phi_{\mathrm{Cl}}^{(K)}(\sigma^{\otimes k})=\mathbb{E}_\sigma\sigma^{\otimes k}=\frac{1}{Z_n}\sum_{\Omega\in\mathcal{P}}\Omega=\frac{1}{Z_n}\sum_{\Omega_P\in\mathcal{P}_P}c_\Omega\Pi_U\Omega_P\Pi_U, \tag{45}$$

where $\Pi_U:=(1/|\mathcal{P}_U|)\sum_{\Omega\in\mathcal{P}_U}\Omega$ and $Z_n=2^n\prod_{i=0}^{k-2}(2^n+2^i)$. The coefficient c_Ω is determined by the following identity:

$$\sum_{\Omega\in\{\Omega_U\Omega_P\Omega_{U'}:\Omega_U,\Omega_{U'}\in\mathcal{P}_U\}}\Omega_P=c_\Omega\Pi_U\Omega_P\Pi_U, \tag{46}$$

which is the multiplicity of the projection. The state frame potential, coinciding with its purity, is given by $\mathrm{Pur}(\Phi_{\mathrm{Cl}}^{(K)}(\sigma^{\otimes k}))=|\mathcal{P}|/Z_n$.

In what follows, we summarize the asymptotic results proven in Ref. [72], that will be particularly useful to derive the expression of the twirling operator for t-doped Clifford circuits.

Lemma 11 (Asymptotics of Clifford-Weingarten functions [72]). Let $\mathcal{W}$ the Gram-Schmidt matrix, and let $\mathcal{W}^{-1}$ be its inverse. Let $n\ge k^2-3k+13$. Then the following properties hold.

(1) $\mathcal{W}$ is symmetric.
(2) $1\le\mathcal{W}_{\Omega,\Omega'}\le d^{k-1}$, $\Omega\ne\Omega'\in\mathcal{P}$.
(3) $\mathcal{W}_{\Omega,\Omega}=d^k$, $\Omega\in\mathcal{P}$.
(4) $d^k-|\mathcal{P}|d^{k-1}\le\lambda(\mathcal{W})\le d^k+|\mathcal{P}|d^{k-1}$ where $\lambda(\mathcal{W})$ is any eigenvalue of $\mathcal{W}$.
(5) $\det(\mathrm{diag}\,\mathcal{W})\left(1-|\mathcal{P}|^2/d\right)\le\det(\mathcal{W})\le\det(\mathrm{diag}\,\mathcal{W})\left(1+2|\mathcal{P}|^2/d\right)$, if $n\ge k^2-3k+13$.
(6) The following asymptotics hold for Clifford-Weingarten functions:

$$\left|(\mathcal{W}^{-1})_{\Omega,\Omega}-\frac{1}{d^k}\right|\le\frac{6|\mathcal{P}|^2}{d^{k+1}},\qquad |(\mathcal{W}^{-1})_{\Omega,\Omega'}|\le\frac{5|\mathcal{P}|^2}{d^{k+1}}. \tag{47}$$

III. TWIRLING OPERATOR OVER THE DOPED-CLIFFORD GROUP

In this section, we define and compute the twirling operator for t-doped Clifford unitaries. t-doped Clifford unitaries are generated by applying Clifford gates augmented with at most t non-Clifford gates, see Fig. 1. It is well known that enhancing the Clifford group with any gate outside the Clifford set yields a universal gate set. Studying t-doped Clifford unitaries thus serves as a framework for studying the transition to the universality regime by controlling the degree of non-Cliffordness introduced in the unitary transformation.

Before proceeding, we emphasize an important distinction: while the twirling operators for both the unitary and Clifford groups can be simply expressed as elements of the commutant of the unitary (or Clifford) group, t-doped Clifford unitaries do not form a group (as can be readily verified). Consequently, we cannot employ the Haar measure, which is the uniform measure over groups. Instead, we must define an appropriate measure over t-doped Clifford circuits, noting that this choice is not unique.

A. Doped Clifford measures

Definition 11 (t-doped Clifford measure). Denote $\mathcal{C}_n^{(t)}$ the ensemble of t-doped Clifford circuits, defined as

$$\mathcal{C}_n^{(t)} := \Bigg\{ U_t \,:\, U_t = \left(\prod_{i=1}^{t} C_i K_i\right) C_0,\ C_i \in \mathcal{C}_n\ \forall i \in [0,t], \\ K_i \in \mathcal{U}_1\ \forall i \in [1,t] \Bigg\}, \tag{48}$$

then a random t-doped Clifford circuit is defined according to the following measure, denoted as $\mu(U_t)$:

$$\mu(U_t) = \mu_{\mathrm{Cl}} \star \mu_1 \star \mu_{\mathrm{Cl}} \cdots \star \mu_{\mathrm{Cl}}, \tag{49}$$

where $\star$ denotes the convolution between measures, while μ_{Cl} is the uniform measure over the n-qubit Clifford group, while μ_1 is *any* measure over the single-qubit unitary group. Notice that μ_1 can also be deterministically a single gate. Let $f(U_t)$ be any function of $U_t \in \mathcal{C}_n^{(t)}$, we denote the average over $\mathcal{C}_n^{(t)}$ as

$$\int \mathrm{d}U_t f(U_t) := \int_{\mathrm{Cl}} \mathrm{d}C_t \int_{\mu_1} \mathrm{d}\mu_1(K_t) \cdots \int_{\mathrm{Cl}} \mathrm{d}C_0 f(U_t), \tag{50}$$

where $\mathrm{d}C := \mathrm{d}\mu_{\mathrm{Cl}}(C)$. Given $O \in \mathcal{B}(\mathcal{H}^{\otimes k})$, the twirling operator $\Phi_t^{(k)}(O)$ reads

$$\Phi_t^{(k)}(O) := \int \mathrm{d}U_t U_t^{\otimes k} O U_t^{\dagger \otimes k}. \tag{51}$$

Before explicitly computing the twirling operator through Weingarten calculus, let us give an alternative expression that we can use later. We can parametrize single-qubit unitaries K_i (up to an irrelevant phase) with a vector $\vec{\alpha} = (\alpha_x, \alpha_y, \alpha_z)$ as $K_i(\vec{\alpha}) = e^{i\vec{\alpha}\cdot\vec{\sigma}}$ where $\vec{\sigma} = (X, Y, Z)$. Hence the measure $\mathrm{d}\mu_1(K_i) = \mathrm{d}\vec{\alpha}$ can be conveniently thought of as a measure on the vector $\vec{\alpha}$. That said, we have the following alternative expression:

$$\Phi_t^{(k)}(O) = \mathbb{E}_{\vec{P}_i} \int \mathrm{d}\vec{\alpha}_1 \cdots \mathrm{d}\vec{\alpha}_t U(\alpha_t) \cdots U(\alpha_1) \Phi_{\mathrm{Cl}}^{(k)}(O) \\ \times U_1^\dagger(\alpha_1) \cdots U_t^\dagger(\alpha_t), \tag{52}$$

where $U_i(\alpha_i) = e^{i\vec{\alpha}\cdot\vec{P}_i}$ where $\vec{P}_i = (P_{i,x}, iP_{i,x}P_{i,z}, P_{i,z})$ and P_x, P_z are a pair of anticommuting Pauli randomly choosen from the Pauli group $\mathbb{P}_n$ (modulo identity), and $\mathbb{E}_{\vec{P}_i}$ denotes the uniform average over this choice. The above expression follows directly from the fact that random Clifford operators map Pauli operators uniformly to other Pauli operators.

That said, Eq. (52) is simply an alternative way to express the twirling operator. While it has the advantage of making the dependence on the Clifford twirling $\Phi_{\mathrm{Cl}}^{(k)}(O)$ explicit, it is not particularly informative on its own. However, we can leverage the Clifford-Weingarten calculus to formally compute $\Phi_t^{(k)}(O)$. In particular, in the next theorem, we derive the explicit form of the doped twirling operator in terms of doped-Clifford-Weingarten functions, introduced in Ref. [77] and also used in Ref. [18].

Theorem 1 (t-doped twirling operator). The t-doped twirling operator $\Phi_t^{(k)}(O)$ reads

$$\Phi_t^{(k)}(O) = \sum_{\Omega \in \mathcal{P}} c_\Omega^{(t)}(O)\Omega, \quad c_\Omega^{(t)}(O) \\ := \sum_{\Omega'} [(\mathcal{W}^{-1}\mathcal{T})^t \mathcal{W}^{-1}]_{\Omega,\Omega'} \mathrm{tr}(\Omega' O), \tag{53}$$

where $\mathcal{W}^{-1}$ are the Clifford-Weingarten functions, $c_{\Omega'}(O) = \sum_{\Omega''} (\mathcal{W}^{-1})_{\Omega'\Omega''} \mathrm{tr}(\Omega'' O)$, see Sec. 9. We defined $\mathcal{T}$ is a $|\mathcal{P}| \times |\mathcal{P}|$ matrix with components $\mathcal{T}_{\Omega,\Omega'} = \mathrm{tr}(\Omega^\dagger \kappa(\Omega'))$, where we defined $\kappa(\Omega) := \int_{\mu_1} \mathrm{d}\mu_1(K)\, K^{\otimes k} \Omega K^{\dagger \otimes k}$ the single twirling with measure μ_1.

Proof. The proof proceeds by induction on t. For $t = 0$, we have $\Phi_t^{(k)}(O) \equiv \Phi_{\mathrm{Cl}}^{(k)}(O)$ by definition, which proves the statement for $t = 0$. Let us suppose that $\Phi_t^{(k)}(O)$ is indeed of the form in Eq. (53), and let us show it for $t + 1$.

By Definition 11, we have

$$\begin{aligned}\Phi_{t+1}^{(k)}(O) &= \int_{\mathrm{Cl}} \mathrm{d}C \int_{\mu_1} \mathrm{d}\mu_1(K)\, C^{\otimes k} K^{\otimes k} \Phi_t^{(k)}(O) K^{\dagger\otimes k} C^{\dagger\otimes k} \\ &= \sum_{\Omega\in\mathcal{P}} c_\Omega^{(t)}(O) \int_{\mathrm{Cl}} \mathrm{d}C \int_{\mu_1} \mathrm{d}\mu_1(K)\, C^{\otimes k} K^{\otimes k} \\ &\quad \times \Omega K^{\dagger\otimes k} C^{\dagger\otimes k} \\ &= \sum_{\Omega} c_\Omega^{(t)}(O) \sum_{\Omega'\Omega''} (\mathcal{W}^{-1})_{\Omega\Omega''} \mathrm{tr}(\Omega''^{\dagger}\kappa(\Omega))\Omega' \\ &= \sum_{\Omega,\Omega'} (\mathcal{W}^{-1}\mathcal{T})_{\Omega,\Omega'} c_\Omega^{(t)}(O)\Omega' \\ &= \sum_{\Omega,\Omega'} [(\mathcal{W}^{-1}\mathcal{T})^{t+1}]_{\Omega,\Omega'} c_\Omega(O)\Omega', \end{aligned} \tag{54}$$

which proves the statement. ■

At first glance, the matrix $(\mathcal{W}^{-1}\mathcal{T})^t\mathcal{W}^{-1}$ may appear as a natural generalization of the Clifford-Weingarten functions discussed in Lemma 9. However, in the following, we present a formulation that is more directly applicable to practical applications, as it explicitly incorporates the dependence on the twirling over the full unitary group.

Lemma 12 (Convergence toward Haar random). Let $\mathcal{W}^{-1}$ be the matrix of Clifford-Weingarten functions and consider the matrix $\mathcal{W}^{-1}\mathcal{T}$. Define $\mathcal{L}$ as a $|\mathcal{P}| \times |\mathcal{P}|$ matrix with components

$$\mathcal{L}_{\Omega,\Omega'} := \begin{cases} \sum_{\sigma\in S_k} (\mathbf{\Lambda}^{-1})_{\pi\sigma} \mathcal{W}_{\sigma\Omega'}, & \pi \in S_k, \\ 0 & \Omega \notin S_k. \end{cases} \tag{55}$$

Define $\Delta := \mathcal{W}^{-1}\mathcal{T} - \mathcal{L}$. The doped-Clifford operator reads

$$\begin{aligned}\Phi_t^{(k)}(O) &= \sum_{\Omega,\Omega'} [\mathbf{\Lambda}^{-1} + \Delta^t \mathcal{W}^{-1}]_{\Omega,\Omega'} \mathrm{tr}(O\Omega)\Omega' = \Phi_{\mathrm{Haar}}^{(k)}(O) \\ &\quad + \sum_{\Omega,\Omega'} [\Delta^t \mathcal{W}^{-1}]_{\Omega,\Omega'} \mathrm{tr}(O\Omega)\Omega'. \end{aligned} \tag{56}$$

The doped-Clifford-Weingarten functions are thus defined through the twirling operator as the components of the matrix $[\Delta^t\mathcal{W}^{-1}]_{\Omega,\Omega'}$.

Proof. For simplicity we denote the non-zero components as $\mathcal{L}_{\pi\Omega} = \sum_\sigma (\mathbf{\Lambda}^{-1})_{\pi,\sigma}\mathcal{W}_{\sigma\Omega}$, because the other components of $\mathcal{L}$ are zero by definition. Using that $\mathcal{T}_{\Omega\pi} = \mathcal{T}_{\pi\Omega} = \mathcal{W}_{\Omega\pi} = \mathcal{W}_{\pi\Omega}$, we now prove four properties of the matrix $\mathcal{L}$.

(1) Let us show that $\mathcal{L}$ is a projector, i.e., $\mathcal{L}^2 = \mathcal{L}$:

$$\begin{aligned}(\mathcal{L}^2)_{\Omega\Omega'} &= \sum_\tau \mathcal{L}_{\Omega\tau}\mathcal{L}_{\tau\Omega'} = \sum_{\pi,\sigma,\tau} (\mathbf{\Lambda}^{-1})_{\Omega\pi} \\ &\quad \times \mathcal{W}_{\pi\tau}(\mathbf{\Lambda}^{-1})_{\tau\sigma}\mathcal{W}_{\sigma\Omega'} \end{aligned} \tag{57}$$

$$\begin{aligned}&= \sum_{\pi,\sigma,\tau} (\mathbf{\Lambda}^{-1})_{\Omega\pi}\mathbf{\Lambda}_{\pi\tau}(\mathbf{\Lambda}^{-1})_{\tau\sigma}\mathcal{W}_{\sigma\Omega'} \\ &= \sum_{\pi,\sigma} \delta_{\pi\sigma}(\mathbf{\Lambda}^{-1})_{\Omega\pi}\mathcal{W}_{\sigma\Omega'} = \mathcal{L}, \end{aligned} \tag{58}$$

where we used the fact that the sum runs only over $\tau \in S_k$ because $\mathcal{L}_{\Omega\Omega'} = 0$ for $\Omega \notin S_k$ and that $\mathcal{W}_{\pi\sigma} = \mathbf{\Lambda}_{\pi\sigma}$, i.e., the Gram-matrix $\mathcal{W}$ equals Λ restricted to permutation space.

(2) Let us show that $\mathcal{L}(\mathcal{W}^{-1}\mathcal{T}) = \mathcal{L}$,

$$\begin{aligned}[\mathcal{L}(\mathcal{W}^{-1}\mathcal{T})]_{\pi\Omega'} &= \sum_{\sigma,\Omega'',\Omega'''} (\mathbf{\Lambda}^{-1})_{\pi\sigma} \\ &\quad \mathcal{W}_{\pi\tau}(\mathcal{W}^{-1})_{\tau\Omega'''}\mathcal{T}_{\Omega'''\Omega'} \end{aligned} \tag{59}$$

$$= \sum_{\sigma,\Omega'''} \delta_{\pi\Omega'''}(\mathbf{\Lambda}^{-1})_{\pi\sigma}\mathcal{T}_{\Omega'''\Omega'} = \sum_\sigma (\mathbf{\Lambda}^{-1})_{\pi\sigma}\mathcal{T}_{\pi\Omega'} \tag{60}$$

$$= \sum_\sigma (\mathbf{\Lambda}^{-1})_{\pi\sigma}\mathcal{W}_{\pi\Omega'} = \mathcal{L}_{\pi\Omega'}, \tag{61}$$

while $[\mathcal{L}(\mathcal{W}^{-1}\mathcal{T})]_{\Omega\Omega'} = 0$ for $\Omega' \notin S_k$ because $\mathcal{L}_{\Omega\Omega'} = 0$ for $\Omega' \notin S_k$.

(3) Let us show that $(\mathcal{W}^{-1}\mathcal{T})\mathcal{L} = \mathcal{L}$,

$$[(\mathcal{W}^{-1}\mathcal{T})\mathcal{L}]_{\Omega\Omega'} = \sum_{\Omega'',\sigma,\tau} (\mathcal{W}^{-1})_{\Omega\Omega''}\mathcal{T}_{\Omega''\sigma}(\mathbf{\Lambda}^{-1})_{\sigma\tau}\mathcal{W}_{\tau\Omega'} \tag{62}$$

$$= \sum_{\Omega'',\sigma,\tau} (\mathcal{W}^{-1})_{\Omega\Omega''}\mathcal{W}_{\Omega''\sigma}(\mathbf{\Lambda}^{-1})_{\sigma\tau}\mathcal{W}_{\tau\Omega'} \tag{63}$$

$$= \sum_{\sigma,\tau} \delta_{\Omega\sigma}(\mathbf{\Lambda}^{-1})_{\sigma\tau}\mathcal{W}_{\tau\Omega'} = \mathcal{L}_{\Omega\Omega'}, \tag{64}$$

where, in the first sum, we restricted the first sum only on $\sigma \in S_k$ because $\mathcal{L}_{\Omega\Omega'} = 0$ for $\Omega \notin S_k$.

(4) Finally, let us show that $\mathcal{L}\mathcal{W}^{-1} = \mathbf{\Lambda}^{-1}$ when restricted to permutation space, otherwise zero,

$$[\mathcal{L}\mathcal{W}^{-1}]_{\pi\Omega} = \sum_{\sigma} (\mathbf{\Lambda}^{-1})_{\pi\sigma} \mathcal{W}_{\sigma\Omega'} \mathcal{W}^{-1}_{\Omega'\Omega} = \delta_{\Omega\sigma} (\mathbf{\Lambda}^{-1})_{\pi\sigma}, \tag{65}$$

while $[\mathcal{L}\mathcal{W}^{-1}]_{\Omega\Omega'} = 0$ is zero because $\Lambda_{\Omega\Omega'} = 0$ for $\Omega \notin S_k$. In what follows, we adopt a slight abuse of notation by using $\mathbf{\Lambda}$ and $\mathbf{\Lambda}^{-1}$ to denote the Gram and Weingarten matrices of the unitary group, viewed as embedded in the Clifford commutant space by padding with zeros outside the permutation subspace.

By definition, we can write $\mathcal{W}^{-1}\mathcal{T} = \mathcal{L} + \Delta$. However, we notice that $\mathcal{L}\Delta = \mathcal{L}(\mathcal{W}^{-1}\mathcal{T}) - \mathcal{L}^2 = \mathcal{L} - \mathcal{L} = 0$, where we used properties (1) and (2) above. Similarly, using (1) and (3), one sees that $\Delta\mathcal{L} = 0$. Consequently, we have that $(\mathcal{W}^{-1}\mathcal{T})^t = \mathcal{L} + \Delta^t$. Hence $(\mathcal{W}^{-1}\mathcal{T})^t\mathcal{W}^{-1} = \mathbf{\Lambda}^{-1} + \Delta^t\mathcal{W}^{-1}$, where we used property (4) above (as well as the abuse of notation we pointed out above). We can insert such identity in the expression of the twirling operator, which reads

$$\Phi_t^{(k)}(O) = \sum_{\Omega\Omega'} [(\mathcal{W}^{-1}\mathcal{T})^t]_{\Omega\Omega'} \mathrm{tr}(\Omega O)\Omega' = \sum_{\Omega\Omega'} [\mathbf{\Lambda}^{-1} + \Delta^t\mathcal{W}^{-1}]_{\Omega\Omega'} \mathrm{tr}(\Omega O)\Omega' \tag{66}$$

$$= \sum_{\pi,\sigma} (\mathbf{\Lambda}^{-1})_{\pi\sigma} \mathrm{tr}(T_\pi O) T_\sigma + \sum_{\Omega\Omega'} [\Delta^t\mathcal{W}^{-1}]_{\Omega\Omega'} \mathrm{tr}(\Omega O)\Omega' \tag{67}$$

$$= \Phi^{(k)}_{\mathrm{Haar}}(O) + \sum_{\Omega\Omega'} [\Delta^t\mathcal{W}^{-1}]_{\Omega\Omega'} \mathrm{tr}(\Omega O)\Omega', \tag{68}$$

which concludes the proof. ∎

Corollary 1. For any choice of μ_1 for which $\|\Delta\|_\infty < 1$, then

$$\lim_{t\to\infty} \Phi_t^{(k)}(O) = \Phi^{(k)}_{\mathrm{Haar}}(O). \tag{69}$$

Proof. The proof immediately follows from Lemma 12. ∎

Now that we have defined the doped-Clifford-Weingarten function, following the same spirit as the Clifford-Weingarten function [72], we aim to study its asymptotic behavior. While a full analytical computation for large k is impractical, we can effectively characterize its behavior in the limit of large n.

B. Asymptotics of doped Clifford-Weingarten functions

Lemma 13 (Properties of $\mathcal{T}$). The $|\mathcal{P}| \times |\mathcal{P}|$ matrix $\mathcal{T}$ with components $\mathcal{T}_{\Omega,\Omega'} := \mathrm{tr}(\Omega^\dagger \kappa(\Omega'))$, obeys the following properties:

(1) $\mathcal{T}$ is Hermitian.
(2) $\mathcal{T}_{\Omega,\Omega} = d^k(\mathrm{tr}(\omega^\dagger\kappa(\omega))/2^k)$, where ω is defined through $\Omega = \omega^{\otimes n}$.
(3) $\mathcal{T}_{\Omega,\Omega'} \le 2d^{k-1}$ for $\Omega \neq \Omega'$.

Proof. Item (1) readily follows from the definition. To show (2) and (3), we can use the tensor product structure of Pauli monomials see Lemma 8, i.e., $\Omega = \omega^{\otimes n}$. Then

$$\mathcal{T}_{\Omega,\Omega'} = \mathrm{tr}(\omega^\dagger\omega')^{n-1}\mathrm{tr}(\omega^\dagger\kappa(\omega')) = (\mathcal{W}|_{n-1})_{\Omega,\Omega'}\mathrm{tr}(\omega^\dagger\kappa(\omega')). \tag{70}$$

Hence, from Lemma 11, we know that $(\mathcal{W}|_{n-1})_{\Omega,\Omega} = (d/2)^k$ and $(\mathcal{W}|_{n-1})_{\Omega,\Omega'} \le (d/2)^{k-1}$. Moreover, it holds that $|\mathrm{tr}(\omega^\dagger\kappa(\omega))| \le \|\omega\|_\infty \|\kappa(\omega)\|_1 \le \|\omega\|_1 \le 2^k$, where we have used the data-processing inequality [100] in the second-to-last inequality and Lemma 8. ∎

Lemma 14 (Properties of $\mathcal{W}^{-1}\mathcal{T}$). Let $n \ge k^2 - 3k + 13$, and $k \ge 3$. Let $\mathcal{W}^{-1}$ be the matrix of Clifford-Weingarten functions, and consider the matrix $\mathcal{W}^{-1}\mathcal{T}$. Let $\mathcal{L}$ be the matrix defined in Eq. (55). Define $\Delta := \mathcal{W}^{-1}\mathcal{T} - \mathcal{L}$. The following facts hold.

(1) $\Delta_{\Omega,\Omega'} = 0$ for $\Omega, \Omega' \in S_k$.
(2) $\left|\Delta_{\Omega,\Omega} - \mathrm{tr}(\omega^\dagger\kappa(\omega))/2^k\right| \le 7|\mathcal{P}|^2/d$, for $\Omega \notin S_k$.
(3) $|\Delta_{\Omega,\Omega'}| \le 7|\mathcal{P}|^2 d$ for $\Omega \neq \Omega'$.

Proof. First note that $\mathcal{T}_{\pi\Omega} = \mathcal{W}_{\pi\Omega}$ for any $\pi \in S_k$ and $\Omega \in \mathcal{P}$. Hence, item (1) follows immediately by definition. Let us now consider we have $\Delta_{\Omega,\Omega} = (\mathcal{W}^{-1}\mathcal{T})_{\Omega,\Omega}$ for $\Omega \notin S_k$ and prove item (2),

$$\Delta_{\Omega,\Omega} = (\mathcal{W}^{-1})_{\Omega,\Omega}\mathcal{T}_{\Omega,\Omega} + \sum_{\Omega'\neq\Omega} (\mathcal{W}^{-1})_{\Omega,\Omega'}\mathcal{T}_{\Omega',\Omega}. \tag{71}$$

Using Lemmas 11 and 13, one can control the second term as

$$\left|\sum_{\Omega'\neq\Omega} (\mathcal{W}^{-1})_{\Omega,\Omega'}\mathcal{T}_{\Omega',\Omega}\right| \le |\mathcal{P}|\frac{6|\mathcal{P}|^2}{d^{k+1}} 2d^{k-1} = \frac{12|\mathcal{P}|^3}{d^2}. \tag{72}$$

It means that we can then use the following chain of inequalities and obtain

$$\begin{aligned}
&\left|\Delta_{\Omega,\Omega} - \frac{\mathrm{tr}(\omega^\dagger \kappa(\omega))}{2^k}\right| \\
&\leq \left|(\mathcal{W}^{-1})_{\Omega,\Omega}\mathcal{T}_{\Omega,\Omega} - \frac{\mathrm{tr}(\omega^\dagger \kappa(\omega))}{2^k}\right| + \frac{12|\mathcal{P}|^3}{d^2} \\
&= \left|(\mathcal{W}^{-1})_{\Omega,\Omega} d^k \frac{\mathrm{tr}(\omega^\dagger \kappa(\omega))}{2^k} - \frac{\mathrm{tr}(\omega^\dagger \kappa(\omega))}{2^k}\right| + \frac{12|\mathcal{P}|^3}{d^2} \\
&= d^k \frac{\mathrm{tr}(\omega^\dagger \kappa(\omega))}{2^k}\left|\mathcal{W}^{-1}_{\Omega,\Omega} - \frac{1}{d^k}\right| + \frac{12|\mathcal{P}|^3}{d^2} \\
&\leq \frac{6|\mathcal{P}|^2}{d} + \frac{12|\mathcal{P}|^3}{d^2} \\
&\leq \frac{6|\mathcal{P}|^2}{d} + \frac{12|\mathcal{P}|^3}{d^2} \\
&\leq \frac{7|\mathcal{P}|^2}{d},
\end{aligned} \tag{73}$$

where we have used again Lemmas 11 and 13, to prove such a bound, and that $|\mathcal{P}|^2 \geq 6|\mathcal{P}|$ for $k \geq 3$, from which item (2) follows. For item (3), we can consider $\Delta_{\Omega,\Omega'}$ with $\Omega \neq \Omega'$, $\Omega \in S_k$ and $\Omega' \notin S_k$ and use the triangle inequality $|\Delta_{\Omega,\Omega'}| \leq |(\mathcal{W}^{-1}T)_{\Omega,\Omega'}| + |\mathcal{L}_{\Omega,\Omega'}|$. Let us bound the two terms separately,

$$\begin{aligned}
&|(\mathcal{W}^{-1}\mathcal{T})_{\Omega,\Omega'}| \\
&\leq |(\mathcal{W}^{-1})_{\Omega,\Omega}||\mathcal{T}_{\Omega,\Omega'}| + \left|(\mathcal{W}^{-1})_{\Omega,\Omega'}\right||\mathcal{T}_{\Omega'\Omega'}| \\
&\quad + \max_{\Omega}|T_{\Omega,\Omega'}| \sum_{\Omega'\neq\Omega}|(\mathcal{W}^{-1})_{\Omega,\Omega'}| \\
&\leq \frac{1}{d^k}\left(1 + \frac{6|\mathcal{P}^2|}{d}\right)2d^{k-1} + d^{k-1}|\mathcal{P}|\frac{10|\mathcal{P}^2|}{d^{k+1}} \\
&\quad + \frac{5d^k|\mathcal{P}|^2}{d^{k+1}}
\end{aligned} \tag{74}$$

and

$$\begin{aligned}
|\mathcal{L}_{\Omega,\Omega'}| &\leq \left(\max_\sigma |\Lambda^{-1}_{\sigma\sigma}| + k! \max_{\sigma\neq\pi}|\Lambda^{-1}_{\sigma\pi}|\right)\max_{\Omega\neq\Omega'}|\mathcal{W}_{\Omega,\Omega'}| \\
&\leq \left(\frac{1}{d^k} + \frac{k!}{d^{k+1}}\right)d^{k-1} \leq \frac{2}{d},
\end{aligned} \tag{75}$$

where we have used the asymptotics of the Weingarten function (see Lemma 2) and Clifford-Weingarten functions in Lemma 11 and that $n \geq k^2 - 3k + 13$ by assumption. Combining Eqs. (74) and (75), we get

$$|\Delta_{\Omega,\Omega'}| \leq \frac{4}{d} + \frac{22|\mathcal{P}^3|}{d^2} + \frac{5|\mathcal{P}^2|}{d} \leq \frac{7|\mathcal{P}|^2}{d}, \tag{76}$$

where we have made use of the fact that $2|\mathcal{P}|^2 \geq 11|\mathcal{P}| + 4$ for $k \geq 3$. This concludes the proof. ■

Theorem 2 (Asymptotics doped-Clifford-Weingarten functions). Assuming that $n \geq \frac{3}{2}(k^2 - 3k + 18 + \log\|\Delta\|_\infty^{-1})$ and $k \geq 3$, the doped-Clifford-Weingarten functions have the following asymptotics:

$$\begin{aligned}
&|[\Delta^t\mathcal{W}^{-1}]_{\Omega,\Omega'}| \leq t\|\Delta\|_\infty^{t-1}\frac{18|\mathcal{P}|^2}{d^{k+1}}, \quad \Omega,\Omega' \in S_k \\
&[\Delta^t\mathcal{W}^{-1}]_{\Omega,\Omega'} = (\Delta_{\Omega,\Omega})^t(\mathcal{W}^{-1})_{\Omega,\Omega'} \pm t\|\Delta\|_\infty^{t-1}\frac{30|\mathcal{P}|^2}{d^{k+1}}.
\end{aligned} \tag{77}$$

Moreover, the following facts hold.

(1) $\left|(\Delta^t)_{\Omega,\Omega} - (\Delta_{\Omega,\Omega})^t\right| \leq t(7|\mathcal{P}|^3/d)\|\Delta\|_\infty^{t-1}$.
(2) $\max|(\Delta^t)_{\Omega,\Omega'}| \leq t\|\Delta\|_\infty^{t-1}7|\mathcal{P}|^2/d$ for $\Omega \neq \Omega'$.

Proof. First, notice that we can write $\mathcal{W}^{-1}\mathcal{T} = \mathcal{L} + \Delta$, and $\mathcal{L}\Delta = 0$, while $\mathcal{L}^2 = \mathcal{L}$. Therefore, we can express the matrix of the doped-Clifford-Weingarten functions as

$$(\mathcal{W}^{-1}\mathcal{T})^t\mathcal{W}^{-1} = \mathbf{\Lambda}^{-1} + \Delta^t\mathcal{W}^{-1}, \tag{78}$$

where $\mathbf{\Lambda}^{-1}$ is only supported on the first $k! \times k!$ block corresponding to permutation. Hence we need to characterize Δ^t. Thus, let $\Omega \notin S_k$. First, the following chain of inequalities holds:

$$\begin{aligned}
(\Delta^t)_{\Omega,\Omega} &\leq (\Delta^{t-1})_{\Omega,\Omega}\Delta_{\Omega,\Omega} + \sum_{\Omega'\neq\Omega}|(\Delta^{t-1})_{\Omega,\Omega'}\Delta_{\Omega',\Omega}| \\
&\leq (\Delta^{t-1})_{\Omega,\Omega}\Delta_{\Omega,\Omega} + |\mathcal{P}|\max_{\Omega\neq\Omega'}|\Delta_{\Omega,\Omega'}|\|\Delta\|_\infty^{t-1} \\
&\leq (\Delta^{t-2}\Delta_{\Omega,\Omega} + |\mathcal{P}|\max_{\Omega\neq\Omega'}|\Delta_{\Omega,\Omega'}|\|\Delta\|_\infty^{t-2})_{\Omega,\Omega}\Delta_{\Omega,\Omega} \\
&\quad + |\mathcal{P}|\max_{\Omega\neq\Omega'}|\Delta_{\Omega,\Omega'}|\|\Delta\|_\infty^{t-1} \\
&= (\Delta^{t-2})_{\Omega,\Omega}\Delta^2_{\Omega,\Omega} + 2|\mathcal{P}|\max_{\Omega\neq\Omega'}|\Delta_{\Omega,\Omega'}|\|\Delta\|_\infty^{t-1}.
\end{aligned} \tag{79}$$

Iterating these inequalities and upper bounding $\max_{\Omega\neq\Omega'}|\Delta_{\Omega,\Omega'}| \leq 7|\mathcal{P}^2|/d$ as in Lemma 14, we finally get

$$\left|(\Delta^t)_{\Omega,\Omega} - (\Delta_{\Omega,\Omega})^t\right| \leq t\frac{7|\mathcal{P}|^3}{d}\|\Delta\|_\infty^{t-1}. \tag{80}$$

While for $\Omega \neq \Omega'$, we have

$$(\Delta^t)_{\Omega,\Omega'} \le (\Delta^{t-1})_{\Omega,\Omega} \max_{\Omega\neq\Omega'} |\Delta_{\Omega,\Omega'}| + \max_{\Omega\neq\Omega'} |(\Delta^{t-1})_{\Omega,\Omega'}| \sum_{\Omega''\neq\Omega} |\Delta_{\Omega''\Omega'}|$$
$$\le \|\Delta\|_\infty^{t-1} \max_{\Omega\neq\Omega'} |\Delta_{\Omega,\Omega'}| + \max_{\Omega\neq\Omega'} |(\Delta^{t-1})_{\Omega,\Omega'}| (\|\Delta\|_\infty + |\mathcal{P}| \max |\Delta_{\Omega,\Omega'}|). \tag{81}$$

Now, let us assume for simplicity that

$$(\|\Delta\|_\infty + |\mathcal{P}| \max |\Delta_{\Omega,\Omega'}|) \le 1, \tag{82}$$

which holds whenever $n \ge \frac{3}{2}(k^2 - 3k + 18 + \log \|\Delta\|_\infty^{-1})$. Since it works for any $\Omega \neq \Omega'$, we have the recurrence relation

$$\max_{\Omega\neq\Omega'} |(\Delta^t)_{\Omega,\Omega'}| \le \max_{\Omega\neq\Omega'} |(\Delta^{t-1})_{\Omega,\Omega'}| + \|\Delta\|_\infty^{t-1} \max_{\Omega\neq\Omega'} |\Delta_{\Omega,\Omega'}|. \tag{83}$$

This gives

$$\max |(\Delta^t)_{\Omega,\Omega'}| \le t\|\Delta\|_\infty^{t-1} \max_{\Omega\neq\Omega'} |\Delta_{\Omega,\Omega'}| \le t\|\Delta\|_\infty^{t-1} \frac{7|\mathcal{P}|^2}{d}, \tag{84}$$

where we have used Lemma 14. Now, putting all together, we can prove the result. Notice that Δ restricted to the first $k! \times k!$ is null. Hence, for $\Omega, \Omega' \in S_k$, we have

$$|[\Delta^t \mathcal{W}^{-1}]_{\Omega,\Omega'}| = \sum_{\Omega''\notin S_k} (\Delta^t)_{\Omega\Omega''} (\mathcal{W}^{-1})_{\Omega''\Omega'}$$
$$\le t\|\Delta\|_\infty^{t-1} \frac{49|\mathcal{P}|^5}{d^{k+2}} \le t\|\Delta\|_\infty^{t-1} \frac{18|\mathcal{P}|^2}{d^{k+1}}, \tag{85}$$

where we have used that $n \ge \frac{3}{2}(k^2 - 3k + 13)$. While for $\Omega \notin S_k$ or $\Omega' \notin S_k$, we find

$$[\Delta^t \mathcal{W}^{-1}]_{\Omega,\Omega'} = \sum_{\Omega''} (\Delta^t)_{\Omega\Omega''} (\mathcal{W}^{-1})_{\Omega''\Omega'}$$
$$\le (\Delta^t_{\Omega,\Omega})(\mathcal{W}^{-1})_{\Omega,\Omega'} + (\Delta^t_{\Omega,\Omega'})(\mathcal{W}^{-1})_{\Omega'\Omega'} + |\mathcal{P}| \max_{\Omega\neq\Omega'} |\mathcal{W}^{-1}| \max_{\Omega\neq\Omega'} \left|\Delta^t_{\Omega,\Omega'}\right|$$
$$\le \|\Delta\|_\infty^t (\mathcal{W}^{-1})_{\Omega,\Omega'} + t\frac{7|\mathcal{P}|^2}{d^{k+1}} \|\Delta\|_\infty^{t-1} \left(1 + \frac{6|\mathcal{P}|^2}{d}\right) + t\frac{49|\mathcal{P}|^5}{d^{k+2}} \|\Delta\|_\infty^{t-1}$$
$$\le \|\Delta\|_\infty^t (\mathcal{W}^{-1})_{\Omega,\Omega'} + t\frac{7|\mathcal{P}|^2}{d^{k+1}} \|\Delta\|_\infty^{t-1} + t\frac{42|\mathcal{P}|^4}{d^{k+2}} \|\Delta\|_\infty^{t-1} + t\frac{18|\mathcal{P}|^2}{d^{k+1}} \|\Delta\|_\infty^{t-1}$$
$$\le \|\Delta\|_\infty^t (\mathcal{W}^{-1})_{\Omega,\Omega'} + t\frac{25|\mathcal{P}|^2}{d^{k+1}} \|\Delta\|_\infty^{t-1} + t\frac{15|\mathcal{P}|}{d^{k+1}} \|\Delta\|_\infty^{t-1}$$
$$\le \|\Delta\|_\infty^t (\mathcal{W}^{-1})_{\Omega,\Omega'} + t\frac{30|\mathcal{P}|^2}{d^{k+1}} \|\Delta\|_\infty^{t-1}, \tag{86}$$

where we have used that $|\mathcal{P}|^2 \ge 3|\mathcal{P}|$ for $k \ge 3$ and used that $n \ge 3/2(k^2 - 3k + 18)$. This concludes the proof. ■

Let us informally describe the content of Theorem 2. On the one hand, Corollary 1 states, quite trivially, that if we dope the Clifford group with any measure described by a perturbation matrix Δ whose operator norm is strictly less than unity, then in the limit of infinite doping, the twirling operator converges to the one over the full unitary group. This is unsurprising, as the Clifford group doped with any gate outside the Clifford group becomes universal. However, this fact provides the confirmation that, among all the valid measures we could have chosen for t-doped Clifford circuits, we selected the one that interpolates between the uniform (Haar) measure over the Clifford group and the uniform (Haar) measure over the full unitary group.

On the other hand, Theorem 2 provides a more fine-grained structure for the twirling over the doped Clifford group. Specifically, it explicitly separates the dependence on the full unitary group twirling from the doped

contributions, which eventually decay as t increases. Informally, we can express the doped-Clifford twirling as

$$\Phi_t^{(k)}(O) \simeq \Phi_{\text{Haar}}^{(k)}(O) + \sum_{\Omega,\Omega'} (\Delta_{\Omega,\Omega})^t (\mathcal{W}^{-1})_{\Omega,\Omega'} \text{tr}(\Omega O)\Omega', \tag{87}$$

where the Haar twirling appears on the left, and the Clifford twirling on the right, with each $\Omega \notin S_k$ term being suppressed by a factor of $(\Delta_{\Omega,\Omega})^t \simeq (2^{-k}\|\kappa(\omega)\|_2^2)^t$ (see Theorem 2). Naturally, for every $\Omega \notin S_k$, we have $(\Delta_{\Omega,\Omega})^t \leq \|\Delta\|_\infty^t$, which indicates that the largest coefficient is suppressed by the operator norm of the perturbation matrix Δ.

This finer structural understanding is useful for determining the amount of doping necessary and sufficient to approximate the twirling operator applied to a given operator O of interest over the full unitary group using the twirling operator over the doped Clifford group. As an application of the tools presented above, we analyze this convergence in the next section.

IV. CONVERGENCE OF RANDOM DOPED CLIFFORD CIRCUITS

In this section, we establish the main result of this work in two complementary ways: first, by analyzing the convergence of the frame potential in Eq. (23); as a corollary we provide improved convergence bound of t-doped Clifford circuits toward state k-designs; second, we analyze the convergence toward relative ε-approximate k-design, by providing tight bounds.

A. Optimal frame potentials convergence

In this section, we prove tight convergence bounds to the frame potential of t-doped Clifford circuits. To this end, we begin by proving some lemmas that are instrumental in deriving the main result.

As a careful reader may infer from Eq. (87), the convergence of doped random Clifford circuits toward the Haar measure is governed by the matrix Δ. Specifically, it depends on the sum $\sum_\Omega (\Delta_{\Omega,\Omega})^t$. In the following key lemma, we establish a lower bound for this coefficient, which will, in turn, provide upper and lower bounds on the convergence rate of the frame potential. As we will see, such a lower bound depends on the spectral form factor of the doping ensemble $\int \text{d}\mu_1(U)|\text{tr}(U)|^{2k}$, which we characterize later for diagonal single-qubit unitaries, as well as Haar-random single-qubit unitaries.

Lemma 15 (Lower bounding convergence via spectral form factor). Let $n \geq 3/2(k^2 - 3k + 18)$ and $k \geq 4$. Let $\Delta := \mathcal{W}^{-1}\mathcal{T} - \mathcal{L}$, where the matrices $\mathcal{W}, \mathcal{T}, \mathcal{L}$ are defined in Sec. III. Then the bound

$$\sum_\Omega (\Delta_{\Omega,\Omega})^t \geq (|\mathcal{P}| - k!) \times \left(\frac{1}{3}\int \text{d}\mu_1(U)\frac{|\text{tr}(U)|^{2k}}{4^k} - \frac{2}{|\mathcal{P}|^{1/10}}\right)^t \tag{88}$$

holds.

Proof. First notice that $\Delta_{\Omega,\Omega} = 0$ for every $\Omega \in S_k$, i.e., for any permutation. Hence, we can restrict the sum over $\Omega \notin S_k$. Using the asymptotics in Lemma 14, we can write the following

$$\sum_{\Omega\notin S_k} (\Delta_{\Omega,\Omega})^t \geq \sum_{\Omega\notin S_k} \left(\frac{\text{tr}(\omega^\dagger\kappa(\omega))}{2^k} - \frac{7|\mathcal{P}|^2}{d}\right)^t \geq (|\mathcal{P}| - k!) \times \left(\frac{1}{|\mathcal{P}| - k!}\sum_{\Omega\notin S_k}\frac{\text{tr}(\omega^\dagger\kappa(\omega))}{2^k} - \frac{7|\mathcal{P}|^2}{d}\right)^t \tag{89}$$

where we have used Jensen's inequality. We can now add and subtract all the remaining contributions coming from permutations, which leads to

$$\sum_{\Omega\notin S_k} (\Delta_{\Omega,\Omega})^t \geq (|\mathcal{P}| - k!)\left(\frac{1}{|\mathcal{P}| - k!}\sum_\Omega \frac{\text{tr}(\omega^\dagger\kappa(\omega))}{2^k} - \frac{k!}{|\mathcal{P}| - k!} - \frac{7|\mathcal{P}|^2}{d}\right)^t. \tag{90}$$

We now focus our attention on the first term. First, by Lemma 8 we note that $\omega^\dagger = \omega^T$. Hence, we have the following identity

$$\sum_\omega \frac{\text{tr}(\omega^\dagger\kappa(\omega))}{2^k} = \sum_\omega \text{tr}(\phi_+^{\otimes k}\mathbb{1}\otimes\kappa[\omega^{\otimes 2}]) = Z_2\mathbb{E}_\sigma \text{tr}(\phi_+^{\otimes k}\mathbb{1}\otimes\kappa[\sigma^{\otimes k}]) \tag{91}$$

where ϕ_+ is the Bell state between two copies of a single qubit. The second equality comes from Lemma 10, i.e., given a 2-qubit stabilizer state $\sigma \in \text{STAB}_2$, then $\mathbb{E}_\sigma \sigma^{\otimes k} = (1/Z_2)\sum_\omega \omega^{\otimes 2}$, where $Z_2 = 4\prod_{i=0}^{k-2}(2^i + 4)$. Noticing that $\phi_+ \in \text{STAB}_2$, to simplify calculation, we can trivially lower bound the average as

$$\begin{aligned}\sum_\omega \frac{\text{tr}(\omega^\dagger\kappa(\omega))}{2^k} &\geq \frac{Z_2}{|\,\text{STAB}_2\,|}\text{tr}(\phi_+^{\otimes k}\mathbb{1}\otimes\kappa[\phi_+^{\otimes k}]) \\ &= \frac{Z_2}{|\,\text{STAB}_2\,|}\int \text{d}\mu_1(U)\text{tr}(\phi_+\mathbb{1}\otimes U\phi_+\mathbb{1}\otimes U^\dagger)^k \\ &= \frac{Z_2}{|\,\text{STAB}_2\,|}\int \text{d}\mu_1(U)\frac{|\text{tr}(U)|^{2k}}{4^k}.\end{aligned} \tag{92}$$

Therefore, the convergence is effectively dictated by the spectral form factor of the ensemble described by the measure μ_1. Before plugging everything together, notice that $|\,\mathrm{STAB}_2\,| = 60$ and that

$$\frac{Z_2}{|\mathcal{P}|-k!} \geq \frac{Z_2}{|\mathcal{P}|} = 4\prod_{i=0}\frac{2^i+4}{2^i+1} = 60\frac{4^{k+1}}{(2^k+8)(2^k+4)}. \tag{93}$$

Then also notice that

$$\frac{4^{k+1}}{(2^k+8)(2^k+4)} \geq \frac{1}{3} \tag{94}$$

for $k \geq 2$. Lastly, we can bound $k!/(|\mathcal{P}|-k!) \leq 1/|\mathcal{P}|^{1/10}$ for every $k \geq 4$. Finally, we have that $7|\mathcal{P}|^2/d \leq 1/|\mathcal{P}|$ for $n \geq 3/2(k^2-3k+18)$. Therefore, in summary, we have the following lower bound

$$\sum_{\Omega\neq S_k}(\Delta_{\Omega,\Omega})^t \geq (|\mathcal{P}|-k!) \times \left(\frac{1}{3}\int \mathrm{d}\mu_1(U)\frac{|\mathrm{tr}(U)|^{2k}}{4^k} - \frac{2}{|\mathcal{P}|^{1/10}}\right)^t. \tag{95}$$

The above equation tells that whenever the average spectral form factor is only polynomially small in k, then the convergence is quadratic $t \sim k^2$. ■

In the following lemma, we lower bound spectral form factors for two interesting ensembles: the diagonal ensemble and the Haar-random ensemble.

Lemma 16 (Lower bounding spectral form factors). The following lower bounds for the spectral form factors of hold for different ensembles.

$$\int \mathrm{d}\mu_1(U)|\mathrm{tr}(U)|^{2k} \geq \frac{2\sqrt{\pi}}{e^2\sqrt{k}}, \quad \text{diagonal ensemble.}$$
$$\int \mathrm{d}\mu_1(U)|\mathrm{tr}(U)|^{2k} \geq \frac{2\sqrt{\pi}}{e^2\sqrt{k}(k+1)}, \quad \text{Haar ensemble.} \tag{96}$$

Proof. First, notice that the spectral form factor only depends on the difference between the eigenphases as

$$|\mathrm{tr}(U)|^2 = |1+e^{i(\theta_1-\theta_2)}| = (1+e^{i(\theta_1-\theta_2)})(1+e^{-i\phi}) = 2+2\cos(\theta_1-\theta_2) = 4\cos^2\left(\frac{\theta_1-\theta_2}{2}\right). \tag{97}$$

Let us begin when μ_1 is the uniform distribution over diagonal unitary matrices,

$$\int \mathrm{d}\mu_1(U)\frac{|\mathrm{tr}(U)|^{2k}}{4^k} = \int_0^{2\pi}\mathrm{d}\theta\cos^{2k}\frac{\theta}{2} = \frac{(2k)!}{(k!)^2 4^k} \geq \frac{2\sqrt{\pi}}{e^2\sqrt{k}}, \tag{98}$$

where, for the lower bound, we have used the following on the factorial $\sqrt{2\pi}k^{k+\frac{1}{2}}e^{-k} \leq k! \leq ek^{k+\frac{1}{2}}e^{-k}$. For the Haar ensemble, we can use the famous eigenvalue probability density on the difference between the eigenphases $p_{\mathrm{Haar}}(\theta) = (1/2\pi)(1-\cos(\theta))$. Therefore, in this case

$$\int_0^{2\pi}\mathrm{d}\theta p_{\mathrm{Haar}}(\theta)\cos^{2k}\frac{\theta}{2} = \frac{1}{2\pi}\int_0^{2\pi}\mathrm{d}\theta(1-\cos(x))\cos^{2k}\frac{\theta}{2} = \frac{(2k)!}{(k!)^2 4^k(k+1)} \geq \frac{2\sqrt{\pi}}{e^2\sqrt{k}(k+1)}. \tag{99}$$

■

Therefore, the spectral form factors for the diagonal ensemble and for single-qubit Haar-random gates are both polynomially small in k. Consequently, they lead to the same convergence result. Below, we, therefore, analyze only the case of random diagonal single-qubit gates. The analysis for single-qubit Haar-random gates is analogous and, most importantly, leads to the same conclusions.

Corollary 2. Let $n \geq 3/2(k^2-3k+18)$ and $k \geq 9$. The following bound holds for the diagonal unitary ensemble:

$$\sum_{\Omega}(\Delta_{\Omega,\Omega})^t \geq \left(\frac{2\sqrt{\pi}}{3e^2\sqrt{k}}\right)^t. \tag{100}$$

Proof. By Lemmas 15 and 16, we have

$$\sum_{\Omega}(\Delta_{\Omega,\Omega})^t \geq (|\mathcal{P}|-k!)\left(\frac{2\sqrt{\pi}}{3e^2\sqrt{k}} - \frac{2}{|\mathcal{P}|^{1/10}}\right). \tag{101}$$

By Lemma 8, we have that for any $k \geq 9$, we have that $2/|\mathcal{P}|^{1/10} \leq \sqrt{\pi}/3e^2\sqrt{k}$, therefore, the result follows. ■

The following lemma establishes upper bounds on the coefficient $\sum_{\Omega}(\Delta_{\Omega,\Omega})^t$.

Lemma 17 (Upper bound [73]). Let $n \geq (k^2-3k+26)$. Let μ_1 be the measure over single-qubit diagonal or Haar-random unitaries. The following upper bound holds:

$$\sum_{\Omega}(\Delta_{\Omega,\Omega})^t \leq (|\mathcal{P}|-k!)\left(\frac{15}{16}\right)^t. \tag{102}$$

Proof. Using the asymptotics in Lemma 14, we can write

$$\sum_{\Omega\notin S_k}(\Delta_{\Omega,\Omega})^t \leq \sum_{\omega\notin S_k}\left(\frac{\mathrm{tr}(\omega^\dagger\kappa(\omega))}{2^k} + \frac{7|\mathcal{P}|^2}{d}\right)^t \leq (|\mathcal{P}|-k!)\left(\max_{\omega\notin S_k}\frac{\mathrm{tr}(\omega^\dagger\kappa(\omega))}{2^k} + \frac{7|\mathcal{P}|^2}{d}\right)^t. \tag{103}$$

For $\mu_1(U)$ be the measure over the diagonal unitary ensemble, then in Ref. [73], the authors have proven

that $\max_{\omega\notin S_k}\mathrm{tr}(\omega^\dagger\kappa(\omega))/2^k \le \frac{7}{8}$. Moreover, as also noted in Ref. [73], the same bound hold for the Haar-random ensemble. Moreover we have that $7|\mathcal{P}|^2/d \le \frac{1}{16}$ for every $k \ge 1$ thanks to the hypothesis of the lemma. Therefore, we finally have

$$\sum_{\Omega\notin S_k}(\Delta_{\Omega,\Omega})^t \le (|\mathcal{P}|-k!)\left(\frac{15}{16}\right)^t. \tag{104}$$

■

Now we established the key lemmas to show the convergence of doped random circuit in multiple flavors. Let us establish the optimal convergence in frame potential in terms of the coefficient of Lemmas 15 and 17.

Theorem 3 (Frame potential convergence). Let μ_1 be the measure over the diagonal unitary ensemble. Let $n \ge f(k,t)$ where $f(k,t) := \frac{7}{4}k^2 - 6k + 19 + 3t + \log t + t\log k$, and $k \ge 9$. The following upper and lower bounds hold for the difference of the frame potentials:

$$2^{\frac{k^2}{4}-6t-t\log k-1} \le \mathcal{F}_t^{(k)} - \mathcal{F}_{\mathrm{Haar}} \le 2^{\frac{k^2}{2}-2t\log\frac{16}{15}+1}. \tag{105}$$

Proof. Let us first recall the definition of frame potential, specifically for the doped random Clifford ensemble $\mathcal{C}_t$, described by the measure μ_t introduced in Definition 11,

$$\mathcal{F}_t^{(k)} = \int \mathrm{d}\mu_t(U)\,\mathrm{d}\mu_t(V)|\mathrm{tr}(U^\dagger V)|^{2k} = \int \mathrm{d}\mu_{2t}(U)|\mathrm{tr}(U)|^{2k}, \tag{106}$$

i.e., it reduces to the average spectral form factor of the ensemble $\mathcal{C}_{2t}$, where the amount of doping is doubled. We can, therefore, express it as

$$\mathcal{F}_t^{(k)} = \int \mathrm{d}\mu_{2t}(U)\mathrm{tr}(U^{\otimes k}\otimes U^{\dagger\otimes k}) = \mathrm{tr}(\Gamma_{2t}^{(k)}) \tag{107}$$

through the frame operator $\Gamma_{2t}^{(k)} := \int \mathrm{d}\mu_t(U)U^{\otimes k}\otimes U^{\dagger\otimes k}$. From Lemma 12, we can express the frame operator as

$$\Gamma_{2t}^{(k)} = \Gamma_{\mathrm{Haar}}^{(k)} + \sum_{\Omega,\Omega'}[\Delta^t\mathcal{W}^{-1}]_{\Omega,\Omega'}\hat{T}\Omega^\dagger\otimes\Omega', \tag{108}$$

where $\hat{T}\prod_{j=1}^k T_{k,k+j}$ and $\Gamma_{\mathrm{Haar}}^{(k)} = \int \mathrm{d}U\,U^{\otimes k}\otimes U^{\dagger\otimes k}$. Given that $\mathcal{F}_t^{(k)} = \mathrm{tr}(\Gamma_{2t}^{(k)})$, below we compute $\mathrm{tr}(\Gamma_t^{(k)})$ for simplicity

$$\begin{aligned}\mathcal{F}_{t/2}^{(k)} &= \mathcal{F}_{\mathrm{Haar}}^{(k)} + \mathrm{tr}\left(\hat{T}\sum_{\Omega,\Omega'}[\Delta^t\mathcal{W}^{-1}]_{\Omega,\Omega'}\Omega^\dagger\otimes\Omega'\right)\\ &= \mathcal{F}_{\mathrm{Haar}}^{(k)} + \sum_{\Omega,\Omega'}[\Delta^t\mathcal{W}^{-1}]_{\Omega,\Omega'}\mathrm{tr}(\hat{T}\Omega^\dagger\otimes\Omega')\\ &= \mathcal{F}_{\mathrm{Haar}}^{(k)} + \sum_{\Omega,\Omega'}[\Delta^t\mathcal{W}^{-1}]_{\Omega,\Omega'}\mathcal{W}_{\Omega',\Omega}\\ &= \mathcal{F}_{\mathrm{Haar}}^{(k)} + \sum_{\Omega}(\Delta^t)_{\Omega,\Omega}.\end{aligned} \tag{109}$$

Now, by assumption $n \ge f(k,t/2)$ above. In order to apply Theorem 2, a sufficient condition is that $\frac{3}{2}(k^2 - 3k + 18 + \log\|\Delta\|_\infty^{-1})$. By Lemma 17, if $n \ge (k^2 - 3k + 26)$ then for the diagonal unitary ensemble $\|\Delta\|_\infty \le \frac{15}{16}$. Hence, a sufficient condition for applying Theorem 2 in this case is $n \ge \frac{3}{2}(k^2 - 3k + 26 + \log\frac{16}{15})$. Note that the condition $n \ge f(k,t/2)$ ensures $n \ge \frac{3}{2}(k^2 - 3k + 26 + \log\frac{16}{15})$ for any $t,k \ge 1$. Applying Theorem 2, we can express $\sum_\Omega(\Delta^t)_{\Omega,\Omega} = \sum_\Omega(\Delta_{\Omega,\Omega})^t \pm (7|\mathcal{P}|^4/d)t\|\Delta\|_\infty^{t-1}$. Now, applying Lemma 15 and Corollary 2, we have

$$\sum_\Omega(\Delta^t)_\Omega \ge (|\mathcal{P}|-k!)\left(\frac{2\sqrt{\pi}}{3e^2\sqrt{k}}\right)^t - \frac{7|\mathcal{P}|^4}{d}t\|\Delta\|_\infty^{t-1}. \tag{110}$$

To derive a nontrivial lower bound we require that $(7|\mathcal{P}|^4/d)t\|\Delta\|_\infty^{t-1} \le \frac{1}{2}(|\mathcal{P}|-k!)\left(2\sqrt{\pi}/3e^2\sqrt{k}\right)^t$. A sufficient condition to ensure it is given by $n \ge f(k,t)$, which explains the reason behind this assumption. To see this, we require $(|\mathcal{P}|-k!) \ge 2^{\frac{k^2}{4}}$ valid for any $k \ge 4$, $|\mathcal{P}| \le 2^{\frac{1}{2}(k^2-3k+6)}$, and $\|\Delta\|_\infty \le \frac{15}{16}$, then we require

$$2^{2(k^2-3k+6)-n+\log 7-(t-1)\log\frac{16}{15}+\log t} \le 2^{\frac{k^2}{4}-t\log_2\frac{3e^2\sqrt{k}}{2\sqrt{\pi}}-1} \tag{111}$$

from which the condition follows. Hence $\sum_\Omega(\Delta^t)_{\Omega,\Omega} \ge 2^{\frac{k^2}{4}-3t-\frac{t}{2}\log k-1}$. The upper bound follows easily. Indeed, we have

$$\begin{aligned}\sum_{\Omega,\Omega'}(\Delta^t)_{\Omega,\Omega'} &\le (|\mathcal{P}|-k!)\left(\frac{15}{16}\right)^t + \frac{7|\mathcal{P}|^4}{d}t\left(\frac{15}{16}\right)^{t-1}\\ &\le (|\mathcal{P}|-k!)\left(\frac{15}{16}\right)^t + \frac{1}{2}(|\mathcal{P}|-k!)\left(\frac{2\sqrt{\pi}}{3e^2\sqrt{k}}\right)^t\\ &\le 2^{\frac{k^2}{2}-t\log\frac{16}{15}} + 2^{k^2/2-2t-\frac{t}{2}\log k} \le 2^{\frac{k^2}{2}-t\log\frac{16}{15}+1},\end{aligned} \tag{112}$$

where we have used the condition $n \geq f(k, t/2)$. We are just left to send $t \mapsto 2t$. This concludes the proof. ■

Remark 1. The bounds appearing in Theorem 3 are obtained via a fully analytical, worst-case analysis and should be interpreted as sufficient guarantees rather than tight conditions. In particular, the constraint on the function $f(n, k)$ is likely an artifact of the techniques used to control the non-Haar contributions within the Clifford commutant, and could potentially be relaxed by a more refined analysis. While these sufficient conditions may place the theorem outside the range of current exact classical simulations, we, nonetheless, expect that the qualitative conclusion of Theorem 3 can be verified numerically for much smaller system sizes.

The above theorem constitutes one of the main results of this work. In particular, it shows that a necessary and sufficient condition to have the frame potential difference to be smaller than ε is having the amount of doping to scale *quadratically* with respect to k, i.e., $t = \tilde{\Theta}(k^2 + \log \varepsilon^{-1})$.

Lemma 18 (Relative state frame potentials and frame potentials are related). Let $\mathcal{C}_t$ be the t-doped Clifford ensemble. Let $\mathcal{S}_t = \{C_t|0\rangle : C \in \mathcal{C}_t\}$, the ensemble of states induced by $\mathcal{C}_t$. Then it holds that

$$\mathcal{R}_{\mathcal{S}_t} = \frac{\mathrm{tr}(\Pi_{\mathrm{sym}})}{Z_n}(\mathcal{F}_t - \mathcal{F}_{\mathrm{Haar}}) + \frac{\mathrm{tr}(\Pi_{\mathrm{sym}})}{Z_n}\sum_{\Omega\neq\Omega'}(\Delta^{2t})_{\Omega',\Omega} \tag{113}$$

where $Z_n = d\prod_{i=0}^{k-2}(d + 2^i)$. Moreover, in the hypothesis of Theorem 3, i.e., $n \geq f(k, t)$, then

$$\left(\frac{1}{2} - c\right)\frac{\mathcal{F}_t^{(k)} - \mathcal{F}_{\mathrm{Haar}}}{\mathcal{F}_{\mathrm{Haar}}} \leq \mathcal{R}_{\mathcal{S}_t} \leq \frac{3}{2}\frac{\mathcal{F}_t^{(k)} - \mathcal{F}_{\mathrm{Haar}}}{\mathcal{F}_{\mathrm{Haar}}}, \tag{114}$$

where $c \leq 2^{-f(k,t)+2k}$ for every t.

Proof. Let us denote as $\Psi_t^{(k)}$ be the average state of the ensemble $\mathcal{S}_t$. Clearly $\Psi_t^{(k)} = \Phi_t^{(k)}(\sigma^{\otimes k})$ for any $\sigma \in \mathrm{STAB}_n$. From Sec. II C, we know that the state frame potential is simply given by the purity of $\Psi_t^{(k)}$, which is given by $\mathrm{Pur}(\Psi_t^{(k)}) = \mathrm{tr}(\sigma^{\otimes k}\Phi_{2t}(\sigma^{\otimes k}))$, since the twirling channel $\Phi_t(\cdot)$ is self-adjoint. However, it also holds that $\mathrm{tr}(\sigma^{\otimes k}\Phi_{2t}(\sigma^{\otimes k})) = \mathbb{E}_\sigma \mathrm{tr}(\sigma^{\otimes k}\Phi_{2t}(\sigma^{\otimes k})) = (1/Z_n)\sum_\Omega \mathrm{tr}(\Omega\Phi_{2t}(\sigma^{\otimes k}))$, see Lemma 10. From Lemma 8, we know that $\mathrm{tr}(\Omega\sigma^{\otimes k}) = 1$ for any $\sigma \in \mathrm{STAB}_n$ and $\Omega \in \mathcal{P}$. Hence, from Lemma 12, we have

$$\begin{aligned}
\mathrm{Pur}(\Psi_t^{(k)}) &= \frac{1}{Z_n}\sum_\Omega \mathrm{tr}(\Omega\Phi_{2t}(\sigma^{\otimes k})) = \mathrm{Pur}(\Psi_{\mathrm{Haar}}^{(k)}) + \frac{1}{Z_n}\sum_{\Omega,\Omega',\Omega''}(\Delta^{2t}\mathcal{W}^{-1})_{\Omega',\Omega''}\mathcal{W}_{\Omega'',\Omega} \\
&= \mathrm{Pur}(\Psi_{\mathrm{Haar}}^{(k)}) + \frac{1}{Z_n}\sum_{\Omega,\Omega',\Omega'',\Omega'''}(\Delta^{2t})_{\Omega',\Omega'''}(\mathcal{W}^{-1})_{\Omega''',\Omega''}\mathcal{W}_{\Omega'',\Omega} = \mathrm{Pur}(\Psi_{\mathrm{Haar}}^{(k)}) + \frac{1}{Z_n}\sum_{\Omega,\Omega',\Omega'''}(\Delta^{2t})_{\Omega',\Omega'''}\delta_{\Omega''',\Omega} \\
&= \mathrm{Pur}(\Psi_{\mathrm{Haar}}^{(k)}) + \frac{1}{Z_n}\sum_{\Omega,\Omega'}(\Delta^{2t})_{\Omega',\Omega} = \mathrm{Pur}(\Psi_{\mathrm{Haar}}^{(k)}) + \frac{1}{Z_n}\sum_{\Omega}(\Delta^{2t})_{\Omega,\Omega} + \sum_{\Omega\neq\Omega'}(\Delta^{2t})_{\Omega',\Omega} \\
&= \mathrm{Pur}(\Psi_{\mathrm{Haar}}^{(k)}) + \frac{1}{Z_n}(\mathcal{F}_t^{(k)} - \mathcal{F}_{\mathrm{Haar}}^{(k)}) + \sum_{\Omega\neq\Omega'}(\Delta^{2t})_{\Omega',\Omega}.
\end{aligned} \tag{115}$$

Noticing that $\mathrm{Pur}(\Psi_{\mathrm{Haar}}^{(k)}) = \mathrm{tr}^{-1}(\Pi_{\mathrm{sym}})$, and using Defintion 9, Eq. (113) follows. Let us now prove Eq. (114). First of all, the following bounds hold:

$$\frac{1}{k!} - \frac{4^k}{dk!} \leq \frac{\mathrm{tr}(\Pi_{\mathrm{sym}})}{Z_n} \leq \frac{1}{k!}. \tag{116}$$

The lower bound follows from $Z_n/d^k \leq 1 + 4^k/d$ in the hypothesis of Theorem 3 and that $1/(1+x) \geq 1 - x$. The upper bound follows by noticing that $\mathrm{tr}(\Pi_{\mathrm{sym}})/Z_n$ is monotonically increasing with d, and therefore,

$$\frac{\mathrm{tr}(\Pi_{\mathrm{sym}})}{Z_n} \leq \lim_d \frac{\mathrm{tr}(\Pi_{\mathrm{sym}})}{Z_n} = \frac{1}{k!}. \tag{117}$$

Moreover, applying Theorem 2, we find $\sum_{\Omega\neq\Omega'}(\Delta^{2t})_{\Omega',\Omega} \leq (7|\mathcal{P}|^4/d)t\|\Delta\|_\infty^{t-1}$ and, in the hypothesis of Theorem 3, we have $(7|\mathcal{P}|^4/d)t\|\Delta\|_\infty^{t-1} \leq \frac{1}{2}(\mathcal{F}_t^{(k)} - \mathcal{F}_{\mathrm{Haar}})$. ■

Corollary 3 (State k-design convergence). Let $\mathcal{S}_t$ be the ensemble of states induced by $\mathcal{C}_t$. Then for $n \geq f(k, t)$

shown in Theorem 3 the bound

$$\|\Psi^{(k)}_{\text{Haar}} - \Psi^{(k)}_t\|_1 \leq 2^{\frac{k^2}{4} - t\log\frac{16}{15} + \frac{1}{2}\log 3} \tag{118}$$

holds. Hence, $t = \Omega(k^2 + \log\varepsilon^{-1})$ are sufficient for $\mathcal{S}_t$ to be an ε-approximate state k-design in trace distance. Similarly,

$$\frac{2^{\frac{k^2}{8} - 3t - \frac{t}{2}\log k - 2}}{\text{tr}(\Pi_{\text{sym}})} \leq \|\Psi^{(k)}_{\text{Haar}} - \Psi^{(k)}_t\|_\infty, \tag{119}$$

and hence $t = \Omega(k^2/\log k + \log\varepsilon^{-1})$ are necessary for $\mathcal{S}_t$ to be a relative ε-approximate state k-design.

Proof. The proof follows from Eq. (37), Lemma 18 and Theorem 3. ■

Corollary 4 (Scrambling of doped Clifford circuits). Let $\text{OTOC}_k(U) := \mathbb{E}_{A_1,\ldots,B_k}(1/d)\text{tr}(A_1\tilde{B}_1\cdots A_k\tilde{B}_k)$, where $\tilde{B}_i := UB_iU^\dagger$. Then, we define the relative error with respect to the Haar value

$$\delta\,\text{OTOC}_k(U) := \frac{\text{OTOC}_k(U) - \mathbb{E}_{U\sim\text{Haar}}\,\text{OTOC}_k(U)}{\mathbb{E}_{U\sim\text{Haar}}\,\text{OTOC}_k(U)}. \tag{120}$$

It holds that $\int \text{d}U_t\,\delta\,\text{OTOC}_k(U_t) = \text{negl}(n)$ if and only if $t = \tilde{\Theta}(k^2 + \log^{1+c} n)$ where $c > 0$ and $\text{negl}(n)$ is a function decreasing faster than any polynomial.

Proof. We employ the relation proven in Ref. [44]: for any ensemble $\mathcal{E}$ of unitaries, it holds that

$$\mathbb{E}_{U\sim\mathcal{E}}\,\text{OTOC}_k(U) = \frac{1}{d^{2k+1}}\mathcal{F}_{\mathcal{E}} \tag{121}$$

where $\mathcal{F}_{\mathcal{E}}$ is the frame potential. Considering the relative ratio, and applying the result of Theorem 3 we find that fixing $t = O(k^2 + \log^{1+c} n)$ for some $c > 0$, then $(\mathcal{F}_t - \mathcal{F}_{\text{Haar}})/\mathcal{F}_{\text{Haar}} = \exp\big(O(\log^{1+c} n)\big)$. Conversely, we have that fixing $t = \Omega(k^2/\log k + \log^{1+c} n)$ $(\mathcal{F}_t - \mathcal{F}_{\text{Haar}})/\mathcal{F}_{\text{Haar}} = \exp\big(\Omega(\log^{1+c} n)\big)$. Noting that $\exp(\log^{1+c} n) = \text{negl}(n)$ proves the desired result. ■

B. Optimal convergence of *t*-doped Clifford circuit toward relative unitary designs

In this section, we establish the optimal convergence of the ensemble of t-doped Clifford circuits through approximate k-design in relative error. The key ingredient for the proof, is the lower bound, as for the upper bound we can import the main result shown in Ref. [73].

Let us first establish the following key lemma.

Lemma 19 (Lower bounding the largest eigenvalue of the t-doped state). Let μ_1 be any measure over the single-qubit unitary group. Let $\Psi^{(k)}_t := \Phi^{(k)}_t(\sigma^{\otimes k})$ for some $\sigma \in \text{STAB}_n$. Then

$$\|\Psi^{(k)}_t\|_\infty \geq B^t_{\mu_1}\frac{d^{k/2}}{2Z_n}, \tag{122}$$

where $B_{\mu_1} := \int \text{d}\vec{\alpha}\,\delta(|\vec{\alpha}| \leq 1/8kt)$ is the volume of a ball of radius $1/8kt$ with the measure $\text{d}\vec{\alpha}$, see Eq. (52).

Proof. Our task is to lower bound the infinity norm of $\Psi^{(k)}_t$. Let us first find tight bounds on $\|\Psi^{(k)}_{\text{Cl}}\|_\infty$. Let $\Omega \in \mathcal{P}_P$ a projective Pauli monomial with $m = k/2$. We know that $\Omega^2 = d^m\Omega$, hence Ω/d^m is a projector and $\|\Omega\|_\infty = d^m$. On the other hand, we also have $\|\Omega\|_1 = \text{tr}(\Omega) = d^{k-m}$. For the choice $m = k/2$, we have that Ω is proportional to a statevector, and that $\|\Omega\|_\infty = \|\Omega\|_1 = d^{k/2}$. We remark that a projective Pauli monomial with $m = k/2$ always exists for $k = 0 \mod 8$, while it is not guaranteed to exist for any k. For example, for $k = 4$ it does not exist. However, a projective Pauli monomial with $m = k/2 - 1$ for k even always exists, see [72] for more details. We use $m = k/2$ for simplicity. We can lower bound

$$\begin{aligned}\|\Psi^{(k)}_{\text{Cl}}\|_\infty &\geq \frac{1}{\|\Omega\|_1}\text{tr}(\Psi^{(k)}_{\text{Cl}}\Omega) = \frac{1}{Z_n\|\Omega\|_1}\sum_{\Omega'\in\mathcal{P}}\text{tr}(\Omega\Omega') \\ &\geq \frac{1}{Z_n\|\Omega\|_1}\text{tr}(\Omega^2) = \frac{d^k}{Z_n\|\Omega\|_1} = \frac{d^{k/2}}{Z_n},\end{aligned} \tag{123}$$

where we used that $\Psi^{(k)}_{\text{Cl}} \propto \sum_{\Omega\in\mathcal{P}}\Omega$ (see Lemma 9), that $\text{tr}(\Omega\Omega') \geq 1$, and that $\text{tr}(\Omega^2) = d^k$.

We are now ready to bound $\|\Psi^{(k)}_t\|_\infty$. Let $|v\rangle$ the normalized eigenstate corresponding to the maximum eigenvalue $\|\Psi^{(k)}_{\text{Cl}}\|_\infty$. We can bound $\|\Psi^{(k)}_t\|_\infty \geq \langle v|\Psi^{(k)}_t|v\rangle$. Writing $\Psi^{(k)}_t = \Phi^{(k)}_t(\Psi^{(k)}_{\text{Cl}})$, we can lower bound it as

$$\begin{aligned}\text{tr}\big(\Phi^{(k)}_t(\Psi^{(k)}_{\text{Cl}})\,|v\rangle\langle v|\big) \geq \mathbb{E}_{\vec{P}_i}\int\prod_i \text{d}\vec{\alpha}_i\delta(|\vec{\alpha}| \leq \varepsilon)\text{tr}\Bigg(|v\rangle\langle v| \\ \left[\prod_i U_i^{\otimes k}(\vec{\alpha}_i)\right]^\dagger\Psi^{(k)}_{\text{Cl}}\left[\prod_i U_i^{\otimes k}(\vec{\alpha}_i)\right]^\dagger\Bigg).\end{aligned} \tag{124}$$

Now, notice that for every i, we have

$$\begin{aligned}\|U(\vec{\alpha}) - \mathbb{1}\|_\infty &= \sum_{l=1}^\infty\frac{1}{l!}\|(i\vec{\alpha}\cdot\vec{P})^l\|_\infty \\ &\leq \sum_{l=1}^\infty\frac{1}{l!}|\vec{\alpha}|^l = e^{|\vec{\alpha}|} - 1 \leq 2\varepsilon.\end{aligned} \tag{125}$$

Since, we restricted the domain of integration up to $|\vec{\alpha}| \leq \varepsilon$. The above inequality holds whenever $\varepsilon \leq 1$. It easily

$$\mathrm{tr}\left(|v\rangle\langle v|\left[\prod_i U_i^{\otimes k}(\vec{\alpha}_i)\right]^\dagger \Psi_{\mathrm{Cl}}^{(k)}\left[\prod_i U_i^{\otimes k}(\vec{\alpha}_i)\right]^\dagger\right) \geq \mathrm{tr}(|v\rangle\langle v|\,\Psi_{\mathrm{Cl}}^{(k)}) - \left|\mathrm{tr}\left(|v\rangle\langle v|\,V\Psi_{\mathrm{Cl}}^{(k)}V^\dagger\right) - \mathrm{tr}\left(|v\rangle\langle v|\,\Psi_{\mathrm{Cl}}^{(k)}\right)\right| \tag{126}$$

Now, notice that

$$\begin{aligned}\left|\mathrm{tr}\left(|v\rangle\langle v|\,V\Psi_{\mathrm{Cl}}^{(k)}V^\dagger\right) - \mathrm{tr}\left(|v\rangle\langle v|\,\Psi_{\mathrm{Cl}}^{(k)}\right)\right| &= \left|\mathrm{tr}\right)|v\rangle\langle v|\,V\Psi_{\mathrm{Cl}}^{(k)}[V^\dagger - \mathbb{1}]) + \mathrm{tr}\left(\Psi_{\mathrm{Cl}}^{(k)}|v\rangle\langle v|\,[V-\mathbb{1}]\right)\right| \\ &\leq \|V\|_\infty \|\Psi_{\mathrm{Cl}}^{(k)}[V^\dagger - \mathbb{1}]\,|v\rangle\langle v|\,\|_1 + \|V-\mathbb{1}\|_\infty \|\Psi_{\mathrm{Cl}}^{(k)}\,|v\rangle\langle v|\,\|_1 \\ &\leq \|\Psi_{\mathrm{Cl}}^{(k)}\|_\infty \|V^\dagger - \mathbb{1}\|_\infty \|\,|v\rangle\langle v|\,\|_1 + \|V-\mathbb{1}\|_\infty \langle v|\Psi_{\mathrm{Cl}}^{(k)}|v\rangle \\ &= 2\|\Psi_{\mathrm{Cl}}^{(k)}\|_\infty \|V-\mathbb{1}\|_\infty. \end{aligned} \tag{127}$$

Noticing that $\|V-\mathbb{1}\|_\infty \leq t\|U^{\otimes k}(\vec{\alpha})-\mathbb{1}\|_\infty \leq 4tk\varepsilon$, we obtain

$$\begin{aligned}&\mathrm{tr}\left(|v\rangle\langle v|\left[\prod_i U_i^{\otimes k}(\vec{\alpha}_i)\right]^\dagger \Psi_{\mathrm{Cl}}^{(k)}\left[\prod_i U_i^{\otimes k}(\vec{\alpha}_i)\right]^\dagger\right) \\ &\quad\geq \|\Psi_{\mathrm{Cl}}^{(k)}\|_\infty(1-4tk\varepsilon). \end{aligned} \tag{128}$$

Choosing $\varepsilon = 1/8kt$, we arrive at

$$\|\Psi_t^{(k)}\|_\infty \geq \frac{1}{2}B_{\mu_1}^t \|\Psi_{\mathrm{Cl}}^{(k)}\|_\infty \geq B_{\mu_1}^t \frac{d^{k/2}}{2Z_n}, \tag{129}$$

concluding the proof. ■

The next theorem establishes the key lower bound to convergence to relative ε-approximate k-design.

Theorem 4 (Lower bound to relative unitary designs). Let $\delta > 0$. Let $\mathcal{C}_t$ be the ensemble of t-doped Clifford circuits. For any $n \geq 2\delta^{-1}k$ and $0 \leq \varepsilon \leq d^{\frac{k}{2}(1-\delta)}$, then $\mathcal{C}_t$ cannot be a relative ε-approximate unitary design unless

$$t \geq \frac{nk}{6\delta^{-1}\log B_{\mu_1}^{-1}}. \tag{130}$$

Proof. According to Definition 6, a necessary condition for $\mathcal{C}_t$ being a relative design is the channel $\delta\Phi = (1+\varepsilon)\Phi_{\mathrm{Haar}}^{(k)} - \Phi_t^{(k)}$ to be a completely positive quantum channel. A necessary condition for complete positivity is simply positivity, i.e., the channel $\delta\Phi$ to output positive matrices when applied on states. We can then apply $\delta\Phi$ on $\sigma^{\otimes k}$ for $\sigma \in \mathrm{STAB}_n$ and get

$$\delta\Phi(\sigma^{\otimes k}) = (1+\varepsilon)\Psi_{\mathrm{Haar}}^{(k)} - \Psi_t^{(k)}. \tag{131}$$

A necessary condition for positivity of the matrix $\delta\Phi(\sigma^{\otimes k})$ is the smallest eigenvalue to be positive. Denoting $\lambda(\delta\Phi(\sigma^{\otimes k}))$ the smallest eigenvalue of $\delta\Phi(\sigma^{\otimes k})$, we can upper bound it as $\lambda(\delta\Phi(\sigma^{\otimes k})) \leq (1+\varepsilon)(1/\mathrm{tr}(\Pi_{\mathrm{sym}})) - B_{\mu_1}^t(d^{k/2}/2Z_n)$, because $[\Psi_{\mathrm{Haar}}^{(k)}, \Psi_t^{(k)}] = 0$. Imposing it larger than 0, from Lemma 19, we arrive at the condition

$$t \geq \frac{1}{\log B_{\mu_1}^{-1}}\left(\frac{nk}{2} - \log(1+\varepsilon) + \log\frac{\mathrm{tr}(\Pi_{\mathrm{sym}})}{2Z_n}\right). \tag{132}$$

Then, for any $n \geq k-1$, we can further lower bound $\log\left(\mathrm{tr}\Pi_{\mathrm{sym}}/Z_n\right)$ using $\log\mathrm{tr}\Pi_{\mathrm{sym}} \geq nk - k\log k$, $\log Z_n \leq nk + \sum_{i=0}^{k-2} \leq nk + 1 - 1/d$, implying $\log\left(\mathrm{tr}\Pi_{\mathrm{sym}}/Z_n\right) \geq -k\log k - 1 + 1/d$. We arrive at

$$t \geq \frac{1}{\log B_{\mu_1}^{-1}}\left(\frac{nk}{2} - \log(1+\varepsilon) - k\log k - 2\right). \tag{133}$$

Let us now discriminate two cases. If $\varepsilon \leq 1$, we have $\log(1+\varepsilon) \leq 1$, hence

$$t \geq \frac{1}{\log B_{\mu_1}^{-1}}\left(\frac{nk}{2} - k\log k - 3\right). \tag{134}$$

Whereas, for any $1 < \varepsilon \leq d^{\frac{k}{2}(1-\delta)}$ and $0 < \delta \leq 1$, we have $\log(1+\varepsilon) \leq \log\varepsilon + (\log e/\varepsilon) \leq nk(1-\delta)/2 + \log e$, which leads to

$$t \geq \frac{1}{\log B_{\mu_1}^{-1}}\left(\frac{nk\delta}{2} - k\log k - 2 - \log e\right). \tag{135}$$

Therefore, for any ε in the range $0 \leq \varepsilon \leq d^{\frac{k(1-\delta)}{2}}$ with $0 < \delta < 1$, the lower bound in Eq. (135) applies being the lowest lower bound. Finally, we have that $\log k + (2 + \log e)/k \leq \frac{2}{3}(n/2)\delta$ for any $n \geq 2\delta^{-1}k$ and $k > 4$, which leads to $t \geq nk/6\delta^{-1}\log B_{\mu_1}^{-1}$ as desired. ■

In the following, we import the result from Ref. [73] to prove an upper bound to relative error unitary designs that matches the lower bound of Theorem 4. We then also

consider the measure μ_1 used in Ref. [73], which is the sum of three delta-functions on U, $U^\dagger$, and I, where U is a generic single-qubit non-Clifford gate.

Corollary 5 (Lower bounding the number of non-Clifford resources). Let $\delta > 0$. Let $\mathcal{C}_t$ be the ensemble of t-doped Clifford circuits and let μ_1 be a single-qubit measure over one of the following ensembles: the discrete set $\{U, U^\dagger, I\}$, the unitary ensemble, or the diagonal unitary ensemble. For any $n \geq 2\delta^{-1}k$ and $0 \leq \varepsilon \leq d^{\frac{k}{2}(1-\delta)}$, then the following lower bounds on the number of non-Clifford resources t hold,

$$t \geq \frac{nk\delta}{18} \qquad \{U, U^\dagger, I\},$$

$$t \geq \frac{nk\delta}{6\log(nk\delta)(17+6\log k)} \qquad \text{Unitary ensemble},$$

$$t \geq \frac{nk\delta}{6\log(nk\delta)(2+\log k)} \qquad \text{Diagonal unitary ensemble}. \tag{136}$$

Proof. Let μ_1 be the discrete measure that samples over $U, U^\dagger, I$. Then for any U such that $\|U - I\|_\infty \geq \Omega(1)$, i.e., that does not depend on k, t, then we simply have $B_{\mu_1} = \frac{1}{3}$, hence we have $t \geq nk\delta/18$.

In the second case, we want to compute $B_{\mu_1} = \int \mathrm{d}\vec{\alpha}\,\delta(\alpha \leq (8kt)^{-1})$ for the Haar measure. Let us start by considering the unitary decomposition $U = \exp(i\vec{\alpha}\cdot\vec{P})$. For a single-qubit unitary, the proposed decomposition can also be written as

$$U = \begin{pmatrix} \cos(|\vec{\alpha}|) + i\frac{\alpha_3}{|\vec{\alpha}|}\sin(|\vec{\alpha}|) & i\frac{\alpha_1}{|\vec{\alpha}|}\sin(|\vec{\alpha}|) + \frac{\alpha_2}{|\vec{\alpha}|}\sin(|\vec{\alpha}|) \\ i\frac{\alpha_1}{|\vec{\alpha}|}\sin(|\vec{\alpha}|) - \frac{\alpha_2}{|\vec{\alpha}|}\sin(|\vec{\alpha}|) & \cos(|\vec{\alpha}|) - i\frac{\alpha_3}{|\vec{\alpha}|}\sin(|\vec{\alpha}|) \end{pmatrix}. \tag{137}$$

Following Ref. [101], we get that the measure over the single-qubit unitary group can be expressed as

$$\int_{|\vec{\alpha}|\leq c}\prod_i \mathrm{d}\vec{\alpha}_i = \frac{1}{\pi^2}\int_0^{2\pi}\mathrm{d}\phi\int_0^{\pi}\mathrm{d}\theta\int_0^{\sin^2 c}\mathrm{d}r\frac{1}{\sqrt{1-r^2}}r^2\sin\theta. \tag{138}$$

The conversion is straightforward: first converting to spinorial coordinates, then to spherical ones. Note that for $c = 2\pi$ the integral returns 1, being the measure normalized. The integral is then equal to

$$\int_{|\vec{\alpha}|\leq c}\prod_i \mathrm{d}\vec{\alpha}_i = \frac{1}{\pi}\left(2\arctan\left(\frac{\sqrt{2}\sin^2(c)}{\cos(c)\sqrt{3-\cos(2c)}}\right) - \sqrt{2}\sin^2(c)\cos(c)\sqrt{3-\cos(2c)}\right). \tag{139}$$

By considering $c = 1/8kt$, we get

$$\int_{|\vec{\alpha}|\leq \frac{1}{8kt}}\prod_i \mathrm{d}\vec{\alpha}_i \geq \frac{4}{3\pi(8kt)^6}\left(1-\frac{1}{(8kt)^2}\right) \geq \frac{63}{48\pi(8kt)^6}, \tag{140}$$

from which we obtain $t\log t \geq nk\delta/6(17+6\log k)$, and, therefore,

$$t \geq \frac{nk}{6\log(nk\delta)(17+6\log k)} \tag{141}$$

is a necessary condition. Let us conclude by computing the same but considering diagonal unitary matrices. We have then

$$B_{\mu_1} = \int_{-\frac{1}{8kt}}^{\frac{1}{8kt}}\mathrm{d}\phi = \frac{1}{4kt}, \tag{142}$$

from which we obtain $t\log t \geq nk\delta/6(2+\log k)$ and, therefore,

$$t \geq \frac{nk\delta}{6\log(nk\delta)(2+\log k)}, \tag{143}$$

concluding the proof. ■

Proposition 3 (Upper bound to relative unitary designs, Corollary 1 in Ref. [73]). Let $\varepsilon > 0$. Let $\mathcal{C}_t$ be the ensemble of t-doped Clifford circuits. Let U be a single-qubit non-Clifford gate. Let μ_1 be the uniform measure over $U, U^\dagger, I$. Then there exist two constants $C_1(U), C_2(U)$ such that for $t \geq C_1(U)\log^2(k)(2nk + \log\varepsilon^{-1})$ then $\mathcal{C}_t$ is a relative ε-approximate k-design, whenever $n \geq C_2(U)k^2$. If instead μ_1 is the single-qubit Haar measure, analogous conclusions can be reached.

Therefore, combining Theorem 4 and Proposition 3, we conclude that, for $k = O(\sqrt{n})$, t-doped Clifford circuits are relative ε-approximate k-design iff $t = \tilde{\Theta}(kn)$.

The convergence lower bound in Theorem 4 holds whenever $k = O(n)$. Since a $k+1$-design must be at least a k-design, we have that for any $k = \Omega(n)$ the number of non-Clifford gates to be applied is at least $t = \Omega(n^2)$, underlying an extremely high cost in terms of non-Clifford resources in producing k-design with large k. In the following proposition, we also cover the regime $k = \Omega(n)$, using the same proof technique of Theorem 4 and a different version of Lemma 19.

Theorem 5 (Lower bound to relative unitary k-design for $k = \Omega(n)$). Let $\delta > 0$. Let $\mathcal{C}_t$ be the ensemble of t-doped Clifford circuit and $U = e^{i\frac{\pi}{8}Z}$ being the T-gate. For any $8n\delta^{-1} \leq k \leq 2^{\frac{n\delta}{2}}$ and $0 \leq \varepsilon \leq d^{\frac{k}{2}(1-\delta)}$, then $\mathcal{C}_t$ cannot be a relative ε-approximate unitary k-design unless

$$t \geq \frac{k\delta}{8}. \tag{144}$$

Proof. Consider $\Psi_t^{(k)} := \Phi_t^{(k)}(\sigma^{\otimes k})$ for $\sigma \in \mathrm{STAB}_n$. We then have the bound

$$\|\Psi_t^{(k)}\|_2 \leq \sqrt{\mathrm{rank}(\Psi_t^{(k)})} \cdot \|\Psi_t^{(k)}\|_\infty. \tag{145}$$

The 2-norm $\|\Psi_t^{(k)}\|_2$ corresponds to (the square root of) the purity of the state vector $\Psi_t^{(k)}$. For an ensemble of pure states $\mathcal{E} = \{|\psi_i\rangle\}$, the average over k-fold tensor powers is given by $\Psi_{\mathcal{E}}^{(k)} = (1/|\mathcal{E}|)\sum_{|\psi\rangle\in\mathcal{E}} |\psi\rangle\langle\psi|^{\otimes k}$, whose purity can be lower bounded

$$\mathrm{tr}((\Psi_{\mathcal{E}}^{(k)})^2) = \frac{1}{|\mathcal{E}|} + \frac{1}{|\mathcal{E}|^2}\sum_{|\psi\rangle\neq|\phi\rangle\in\mathcal{E}} |\langle\psi|\phi\rangle|^2 \geq \frac{1}{|\mathcal{E}|}. \tag{146}$$

This implies that the 2-norm is lower bounded by the inverse square root of the size of the ensemble, i.e., $\|\Psi_{\mathcal{E}}^{(k)}\|_2 \geq |\mathcal{E}|^{-1/2}$. Choosing $\mathcal{E} = \mathcal{C}_t$, the ensemble of t-doped Clifford circuits, we can crudely upper bound the ensemble size by $|\mathcal{C}_t| \leq |\mathcal{C}_n| \cdot 4^{nt}$, where $|\mathcal{C}_n| \leq 2^{2n^2+n}$ is the size of the Clifford group, and 4^{nt} accounts for the number of possible rotations $e^{i\frac{\pi}{8}P}$ with $P \in \mathbb{P}_n$. Therefore, we obtain the lower bound

$$\|\Psi_t^{(k)}\|_\infty \geq \frac{2^{-nt-n^2-n/2}}{\sqrt{\mathrm{tr}(\Pi_{\mathrm{sym}})}}, \tag{147}$$

where we have used that $\mathrm{rank}(\Psi_t^{(k)}) = \mathrm{tr}(\Pi_{\mathrm{sym}})$. Now, using the same strategy as in Theorem 4, we can upper bound the smallest eigenvalue of $\delta\Phi(\sigma^{\otimes k}) = (1+\varepsilon)\Psi_{\mathrm{Haar}}^{(k)} - \Psi_t^{(k)}$ as

$$\lambda_{\min}(\delta\Phi(\sigma^{\otimes k})) \leq \frac{1+\varepsilon}{\mathrm{tr}(\Pi_{\mathrm{sym}})} - \frac{2^{-nt-n^2-n/2}}{\sqrt{\mathrm{tr}(\Pi_{\mathrm{sym}})}}. \tag{148}$$

Requiring this quantity to be non-negative and applying the bounds from Theorem 4, we arrive at

$$t \geq \frac{k}{2} - n - \frac{k\log k}{2n} - \frac{\log(1+\varepsilon)}{n} - \frac{1}{2}. \tag{149}$$

For any $\varepsilon \leq d^{\frac{k}{2}(1-\delta)}$ with $0 < \delta \leq 1$, we obtain

$$t \geq \frac{k\delta}{2} - n - \frac{k\log k}{2n} - \frac{\log e}{n} - \frac{1}{2}, \tag{150}$$

which implies $t \geq k\delta/8$, for any $8n\delta^{-1} \leq k \leq 2^{n\delta/2}$, thus concluding the proof. ■

C. The non-Clifford cost of generating pseudorandom unitaries

In this section, we provide a proof of Proposition 1 by synthesizing results from Refs. [58,91]. A similar proof is also presented in Ref. [102]. First, Ref. [58] shows that it is possible to construct a pseudorandom unitary ensemble using a shallow quantum circuit of depth $O(\mathrm{poly}\log\log n)$, assuming the hardness of the *learning with errors* (LWE) problem for quantum computers. We refer the reader to Ref. [58] for full technical details. A circuit of depth D on n qubits contains at most $O(nD)$ gates, and each gate can be implemented using $O(1)$ single-qubit non-Clifford gates [100]. Hence, the construction in Ref. [58] requires at most $O(n\,\mathrm{poly}\log\log n)$ non-Clifford gates, which yields the upper bound in Proposition 1.

To establish the lower bound, consider a quantum state vector $|\psi\rangle$ on n qubits. Its associated stabilizer group is defined as

$$G_\psi := \{P \in \mathbb{P}_n : \langle\psi|P|\psi\rangle = \pm 1\}. \tag{151}$$

According to Refs. [91,103], if $|\psi\rangle$ is prepared using t single-qubit non-Clifford gates, then $\dim G_\psi \geq n - 2t$. Furthermore, Ref. [91] introduced a polynomial-time quantum algorithm that, given a state vector $|\psi\rangle$ and parameters $k \in [n]$ $\varepsilon = \Omega(\mathrm{poly}^{-1} n)$, can determine whether

(1) $\dim G_\psi \geq n - k$ or
(2) $|\psi\rangle$ is ε-far in trace distance from any state satisfying (1).

Now, assume for contradiction that there exists a pseudorandom ensemble $\mathcal{U}$ generated using only $\leq \alpha n$ with

$\alpha < 1$ non-Clifford gates (with high probability). Then, we can build a polynomial-time distinguisher as follows: apply the algorithm from Ref. [91] to states generated (from $|0\rangle^{\otimes n}$) by unitaries from $\mathcal{U}$ and from the Haar measure. For unitaries in $\mathcal{U}$, the algorithm will output case (1). We now show that for Haar-random unitaries, the algorithm almost surely outputs case (2).

To do so, it suffices to prove that a Haar-random state is far in trace distance from any state with stabilizer group dimension $\geq n-k$, for $k \leq \alpha n$. By the compression theorem [103], any such state vector can be expressed as $|\psi\rangle = C(|0\rangle^{\otimes n-k} \otimes |\phi_k\rangle)$, where $|\phi_k\rangle$ is a k-qubit state vector and C is a Clifford circuit. Thus, the number of such states is at most $|\mathcal{C}_n| \cdot |S_{k,\varepsilon}|$, where $S_{k,\varepsilon}$ is an ε-net covering of the k-qubit pure state space.

Applying Lévy's lemma and a union bound, we obtain

$$\begin{aligned}\Pr_{|\psi\rangle\sim\text{Haar}}\big(\exists\,|\psi_{n-k}\rangle \text{ with } \dim G_{\psi_{n-k}} \geq n-k \text{ and} \\ \operatorname{tr}(\psi\psi_{n-k}) \geq d^{-1/3}\big) \leq |\mathcal{C}_n| \cdot |S_{k,\varepsilon}| \\ \cdot e^{-\Omega(2^{n/3})} = \operatorname{negl}(n). \end{aligned} \tag{152}$$

The last step of Eq. (152) follows from the fact that $|C_n| \leq 2^{2n^2+n}$, while

$$S_{k,\varepsilon} = \left(\frac{1}{\varepsilon}\right)^{O(2^k)} = 2^{O(2^{n\alpha})} \tag{153}$$

for any $\varepsilon = \Theta(1)$. Hence, fixing $\varepsilon = 1/4$ and choosing α appropriately, Eq. (152) follows. Hence, Eq. (152) proves that, with overwhelming probability, a Haar-random state is at least $\sqrt{1-d^{-1/3}}$-far in trace distance from any state vector $|\phi\rangle = C(|0\rangle \otimes |\phi_k\rangle)$ with $|\phi_k\rangle \in S_{k,1/4}$. Since $S_{k,1/4}$ is an ε-covering net with $\varepsilon = 1/4$, this further implies that the Haar-random state is at least $\sqrt{1-d^{-1/3}} - 1/4 \geq 0.2$-far from any state with $\dim G \geq n-k$. We conclude that the algorithm from Ref. [91] serves as a valid distinguisher unless $t > \alpha n = \Omega(n)$, completing the proof of Proposition 1.

D. Implications for ε-nets

In this work, we have quantified in what precise sense the ensemble of doped Clifford circuits is spread within the unitary group. We here explore finally a different reading of that notion, i.e., to what extent the quantum circuits prepared by doped Clifford circuits approximate all quantum states, in the sense of constituting an ϵ-net, defined below.

Definition 12 (ε-net). Let $\mathcal{E}$ be a discrete ensemble of unitary operators. $\mathcal{E}$ is a ε covering net iff for any $U \in \mathcal{U}_n$, there exists $U' \in \mathcal{E}$ such that $\|U - U'\|_\diamond \leq \varepsilon$, where $\|\cdot\|_\diamond$ is the diamond norm.

Below, we show that, as one would expect, creating an ε-net requires an exponential amount of non-Clifford resources, as it also requires an exponential gate complexity.

Proposition 4 (ε-net via t-doped Clifford circuits). Let $\mathcal{C}_t$ be the ensemble of t-doped Clifford circuits doped with single-qubit T gates. Then $\mathcal{C}_t$ cannot form a ε-net with respect to the diamond norm on unitaries unless $t = \Omega(4^n/(n\log\varepsilon^{-1}))$.

Proof. The proof is done via a simple counting argument, and using some steps of the proof of Theorem 5. It is known [104, Proposition 7] (see also Ref. [105, Lemma 9]) that for an ensemble of unitaries $\mathcal{E}$ to form a ε-net with respect to the diamond norm the size $|\mathcal{E}|$ must be lower bounded by

$$|\mathcal{E}| \geq \left(\frac{c_\diamond}{\varepsilon}\right)^{d^2}, \tag{154}$$

with $c_\diamond > 0$ being a universal constant. Hence, a necessary condition for $\mathcal{C}_t$ to form a ε-net is

$$|\mathcal{C}_t| \geq \left(\frac{c_\diamond}{d}\right)^{d^2}. \tag{155}$$

We can upper bound $|\mathcal{C}_t| \leq 2^{2n^2+n+4nt}$, which immediately gives the necessary condition on the doping parameter t as

$$t \geq \frac{d^2}{n}\log\frac{c_\diamond}{\varepsilon} - n - 1, \tag{156}$$

proving the statement. ■

These statements provide further evidence that exploring the full set of unitaries—for example, by constructing an ϵ-net or by generating unitaries at random—is extremely costly in terms of non-Clifford resources.

ACKNOWLEDGMENTS

This work has been inspired by the Ph.D. thesis of L.L. [106]. We thank Toshihiro Yada for having identified a flaw in an earlier version of the manuscript. This work has been supported by the DFG (CRC 183, FOR 2724), by the BMBF (Hybrid++, QuSol), the BMWK (EniQmA), the Munich Quantum Valley (K-8), the QuantERA (HQCC), the Alexander-von-Humboldt Foundation, and Berlin Quantum. This work has also been funded by the DFG under Germany's Excellence Strategy—The Berlin Mathematics Research Center MATH+ (EXC-2046/1, Project ID: 390685689). S.F.E.O. acknowledges support from PNRR MUR project (No. PE0000023-NQSTI). A.H. acknowledges support from the PNRR MUR project (No. PE0000023-NQSTI) and the PNRR MUR project (No. CN 00000013-ICSC).

DATA AVAILABILITY

No data were created or analyzed in this study.

[1] A. Ambainis and A. Smith, in *Proceedings of RANDOM'04*, Lecture Notes in Computer Science Vol. 3122 (Springer, Cambridge, MA, USA, 2004), pp. 249–260.

[2] P. Hayden, D. Leung, P. W. Shor, and A. Winter, Randomizing quantum states: Constructions and applications, Commun. Math. Phys. **250**, 371 (2004).

[3] W. Kretschmer, Quantum Pseudorandomness and Classical Complexity, in *Proceedings of the 16th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2021)*, Leibniz International Proceedings in Informatics (LIPIcs) (Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2021), Vol. 197, pp. 2:1–2:20.

[4] T. Morimae and T. Yamakawa, in *CRYPTO 2022* (2022), pp. 269–295.

[5] P. Ananth, L. Qian, and H. Yuen, in *CRYPTO 2022* (2022), pp. 208–236.

[6] P. Sen, in *21st Annual IEEE Conference on Computational Complexity (CCC'06)* (IEEE, Prague, Czech Republic, 2006), pp. 14–287.

[7] F. G. S. L. Brandão and M. Horodecki, Exponential quantum speed-ups are generic, Quantum Inf. Comput. **13**, 901 (2013).

[8] S. Boixo, S. V. Isakov, V. N. Smelyanskiy, R. Babbush, N. Ding, Z. Jiang, M. J. Bremner, J. M. Martinis, and H. Neven, Characterizing quantum supremacy in near-term devices, Nat. Phys. **14**, 595 (2018).

[9] F. Arute *et al.*, Quantum supremacy using a programmable superconducting processor, Nature **574**, 505 (2019).

[10] A. Bouland, B. Fefferman, C. Nirkhe, and U. Vazirani, On the complexity and verification of quantum random circuit sampling, Nat. Phys. **15**, 159 (2019).

[11] R. Kueng, H. Rauhut, and U. Terstiege, Low rank matrix recovery from rank one measurements, Appl. Comput. Harmon. Anal. **42**, 88 (2017).

[12] S. Kimmel and Y. Liu, in *2017 International Conference on Sampling Theory and Applications (SampTA)* (IEEE, Tallinn, Estonia, 2017), pp. 345–349.

[13] R. Kueng, H. Zhu, and D. Gross, Distinguishing quantum states using Clifford orbits, arXiv:1609.08595.

[14] M. Oszmaniec, R. Augusiak, C. Gogolin, J. Kolodynski, A. Acin, and M. Lewenstein, Random bosonic states for robust quantum metrology, Phys. Rev. X **6**, 041044 (2016).

[15] J. Eisert, D. Hangleiter, N. Walk, I. Roth, D. Markham, R. Parekh, U. Chabaud, and E. Kashefi, Quantum certification and benchmarking, Nat. Rev. Phys. **2**, 382 (2020).

[16] C. Dankert, R. Cleve, J. Emerson, and E. Livine, Exact and approximate unitary 2-designs and their application to fidelity estimation, Phys. Rev. A **80**, 012304 (2009).

[17] E. Magesan, J. M. Gambetta, and J. Emerson, Scalable and robust randomized benchmarking of quantum processes, Phys. Rev. Lett. **106**, 180504 (2011).

[18] B. Magni, A. Christopoulos, A. D. Luca, and X. Turkeshi, Anticoncentration in Clifford circuits and beyond: From random tensor networks to pseudo-magic states, Phys. Rev. X **15**, 031071 (2025).

[19] I. Devetak, The private classical capacity and quantum capacity of a quantum channel, IEEE Trans. Inf. Theory **51**, 44 (2005).

[20] I. Devetak and A. Winter, Relating quantum privacy and quantum coherence: An operational approach, Phys. Rev. Lett. **93**, 080501 (2004).

[21] B. Groisman, S. Popescu, and A. Winter, Quantum, classical, and total amount of correlations in a quantum state, Phys. Rev. A **72**, 032317 (2005).

[22] A. Abeyesinghe, I. Devetak, P. Hayden, and A. Winter, The mother of all protocols: Restructuring quantum information's family tree, Proc. R. Soc. A **465**, 2537 (2009).

[23] F. Dupuis, M. Berta, J. Wullschleger, and R. Renner, One-shot decoupling, Commun. Math. Phys. **328**, 251 (2014).

[24] O. Szehr, F. Dupuis, M. Tomamichel, R. Renner *et al.*, Decoupling with unitary approximate two-designs, New J. Phys. **15**, 053022 (2013).

[25] M. Horodecki, J. Oppenheim, and A. Winter, Partial quantum information, Nature **436**, 673 (2005).

[26] M. Horodecki, J. Oppenheim, and A. Winter, Quantum state merging and negative information, Commun. Math. Phys. **269**, 107 (2006).

[27] Y. Nakata, E. Wakakuwa, and H. Yamasaki, One-shot quantum error correction of classical and quantum information, Phys. Rev. A **104**, 012408 (2021).

[28] E. Wakakuwa and Y. Nakata, One-shot triple-resource trade-off in quantum channel coding, IEEE Trans. Inf. Theory **69**, 2400 (2023).

[29] S. Popescu, A. J. Short, and A. Winter, Entanglement and the foundations of statistical mechanics, Nat. Phys. **2**, 754 (2006).

[30] N. Linden, S. Popescu, A. J. Short, and A. Winter, Quantum mechanical evolution towards thermal equilibrium, Phys. Rev. E **79**, 061103 (2009).

[31] L. del Rio, A. Hutter, R. Renner, and S. Wehner, Relative thermalization, Phys. Rev. E **94**, 022104 (2016).

[32] K. Kaneko, E. Iyoda, and T. Sagawa, Characterizing complexity of many-body quantum dynamics by higher-order eigenstate thermalization, Phys. Rev. A **101**, 042126 (2020).

[33] C. Gogolin and J. Eisert, Equilibration, thermalisation, and the emergence of statistical mechanics in closed quantum systems, Rep. Prog. Phys. **79**, 56001 (2016).

[34] M. Ippoliti and W. W. Ho, Solvable model of deep thermalization with distinct design times, Quantum **6**, 886 (2022).

[35] P. Hayden and J. Preskill, Black holes as mirrors: Quantum information in random subsystems, J. High Energy Phys. **2007**, 120 (2007).

[36] Y. Nakata, E. Wakakuwa, and M. Koashi, Black holes as clouded mirrors: The Hayden-Preskill protocol with symmetry, Quantum **7**, 928 (2023).

[37] E. Onorati, O. Buerschaper, M. Kliesch, W. Brown, A. H. Werner, and J. Eisert, Mixing properties of stochastic quantum Hamiltonians, Commun. Math. Phys. **355**, 905 (2017).

[38] Y. Nakata, Y. Takeuchi, M. Kliesch, and A. Darmawan, On computational complexity of unitary and state design properties, PRX Quantum **6**, 030345 (2025).

[39] S. F. E. Oliviero, L. Leone, S. Lloyd, and A. Hamma, Unscrambling quantum information with Clifford decoders, Phys. Rev. Lett. **132**, 080402 (2024).

[40] L. Leone, S. F. E. Oliviero, S. Lloyd, and A. Hamma, Learning efficient decoders for quasichaotic quantum scramblers, Phys. Rev. A **109**, 022429 (2024).

[41] Y. Sekino and L. Susskind, Fast scramblers, J. High Energy Phys. **2008**, 065 (2008).

[42] N. Lashkari, D. Stanford, M. Hastings, T. Osborne, and P. Hayden, Towards the fast scrambling conjecture, J. High Energy Phys. **2013**, 22 (2013).

[43] J. Maldacena, S. H. Shenker, and D. Stanford, A bound on chaos, J. High Energy Phys. **2016**, 106 (2016).

[44] D. A. Roberts and B. Yoshida, Chaos and complexity by design, J. High Energy Phys. **2017**, 121 (2017).

[45] J. S. Cotler, D. K. Mark, H-Y. Huang, F. Hernández, J. Choi, A. L. Shaw, M. Endres, and S. Choi, Emergent quantum state designs from individual many-body wave functions, PRX Quantum **4**, 010311 (2023).

[46] D. K. Mark, D.K. Mark, H.Y. Huang, F. Hernandez, J. Choi, A.L. Shaw, M. Endres, and S. Choi, Maximum entropy principle in deep thermalization and in Hilbert-space ergodicity, Phys. Rev. X **14**, 041051 (2024).

[47] S. Pilatowsky-Cameo, I. Marvian, S. Choi, and W. W. Ho, Hilbert-space ergodicity in driven quantum systems: Obstructions and designs, Phys. Rev. X **14**, 041059 (2024).

[48] A. Munson, N. B. T. Kothakonda, J. Haferkamp, N. Yunger Halpern, J. Eisert, and P. Faist, Complexity-constrained quantum thermodynamics, PRX Quantum **6**, 010346 (2025).

[49] J. Emerson, Y. S. Weinstein, M. Saraceno, S. Lloyd, and D. G. Cory, Pseudo-random unitary operators for quantum information processing, Science **302**, 2098 (2003).

[50] D. Gross, K. Audenaert, and J. Eisert, Evenly distributed unitaries: On the structure of unitary designs, J. Math. Phys. **48**, 052104 (2007).

[51] J. Haferkamp, F. Montealegre-Mora, M. Heinrich, J. Eisert, D. Gross, and I. Roth, Efficient unitary designs with a system-size independent number of non-Clifford gates, Commun. Math. Phys. **397**, 995 (2023).

[52] F. G. S. L. Brandão, A. W. Harrow, and M. Horodecki, Local random quantum circuits are approximate polynomial-designs, Commun. Math. Phys. **346**, 397 (2016).

[53] A. W. Harrow and R. A. Low, Random quantum circuits are approximate 2-designs, Commun. Math. Phys. **291**, 257 (2009).

[54] A. W. Harrow and S. Mehraban, Approximate unitary t-designs by short random quantum circuits using nearest-neighbor and long-range gates, Commun. Math. Phys. **401**, 1531 (2023).

[55] N. LaRacuente and F. Leditzky, Approximate unitary *k*-designs from shallow, low-communication circuits, Commun. Math. Phys. **407**, 51 (2026).

[56] Z. Ji, Y.-K. Liu, and F. Song, in *Advances in Cryptology – CRYPTO 2018* (Springer International Publishing, Santa Barbara, CA, USA, 2018), pp. 126–152.

[57] T. Metger, A. Poremba, M. Sinha, and H. Yuen, Pseudorandom unitaries with non-adaptive security, arXiv:2402.14803.

[58] T. Schuster, J. Haferkamp, and H.-Y. Huang, Random unitaries in extremely low depth, Science **389**, 92 (2025).

[59] F. Ma and H.-Y. Huang, How to Construct Random Unitaries, in *Proceedings of the 57th Annual ACM Symposium on Theory of Computing (STOC '25)* (Association for Computing Machinery, New York, NY, USA, 2025), pp. 806–809.

[60] T. Haug, K. Bharti, and D. E. Koh, Pseudorandom unitaries are neither real nor sparse nor noise-robust, Quantum **9**, 1759 (2025).

[61] B. Eastin and E. Knill, Restrictions on transversal encoded quantum gate sets, Phys. Rev. Lett. **102**, 110502 (2009).

[62] D. Gottesman, Stabilizer codes and quantum error correction, Ph.D. thesis, California Institute of Technology, 1997.

[63] Z. Webb, The Clifford group forms a unitary 3-design, QIC **16**, 1379 (2016).

[64] H. Zhu, Multiqubit Clifford groups are unitary 3-designs, Phys. Rev. A **96**, 062336 (2017).

[65] A. J. Scott, Optimizing quantum process tomography with unitary 2-designs, J. Phys. A **41**, 055308 (2008).

[66] I. Roth, R. Kueng, S. Kimmel, Y.-K. Liu, D. Gross, J. Eisert, and M. Kliesch, Recovering quantum gates from few average gate fidelities, Phys. Rev. Lett. **121**, 170502 (2018).

[67] J. J. Wallman and S. T. Flammia, Randomized benchmarking with confidence, New J. Phys. **16**, 103032 (2014).

[68] J. J. Wallman, Randomized benchmarking with gate-dependent noise, Quantum **2**, 47 (2018).

[69] H. Zhu, R. Kueng, M. Grassl, and D. Gross, The Clifford group fails gracefully to be a unitary 4-design, arXiv:1609.08172.

[70] A. Deshpande, B. Fefferman, S. Ghosh, M. Gullans, and D. Hangleiter, Peaked quantum advantage using error correction, arXiv:2510.05262 [quant-ph].

[71] Y. Zhang, S. Vijay, Y. Gu, and Y. Bao, Designs from magic-augmented clifford circuits, arXiv:2507.02828 [quant-ph].

[72] L. Bittel, J. Eisert, L. Leone, A. A. Mele, and S. F. E. Oliviero, A complete theory of the Clifford commutant, arXiv:2504.12263.

[73] J. Haferkamp, Random quantum circuits are approximate unitary *t*-designs in depth $O(nt^{5+o(1)})$, Quantum **6**, 795 (2022).

[74] S.-K. Jian, G. Bentsen, and B. Swingle, Linear growth of circuit complexity from Brownian dynamics, arXiv:2206.14205.

[75] J. Haferkamp, P. Faist, N. B. T. Kothakonda, J. Eisert, and N. Yunger Halpern, Linear growth of quantum circuit complexity, Nat. Phys. **18**, 528 (2022).

[76] A. Kitaev, in *Talk Given at the Fundamental Physics Prize Symposium* (Caltech, Pasadena, CA, USA, 2014), Vol. 10.

[77] L. Leone, S. F. E. Oliviero, Y. Zhou, and A. Hamma, Quantum chaos is quantum, Quantum **5**, 453 (2021).
[78] S. F. E. Oliviero, L. Leone, and A. Hamma, Transitions in entanglement complexity in random quantum circuits by measurements, Phys. Lett. A **418**, 127721 (2021).
[79] R. Bhatia, *Matrix Analysis* (Springer New York, 1997), Vol. 169.
[80] J. Cotler, N. Hunter-Jones, J. Liu, and B. Yoshida, Chaos, complexity, and random matrices, J. High Energy Phys. **2017**, 48 (2017).
[81] J. Liu, Spectral form factors and late time quantum chaos, Phys. Rev. D **98**, 086026 (2018).
[82] L. Leone, S. F. E. Oliviero, and A. Hamma, Isospectral twirling and quantum chaos, Entropy **23**, 1073 (2021).
[83] S. F. E. Oliviero, L. Leone, F. Caravelli, and A. Hamma, Random matrix theory of the isospectral twirling, SciPost Phys. **10**, 76 (2021).
[84] H. Gharibyan, M. Hanada, S. H. Shenker, and M. Tezuka, Onset of random matrix behavior in scrambling systems, J. High Energy Phys. **2018**, 124 (2018).
[85] M. Winer and B. Swingle, Reappearance of thermalization dynamics in the late-time spectral form factor, arXiv:2307.14415.
[86] J. P. Keating and N. C. Snaith, Random matrix theory and $\zeta(1/2 + \mathrm{it})$, Commun. Math. Phys. **214**, 57 (2000).
[87] Y. Imry, Active transmission channels and universal conductance fluctuations, Eur. Phys. Lett. **1**, 249 (1986).
[88] S. Bravyi and D. Gosset, Improved classical simulation of quantum circuits dominated by Clifford gates, Phys. Rev. Lett. **116**, 250501 (2016).
[89] A. Gu, S. F. E. Oliviero, and L. Leone, Magic-induced computational separation in entanglement theory, PRX Quantum **6**, 020324 (2025).
[90] Z. Ji, Y.-K. Liu, and F. Song, in *Advances in Cryptology – CRYPTO 2018*, edited by H. Shacham and A. Boldyreva, Lecture Notes in Computer Science (Springer International Publishing, Cham, 2018), pp. 126–152.
[91] S. Grewal, V. Iyer, W. Kretschmer, and D. Liang, Efficient learning of quantum states prepared with few non-Clifford gates, Quantum **9**, 1907 (2025).
[92] B. Skinner, J. Ruhman, and A. Nahum, Measurement-induced phase transitions in the dynamics of entanglement, Phys. Rev. X **9**, 031009 (2019).
[93] A. F. Mello, A. Santini, G. Lami, J. De Nardis, and M. Collura, Clifford dressed time-dependent variational principle, Phys. Rev. Lett. **134**, 150403 (2025).
[94] D. Weingarten, Asymptotic behavior of group integrals in the limit of infinite rank, J. Math. Phys. **19**, 999 (1978).
[95] Z.-W. Liu, S. Lloyd, E. Zhu, and H. Zhu, Entanglement, quantum randomness, and complexity beyond scrambling, J. High Energy Phys. **2018**, 41 (2018).
[96] R. A. Low, Pseudo-randomness and learning in quantum computation, Ph.D. thesis, University of Bristol, UK, 2010.
[97] M. Oszmaniec, A. Sawicki, and M. Horodecki, Epsilon-nets, unitary designs and random quantum circuits, IEEE Trans. Inf. Therory **68**, 989 (2022).
[98] G. Lami, T. Haug, and J. De Nardis, Quantum state designs with Clifford-enhanced matrix product states, PRX Quantum **6**, 010345 (2025).
[99] D. Gross, S. Nezami, and M. Walter, Schur–Weyl duality for the Clifford group with applications: Property testing, a robust Hudson theorem, and de Finetti representations, Commun. Math. Phys. **385**, 1325 (2021).
[100] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, 2000).
[101] A. Baker, in *Matrix groups* (Springer, London, England, UK, 2002), pp. 181–209.
[102] S. Grewal, V. Iyer, W. Kretschmer, and D. Liang, Pseudoentanglement ain't cheap, arXiv:2404.00126 [quant-ph].
[103] L. Leone, S. F. E. Oliviero, S. Lloyd, and A. Hamma, Learning efficient decoders for quasi-chaotic quantum scramblers, Phys. Rev. A **109**, 022429 (2024).
[104] S. J. Szarek, in *Proceedings of Research Workshop on Banach Space Theory* (University of Iowa Press, Iowa City, Iowa, 1982), pp. 169–185.
[105] H. Zhao, L. Lewis, I. Kannan, Y. Quek, H.-Y. Huang, and M. C. Caro, Learning quantum states and unitaries of bounded gate complexity, PRX Quantum **5**, 040306 (2024).
[106] L. Leone, Clifford group and beyond: Theory and applications in quantum information, Ph.D. thesis, University of Massachusetts Boston, 2023.