Moments of quantum channel ensembles

Matthew Duschenes,[1, 2, 3, *] Diego García-Martín,[4] Zoë Holmes,[5, 6] and M. Cerezo[4, †]

[1] *Theoretical Division, Los Alamos National Laboratory, 87545 NM, USA*
[2] *Institute for Quantum Computing, University of Waterloo, N2L 3G1 ON, Canada*
[3] *Perimeter Institute for Theoretical Physics, Waterloo, N2L 2Y5 ON, Canada*
[4] *Information Sciences, Los Alamos National Laboratory, 87545 NM, USA*
[5] *Institute of Physics, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland*
[6] *Centre for Quantum Science and Engineering, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland*

Moments of ensembles of unitaries play a central role in quantum information theory as they capture the statistical properties of dynamics of systems with some form of randomness. Indeed, concepts such as approximate t-designs arise when comparing how close an associated moment operator of a given unitary ensemble is to that of another, reference ensemble. Despite the importance of moment operators, their properties have not been as explored for quantum channels. In this work we develop a theoretical framework to compute moment operators for ensembles of quantum channels, for all moment orders t, with a special focus on determining ensembles that can be used as points of reference. By deriving hierarchies between ensembles, via inequalities of their moment operator norms, we give them operational meaning, and define useful concepts such as that of channel t-designs. Finally, we perform theoretical and numerical studies which show that different types of noise can decrease the norm of the moment operators (e.g., depolarizing noise), as well as increase it (e.g., amplitude damping), and generalize noise-induced concentration phenomena to channel-design-induced phenomena. Along the way, we find a block-orthogonal basis for permutations, which greatly simplifies our analyses, and may be of independent interest.

I. INTRODUCTION

Quantum experiments, whether intrinsically or by design, can involve randomness. Thus, each repetition samples a distinct evolution from an ensemble and applies it to the system. The statistical properties of such experiments are captured by the moment operator, i.e., the superoperator given by the ensemble-averaged adjoint action (the twirl). Focusing on the moment operator allows one to study ensembles independently of initial states and final measurements. Moment operators arise widely in quantum information science, including quantum-supremacy experiments [1–7], quantum chaos [8–11], transitions of quantum correlations [12–14], variational quantum algorithms [15–24], concentration phenomena [15–20], randomized benchmarking [25–30], error mitigation [31, 32], and as Markov process matrices in analyses of fixed points of ensembles [9, 10, 21, 33–51].

Within this context, given an ensemble of channels, it is often useful to quantify its proximity to a reference ensemble. Mathematically, this reduces to comparing their moment operators under a chosen distance measure. This problem has been extensively studied for ensembles of unitaries, where we say that an ensemble forms a t-design with respect to a reference if their moments match up to the t-th order [38, 40, 43–45, 47, 52–61]. A priori, computing and comparing moment operators are very demanding–often intractable–tasks. The situation simplifies when the reference ensemble is a group as the associated moment operator is the projector onto the group's commutant, and its properties admit analytic characterization via Weingarten calculus [62, 63] (see [64] for a quantum information-oriented overview). Of particular relevance is the ensemble drawn from the Haar measure on the unitary group. Not only is this a natural comparison point for other ensembles, it also enjoys desirable extremal properties: there is a hierarchy between unitary ensembles arising from inequalities between moment-operator Hilbert–Schmidt norms, with the Haar ensemble at the base [65]. Equivalently, the Haar ensemble minimizes frame potentials [66], which quantify how evenly spread a set of unitaries is.

The previous studies, however, have been almost entirely restricted to ensembles of strictly unitaries. In contrast, ensembles of completely-positive trace-preserving maps, or quantum channels, have been largely unexplored [67–72]. For instance, it is still unknown whether hierarchies and desirable properties for reference ensembles of more general operators can be found. Ultimately, exact forms of moment operators are necessary to determine the most appropriate reference ensemble. Indeed, there are many important open questions concerning ensembles that must be addressed, such as: *What is the most natural generalization of the uniform (Haar) distribution of unitaries to channels?*, or, *Is there a reference ensemble of channels that plays a privileged role over other ensembles?*

In this work we take steps towards answering the previous questions by presenting a framework to study moments of ensembles of quantum channels (see Fig. 1). In Section II, we introduce our formalism and notation, which captures as a special case ensembles of strictly unitaries, thus allowing us to borrow inspiration from the standard literature. In Section III, we start by generalizing the archetypal reference ensemble of the Haar measure over the unitary group, i.e., sampling uniformly at random from the set of all unitaries, and we consider the ensemble of the Lebesgue measure [70] over quantum channels, i.e., sampling uniformly at random from the set of all quantum channels. This study prompted us to consider a parametrized family of reference channel ensembles [71] which we dub the "channel Haar" (cHaar) ensemble. The cHaar ensemble is

* mduschen@uwaterloo.ca
† cerezo@lanl.gov

shown to interpolate between two reference ensembles of interest, the Haar reference ensemble of strictly unitaries, and what we dub as the Depolarize reference ensemble, where the maximally depolarizing channel is applied with unit probability. By deriving the exact forms and spectral properties for the associated moment operators, we compute deviations of the cHaar ensemble from the Haar and Depolarize ensembles. Such calculations indicate that random quantum channels are on average, inherently depolarizing, with perturbative non-unital behavior, and interpolate with environment dimension between unitary and depolarizing behavior. Finally, given these insights, we deduce a hierarchy between these ensembles, with the Depolarize ensemble found to be at the base of the hierarchy. Thus, we argue that the Depolarize ensemble is the most appropriate reference ensemble for quantum channels.

Expanding on these studies, we explicitly calculate the extent to which quantum circuits consisting of random unitaries, and fixed unital or non-unital noise, form approximate Depolarize t-designs. Here, we observe that certain types of noise can decrease the norm of the moment operator (e.g., unital Pauli noise), or can increase it (e.g., non-unital amplitude damping). Finally, we study concentration of expectation values. We find that what has previously been attributed to be noise-induced concentration phenomena [19], can be more generally thought of as channel-design-induced concentration phenomena. To complement our theoretical studies, in Section IV we conduct numerical studies of several noisy quantum circuits of interest, consisting of layers of parametrized unitary and fixed noise components. In this setting, we demonstrate the convergence of a circuit's statistical properties as a function of circuit depth. Finally, in Section V, we interpret the derived average behaviors of random quantum channels, and discuss several future applications of our formalisms.

II. PRELIMINARIES

In this section we present the notation used throughout this work, with additional details in Appendix A.

A. Moments of Ensembles of Channels

Here, we denote d-dimensional quantum Hilbert spaces as $\mathcal{H} = \mathbb{C}^d$, the space of linear operators acting on $\mathcal{H}$ as $\mathcal{B}[\mathcal{H}]$, and the space of linear superoperators acting on $\mathcal{B}[\mathcal{H}]$ as $\mathcal{B}[\mathcal{B}[\mathcal{H}]]$. We also make use of the linear vectorization map which transforms operators to vectors, vec : $\mathcal{B}[\mathcal{H}] \to \mathcal{H}^{\otimes 2}$, and whose action on $X \in \mathcal{B}[\mathcal{H}]$ is

$$X = \sum_{\alpha,\beta\in[d]} X_{\alpha\beta}|\alpha\rangle\langle\beta| \xrightarrow[\text{vec}]{} |X\rangle\rangle = \sum_{\alpha,\beta\in[d]} X_{\alpha\beta}|\alpha\beta\rangle \ , \quad (1)$$

where $[d] \equiv \{0, 1, \ldots, d-1\}$. The (Hilbert-Schmidt) inner product between two vectorized operators $X, Y \in \mathcal{B}[\mathcal{H}]$ is $\langle\langle X|Y\rangle\rangle = \mathrm{Tr}\left[X^\dagger Y\right]$, and given the identity operator $I_d \in \mathcal{B}[\mathcal{H}]$ on $\mathcal{H}$, we normalize it as

$$\chi_d(X,Y) := \frac{\langle\langle X|Y\rangle\rangle}{d} \quad , \quad \chi_d(X) := \chi_d(I_d, X) \ . \quad (2)$$

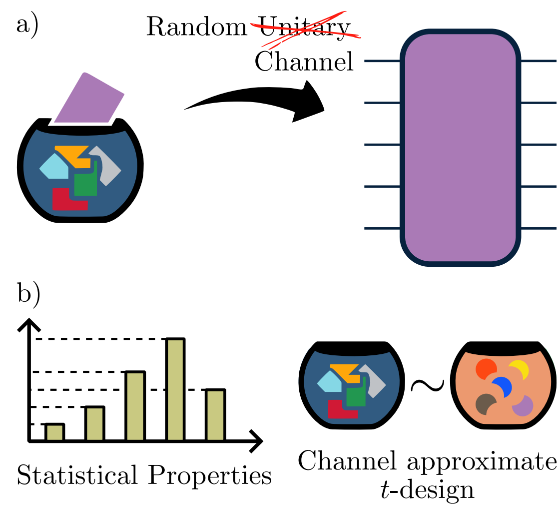

Figure 1. **Schematic representation of our results.** (a) We present a framework to study the statistical properties of an experimental setup where, at every run, a channel is sampled from some set according to a given probability distribution. (b) Our formalism centers around the moment operator, as it encodes the statistical properties arising from evolving states according to the ensemble, and performing a measurement at its output; and is also the central object used to define t-designs.

The vectorization map induces an action over linear maps, vec : $\mathcal{B}[\mathcal{B}[\mathcal{H}]] \to \mathcal{B}[\mathcal{H}^{\otimes 2}]$, which allows us to represent superoperators as matrices. Given a quantum channel $\Lambda \in \mathcal{B}[\mathcal{B}[\mathcal{H}]]$ (a completely positive, trace-preserving linear map), its action in terms of Kraus operators $\{K_\Lambda\}$ is

$$\Lambda(\cdot) = \sum_{K_\Lambda} K_\Lambda(\cdot)K_\Lambda^\dagger \xrightarrow[\text{vec}]{} \widehat{\Lambda} = \sum_{K_\Lambda} K_\Lambda \otimes K_\Lambda^* \ , \quad (3)$$

and superoperators satisfy $|\Lambda(X)\rangle\rangle = \widehat{\Lambda}|X\rangle\rangle$. Finally, t-copies of channels have corresponding superoperators,

$$\widehat{\Lambda^{\otimes t}} = \sum_{K_{\Lambda_0},\ldots,K_{\Lambda_{t-1}}} K_{\Lambda_0} \otimes \cdots \otimes K_{\Lambda_{t-1}} \otimes K_{\Lambda_0}^* \otimes \cdots \otimes K_{\Lambda_{t-1}}^* \ . \quad (4)$$

In this work, we will analyze ensembles of quantum channels $\mathcal{C} \subseteq \mathcal{B}[\mathcal{B}[\mathcal{H}]]$, namely sets of channels $\{\Lambda \in \mathcal{B}[\mathcal{B}[\mathcal{H}]]\}$ accompanied with a probability distribution $d\Lambda$, and ensemble averages of channel-dependent quantities $\mathcal{F}(\Lambda)$,

$$\langle\mathcal{F}(\Lambda)\rangle_{\mathcal{C}} = \int_{\mathcal{C}} d\Lambda \ \mathcal{F}(\Lambda) \ . \quad (5)$$

Note that here we have assumed that the ensemble is continuous, leading to an integral over $\mathcal{C}$. The discrete case follows by simply replacing the integral by a summation.

In particular, we will consider scenarios where an initial state ρ is evolved under the action of a channel Λ, sampled from $\mathcal{C}$, at whose output we measure an observable O. The statistical properties of this setup are captured by t-th order moments of this expectation value, of the form

$$\left\langle \mathrm{Tr}\left[\Lambda(\rho)O\right]^t \right\rangle_{\mathcal{C}} \ . \quad (6)$$

For example, expectation value means and variances are $\mu_{\mathcal{C}} = \langle\mathrm{Tr}\left[\Lambda(\rho)O\right]\rangle_{\mathcal{C}}$ and $\sigma^2_{\mathcal{C}} = \left\langle \mathrm{Tr}\left[\Lambda(\rho)O\right]^2 \right\rangle_{\mathcal{C}} - \left\langle \mathrm{Tr}\left[\Lambda(\rho)O\right] \right\rangle^2_{\mathcal{C}}$.

General t-th order moments of expectation values can be more conveniently rewritten as

$$\begin{aligned}\left\langle \mathrm{Tr}\left[\Lambda(\rho)O\right]^t\right\rangle_{\mathcal{C}} &= \left\langle \mathrm{Tr}\left[\Lambda^{\otimes t}(\rho^{\otimes t})O^{\otimes t}\right]\right\rangle_{\mathcal{C}} \\ &= \mathrm{Tr}\left[\left\langle\Lambda^{\otimes t}(\rho^{\otimes t})O^{\otimes t}\right\rangle_{\mathcal{C}}\right] \\ &= \mathrm{Tr}\left[\mathcal{T}_{\mathcal{C}}^{(t)}(\rho^{\otimes t})O^{\otimes t}\right] \ , \end{aligned} \tag{7}$$

in terms of the t-th order twirl operator $\mathcal{T}_{\mathcal{C}}^{(t)} : \mathcal{B}[\mathcal{H}^{\otimes t}] \to \mathcal{B}[\mathcal{H}^{\otimes t}]$, whose action on operators $X \in \mathcal{B}[\mathcal{H}^{\otimes t}]$ is

$$\mathcal{T}_{\mathcal{C}}^{(t)}(X) = \left\langle \Lambda^{\otimes t}(X)\right\rangle_{\mathcal{C}} = \int_{\mathcal{C}} d\Lambda \ \Lambda^{\otimes t}(X) \ . \tag{8}$$

We can then define the t-th order moment operator as the superoperator of the t-th order twirl,

$$\widehat{\mathcal{T}}_{\mathcal{C}}^{(t)} = \int_{\mathcal{C}} d\Lambda \ \widehat{\Lambda^{\otimes t}} \ , \tag{9}$$

such that

$$\left\langle\left[\mathrm{Tr}\left[\Lambda(\rho)O\right]^t\right]\right\rangle_{\mathcal{C}} = \langle\!\langle O^{\otimes t}|\widehat{\mathcal{T}}_{\mathcal{C}}^{(t)}|\rho^{\otimes t}\rangle\!\rangle \ . \tag{10}$$

The moment operators of Eq. (9) will be central in our theory as they allow us to study the statistical properties of ensembles independently from the initial state and measurement operator. Moreover, we can use them to readily tackle experiments such as one where the quantum evolution arises from k concatenations of several random channels Λ, each sampled identically and independently from the set $\mathcal{C}$ according to $d\Lambda$. In this case, we obtain expectation values such as in Eq. (10), but where we replace the moment operator by its k-th power $\widehat{\mathcal{T}}_{\mathcal{C}}^{(t)} \to \widehat{\mathcal{T}}_{\mathcal{C}}^{(t)k}$.

At this point, we find it important to remark that we can always express the moment operator as

$$\widehat{\mathcal{T}}_{\mathcal{C}}^{(t)} = \widehat{\mathcal{D}}_d^{\otimes t} \ + \ \widehat{\Delta}_{\mathcal{C}}^{(t)} \ , \tag{11}$$

where we have defined a trace-preserving term and a deviation, respectively as

$$\widehat{\mathcal{D}}_d^{\otimes t} = \frac{1}{d^t}|I_d^{\otimes t}\rangle\!\rangle\langle\!\langle I_d^{\otimes t}| \ , \tag{12}$$

$$\widehat{\Delta}_{\mathcal{C}}^{(t)} = \frac{1}{d^t}\sum_{\substack{P\in\mathcal{S}_{\mathcal{C}}^{(t)}\setminus\{I_d^{\otimes t}\} \\ S\in\mathcal{S}_{\mathcal{C}}^{(t)}}} \tau_{\mathcal{C}}^{(t)}(P,S)\ |P\rangle\!\rangle\langle\!\langle S| \ , \tag{13}$$

in terms of an appropriate ensemble-dependent basis of operators $P, S \in \mathcal{S}_{\mathcal{C}}^{(t)}$. The ensemble-dependent coefficients $\tau_{\mathcal{C}}^{(t)}(P,S)$ correspond to the entries of a transfer matrix that represent how an element of the basis is mapped to different elements of the basis by the moment operator.

To finish, we note that analytically evaluating the moment operator can lead to intractable calculations. In these cases, one can instead opt to study how ε-close in norm the moment operator of $\mathcal{C}$ is to that of a reference ensemble $\mathcal{C}'$. This allows us to introduce the notion of an (additive) ε-approximate t-channel design, with respect to $\mathcal{C}'$, as the case when the difference of moment operators satisfies

$$\left\|\widehat{\mathcal{T}}_{\mathcal{C}}^{(t)} - \widehat{\mathcal{T}}_{\mathcal{C}'}^{(t)}\right\| \le \varepsilon \ . \tag{14}$$

The choice of norm $\|\cdot\|$ is task-dependent, and unless noted, we use the Hilbert-Schmidt norm. Moreover, we consider ε to be the additive error, but our definition can readily be adapted to the relative error. Importantly, we will also see below that in many cases of interest the evaluation of Eq. (14) leads to analysis of the moment operator norm,

$$\left\|\widehat{\mathcal{T}}_{\mathcal{C}}^{(t)}\right\| \ . \tag{15}$$

A moment operator can be though of as a Markov transfer matrix like process [9, 10, 21, 33–51], and whose operational meaning will be elucidated as we study its properties throughout this work. As an aside, other than explicit relevance to statistics, we note an immediate interpretation of moment operators that is relevant to randomized benchmarking [73], and follows from the relationship between superoperators and their associated Choi states. In particular for pairs of channels, inner products between their superoperators correspond to inner products between their Choi states. It follows that the norm and trace of moment operators correspond respectively to the average overlap between the Choi states, and the average entanglement fidelity of the Choi states [73], of the channels from its ensemble. Concurrently, design properties of ensembles of unitaries have been attributed to their expressive power within variational quantum algorithms [18, 66, 74, 75]. However it remains to be seen whether similar notions of capability or utility of ensembles of quantum channels directly follow from their design properties. Other questions regarding operational interpretations will arise within this work once exact forms for moment operators over specific ensembles of quantum channels are derived.

B. Special Case: Ensembles of Unitaries

Since unitaries represent the evolution of closed quantum systems, the study of moments of ensembles of unitaries, has received considerable attention within the field of quantum information [62, 76–79]. Here we review some well-known results concerning ensembles of unitaries.

First, we recall that the t-th order moment operator over unitary ensembles $\mathcal{U}$ can now be expressed as

$$\widehat{\mathcal{T}}_{\mathcal{U}}^{(t)} = \int_{\mathcal{U}} dU \ U^{\otimes t} \otimes U^{\otimes t\,*} \ , \tag{16}$$

where dU denotes the probability distribution over $\mathcal{U}$. For the simple case when $\mathcal{U}$ contains a single unitary U (which is sampled with probability one), we readily find

$$\widehat{\mathcal{T}}_{U}^{(t)} = U^{\otimes t} \otimes U^{\otimes t\,*} \ , \quad \text{and} \quad \left\|\widehat{\mathcal{T}}_{U}^{(t)}\right\|^2 = d^{2t} \ . \tag{17}$$

For the case when $\mathcal{U} = G$ is a group, and $dU = d\mu(U)$ its associated Haar measure, one can use Weingarten calculus [63, 64, 80] to evaluate moment operators. Now, $\widehat{\mathcal{T}}_{G}^{(t)}$ is a projector onto the t-th order commutant, defined as

$$\mathrm{com}^{(t)}(G) = \{X \in \mathcal{B}[\mathcal{H}^{\otimes t}] : [U^{\otimes t}, X] = 0, \forall\ U \in G\} \ , \tag{18}$$

and hence satisfies

$$\widehat{\mathcal{T}}_{G}^{(t)}\widehat{\mathcal{T}}_{G}^{(t)} = \widehat{\mathcal{T}}_{G}^{(t)} \quad \text{and} \quad \left\|\widehat{\mathcal{T}}_{G}^{(t)}\right\|^2 = \dim(\mathrm{com}^{(t)}(G)) \ . \tag{19}$$

Equation (19) is extremely useful for bounding the difference between moment operators with respect to some ensemble $\mathcal{U}$, and with respect to a reference ensemble of the Haar-distributed group G, whenever $\mathcal{U} \subseteq G$. Note that $\mathcal{U}$ does not need to form a group itself. Such comparisons are central when studying approximate ε-t-designs. Indeed, we say that $\mathcal{U}$ is an ε-t-design over G if and only if $\|\widehat{\mathcal{T}}_{\mathcal{U}}^{(t)} - \widehat{\mathcal{T}}_{G}^{(t)}\| \leq \varepsilon$, where the choice of norm depends on the task at hand [38, 81]. In the 2-norm case, combining Eq. (19) along with the left- and right- invariance of the Haar measure, we find that for all ensembles $\mathcal{U} \subseteq G$,

$$\widehat{\mathcal{T}}_{G}^{(t)}\widehat{\mathcal{T}}_{\mathcal{U}}^{(t)} = \widehat{\mathcal{T}}_{\mathcal{U}}^{(t)}\widehat{\mathcal{T}}_{G}^{(t)} = \widehat{\mathcal{T}}_{G}^{(t)} \ , \tag{20}$$

from where it then follows that

$$\left\|\widehat{\mathcal{T}}_{\mathcal{U}}^{(t)} - \widehat{\mathcal{T}}_{G}^{(t)}\right\|^2 = \left\|\widehat{\mathcal{T}}_{\mathcal{U}}^{(t)}\right\|^2 - \left\|\widehat{\mathcal{T}}_{G}^{(t)}\right\|^2 \ , \tag{21}$$

and hence,

$$\dim(\mathrm{com}^{(t)}(G)) \leq \left\|\widehat{\mathcal{T}}_{\mathcal{U}}^{(t)}\right\|^2 \ . \tag{22}$$

In addition, one can also readily find that any k concatenations of the Haar ensemble are invariant,

$$\widehat{\mathcal{T}}_{\mathcal{U}}^{(t)k} = \widehat{\mathcal{T}}_{\mathcal{U}}^{(t)} \ . \tag{23}$$

Notably, Eq. (22) can be used to establish a partial ordering between the norms of the moment operators for the Haar measure over groups. In particular, given groups $G_1, G_2, \ldots$ such that $\{U\} \subseteq G_1 \subseteq G_2 \subseteq \cdots \subseteq \mathcal{U}(d)$ then

$$\left\|\widehat{\mathcal{T}}_{\mathcal{U}(d)}^{(t)}\right\|^2 \leq \cdots \leq \left\|\widehat{\mathcal{T}}_{G_2}^{(t)}\right\|^2 \leq \left\|\widehat{\mathcal{T}}_{G_1}^{(t)}\right\|^2 \leq \left\|\widehat{\mathcal{T}}_{U}^{(t)}\right\|^2 = d^{2t} \tag{24}$$

Here, we denote $\mathcal{U}(d)$ as the unitary group of all unitaries acting on $\mathcal{H}$ with uniform Haar measure, henceforth referred to as the Haar ensemble. Equation (24) implies that the more "unstructured" a group G is (i.e., the less symmetries it has) the smaller the norm of its moment operator. Thus, the Haar ensemble $\mathcal{U}(d)$, at the base of such a hierarchy of unitary ensembles, serves as an appropriate reference ensemble for other subsets of unitaries. Moreover, we note that such norm inequalities appear within the context of variational quantum algorithms [82, 83] and quantum machine learning [84, 85] as a measure of the expressiveness of the ensemble of unitaries arising from a parametrized circuit [15, 40, 47, 51, 66, 72]. In this context, it has been shown that the smaller the norm, the greater the concentration (and the more difficult the optimization) of the variational landscape [17, 18, 66, 74, 86, 87].

Given the privileged role that the Haar ensemble of unitaries $\mathcal{U}(d)$ plays, we find it convenient to recall some of its properties [64]. To begin, the t-th order commutant of $\mathcal{U}(d)$ is given by the system permuting representation $\mathrm{com}^{(t)}(\mathcal{U}(d)) = \mathcal{S}_d^{(t)}$ of the Symmetric group $\mathcal{S}_t$, namely all permutations of t copies a space $\mathcal{H}^{\otimes t}$. Moreover, given a permutation $\sigma \in \mathcal{S}_t$, the system permuting representation $V_d \ : \ \mathcal{S}_t \to \mathcal{S}_d^{(t)} \subset \mathcal{B}[\mathcal{H}^{\otimes t}]$ is,

$$V_d(\sigma) = \sum_{\alpha_0,\ldots,\alpha_{t-1}\in[d]} |\alpha_{\sigma^{-1}(0)},\ldots,\alpha_{\sigma^{-1}(t-1)}\rangle\langle\alpha_0,\ldots,\alpha_{t-1}| \ , \tag{25}$$

with associated normalized inner products, or characters,

$$\chi_d^{(t)}(\sigma,\pi) = \frac{\langle\langle V_d(\sigma)|V_d(\pi)\rangle\rangle}{d^t} \quad : \quad \chi_d^{(t)}(\sigma) = \chi_d^{(t)}(e,\sigma) \ , \tag{26}$$

where e denotes the identity permutation with representation $V_d(e) = I_d^{\otimes t}$. For $t \leq d$ then $\dim(\mathrm{com}^{(t)}(\mathcal{U}(d))) = t!$, and for any unitary group G, we have the bounds

$$t! \ \leq \ \left\|\widehat{\mathcal{T}}_{G}^{(t)}\right\|^2 \ \leq \ d^{2t} \ . \tag{27}$$

To finish, we note that since $\widehat{\mathcal{T}}_{\mathcal{U}(d)}^{(t)}$ is a projector [51, 64], we can express the Haar ensemble moment operator as,

$$\widehat{\mathcal{T}}_{\mathcal{U}(d)} = \frac{1}{d^t}\sum_{\sigma,\pi\in\mathcal{S}_t} W_d^{(t)-1}(\sigma,\pi)\ |V_d(\sigma)\rangle\rangle\langle\langle V_d(\pi)| \ , \tag{28}$$

where the $t! \times t!$ transfer matrix $W_d^{(t)-1}$ is known as the Weingarten matrix, the inverse of the Gram matrix $W_d^{(t)}$ with entries of the overlaps, $W_d^{(t)}(\sigma,\pi) = \chi_d^{(t)}(\sigma,\pi)$. Moreover, it is not hard to see that for any unitary ensemble $\mathcal{U} \subseteq \mathcal{U}(d)$ one has

$$\widehat{\mathcal{T}}_{\mathcal{U}}^{(t)} = \widehat{\mathcal{T}}_{\mathcal{U}(d)} + \cdots \ , \tag{29}$$

which follows from Eq. (20) since $\widehat{\mathcal{T}}_{\mathcal{U}(d)}^{(t)}$ is a projector.

III. RESULTS

In this section, we present the main results of this work, with additional details in the appendices. Specifically, we define reference ensembles of quantum channels, derive their relationships and properties (Appendix A and Appendix B), calculate moments of specific ensembles of quantum channels (Appendix C), and investigate concentration phenomena arising in the context of noisy variational quantum algorithms (Appendix D).

A. Reference Ensembles

As previously discussed, it is convenient to introduce reference ensembles to define concepts such as approximate t-designs. Here, we consider three candidate reference ensembles for quantum channels and derive their properties and relationships. Ultimately, such properties allow us to determine whether a hierarchy exists between these ensembles that suggests the most appropriate reference ensemble for quantum channels.

Our first candidate reference ensemble is the *Haar* ensemble given its place at the base of the unitary ensemble hierarchy. We will denote it as $\mathcal{U}(d)$, with its t-th order twirl given by

$$\mathcal{T}_{\mathcal{U}(d)}^{(t)}(X) = \int_{\mathcal{U}(d)} dU\ U^{\otimes t}\ X\ U^{\otimes t\dagger} \ , \tag{30}$$

and associated moment operator

$$\widehat{\mathcal{T}}_{\mathcal{U}}^{(t)} = \int_{\mathcal{U}} dU\ U^{\otimes t} \otimes U^{\otimes t\,*} \ . \tag{31}$$

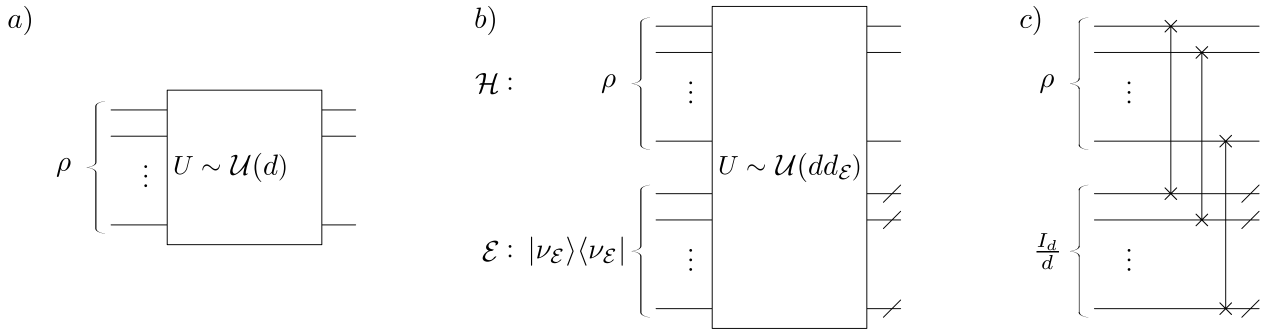

Figure 2. **Circuits implementing the reference channel ensembles.** (a) The Haar ensemble can be implemented by sending a state ρ through a random unitary uniformly sampled from $\mathcal{U}(d)$. b) The cHaar ensemble can be implemented via its Stinespring dilation: Initialize the joint state of the tensor product Hilbert space $\mathcal{H}\otimes\mathcal{E}$ in $\rho\otimes|\nu_{\mathcal{E}}\rangle\langle\nu_{\mathcal{E}}|$, apply a joint unitary uniformly sampled from $\mathcal{U}(dd_{\mathcal{E}})$, and trace out the environment. c) The Depolarize ensemble can be applied with a circuit that swaps the state ρ with the maximally mixed state $\frac{I_d}{d}$, and then traces out the registers containing ρ.

For our second candidate reference ensemble, we recall the definition of the Stinespring dilation [88], which states that any quantum channel $\Lambda \in \mathcal{B}[\mathcal{H}] \to \mathcal{B}[\mathcal{H}]$, can expressed as a unitary $U_\Lambda : \mathcal{H}\otimes\mathcal{E} \to \mathcal{H}\otimes\mathcal{E}$ acting on the tensor product of the system Hilbert space $\mathcal{H}$ of dimension d and an environment Hilbert space $\mathcal{E}$ of dimension $d_{\mathcal{E}}$, with the environment being in a pure state $\nu_{\mathcal{E}} = |\nu_{\mathcal{E}}\rangle\langle\nu_{\mathcal{E}}| \in \mathcal{B}[\mathcal{E}]$. Thus the action of any quantum channel is

$$\Lambda(X) = \mathrm{Tr}_{\mathcal{E}}\left[U_\Lambda\left(X\otimes\nu_{\mathcal{E}}\right)U_\Lambda^\dagger\right] \;, \tag{32}$$

where the environment $\mathcal{E}$ is partially traced out. A key advantage of working with the Stinespring dilation is that we can now analyze ensembles of random quantum channels in terms of the random Stinespring unitaries U_Λ that generate such channels. Sampling quantum channels can thus be directly related to sampling unitaries in larger composite spaces. Notably, it has been shown that a consistent measure $d\Lambda$ for quantum channels exists that is equivalent to sampling unitaries U_Λ according to the Haar measure dU over $\mathcal{U}(dd_{\mathcal{E}})$ [71]. In fact, when $d_{\mathcal{E}} = d^2$, this measure $d\Lambda$ is the Lebesgue measure over quantum channels.

Given this Stinespring dilation-based approach, we define our second candidate reference ensemble as the environment dimension $d_{\mathcal{E}}$-parametrized family of ensembles of channels, which we denote as the channel-Haar or *cHaar* ensemble $\mathcal{C}(d,d_{\mathcal{E}})$, with a t-th order twirl of,

$$\mathcal{T}^{(t)}_{\mathcal{C}(d,d_{\mathcal{E}})}(X) = \int_{\mathcal{U}(dd_{\mathcal{E}})} dU\; \mathrm{Tr}_{\mathcal{E}^{\otimes t}}\left[U^{\otimes t}(X\otimes\nu_{\mathcal{E}}^{\otimes t})U^{\otimes t\dagger})\right] \;, \tag{33}$$

and an associated moment operator of,

$$\widehat{\mathcal{T}}^{(t)}_{\mathcal{C}(d,d_{\mathcal{E}})} = \left(I_d^{\otimes 2t}\otimes\langle\!\langle I_{d_{\mathcal{E}}}^{\otimes t}|\right)\;\widehat{\mathcal{T}}^{(t)}_{\mathcal{U}(dd_{\mathcal{E}})}\;\left(I_d^{\otimes 2t}\otimes|\nu_{\mathcal{E}}^{\otimes t}\rangle\!\rangle\right) \;. \tag{34}$$

Finally, we can readily use our expansions for Haar moment operators in terms of the Haar ensemble commutant of permutations in Eq. (28) to find,

$$\widehat{\mathcal{T}}^{(t)}_{\mathcal{C}(d,d_{\mathcal{E}})} = \frac{1}{d^t}\sum_{\sigma,\pi\in\mathcal{S}_t}\chi^{(t)}_{d_{\mathcal{E}}}(\sigma)\;W^{-1}_{dd_{\mathcal{E}}}(\sigma,\pi)\;|V_d(\sigma)\rangle\!\rangle\langle\!\langle V_d(\pi)| \,. \tag{35}$$

Above we have used the fact that under a simple reordering of the Hilbert space $(\mathcal{H}\otimes\mathcal{E})^{\otimes t} \to \mathcal{H}^{\otimes t}\otimes\mathcal{E}^{\otimes t}$ [46] we can express $|V_{dd_{\mathcal{E}}}(\sigma)\rangle\!\rangle \to |V_d(\sigma)\rangle\!\rangle\otimes|V_{d_{\mathcal{E}}}(\sigma)\rangle\!\rangle$, as well as the fact that $\langle\!\langle V_{d_{\mathcal{E}}}(\pi)|\nu_{\mathcal{E}}^{\otimes t}\rangle\!\rangle = 1$, $\forall\; \pi\in S_t$ and $\forall\; |\nu_{\mathcal{E}}\rangle \in \mathcal{E}$.

For the third candidate reference ensemble, we recall the depolarizing channel

$$\mathcal{D}_d(X) = \frac{\mathrm{Tr}\,[X]}{d} I_d \;, \tag{36}$$

that has been shown to describe the trace-preservation component in all ensembles, and will be key to our analysis. We thus define our third candidate reference ensemble as the ensemble containing only the depolarizing channel with unit probability, which we denote as the *Depolarize* ensemble $\mathcal{D}(d)$, with a t-th order twirl of,

$$\mathcal{T}^{(t)}_{\mathcal{D}(d)}(X) = \mathcal{D}_d^{\otimes t}(X) = \frac{\mathrm{Tr}\,[X]}{d^t} I_d^{\otimes t} \;, \tag{37}$$

and an associated moment operator of,

$$\widehat{\mathcal{T}}^{(t)}_{\mathcal{D}(d)} = \widehat{\mathcal{D}}_d^{\otimes t} = \frac{1}{d^t}|I_d^{\otimes t}\rangle\!\rangle\langle\!\langle I_d^{\otimes t}| \;. \tag{38}$$

Quantum circuits to experimentally implement each of the three candidate reference ensembles are shown in Fig. 2.

In the following sections, after deriving explicit moments of each ensemble, we will compare the Haar, cHaar and Depolarize reference ensembles, and deduce their hierarchy.

B. Properties of Moment Operators

Having defined the Haar, cHaar, and Depolarize ensembles, we now proceed to analyze their moment operators. First, simply from their definitions, we can immediately prove the following properties regarding their relationships:

Proposition 1 (Properties of Reference Ensembles)**.** *Let $\widehat{\mathcal{T}}^{(t)}_{\mathcal{U}(d)}$, $\widehat{\mathcal{T}}^{(t)}_{\mathcal{C}(d,d_{\mathcal{E}})}$ and $\widehat{\mathcal{T}}^{(t)}_{\mathcal{D}(d)}$ respectively denote the moment operators for the Haar, cHaar and Depolarize ensembles. Then, for $t=1$ and any $d_{\mathcal{E}}$, the ensembles are identical,*

$$\widehat{\mathcal{T}}^{(1)}_{\mathcal{U}(d)} = \widehat{\mathcal{T}}^{(1)}_{\mathcal{C}(d,d_{\mathcal{E}})} = \widehat{\mathcal{T}}^{(1)}_{\mathcal{D}(d)} \;, \tag{39}$$

with such equalities not holding for $t>1$. While $\widehat{\mathcal{T}}^{(t)}_{\mathcal{U}(d)}$ and $\widehat{\mathcal{T}}^{(t)}_{\mathcal{D}(d)}$ are self-adjoint projectors for all t and $d_{\mathcal{E}}$,

$$\widehat{\mathcal{T}}^{(t)}_{\mathcal{U}(d)}\widehat{\mathcal{T}}^{(t)}_{\mathcal{U}(d)}=\widehat{\mathcal{T}}^{(t)}_{\mathcal{U}(d)}\ ,\quad \widehat{\mathcal{T}}^{(t)}_{\mathcal{D}(d)}\widehat{\mathcal{T}}^{(t)}_{\mathcal{D}(d)}=\widehat{\mathcal{T}}^{(t)}_{\mathcal{D}(d)}\ , \tag{40}$$

$\widehat{\mathcal{T}}^{(t)}_{\mathcal{C}(d,d_{\mathcal{E}})}$ is not a self-adjoint projector for $t>1$ and $d_{\mathcal{E}}>1$,

$$\widehat{\mathcal{T}}^{(t)}_{\mathcal{C}(d,d_{\mathcal{E}})}\widehat{\mathcal{T}}^{(t)}_{\mathcal{C}(d,d_{\mathcal{E}})}\neq\widehat{\mathcal{T}}^{(t)}_{\mathcal{C}(d,d_{\mathcal{E}})}\quad ,\quad \widehat{\mathcal{T}}^{(t)\dagger}_{\mathcal{C}(d,d_{\mathcal{E}})}\neq\widehat{\mathcal{T}}^{(t)}_{\mathcal{C}(d,d_{\mathcal{E}})}\ . \tag{41}$$

Moreover, for any t and any $d_{\mathcal{E}}$, the cHaar ensemble is invariant under any ensemble of unitaries $\mathcal{U}$,

$$\widehat{\mathcal{T}}^{(t)}_{\mathcal{C}(d,d_{\mathcal{E}})}\widehat{\mathcal{T}}^{(t)}_{\mathcal{U}}=\widehat{\mathcal{T}}^{(t)}_{\mathcal{U}}\widehat{\mathcal{T}}^{(t)}_{\mathcal{C}(d,d_{\mathcal{E}})}=\widehat{\mathcal{T}}^{(t)}_{\mathcal{C}(d,d_{\mathcal{E}})}\ . \tag{42}$$

The Depolarize ensemble is right-invariant under any ensemble of channels $\mathcal{C}$ and is left-invariant under adjoints,

$$\widehat{\mathcal{T}}^{(t)}_{\mathcal{D}(d)}\widehat{\mathcal{T}}^{(t)}_{\mathcal{C}}=\widehat{\mathcal{T}}^{(t)}_{\mathcal{D}(d)}\quad ,\quad \widehat{\mathcal{T}}^{(t)\dagger}_{\mathcal{C}}\widehat{\mathcal{T}}^{(t)}_{\mathcal{D}(d)}=\widehat{\mathcal{T}}^{(t)}_{\mathcal{D}(d)}\ , \tag{43}$$

and is left-invariant under unital ensembles of channels $\mathcal{C}_{\mathcal{U}}$,

$$\widehat{\mathcal{T}}^{(t)}_{\mathcal{C}_{\mathcal{U}}}\widehat{\mathcal{T}}^{(t)}_{\mathcal{D}(d)}=\widehat{\mathcal{T}}^{(t)}_{\mathcal{D}(d)}\ . \tag{44}$$

These relationships between ensembles already suggest that the Depolarize ensemble, although not corresponding to uniformly sampling all quantum channels, is potentially an appropriate reference ensemble. The relevance of this Depolarize ensemble for quantum channels follows partially from its right, and often left invariance under concatenation with other ensembles. It is also important to note that any moment operator of any trace-preserving ensemble involves a Depolarize moment operator term (which is reminiscent of Eq. (29) and how the moment operator for any unitary ensembles contain the term $\widehat{\mathcal{T}}^{(t)}_{\mathcal{U}(d)}$). In contrast, the cHaar ensemble is found to not be left- or right- invariant under concatenations with other ensembles. Such properties make the cHaar ensemble potentially less relevant as a reference ensemble for quantum channels.

It is perhaps surprising that the cHaar ensemble, does not appear to be a natural generalization of the Haar ensemble, and does not immediately appear to be the most appropriate reference ensemble for quantum channels. After all, the Depolarize ensemble only contains a single channel! To fully support this proposal, we must derive exact properties of the cHaar ensemble and deduce its place in the hierarchy of ensembles, as a function of environment dimension $d_{\mathcal{E}}$, order t, and concatenations k.

Theorem 1 (Hierarchy of moment operator norms)**.** *The following dimension-dependent relationships between the Depolarize, cHaar and Haar ensembles hold:*
1) The cHaar ensemble, and its k-concatenations, interpolates with system and environment dimension between the Haar and Depolarize ensembles. That is, for all $k\geq 0$:

$$\lim_{dd_{\mathcal{E}}\to d}\widehat{\mathcal{T}}^{(t)k}_{\mathcal{C}(d,d_{\mathcal{E}})}=\widehat{\mathcal{T}}^{(t)}_{\mathcal{U}(d)}\ ,\quad \lim_{dd_{\mathcal{E}}\to\infty}\widehat{\mathcal{T}}^{(t)k}_{\mathcal{C}(d,d_{\mathcal{E}})}=\widehat{\mathcal{T}}^{(t)}_{\mathcal{D}(d)}\ . \tag{45}$$

In fact, the cHaar ensemble moment operator has an exact closed form expression, consisting of trace-preserving, non-unital, and unital terms, expressed for simplicity in terms of some traceless operators $T_d^{(t,k)},T_d^{(t,k)\prime}\in\mathcal{B}[\mathcal{H}^{\otimes t}]$, and asymptotic scalings of transfer matrix elements,

$$\begin{aligned}\lim_{dd_{\mathcal{E}}\to\infty}\widehat{\mathcal{T}}^{(t)k}_{\mathcal{C}(d,d_{\mathcal{E}})}=&\ \frac{1}{d^t}|I_d^{\otimes t}\rangle\!\rangle\langle\!\langle I_d^{\otimes t}| \\ &+\ \frac{1}{d^t}\mathcal{O}\left(\frac{1}{dd_{\mathcal{E}}}\right)\ |T_d^{(t,k)}\rangle\!\rangle\langle\!\langle I_d^{\otimes t}| \\ &+\ \frac{1}{d^t}\mathcal{O}\left(\frac{1}{d^2d_{\mathcal{E}}^k}\right)|T_d^{(t,k)}\rangle\!\rangle\langle\!\langle T_d^{(t,k)\prime}|\ .\end{aligned} \tag{46}$$

2) The cHaar ensemble forms an ε-approximate Depolarize t-design (as introduced in Eq. (14)), for $\varepsilon\in\mathcal{O}\left(\frac{1}{d_{\mathcal{E}}}\right)$, in the fixed t and the asymptotic $dd_{\mathcal{E}}\to\infty$ limit.
3) In the asymptotic $dd_{\mathcal{E}}\to\infty$ limit, there exists a strict hierarchy between the k-concatenated t-th order ensembles, via their moment operator norms, with $k\geq k'$,

$$\left\|\widehat{\mathcal{T}}^{(t)}_{\mathcal{D}(d)}\right\|^2\leq\left\|\widehat{\mathcal{T}}^{(t)k}_{\mathcal{C}(d,d_{\mathcal{E}})}\right\|^2\leq\left\|\widehat{\mathcal{T}}^{(t)k'}_{\mathcal{C}(d,d_{\mathcal{E}})}\right\|^2\leq\left\|\widehat{\mathcal{T}}^{(t)}_{\mathcal{U}(d)}\right\|^2\ . \tag{47}$$

As per part 1) and 2) of Theorem 1, in the appropriate limits, the cHaar ensemble reduces to either the Depolarize ensemble or the Haar ensemble, and can be thought as interpolating with dimension between these ensembles.

More concretely, for large enough systems and environments such that $dd_{\mathcal{E}}\to\infty$ (for example if $d_{\mathcal{E}}$ is a function of d, i.e., the maximal $d_{\mathcal{E}}=d^2$ necessary to implement any quantum channel) the cHaar ensemble approaches the Depolarize ensemble, and evenly dissipates quantum information in all directions of the Hilbert space. Moreover, even with finite-sized systems and environments, the perturbative next-leading-order non-unital contributions to the moment operator vanish inversely with system and environment dimensions, and are intriguingly independent of concatenations. The exact forms for cHaar moment operators generalize previous derivations of random quantum channels being asymptotically depolarizing [72], and offer a complete depiction of their average behavior.

On the other hand, for trivial environments such that $d_{\mathcal{E}}\to 1$, the cHaar ensemble reduces to the Haar ensemble. The hierarchy between ensembles in part 3) of Theorem 1 then implies that concatenating cHaar ensembles decreases their norm, tending towards the base of the hierarchy set by the Depolarize ensemble. Thus, as a consequence of the dissipative nature of tracing out the environment, whereas cHaar ensembles with a finite environment are not maximally dissipative, their behavior can be made more so via concatenation. (We note that while the hierarchy between ensembles in Theorem 1 holds asymptotically when $dd_{\mathcal{E}}\to\infty$, we have explicitly verified it for $t=2,3,4$, $k=1,3$, and various $d,d_{\mathcal{E}}$ in Fig. 10 of Appendix B, and we conjecture that it holds for all t, k, and $d,d_{\mathcal{E}}$.)

Theorem 1 is derived explicitly in Appendix B by studying the spectrum of the cHaar moment operator. Namely, since the cHaar moment operator fails to be a projector, unlike the Haar and Depolarize moment operators, its eigenvalues indicate the inherent average dissipation of information of random quantum channels. This realization, as well as the derived concatenation-independent behavior of its next-leading-order non-unital contributions, prompted us to study the spectral properties of $\widehat{\mathcal{T}}^{(t)}_{\mathcal{C}(d,d_{\mathcal{E}})}$.

Theorem 2 (Spectrum of cHaar Moment Operator)**.** *The t-th order cHaar moment operator has the exact leading term of a projector,*

$$\widehat{\mathcal{T}}^{(t)}_{\mathcal{C}(d,d_{\mathcal{E}})} = \frac{|\psi\rangle\rangle\langle\langle\varphi|}{\langle\langle\varphi|\psi\rangle\rangle} + \cdots \ , \tag{48}$$

which implies that it has a leading eigenvalue $\lambda = 1$ with associated right- and left-eigenvectors, given by,

$$|\psi\rangle\rangle = \sum_{\sigma\in\mathcal{S}_t} \chi^{(t)}_{dd_{\mathcal{E}}}(\sigma)\ |V_d(\sigma)\rangle\rangle \quad , \quad |\varphi\rangle\rangle = |V_d(e)\rangle\rangle \ . \tag{49}$$

In the fixed t, asymptotic $dd_{\mathcal{E}} \to \infty$, the t-th order cHaar moment operator has a single leading eigenvalue $\lambda = 1$ from the above contribution, and all non-leading eigenvalues are, up-to t-dependent factors, $\lambda \in \mathcal{O}\left(\frac{1}{d_{\mathcal{E}}}\right) < 1$.

We note that a direct consequence of the previous theorem is that k-concatenations of the cHaar ensemble will have moment operators with the fixed points,

$$\widehat{\mathcal{T}}^{(t)k}_{\mathcal{C}(d,d_{\mathcal{E}})} = \frac{|\psi\rangle\rangle\langle\langle\varphi|}{\langle\langle\varphi|\psi\rangle\rangle} + \cdots \ . \tag{50}$$

Moreover, Theorem 2 implies that the cHaar ensemble moment operator's only fixed point is its eigenvector $|\psi\rangle\rangle$ of Eq. (49). In turn, this leading eigenvector converges with dimension to the Depolarize channel's single fixed point

$$\lim_{dd_{\mathcal{E}}\to\infty} |\psi\rangle\rangle \to |\varphi\rangle\rangle \quad \leftrightarrow \quad \lim_{dd_{\mathcal{E}}\to\infty} \frac{|\psi\rangle\rangle\langle\langle\varphi|}{\langle\langle\varphi|\psi\rangle\rangle} \to \widehat{\mathcal{D}}_d^{\otimes t} \ . \tag{51}$$

This observation regarding the eigenvectors and the single, depolarization fixed point of the cHaar ensemble further underlies the hierarchy between ensembles in Theorem 1.

Proposition 1, and Theorems 1 and 2 have several important implications, that culminate in the Depolarize ensemble being found to be a natural reference ensemble for ensembles of quantum channels.

Claim 1. *The Depolarize ensemble $\mathcal{D}(d)$ is a natural generalization of the Haar ensemble. The invariance properties of $\mathcal{D}(d)$ impose that its moment operator is present in the moment operators of all quantum channel ensembles, and place the Depolarize ensemble at the bottom of the moment operator norm hierarchy of all quantum channel ensembles. Therefore, any ensemble of channels $\mathcal{C}$ satisfies*

$$\left\|\widehat{\mathcal{T}}^{(t)}_{\mathcal{D}(d)}\right\|^2 \leq \left\|\widehat{\mathcal{T}}^{(t)}_{\mathcal{C}}\right\|^2 \ . \tag{52}$$

Claim 2. *The Depolarize ensemble therefore serves as a natural reference ensemble for quantum channels for defining t-designs. Concretely, an ensemble $\mathcal{C}$ is defined to be an ε-approximate Depolarize t-design, if and only if*

$$\left\|\widehat{\mathcal{T}}^{(t)}_{\mathcal{C}} - \widehat{\mathcal{T}}^{(t)}_{\mathcal{D}(d)}\right\| \leq \varepsilon \ . \tag{53}$$

It admittedly remains surprising that from our criteria of positions within hierarchies of moment operator norms, the most appropriate reference ensemble for quantum channels is ultimately found to be the Depolarize ensemble and not the cHaar ensemble. To explain this outcome, we recall that only when $d_{\mathcal{E}} = d^2$, does the cHaar ensemble have a uniform Lebesgue measure over quantum channels, and only when $d_{\mathcal{E}} = d$, is there a resulting uniform distribution of quantum states output by the cHaar ensemble [71]. However, even at such special dimensions, the cHaar ensemble still does not share the defining characteristics of the Haar or Depolarize ensembles that make them the most appropriate reference ensembles, namely invariance under concatenations. In fact, cHaar ensembles for sufficiently large environment dimension $d_{\mathcal{E}}$, with non-leading eigenvalues $|\lambda| < 1$, will dissipate information upon concatenation. It is therefore proposed that criteria of ensemble hierarchies and ensemble invariance, as opposed to (the admittedly desirable property of) uniformity of measure, are used when considering future reference ensembles.

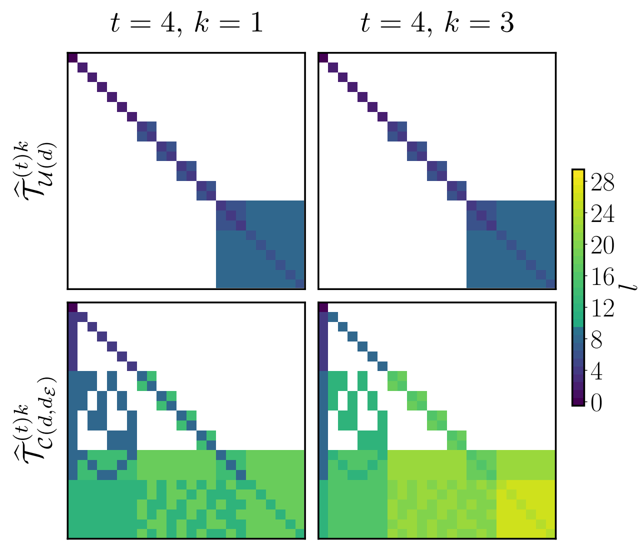

Figure 3. **Moment operators in a basis that block-diagonalizes the Haar moment operator as a permutation transfer matrix.** We show the scaling of the k-concatenated t-th order moment operators for the Haar (top) and cHaar (bottom) ensembles for $d_{\mathcal{E}} = d^2$, $k = 1$ (left) , $k = 3$ (right), and $t = 4$. In this basis, the $t!$ elements are sorted by their support, from the smallest identity (top-left corner), to the largest cycle (bottom-right corner). Colors represent the leading-order scaling l with $1/d$ of elements $\tau^{(t,k)}_{\mathcal{C}} \sim \mathcal{O}(1/d^l)$.

A possible explanation for why the Depolarize ensemble sits at the bottom of the norm hierarchy for quantum channels can be offered by the distribution of its Kraus operators. In particular, the squared Hilbert-Schmidt norm of the t-th order moment operator can be interpreted as the t-th moment of the overlaps between the Kraus operators in the ensemble, or frame potentials,

$$\left\|\widehat{\mathcal{T}}^{(t)}_{\mathcal{C}}\right\|^2 = \int_{\mathcal{C}\times\mathcal{C}} d\Lambda\ d\Upsilon \sum_{K_\Lambda, K_\Upsilon} |\langle\langle K_\Upsilon|K_\Lambda\rangle\rangle|^{2t} \ , \tag{54}$$

where we used Eq. (4). In the unitary case, the Haar reference ensemble minimizes the frame potentials, and hence maximizes the spread of the corresponding (unitary) Kraus operators. In the case of more general channels, it is precisely the Depolarize ensemble the one that maximizes the spread of the Kraus operators, offering a geometric interpretation of its privileged role as a reference ensemble.

Our derivations of the properties of the cHaar moment operator are made possible via a novel analysis of the transfer matrices $\tau_{\mathcal{C}}^{(t)}$ of the Haar and cHaar moment operator in the permutation basis, namely the $t!\times t!$ matrices that map permutations to permutations. In Appendices A and B we derive a crucial change of basis, from non-orthogonal permutations [89], to so-called localized permutations. This novel localized basis consists of operators, each with support over a strict subset of copies of the Hilbert space, that are orthogonal to any operators with non-identical support. Due to its well-defined support and block-orthogonality, this basis block diagonalizes the Haar transfer matrix and block lower-triangularizes the cHaar transfer matrix[1].

The resulting transfer matrices in Fig. 3 for $t = 4$ and $k = 1, 3$ concatenations, depict their leading-order scaling with d, obtained from closed-form expressions and symbolic methods [91]. The Haar ensemble transfer matrix is block-diagonal, with elements that are invariant under concatenations. Such structure reflects that its moment operator is a projector, and is support-preserving. The cHaar ensemble transfer matrix is block-lower-triangular, with elements that decay with concatenations, except for the first column which corresponds to its fixed-point leading eigenvector of Eq. (48). Such structure reflects that its moment operator is non-unital, and is support-non-decreasing, given it is generated by a support-preserving Haar ensemble, with its output environment traced over. Given the exact forms of the transfer matrices, in particular the k-concatenated transfer matrices $\tau_{\mathcal{C}}^{(t,k)}$, the spectral properties, norms and traces of the associated moment operators may be derived in Theorem 2.

C. Moment Operators of Noisy Quantum Circuits

Here, we use our formalisms to study the moment operators of noisy quantum circuits. As per Fig. 4, we will consider k-layers of random unitaries sampled from a unitary ensemble $\mathcal{U}$, followed by a fixed noise channel $\mathcal{N}$.

The t-th order moment operator for this evolution is $(\widehat{\mathcal{N}}^{\otimes t}\widehat{\mathcal{T}}_{\mathcal{U}}^{(t)})^k$. To describe these operators, we define an orthogonal basis $\mathcal{P}_d = \{P^{(\mu)}\}_{\mu\in[d^2]}$ for $\mathcal{B}[\mathcal{H}]$, with $P^{(0)} = I_d$, which can be generalized to $\mathcal{P}_d^t = \{P = \otimes_{i\in\Gamma_P} P_i\}$ for $\mathcal{B}[\mathcal{H}^{\otimes t}]$, with operators of support $\Gamma_P \subseteq [t]$ and locality $|P| = |\Gamma_P|$. In such a basis, we consider a constant noise model, with a superoperator of the following form,

$$\widehat{\mathcal{N}}_{\gamma\eta} = \frac{1}{d}\begin{bmatrix} 1 & 0 & \cdots & 0 \\ \eta_d(P^{(1)}) & 1-\gamma_d(P^{(1)}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \eta_d(P^{(d^2-1)}) & 0 & \cdots & 1-\gamma_d(P^{(d^2-1)}) \end{bmatrix}, \quad (55)$$

where $\gamma_d(P^{(\mu)}) \geq \gamma$, $\eta_d(P^{(\mu)}) \geq \eta$, $\forall\ \mu \in [d^2]$, with appropriate additional constraints on γ_d, η_d such that $\mathcal{N}$ is

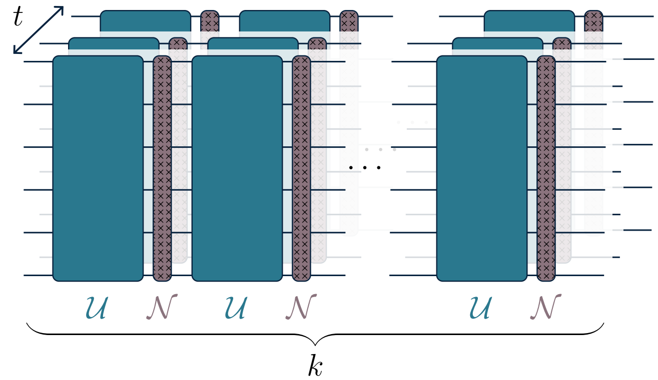

Figure 4. **Concatenation of random unitaries and noisy channels.** We study the t-th moment operator of a noisy circuit where we concatenate k layers of random unitaries sampled from $\mathcal{U}$ (blue) with a fixed noise channel $\mathcal{N}$ (brown).

completely positive [92]. This noise model is composed of a diagonal part which reduces the magnitude of operators, and a non-unital part that maps identity to non-identity operators, and thus encompasses both unital and non-unital Pauli noise [88, 93]. The t-th fold tensor product superoperator for the noise noise component is thus

$$\widehat{\mathcal{N}}_{\gamma\eta}^{\otimes t} = \frac{1}{d^t}\sum_{P,S\in\mathcal{S}_{\mathcal{N}_{\gamma\eta}}^{(t)}} \tau_{\mathcal{N}_{\gamma\eta}}^{(t)}(P,S)\ |P\rangle\!\rangle\langle\!\langle S|\ , \quad (56)$$

with transfer matrix elements

$$\tau_{\mathcal{N}_{\gamma\eta}}^{(t)}(P,S) \in \mathcal{O}\Big((1-\gamma)^{|S|}\ \eta^{|P|-|S|}\Big)\ \delta_{\Gamma_P\supseteq\Gamma_S}\ . \quad (57)$$

Under this framework, the following theorem holds, regarding the effect of noise on channel design properties.

Theorem 3 (Noise-Induced Depolarize t-Designs)**.** *For k-concatenated t-th order unitary ensembles, diagonal unital noise decreases, and non-unital noise increases moment operator norms. The respective ensembles approach and deviate from being Depolarize t-designs.*

More explicitly, we can compute the leading-order scaling of composite unitary and noise moment operators, given Haar random unitaries $\mathcal{U}(d)$ or randomly parametrized unitaries $\mathcal{G}_\Theta = \{U_\theta^G = e^{-i\theta G}, G \in \mathcal{P}_d, \theta \sim \Theta\}$ generated by a single involutory G, with a parameter distribution $\theta \sim \Theta$, and a commutant $\mathcal{S}_G$ of G, yielding,

$$\left\|\left(\widehat{\mathcal{N}}_{\gamma}^{\otimes t}\widehat{\mathcal{T}}_{\mathcal{U}(d)}^{(t)}\right)^k\right\|^2 = 1 + \binom{t}{2}\ \mathcal{O}\big((1-\gamma)^{4k}\big) \quad (58)$$

$$\left\|\left(\widehat{\mathcal{N}}_{\gamma\eta}^{\otimes t}\widehat{\mathcal{T}}_{\mathcal{U}(d)}^{(t)}\right)^k\right\|^2 = 1 + t\ |\mathcal{P}_d\backslash\{I_d\}|\ \mathcal{O}\big(\eta^2\big) \quad (59)$$

$$\left\|\left(\widehat{\mathcal{N}}_{\gamma}^{\otimes t}\widehat{\mathcal{T}}_{\mathcal{G}_\Theta}^{(t)}\right)^k\right\|^2 = 1 + t\ |\mathcal{S}_G\backslash\{I_d\}|\ \mathcal{O}\big((1-\gamma)^{2k}\big)\ . \quad (60)$$

Here, the leading-order scaling depends strictly on the noise model, with no explicit dimension-dependence from the unitary ensemble. Unital noise is evidently inherently depolarizing, decreasing moment operator norms exponentially with concatenations and making the noisy ensemble converge towards a Depolarize t-design.

[1] We note that this change of basis in terms of $t!\times t!$ matrices is not a Schur transformation [90], as the latter is a $d^t \times d^t$ matrix.

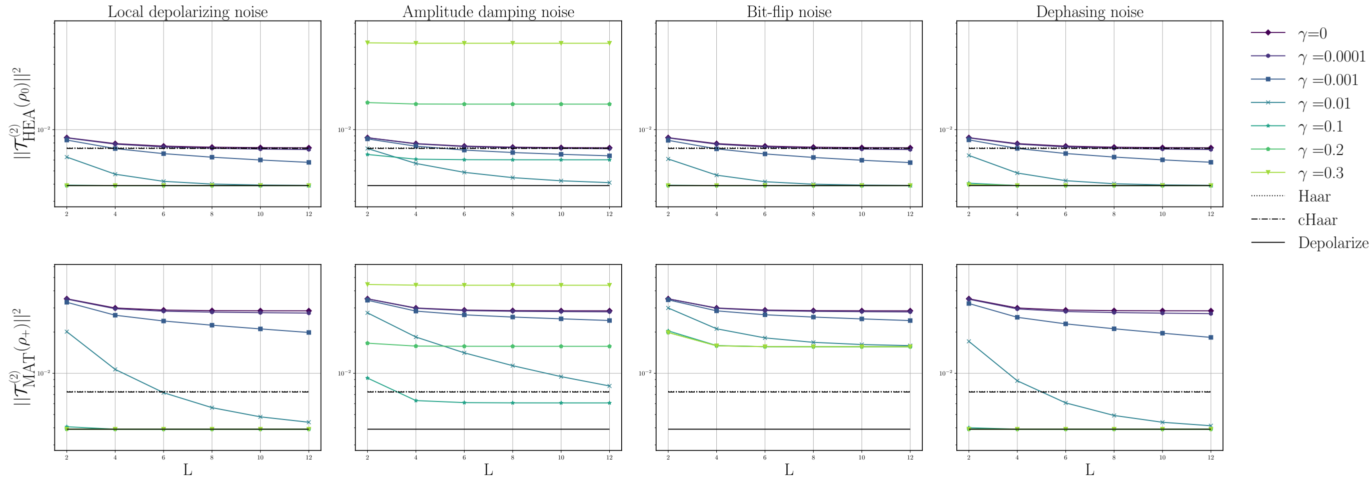

Figure 5. **Effects of noise on $||\mathcal{T}^{(2)}_{\mathcal{C}}(\rho)||^2$.** We study two parametrized quantum circuit architectures: a hardware-efficient ansätz (HEA) and a matchgate ansätz (MAT), on $n = 7$ qubits and $L \in [L]$ layers. We take the variational parameters uniformly in $[0, 2\pi]$, and we simulate the effect of four different types of single-qubit noise channels, namely, local depolarizing, amplitude damping, bit-flip and dephasing nose, with noise strengths $\gamma \in [0, 0.3]$. The input states for the HEA and MAT circuits are ρ_0 and ρ_+, respectively. We also plot the value for the Haar, cHaar with $d_{\mathcal{E}} = d^2$ (i.e., the Lebesgue measure) and Depolarize reference ensembles. These simulations show that certain types of noise can decrease the Hilbert-Schmidt norm of the 2-order moment operator (e.g., depolarizing), while others can increase it (e.g., amplitude damping).

D. Channel Designs and Expectation Value Concentration

In this section, we relate concentration of expectation values to channel design properties. As previously discussed, moment operator norms for ensembles of unitaries have been given operational meaning in the context of variational algorithms, as they can be directly related to the concentration phenomena [18]. Here, we extend such intuition to the case of channels, and show that similar phenomena are exhibited by ensembles of quantum channels.

In particular, let us define the expectation value

$$\mathcal{L}_\Lambda(\rho, O) = \mathrm{Tr}\left[\Lambda(\rho)O\right] \ , \tag{61}$$

where ρ is a quantum state, O an observable, and Λ is a channel from an ensemble $\mathcal{C}$. We can bound the probability of $\mathcal{L}_\Lambda$ deviating from its mean $\mu_{\mathcal{C}} = \langle \mathrm{Tr}\left[\Lambda(\rho)O\right]\rangle_{\mathcal{C}}$ by a quantity $\epsilon \geq 0$, in terms of its variance $\sigma^2_{\mathcal{C}} = \langle \mathrm{Tr}\left[\Lambda(\rho)O\right]^2\rangle_{\mathcal{C}} - \langle \mathrm{Tr}\left[\Lambda(\rho)O\right]\rangle^2_{\mathcal{C}}$, via Chebyshev's inequality

$$\mathrm{Probability}\left[|\mathcal{L}_\Lambda - \mu_{\mathcal{C}}| \geq \epsilon\right] \ \leq \ \frac{\sigma^2_{\mathcal{C}}}{\epsilon^2} \ . \tag{62}$$

Since computing variances $\sigma^2_{\mathcal{C}}$ for arbitrary ensembles of quantum channels $\mathcal{C}$ may be difficult, we can instead study concentration phenomena by comparing how close $\mathcal{C}$ is to a reference ensemble $\mathcal{C}'$. In particular, one can show that the following theorem holds, regarding bounds on the variance of expectation values.

Theorem 4 (Variance bounds in terms of reference ensemble)**.** *Let $\mathcal{L}_\Lambda$ be an expectation value as in Eq. (61), and $\mathcal{C}, \mathcal{C}'$ be two ensembles of channels. The variance of $\mathcal{L}_\Lambda$ over $\mathcal{C}$ is upper bounded as*

$$\sigma^2_{\mathcal{C}} \ \leq \ \sigma^2_{\mathcal{C}'} \ + \ \|\rho\|_1^2 \ \|O\|_\infty^2 \ \|\mathcal{T}^{(2)}_{\mathcal{C}} - \mathcal{T}^{(2)}_{\mathcal{C}'}\|_\diamond \ . \tag{63}$$

In particular, when ρ is a quantum state, when the observable O satisfies $\mathrm{Tr}[O] = 1$ (e.g., a projector) or $\mathrm{Tr}[O] = 0$ (e.g., a Pauli operator), and the reference ensemble $\mathcal{C}' \in \{\mathcal{C}(d, d_{\mathcal{E}}), \mathcal{D}(d)\}$, we can upper bound the variance $\sigma^2_{\mathcal{C}'}$ as

$$\sigma^2_{\mathcal{C}'} \ \leq \ \begin{cases} \mathcal{O}\left(\dfrac{1}{dd_{\mathcal{E}}}\right) \|\rho\|_2^2 & \begin{array}{l}\mathrm{Tr}[O] = 0 \\ \mathcal{C}' = \mathcal{C}(d, d_{\mathcal{E}})\end{array} \\ \mathcal{O}\left(\dfrac{1}{d^2}\right) & \begin{array}{l}\mathrm{Tr}[O] = 1 \\ \mathcal{C}' = \mathcal{C}(d, d_{\mathcal{E}})\end{array} \\ 0 & \begin{array}{l}\mathrm{Tr}[O] = 0 \\ \mathcal{C}' = \mathcal{D}(d)\end{array} \\ \dfrac{1}{d^2} & \begin{array}{l}\mathrm{Tr}[O] = 1 \\ \mathcal{C}' = \mathcal{D}(d)\end{array} \end{cases} . \tag{64}$$

Theorem 4 encompasses the results for unitary ensembles of Ref. [18], and allows us to study non-unitary concentration phenomena. For instance, noise-induced barren plateaus described in Ref. [19], can now be more generally thought of as arising from noisy circuits becoming a t-design with respect to the Depolarize ensemble.

IV. NUMERICAL EXPERIMENTS

In this section we present numerical simulations for ensembles of noisy parametrized quantum circuits. Specifically, we study the behavior of the norm of $\mathcal{T}^{(2)}_{\mathcal{C}}$ twirls applied to an initial pure state ρ, for different circuit depths L and noise strengths γ.

We consider the following (noiseless) unitary ansätze for n qubits: hardware efficient ansätze (HEA) [94] and matchgate ansätze (MAT) [95], each with a set of generators $\mathcal{G}$ in terms of the X_i, Y_i, Z_i local Pauli operators acting on qubit $i \in [n]$,

$$\mathcal{G}_{\text{HEA}} = \{X_i\}_{i=0}^{n-1} \cup \{Y_i\}_{i=0}^{n-1} \cup \{Z_i Z_{i+1}\}_{i=0}^{n-2} \ , \tag{65}$$
$$\mathcal{G}_{\text{MAT}} = \{X_i\}_{i=0}^{n-1} \cup \{Z_i Z_{i+1}\}_{i=0}^{n-2} \ . \tag{66}$$

The noiseless parametrized quantum circuits $\{U(\vec{\theta})\}_{\vec{\theta}}$ have a layered structure, obtained from the generators $\{G_k\}$ via exponentiation as

$$U(\vec{\theta}) = \prod_{l=0}^{L-1} \prod_{k=0}^{K-1} e^{-i\theta_{lk} G_k} \ , \tag{67}$$

where $\vec{\theta} = (\theta_0, \dots, \theta_{LK-1})$ are variational parameters, distributed uniformly in $[0, 2\pi]$, and $G_k \in \mathcal{G}_{HEA}$ or $\mathcal{G}_{MAT}$. As initial states, for HEA we choose $\rho_0 \equiv (|0\rangle\langle 0|)^{\otimes n}$, and for MAT we choose $\rho_+ \equiv \left(\frac{(|0\rangle+|1\rangle)(\langle 0|+\langle 1|])}{2}\right)^{\otimes n}$.

We also consider the following noise models consisting of local noise channels acting on qubit $i \in [n]$ after each unitary gate $e^{-i\theta_{lk}G_k}$. In particular, we consider bit-flip (Λ^i_{BF}), dephasing (Λ^i_D), local-depolarizing (Λ^i_{LD}) and amplitude damping (Λ^i_{AD}) noise. Given a probability γ of noise occurring, also referred to as the noise strength, the action of these local noise channels is,

$$\begin{aligned}
\Lambda^i_{BF}(\rho) &= (1-\gamma)\rho + \gamma \ X_i \rho X_i \ , \\
\Lambda^i_{D}(\rho) &= (1-\gamma)\rho + \gamma \ Z_i \rho Z_i \ , \\
\Lambda^i_{LD}(\rho) &= (1-\gamma)\rho + \frac{\gamma}{3}\left(X_i\rho X_i + Y_i \rho Y_i + Z_i \rho Z_i\right) \ , \\
\Lambda^i_{AD}(\rho) &= K^i_{AD_1}\rho K^{i\dagger}_{AD_1} + K^i_{AD_2}\rho K^{i\dagger}_{AD_2} \ ,
\end{aligned} \tag{68}$$

where the amplitude damping channel Kraus operators are,

$$K^i_{AD_1} \equiv \sqrt{\gamma}\ |0\rangle_i\langle 1| \ , K^i_{AD_2} \equiv |0\rangle_i\langle 0| + \sqrt{1-\gamma}\ |1\rangle_i\langle 1| . \tag{69}$$

In Fig. 5, we plot the (squared) Hilbert-Schmidt norm, or the purity, of the average of $t = 2$ copies of states acted on by channels from the ensemble generated by noisy circuits, as a function of circuit depth L and noise strength γ. From these results, we can compare the behavior of the HEA and MAT ensembles, in both the noiseless and noisy cases.

First, we consider the noiseless case. We observe that $||\mathcal{T}^{(2)}_{\text{HEA}}(\rho_0)||^2$ converges with depth towards the Haar random value, meaning the circuit becomes an ε-approximate state 2-design [15]. In contrast, $||\mathcal{T}^{(2)}_{\text{MAT}}(\rho_+)||^2$ does not converge towards the Haar random value, but instead plateaus at a higher value and does not become a ε-approximate state 2-design. Intuitively, this is a consequence of the fact that the HEA generators in Eq. (65) produce controllable circuits (meaning that any unitary from $\mathbb{SU}(2^n)$ can be implemented given sufficient depth and an appropriate choice of parameters). Conversely, the MAT generators in Eq. (66) produce free-fermionic evolutions, which are a small subgroup of $\mathbb{SU}(2^n)$ that do not form state designs.

Next, we consider the noisy case. Under local depolarizing noise channels, both HEA and MAT norms converge with depth to the absolute minimum norm, namely, the one given by the Depolarize ensemble (as is expected from previous studies of noise-induced phenomena in Ref. [19]). Similarly, under local dephasing noise, both HEA and MAT norms also converge with depth to the absolute minimum Depolarize ensemble norm. However, under local amplitude damping noise, interestingly both HEA and MAT norms decrease when the noise strength is small, and increases with increasing noise strength. Finally, under local bit-flip errors, both HEA and MAT norms are decreased, although the HEA norm is driven towards the minimum Depolarize ensemble norm, whereas the MAT norm converges to a larger value.

These numerical experiments showcase that certain types of noise can decrease or increase the norm of the moment operator, and reveal a subtle interplay between the exact circuit that is run on a quantum computer, and the type and strength of noise present in the device.

The simulations have been performed with the open-source library `Qibo` [96, 97]. One of the main technical issues in numerically evaluating twirl norms, via sampling from an ensemble, is that the error in such estimations is often large compared to the norm itself, unless a prohibitively large number of samples are simulated. Hence, we employ a method to evaluate twirl norms for $t = 2$ exactly, avoiding any sampling error. The method is similar to that in Ref. [98] and is based on the following proposition, whose proof can be found in Appendix C 3.

Proposition 2. *The average adjoint action of $U^{\otimes 2}$, with $U \equiv e^{i\theta G} \sim \mathcal{G}_\Theta$, for fixed generators $G \in \mathcal{G}$, $G^2 = I_d$, and a uniform distribution of parameters $\theta \in \Theta = [0, 2\pi]$, is*

$$\mathcal{T}^{(2)}_{\mathcal{G}_\Theta}(\cdot) = \ \frac{3}{8}\left(\cdot + G^{\otimes 2}\cdot G^{\otimes 2}\right) - \frac{1}{8}\left\{\cdot\ , G^{\otimes 2}\right\} \ + \frac{1}{2} G^{(2)} \cdot G^{(2)} \ ,$$

where $G^{(2)} = G \otimes I_d + I_d \otimes G$.

V. CONCLUSIONS

In this work, we introduced a framework for studying moment operators of ensembles of random quantum channels. A key contribution is the proposal of three reference ensembles: Haar, cHaar, and Depolarize, for which we derived exact expressions for their spectra, traces, and norms, and obtained a hierarchy between these ensembles. In particular, we showed that the cHaar ensemble interpolates (via the environment dimension) between the unitary and depolarizing limits. Indeed, we found that its dominant fixed point is predominantly depolarizing with a perturbatively non-unital component, highlighting persistent non-unital effects and the intrinsic dissipative nature of quantum channels [99–102]. Finally, instead of uniformity of measure, derived hierarchies between ensembles, generally arising from invariance properties of ensembles, are shown to be more versatile criteria to compare ensembles in these more general settings. Based on these criteria, we have shown that the Depolarize ensemble is the most natural reference ensemble for ensembles of quantum channels.

The norm of the ensemble moment operator plays a central role in describing relationships between ensembles. It is natural to ask for further operational interpretation of this quantity, particularly to describe the capabilities of random quantum channels. In the case of ensembles of random unitaries, this quantity has been interpreted as the expressive power of the ensemble [66]. However, in the case of ensembles of random quantum channels, given their inherent depolarizing and therefore not necessarily useful behavior, this feels less applicable to describe channel-centric algorithms [103]. The Depolarize or asymptotic dimension cHaar reference ensembles do capture in some sense the 'uniformity' or 'symmetry' of an ensemble, in terms of its ability to dissipate information evenly in all directions, however neither term is quite satisfactory. Perhaps, as channel-centric algorithms become more prevalent, similarly to how the laser was presented originally as a 'solution looking for a problem', here we have a 'property looking for a name'. We leave finding a nice, operationally meaningful, name for this important quantity as future work.

Applying the formalism to noisy parametrized circuits with diagonal unital and non-unital Pauli-like noise, we observed opposite trends for moment-operator norms. Unital noise typically decreases such norms, while non-unital amplitude damping increases them. Circuit depth drives ensembles toward the appropriate reference designs, and concentration of observables emerges. Such analysis ultimately offers a channel-design-centric view on noisy barren-plateau–type phenomena [17–19, 64, 99, 101].

By establishing the previous clear connection to quantum machine learning, we highlight that our framework can also be directly relevant to more general hybrid classical–quantum protocols that rely on sampling and inversion (e.g., randomized benchmarking, classical shadows, and error-mitigation schemes). Known design parameters and exact moment forms can potentially tighten sample-complexity bounds and guide experimental choices, particularly when noise models are difficult to characterize within quantum devices [25, 104–108].

To finish, we note that our work paves the way for several future research directions that merit follow-up: (i) non-asymptotic spectral analyses and representation-theoretic proofs for the cHaar–Haar hierarchy and related block structures, with links to Weingarten calculus and partition algebras [72, 89, 109]; (ii) characterization of more specific, physically motivated, and experimentally relevant ensembles of channels (e.g., Stinespring unitaries beyond Haar, structured entanglement, composite ansätze), and principled criteria for channel designs [24, 48, 51, 110–112]; and (iii) targeted experiments that probe higher-order moments and benchmark predicted concentration trends, alongside implications for simulability and noise-assisted or active-error-mitigation strategies [100, 113–116].

ACKNOWLEDGMENTS

The authors would like to thank Martin Larocca, Zachary Mann, and Mark M. Wilde, for their valuable insights into noisy quantum processes. MD would like to acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), and of the U.S. Department of Energy (DOE) through a quantum computing program sponsored by the Los Alamos National Laboratory (LANL) Information Science & Technology Institute. Research at the Perimeter Institute is supported in part by the Government of Canada through the Department of Innovation, Science and Economic Development Canada and by the Province of Ontario through the Ministry of Economic Development, Job Creation and Trade. DGM acknowledges support by the Laboratory Directed Research and Development (LDRD) program of LANL under project number 20260043DR, as well as by LANL's ASC Beyond Moore's Law project. ZH acknowledges support from the Sandoz Family Foundation-Monique de Meuron program for Academic Promotion. MC was supported by the U.S. DOE, Office of Science, Office of Advanced Scientific Computing Research through the Accelerated Research in Quantum Computing Program MACH-Q project. This research used resources provided by the LANL Institutional Computing Program, which is supported by the U.S. DOE National Nuclear Security Administration under Contract No. 89233218CNA000001

[1] 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, Nature Physics **14**, 595 (2018).

[2] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, B. Burkett, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney, A. Dunsworth, E. Farhi, B. Foxen, A. Fowler, C. Gidney, M. Giustina, R. Graff, K. Guerin, S. Habegger, M. P. Harrigan, M. J. Hartmann, A. Ho, M. Hoffmann, T. Huang, T. S. Humble, S. V. Isakov, E. Jeffrey, Z. Jiang, D. Kafri, K. Kechedzhi, J. Kelly, P. V. Klimov, S. Knysh, A. Korotkov, F. Kostritsa, D. Landhuis, M. Lindmark, E. Lucero, D. Lyakh, S. Mandrà, J. R. McClean, M. McEwen, A. Megrant, X. Mi, K. Michielsen, M. Mohseni, J. Mutus, O. Naaman, M. Neeley, C. Neill, M. Y. Niu, E. Ostby, A. Petukhov, J. C. Platt, C. Quintana, E. G. Rieffel, P. Roushan, N. C. Rubin, D. Sank, K. J. Satzinger, V. Smelyanskiy, K. J. Sung, M. D. Trevithick, A. Vainsencher, B. Villalonga, T. White, Z. J. Yao, P. Yeh, A. Zalcman, H. Neven, and J. M. Martinis, Quantum supremacy using a programmable superconducting processor, Nature **574**, 505 (2019).

[3] Y. Wu, W.-S. Bao, S. Cao, F. Chen, M.-C. Chen, X. Chen, T.-H. Chung, H. Deng, Y. Du, D. Fan, *et al.*, Strong quantum computational advantage using a superconducting quantum processor, Physical Review Letters **127**, 180501 (2021).

[4] A. M. Dalzell, N. Hunter-Jones, and F. G. S. L. Brandão, Random quantum circuits anticoncentrate in log depth, PRX Quantum **3**, 010333 (2022).

[5] M. Oszmaniec, N. Dangniam, M. E. Morales, and Z. Zimborás, Fermion sampling: a robust quantum computational advantage scheme using fermionic linear optics and magic input states, PRX Quantum **3**, 020328 (2022).

[6] H.-Y. Huang, R. Kueng, G. Torlai, V. V. Albert, and J. Preskill, Provably efficient machine learning for quantum many-body problems, Science **377**, eabk3333 (2022).

[7] E. Knill, D. Leibfried, R. Reichle, J. Britton, R. B.

Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland, Randomized benchmarking of quantum gates, Physical Review A **77**, 012307 (2008).

[8] A. Nahum, J. Ruhman, S. Vijay, and J. Haah, Quantum entanglement growth under random unitary dynamics, Physical Review X **7**, 031016 (2017).

[9] C. W. von Keyserlingk, T. Rakovszky, F. Pollmann, and S. L. Sondhi, Operator hydrodynamics, otocs, and entanglement growth in systems without conservation laws, Physical Review X **8**, 021013 (2018).

[10] A. Nahum, S. Vijay, and J. Haah, Operator spreading in random unitary circuits, Physical Review X **8**, 021014 (2018).

[11] W. W. Ho and S. Choi, Exact emergent quantum state designs from quantum chaotic dynamics, Physical Review Letters **128**, 060601 (2022).

[12] Y. Li, X. Chen, and M. P. Fisher, Quantum zeno effect and the many-body entanglement transition, Physical Review B **98**, 205136 (2018).

[13] B. Skinner, J. Ruhman, and A. Nahum, Measurement-induced phase transitions in the dynamics of entanglement, Physical Review X **9**, 031009 (2019).

[14] C.-M. Jian, Y.-Z. You, R. Vasseur, and A. W. Ludwig, Measurement-induced criticality in random quantum circuits, Physical Review B **101**, 104302 (2020).

[15] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, Barren plateaus in quantum neural network training landscapes, Nature Communications **9**, 1 (2018).

[16] M. Cerezo, A. Sone, T. Volkoff, L. Cincio, and P. J. Coles, Cost function dependent barren plateaus in shallow parametrized quantum circuits, Nature Communications **12**, 1 (2021).

[17] M. Larocca, S. Thanasilp, S. Wang, K. Sharma, J. Biamonte, P. J. Coles, L. Cincio, J. R. McClean, Z. Holmes, and M. Cerezo, A review of barren plateaus in variational quantum computing, Nature Reviews Physics **3**, 625–644 (2025).

[18] Z. Holmes, K. Sharma, M. Cerezo, and P. J. Coles, Connecting ansatz expressibility to gradient magnitudes and barren plateaus, PRX Quantum **3**, 010313 (2022).

[19] S. Wang, E. Fontana, M. Cerezo, K. Sharma, A. Sone, L. Cincio, and P. J. Coles, Noise-induced barren plateaus in variational quantum algorithms, Nature Communications **12**, 1 (2021).

[20] D. Stilck França and R. Garcia-Patron, Limitations of optimization algorithms on noisy quantum devices, Nature Physics **17**, 1221 (2021).

[21] J. Napp, Quantifying the barren plateau phenomenon for a model of unstructured variational ansätze, arXiv preprint arXiv:2203.06174 (2022).

[22] M. Ragone, B. N. Bakalov, F. Sauvage, A. F. Kemper, C. Ortiz Marrero, M. Larocca, and M. Cerezo, A lie algebraic theory of barren plateaus for deep parameterized quantum circuits, Nature Communications **15**, 7172 (2024).

[23] E. Fontana, D. Herman, S. Chakrabarti, N. Kumar, R. Yalovetzky, J. Heredge, S. H. Sureshbabu, and M. Pistoia, Characterizing barren plateaus in quantum ansätze with the adjoint representation, Nature Communications **15**, 7171 (2024).

[24] N. L. Diaz, D. García-Martín, S. Kazi, M. Larocca, and M. Cerezo, Showcasing a barren plateau theory beyond the dynamical lie algebra, arXiv preprint arXiv:2310.11505 (2023).

[25] A. Elben, S. T. Flammia, H.-Y. Huang, R. Kueng, J. Preskill, B. Vermersch, and P. Zoller, The randomized measurement toolbox, Nature Review Physics 10.1038/s42254-022-00535-2 (2022).

[26] H.-Y. Huang, R. Kueng, and J. Preskill, Predicting many properties of a quantum system from very few measurements, Nature Physics **16**, 1050 (2020).

[27] A. Zhao, N. C. Rubin, and A. Miyake, Fermionic partial tomography via classical shadows, Physical Review Letters **127**, 110504 (2021).

[28] K. H. Wan, O. Dahlsten, H. Kristjánsson, R. Gardner, and M. Kim, Quantum generalisation of feedforward neural networks, npj Quantum information **3**, 1 (2017).

[29] F. Sauvage and M. Larocca, Classical shadows with symmetries, arXiv preprint arXiv:2408.05279 10.48550/arXiv.2408.05279 (2024).

[30] M. West, A. A. Mele, M. Larocca, and M. Cerezo, Real classical shadows, arXiv preprint arXiv:2410.23481 (2024).

[31] H.-Y. Hu, A. Gu, S. Majumder, H. Ren, Y. Zhang, D. S. Wang, Y.-Z. You, Z. Minev, S. F. Yelin, and A. Seif, Demonstration of robust and efficient quantum property learning with shallow shadows, arXiv preprint arXiv:2402.17911 (2024).

[32] Z. Cai, R. Babbush, S. C. Benjamin, S. Endo, W. J. Huggins, Y. Li, J. R. McClean, and T. E. O'Brien, Quantum error mitigation, Reviews of Modern Physics **95**, 045005 (2023).

[33] P. Hayden and J. Preskill, Black holes as mirrors: quantum information in random subsystems, Journal of High Energy Physics **9**, 120 (2007).

[34] Y. Sekino and L. Susskind, Fast scramblers, Journal of High Energy Physics **2008**, 065 (2008).

[35] W. Brown and O. Fawzi, Scrambling speed of random quantum circuits, arXiv preprint arXiv:1210.6644 https://doi.org/10.48550/arXiv.1210.6644 (2012).

[36] N. Lashkari, D. Stanford, M. Hastings, T. Osborne, and P. Hayden, Towards the fast scrambling conjecture, Journal of High Energy Physics **2013**, 1 (2013).

[37] P. Hosur, X.-L. Qi, D. A. Roberts, and B. Yoshida, Chaos in quantum channels, Journal of High Energy Physics **2016**, 1 (2016).

[38] N. Hunter-Jones, Unitary designs from statistical mechanics in random quantum circuits, arXiv preprint arXiv:1905.12053 (2019).

[39] B. Barak, C.-N. Chou, and X. Gao, Spoofing linear cross-entropy benchmarking in shallow quantum circuits, arXiv preprint arXiv:2005.02421 (2020).

[40] A. W. Harrow and S. Mehraban, Approximate unitary t-designs by short random quantum circuits using nearest-neighbor and long-range gates, Communications in Mathematical Physics **401**, 1531 (2023).

[41] A. Letcher, S. Woerner, and C. Zoufal, Tight and efficient gradient bounds for parameterized quantum circuits, Quantum **8**, 1484 (2024).

[42] P. Hayden, S. Nezami, X.-L. Qi, N. Thomas, M. Walter, and Z. Yang, Holographic duality from random tensor networks, Journal of High Energy Physics **2016**, 1 (2016).

[43] D. Belkin, J. Allen, S. Ghosh, C. Kang, S. Lin, J. Sud, F. Chong, B. Fefferman, and B. K. Clark, Approximate t-designs in generic circuit architectures, PRX Quantum **5**, 040344 (2024).

[44] S. Mittal and N. Hunter-Jones, Local random quantum circuits form approximate designs on arbitrary architectures, arXiv preprint arXiv:2310.19355 (2023).

[45] T. Schuster, J. Haferkamp, and H.-Y. Huang, Random unitaries in extremely low depth, arXiv preprint arXiv:2407.07754 (2024).

[46] P. Braccia, P. Bermejo, L. Cincio, and M. Cerezo, Computing exact moments of local random quantum circuits

via tensor networks, Quantum Machine Intelligence **6**, 54 (2024).

[47] A. E. Deneris, P. Bermejo, P. Braccia, L. Cincio, and M. Cerezo, Exact spectral gaps of random one-dimensional quantum circuits, arXiv preprint arXiv:2408.11201 10.48550/arXiv.2408.11201 (2024).

[48] D. García-Martín, P. Braccia, and M. Cerezo, Architectures and random properties of symplectic quantum circuits, arXiv preprint arXiv:2405.10264 (2024).

[49] M. West, A. A. Mele, M. Larocca, and M. Cerezo, Random ensembles of symplectic and unitary states are indistinguishable, arXiv preprint arXiv:2409.16500 10.48550/arXiv.2409.16500 (2024).

[50] L. Schatzki, Random real valued and complex valued states cannot be efficiently distinguished, arXiv preprint arXiv:2410.17213 (2024).

[51] D. García-Martín, M. Larocca, and M. Cerezo, Quantum neural networks form gaussian processes, Nature Physics **21**, 1153 (2025).

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

[53] A. W. Harrow and R. A. Low, Random quantum circuits are approximate 2-designs, Communications in Mathematical Physics **291**, 257 (2009).

[54] F. G. Brandao, A. W. Harrow, and M. Horodecki, Local random quantum circuits are approximate polynomial-designs, Communications in Mathematical Physics **346**, 397 (2016).

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

[56] J. Haferkamp and N. Hunter-Jones, Improved spectral gaps for random quantum circuits: Large local dimensions and all-to-all interactions, Physical Review A **104**, 022417 (2021).

[57] W. G. Brown and L. Viola, Convergence rates for arbitrary statistical moments of random quantum circuits, Phys. Rev. Lett. **104**, 250501 (2010).

[58] Y. Nakata, C. Hirche, M. Koashi, and A. Winter, Efficient quantum pseudorandomness with nearly time-independent hamiltonian dynamics, Physical Review X **7**, 021006 (2017).

[59] C.-F. Chen, J. Docter, M. Xu, A. Bouland, and P. Hayden, Efficient unitary t-designs from random sums, arXiv preprint arXiv:2402.09335 (2024).

[60] C.-F. Chen, J. Haah, J. Haferkamp, Y. Liu, T. Metger, and X. Tan, Incompressibility and spectral gaps of random circuits, arXiv preprint arXiv:2406.07478 (2024).

[61] T. Yada, R. Suzuki, Y. Mitsuhashi, and N. Yoshioka, Non-haar random circuits form unitary designs as fast as haar random circuits, arXiv preprint arXiv:2504.07390 (2025).

[62] B. Collins and P. Śniady, Integration with respect to the haar measure on unitary, orthogonal and symplectic group, Communications in Mathematical Physics **264**, 773 (2006).

[63] B. Collins, S. Matsumoto, and J. Novak, The weingarten calculus, Notices Of The American Mathematical Society **69**, 734 (2022).

[64] A. A. Mele, Introduction to haar measure tools in quantum information: A beginner's tutorial, Quantum **8**, 1340 (2024).

[65] D. Gross, K. Audenaert, and J. Eisert, Evenly distributed unitaries: On the structure of unitary designs, Journal of mathematical physics **48**, 052104 (2007).

[66] S. Sim, P. D. Johnson, and A. Aspuru-Guzik, Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms, Advanced Quantum Technologies **2**, 1900070 (2019).

[67] I. Bengtsson and K. Zyczkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement* (Cambridge University Press, 2006).

[68] B. Collins and I. Nechita, Random quantum channels i: graphical calculus and the bell state phenomenon, Communications in Mathematical Physics **297**, 345 (2009).

[69] B. Collins and I. Nechita, Random matrix techniques in quantum information theory, Journal of Mathematical Physics **57**, 15215 (2016).

[70] W. Bruzda, V. Cappellini, H.-J. Sommers, and K. Życzkowski, Random quantum operations, Physics Letters A **373**, 320 (2009).

[71] R. Kukulski, I. Nechita, L. Pawela, Z. Puchala, and K. Zyczkowski, Generating random quantum channels, Journal of Mathematical Physics **62**, pages = 062201 (2021).

[72] J. Bai, J. Wang, and Z. Yin, Primitivity for random quantum channels, Quantum Information Processing **23**, 1 (2024).

[73] C. J. Wood, J. D. Biamonte, and D. G. Cory, Tensor networks and graphical calculus for open quantum systems, arXiv preprint arXiv:1111.6950 (2011).

[74] Y. Du, M. H. Hsieh, T. Liu, and D. Tao, Expressive power of parametrized quantum circuits, Physical Review Research **2**, 033125 (2020).

[75] L.-W. Yu, W. Li, Q. Ye, Z. Lu, Z. Han, and D.-L. Deng, Expressibility-induced concentration of quantum neural tangent kernels, arXiv preprint arXiv:2311.04965 (2023).

[76] N. LaRacuente and F. Leditzky, Approximate unitary k-designs from shallow, low-communication circuits, arXiv preprint arXiv:2407.07876 (2024).

[77] L. Grevink, J. Haferkamp, M. Heinrich, J. Helsen, M. Hinsche, T. Schuster, and Z. Zimborás, Will it glue? on short-depth designs beyond the unitary group, arXiv preprint arXiv:2506.23925 (2025).

[78] H. Zhu, R. Kueng, M. Grassl, and D. Gross, The clifford group fails gracefully to be a unitary 4-design, arXiv preprint arXiv:1609.08172 (2016).

[79] D. Petz and J. Réffy, On asymptotics of large haar distributed unitary matrices, Periodica Mathematica Hungarica **49**, 103 (2004).

[80] M. Ragone, Q. T. Nguyen, L. Schatzki, P. Braccia, M. Larocca, F. Sauvage, P. J. Coles, and M. Cerezo, Representation theory for geometric quantum machine learning, arXiv preprint arXiv:2210.07980 (2022).

[81] S. N. Hearth, M. O. Flynn, A. Chandran, and C. R. Laumann, Unitary k-designs from random number-conserving quantum circuits, Physical Review X **15**, 021022 (2025).

[82] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles, Variational quantum algorithms, Nature Reviews Physics **3**, 625–644 (2021).

[83] K. Bharti, A. Cervera-Lierta, T. H. Kyaw, T. Haug, S. Alperin-Lea, A. Anand, M. Degroote, H. Heimonen, J. S. Kottmann, T. Menke, *et al.*, Noisy intermediate-scale quantum algorithms, Reviews of Modern Physics **94**, 015004 (2022).

[84] Y. Gujju, A. Matsuo, and R. Raymond, Quantum machine learning on near-term quantum devices: Current state of supervised and unsupervised techniques for real-world applications, Physical Review Applied **21**, 067001 (2024).

[85] Y. Wang and J. Liu, A comprehensive review of quantum machine learning: from nisq to fault tolerance, Reports

on Progress in Physics 10.1088/1361-6633/ad7f69 (2024).
[86] K. Nakaji and N. Yamamoto, Expressibility of the alternating layered ansatz for quantum computation, Quantum **5**, 434 (2021).
[87] R. J. Garcia, Y. Zhou, and A. Jaffe, Quantum scrambling with classical shadows, Physical Review Research **3**, 033155 (2021).
[88] M. M. Wilde, *Quantum information theory* (Cambridge University Press, 2013).
[89] A. W. Harrow, Approximate orthogonality of permutation operators, with application to quantum information, Letters in Mathematical Physics **114**, 1 (2024).
[90] A. W. Harrow, Applications of coherent classical communication and the schur transform to quantum information theory, arXiv preprint quant-ph/0512255 (2005).
[91] Y. Cardin, H. de Guise, and N. Quesada, Haarpy, a python library for the symbolic calculation of weingarten functions (2024).
[92] D. Greenbaum, Introduction to quantum gate set tomography, arXiv preprint arXiv:1509.02921 (2015).
[93] K. Sharma, S. Khatri, M. Cerezo, and P. J. Coles, Noise resilience of variational quantum compiling, New Journal of Physics **22**, 043006 (2020).
[94] M. Larocca, P. Czarnik, K. Sharma, G. Muraleedharan, P. J. Coles, and M. Cerezo, Diagnosing Barren Plateaus with Tools from Quantum Optimal Control, Quantum **6**, 824 (2022).
[95] B. M. Terhal and D. P. DiVincenzo, Classical simulation of noninteracting-fermion quantum circuits, Physical Review A **65**, 032325 (2002).
[96] S. Efthymiou, S. Ramos-Calderer, C. Bravo-Prieto, A. Pérez-Salinas, D. García-Martín, A. Garcia-Saez, J. I. Latorre, and S. Carrazza, Qibo: a framework for quantum simulation with hardware acceleration, Quantum Science and Technology **7**, 015018 (2021).
[97] S. Efthymiou, M. Lazzarin, A. Pasquale, and S. Carrazza, Quantum simulation with just-in-time compilation, Quantum **6**, 814 (2022).
[98] V. Heyraud, Z. Li, K. Donatella, A. L. Boité, and C. Ciuti, Efficient estimation of trainability for variational quantum circuits, PRX Quantum **4**, 040335 (2023).
[99] B. Fefferman, S. Ghosh, M. Gullans, K. Kuroiwa, and K. Sharma, Effect of non-unital noise on random circuit sampling, PRX Quantum **5**, 030317 (2024).
[100] A. A. Mele, A. Angrisani, S. Ghosh, S. Khatri, J. Eisert, D. Stilck França, and Y. Quek, Noise-induced shallow circuits and absence of barren plateaus, arXiv preprint arXiv:2403.13927 (2024).
[101] P. Singkanipa and D. A. Lidar, Beyond unital noise in variational quantum algorithms: noise-induced barren plateaus and fixed points, arXiv preprint arXiv:2402.08721 (2024).
[102] C. Hirche, C. Rouzé, and D. Stilck França, On contraction coefficients, partial orders and approximation of capacities for quantum channels, Quantum **6**, 862 (2022).
[103] S. Lohani, J. M. Lukens, D. E. Jones, T. A. Searles, R. T. Glasser, and B. T. Kirby, Improving application performance with biased distributions of quantum states, Physical Review Research **3**, 043145 (2021).
[104] E. Magesan, J. M. Gambetta, and J. Emerson, Scalable and robust randomized benchmarking of quantum processes, Physical review letters **106**, 180504 (2011).
[105] J. Kunjummen, M. C. Tran, D. Carney, and J. M. Taylor, Shadow process tomography of quantum channels, Physical Review A **107**, 042403 (2023).
[106] Y. Hama and H. Nishi, Quantum-error-mitigation circuit groups for noisy quantum metrology, arXiv preprint arXiv:2303.01820 (2023).
[107] A. Ijaz, C. H. Alderete, F. Sauvage, L. Cincio, M. Cerezo, and M. L. Goh, More buck-per-shot: Why learning trumps mitigation in noisy quantum sensing, Materials Today Quantum , 100042 (2025).
[108] E. Van Den Berg, Z. K. Minev, A. Kandala, and K. Temme, Probabilistic error cancellation with sparse pauli–lindblad models on noisy quantum processors, Nature Physics , 1 (2023).
[109] B. Collins and S. Matsumoto, Weingarten calculus via orthogonality relations: new applications, arXiv preprint arXiv:1701.04493 (2017).
[110] Y. Ge and J. Eisert, Area laws and efficient descriptions of quantum many-body states, New Journal of Physics **18**, 083026 (2016).
[111] J. Czartowski and K. Zyczkowski, Quantum pushforward designs, Physical Review A **111**, 032433 (2025).
[112] Z. Webb, The clifford group forms a unitary 3-design, Quantum Information and Computation **16**, 1379 (2016).
[113] J. D. Guimaraes, J. Lim, M. I. Vasilevskiy, S. F. Huelga, and M. B. Plenio, Noise-assisted digital quantum simulation of open systems using partial probabilistic error cancellation, PRX Quantum **4**, 040329 (2023).
[114] D. A. Lidar, Review of decoherence-free subspaces, noiseless subsystems, and dynamical decoupling, Quantum information and computation for chemistry , 295 (2014).
[115] M.-J. So and M.-S. Choi, Symmetry and liouville space formulation of decoherence-free subsystems, arxiv preprint arXiv:2507.15506 doi.org/10.48550/arXiv.2507.15506 (2025).
[116] C. Strydom and M. Tame, Implementation of single-qubit measurement-based t-designs using ibm processors, Scientific Reports **12**, 1 (2022).
[117] B. Sagan, *The symmetric group: representations, combinatorial algorithms, and symmetric functions*, Vol. 203 (Springer Science & Business Media, 2001).
[118] T. Brady, Partial order on the symmetric group and new k(π,1)'s for the braid groups, Advances in Mathematics **161**, 20 (2001).
[119] S. Linton, N. Ruškuc, and V. Vatter, *Permutation Patterns*, London Mathematical Society Lecture Note Series (Cambridge University Press, 2010).
[120] M. Schroeder, Permutations Cycles and Derangements, in *Number Theory in Science and Communication: With Applications in Cryptography, Physics, Digital Information, Computing, and Self-Similarity* (Springer Berlin Heidelberg, Berlin, Heidelberg, 2009) pp. 147–158.
[121] A. Nica and R. Speicher, *Lectures on the Combinatorics of Free Probability*, London Mathematical Society Lecture Note Series (Cambridge University Press, 2006).
[122] R. Rama, Inclusion exclusion principle, in *Topics in Combinatorics and Graph Theory* (Springer Nature Switzerland, Cham, Switzerland, 2025) pp. 111–129.
[123] D. M. Topkis, The structure of sublattices of the product of n lattices, Pacific Journal of Mathematics **65**, 525 (1976).
[124] H. Muhle, The core label order of a congruence-uniform lattice, Algebra Universalis **80**, 1 (2019).
[125] P. Zinn-Justin, Jucys-murphy elements and weingarten matrices, Lett. Math. Phys. **91**, 119 (2009).
[126] R. A. Horn and C. R. Johnson, *Matrix analysis* (Cambridge university press, 2012).
[127] D. Chruściński, Dynamical maps beyond markovian regime, Physics Reports **992**, 1 (2022).
[128] B. Baumgartner, An inequality for the trace of matrix products, using absolute values, arXiv preprint arXiv:1106.6189 (2011).

Appendix A: Spaces, Operators, and Permutations

In this appendix, we present preliminary definitions, including relevant spaces, operators, norms, and permutation operators, that will be useful throughout the rest of this work.

In what follows, we will denote the d-dimensional Hilbert space as $\mathcal{H} = \mathbb{C}^d$, the set of bounded linear operators acting on $\mathcal{H}$ as $\mathcal{B}[\mathcal{H}]$, and the set of bounded linear superoperators acting on operators in $\mathcal{B}[\mathcal{H}]$ as $\mathcal{B}[\mathcal{B}[\mathcal{H}]]$. For convenience, we also recall the vectorization map, where given an operator $X \in \mathcal{B}[\mathcal{H}]$, its vectorization is $X = \sum_{\alpha,\beta\in[d]} X_{\alpha\beta}|\alpha\rangle\langle\beta| \underset{\text{vec}}{\longrightarrow} |X\rangle\!\rangle = \sum_{\alpha,\beta\in[d]} X_{\alpha\beta}|\alpha\beta\rangle$. Given two operators $X, Y \in \mathcal{B}[\mathcal{H}]$, their normalized inner product, or overlap is

$$\chi_d(X,Y) = \frac{1}{d}\mathrm{Tr}\left[X^\dagger Y\right] = \frac{\langle\!\langle X|Y\rangle\!\rangle}{d} \qquad , \qquad \chi_d(X) = \frac{1}{d}\mathrm{Tr}\left[X\right] = \frac{\langle\!\langle I_d|X\rangle\!\rangle}{d} \ , \tag{A1}$$

We will also make use of an orthonormal basis for the space of operators, denoted as $\mathcal{P}_d = \{P, S \in \mathcal{B}[\mathcal{H}] \ : \ \chi_d(P,S) = \delta_{PS}\}$, and which satisfies $\mathcal{B}[\mathcal{H}] = \mathrm{span}_{\mathbb{C}}\{\mathcal{P}_d\}$. To simplify analyzes, we assume $I_d \in \mathcal{P}_d$, where I_d denotes the $d \times d$ dimensional identity, such that non-identity basis operators $P \in \mathcal{P}_d \backslash \{I_d\}$ are traceless, $\chi_d(P) = 0$.

The t-fold tensor product of the Hilbert space is denoted as $\mathcal{H}^{\otimes t}$, and we define the set $[t] = \{0, 1, \ldots, t-1\}$. Moreover, from $\mathcal{P}_d$ we define a basis for the space of linear operators $\mathcal{B}[\mathcal{H}^{\otimes t}]$, which is simply given by $\mathcal{P}_d^t = \mathcal{P}_d^{\otimes t}$. Given two operators $X, Y \in \mathcal{B}[\mathcal{H}^{\otimes t}]$, their normalized inner product is denoted as $\chi_d^{(t)}(X,Y) = \langle\!\langle X|Y\rangle\!\rangle/d^t$. Note that in what follows, we indicate t-fold tensor product generalizations with t or (t) superscripts, and particular copies are denoted with indices $i \in [t]$. In addition, given an operator $X = \otimes_{i\in[t]} X_i \in \mathcal{B}[\mathcal{H}^{\otimes t}]$ with tensor product structure, we define its support Γ_X and locality as the indices of non-identity operators over the t copies of the space,

$$\Gamma_X = \{i \in [t] \ : \ X_i \notin \{I_d\}\} \qquad , \qquad |X| = |\Gamma_X| \in [t+1] \ . \tag{A2}$$

Operators $X = \otimes_{i\in\Gamma_X} X_i$, restricted to a subset of their support $\Gamma \subseteq \Gamma_X$, are denoted as $X_\Gamma = \otimes_{i\in\Gamma} X_i$. For general operators $X = \sum_x X_x$ with support $\Gamma_X = \cup_x \Gamma_{X_x}$, we refer to such operators as localized with definitive support, if all operators within their expansion have identical support $\Gamma_{X_x} = \Gamma_X$. We also refer to operators $X, Y \in \mathcal{B}[\mathcal{H}^{\otimes t}]$ as being orthogonal with respect to support if they are orthogonal due to having non-identical support, $\chi_d^{(t)}(X,Y) \propto \delta_{\Gamma_X\Gamma_Y}$.

Next, we will denote the Symmetric group as $\mathcal{S}_t = \{\sigma : [t] \to [t]\}$ [117]. We will denote the identity permutation as $e = ()$, and transpositions acting on indices $i \neq j \in [t]$ as $\tau = (ij)$. Here we find it important to recall that any permutation $\sigma \in \mathcal{S}_t$ can be described by the support $\Gamma_\sigma \subseteq [t]$ and locality $l_\sigma = |\Gamma_\sigma| \leq t$ where they act non-trivially; by their size $|\sigma|$ of the number of transpositions $\sigma = \prod_{\tau\in\sigma} \tau$; or by their cycle structure of disjoint cycles $\sigma = \prod_{\lambda\in\sigma} \lambda$. In addition, permutations can be related in terms of the size of their overlap $\left|\pi^{-1}\sigma\right|$, or in terms of their sub-permutations $\pi \subseteq \sigma$, a partial ordering between permutations [118], that corresponds to factorizations $\sigma = (\pi)(\pi^{-1}\sigma)$ such that,

$$\pi \subseteq \sigma \quad \leftrightarrow \quad \left|\pi^{-1}\sigma\right| = |\sigma| - |\pi| \ . \tag{A3}$$

Further, the size of permutations satisfies a triangle inequality, with upper and lower bounds satisfied by sub-permutations. Indeed, for $\sigma, \pi, \xi \in \mathcal{S}_t$, one has

$$\big||\sigma| - |\pi|\big| \ \leq \ \left|\pi^{-1}\sigma\right| \ = \ \left|(\xi^{-1}\sigma)^{-1}(\xi^{-1}\pi)\right| \ \leq \ \left|\xi^{-1}\sigma\right| + \left|\xi^{-1}\pi\right| \ \leq \ |\sigma| + |\pi| - 2|\xi| \ \leq \ \big||\sigma| + |\pi|\big| \ . \tag{A4}$$

The representation of the Symmetric group which permutes copies of the Hilbert space is denoted as $V_d \ : \ \sigma \in \mathcal{S}_t \to V_d(\sigma) \in \mathcal{S}_d^{(t)} \subset \mathcal{B}[\mathcal{H}^{\otimes t}]$. From here, we can define the overlap between the representations of permutations $\sigma, \pi \in \mathcal{S}_t$ as

$$\chi_d^{(t)}(\sigma,\pi) = \frac{1}{d^{|\sigma^{-1}\pi|}} \ . \tag{A5}$$

These overlaps can be be used to form a $t! \times t!$ Gram matrix $W_d^{(t)}$, with elements, $W_d^{(t)}(\sigma,\pi) = \chi_d^{(t)}(\sigma,\pi)$, whose inverse $W_d^{(t)-1}$, dubbed the Weingarten matrix, has elements (which are only known to leading-order [62]) of,

$$W_d^{(t)-1}(\sigma,\pi) = c(\sigma^{-1}\pi) \ \frac{1}{d^{|\sigma^{-1}\pi|}} \left[1 + \mathcal{O}\left(\frac{1}{d}\right)\right] \leq \chi_t \ c(\sigma^{-1}\pi) \ \frac{1}{d^{|\sigma^{-1}\pi|}} \ . \tag{A6}$$

Here, we have defined the Möbius functions over a permutation's cycles as $c(\sigma) = \prod_{\lambda\in\sigma} c(\lambda)$, $c(\lambda) = (-1)^{|\lambda|} c_{|\lambda|}$, where the Catalan numbers are $c_l = \frac{1}{l+1}\binom{2l}{l}$, and χ_t is a t-dependent upper bound [109].

To simplify our analysis, we will define the set of character-normalized permutations $\hat{\mathcal{S}}_t$ given by,

$$\hat{\mathcal{S}}_t = \{\hat{\sigma} = V_d(\sigma)/\chi_d^{(t)}(\sigma)\}_{\sigma\in\mathcal{S}_t} \qquad : \qquad \chi_d^{(t)}(\hat{\sigma},\hat{\pi}) = \frac{\chi_d^{(t)}(\sigma,\pi)}{\chi_d^{(t)}(\sigma)\ \chi_d^{(t)}(\pi)} \ . \tag{A7}$$

Here, it is important to note that while the elements of $\mathcal{S}_t$ are linearly independent for $d \geq t$ [64], they are not orthogonal [89]. Hence, we find it convenient to derive an alternative basis denoted as the localized permutations. The elements of this basis are in one-to-one correspondence with the original permutations, are localized with definitive support, and crucially are orthogonal with respect to support.

We can motivate our localized basis by considering a simple example on $t = 3$. First, let us first recall that the normalized operator for a transposition $\tau = (ij)$ acting on indices $i \neq j \in [t]$ is

$$\hat{\tau} = \sum_{P\in\mathcal{P}_d} P_i \otimes P_j^\dagger \ . \tag{A8}$$

Then, given the $l = 3$-cycle $\lambda = (012)$, we can decompose it as $(012) = (01)(02) = (12)(01)$. As such, we can express

$$\begin{aligned}\hat{\lambda} &= \sum_{P\in\mathcal{P}_d} P_0\otimes P_1^\dagger \sum_{S\in\mathcal{P}_d} S_0\otimes S_2^\dagger \ = \sum_{P,S\in\mathcal{P}_d} P_0S_0\otimes P_1^\dagger\otimes S_2^\dagger \\ &= \underbrace{I_d\otimes I_d\otimes I_d}_{()} + \sum_{P\in\mathcal{P}_d\backslash\{I_d\}}\underbrace{P\otimes P^\dagger\otimes I_d}_{(01)} + \sum_{P\in\mathcal{P}_d\backslash\{I_d\}}\underbrace{P\otimes I_d\otimes P^\dagger}_{(02)} + \sum_{P\in\mathcal{P}_d\backslash\{I_d\}}\underbrace{I_d\otimes P\otimes P^\dagger}_{(12)} + \sum_{P\neq S\in\mathcal{P}_d\backslash\{I_d\}}\underbrace{PS\otimes P^\dagger\otimes S^\dagger}_{(012)} \ .\end{aligned}$$

The brackets in the previous equation showcase that the final decomposition contains grouping of tensor products of operators that match the expansion of λ into its sub-permutations, i.e., $\{\pi = (), (01), (02), (12), (012) \subseteq (012) = \lambda\}$. Importantly, since the decomposition of a permutation into transpositions is not unique, the previous decomposition is also not unique. However, such expansions suggest there exist general relationships between representations of permutations, and their sub-permutations. We seek to exploit such relationships to define the basis of localized permutations.

Theorem 5 (Localized Basis for Permutations). *Let $\hat{\mathcal{S}}_t$ be the set of character normalized permutations. There exists a change of basis, $\phi_t : \hat{\mathcal{S}}_t \to [\hat{\mathcal{S}}_t]$, to a localized permutation basis $[\hat{\mathcal{S}}_t]$,*

$$\hat{\sigma} \ = \ \sum_{\pi\subseteq\sigma} [\hat{\pi}] \quad \leftrightarrow \quad [\hat{\sigma}] \ = \ \sum_{\pi\subseteq\sigma} \phi_t(\sigma,\pi)\ \hat{\pi} \ , \tag{A9}$$

where the elements of $[\hat{\mathcal{S}}_t] = \{[\hat{\sigma}]\}_{\sigma\in\mathcal{S}_t}$ depend on sub-permutation relationships $\pi \subseteq \sigma$, and obey the following properties:

i). Number of sub-permutations, of l-length cycles and of permutations σ: $c_l = \dfrac{1}{l+1}\dbinom{2l}{l} \leq l! \ \to \ c_\sigma = \displaystyle\prod_{\lambda\in\sigma} c_{l_\lambda} \leq c_{l_\sigma}$

Locality-l blocks of non-orthogonal operators: $\dbinom{t}{l}$ *number of* $!l = \left\lfloor \dfrac{l!}{e} + \dfrac{1}{2}\right\rfloor$ *sized blocks, such that* $t! = \displaystyle\sum_{l\in[t+1]\backslash\{1\}} \binom{t}{l}\, !l$

ii). Change of basis transformation: $\phi_t : \hat{\mathcal{S}}_t \to [\hat{\mathcal{S}}_t]$ is invertible, lower-triangular as per the sub-permutation ordering, is dimension-d-independent, and is solely dependent on the structure of the t-th order permutations:

$$\phi_t(\sigma,\pi) \ = \ c(\pi^{-1}\sigma)\ \delta_{\sigma\supseteq\pi} \qquad , \qquad c(\sigma) = \prod_{\lambda\in\sigma} c(\lambda) \quad , \quad c(\lambda) = (-1)^{|\lambda|}\ c_{|\lambda|} \qquad , \qquad \phi_t^{-1}(\sigma,\pi) \ = \ \delta_{\sigma\supseteq\pi} \ . \tag{A10}$$

iii). Linearly independent for $d \geq t$, and in one-to-one correspondence operators with permutations: $[\hat{\sigma}] \in [\hat{\mathcal{S}}_t] \leftrightarrow \sigma \in \mathcal{S}_t$

iv). Orthogonal with respect to support, and localized with definitive support $\Gamma_{[\hat{\sigma}]} = \Gamma_\sigma$:

$$\chi_d^{([\hat{\mathcal{S}}_t])} \ = \ \bigoplus_{\Gamma\subseteq[t]} \chi_d^{(\Gamma)} \ : \ \chi_d^{(\Gamma)} = \begin{bmatrix} \chi_d^{(t)}([\hat{\sigma}],[\hat{\sigma}]) & \cdots & \chi_d^{(t)}([\hat{\sigma}],[\hat{\pi}]) \\ \vdots & \ddots & \vdots \\ \chi_d^{(t)}([\hat{\pi}],[\hat{\sigma}]) & \cdots & \chi_d^{(t)}([\hat{\pi}],[\hat{\pi}]) \end{bmatrix}_{\substack{\sigma,\pi\in\mathcal{S}_t \\ \Gamma_\sigma=\Gamma_\pi=\Gamma}} \quad : \quad \left|\chi_d^{(t)}([\hat{\sigma}],[\hat{\pi}])\right| \ \leq \ c_\sigma^2\ c_\pi^2\ d^{|\sigma|+|\pi|}\ \delta_{\Gamma_\sigma\Gamma_\pi} \tag{A11}$$

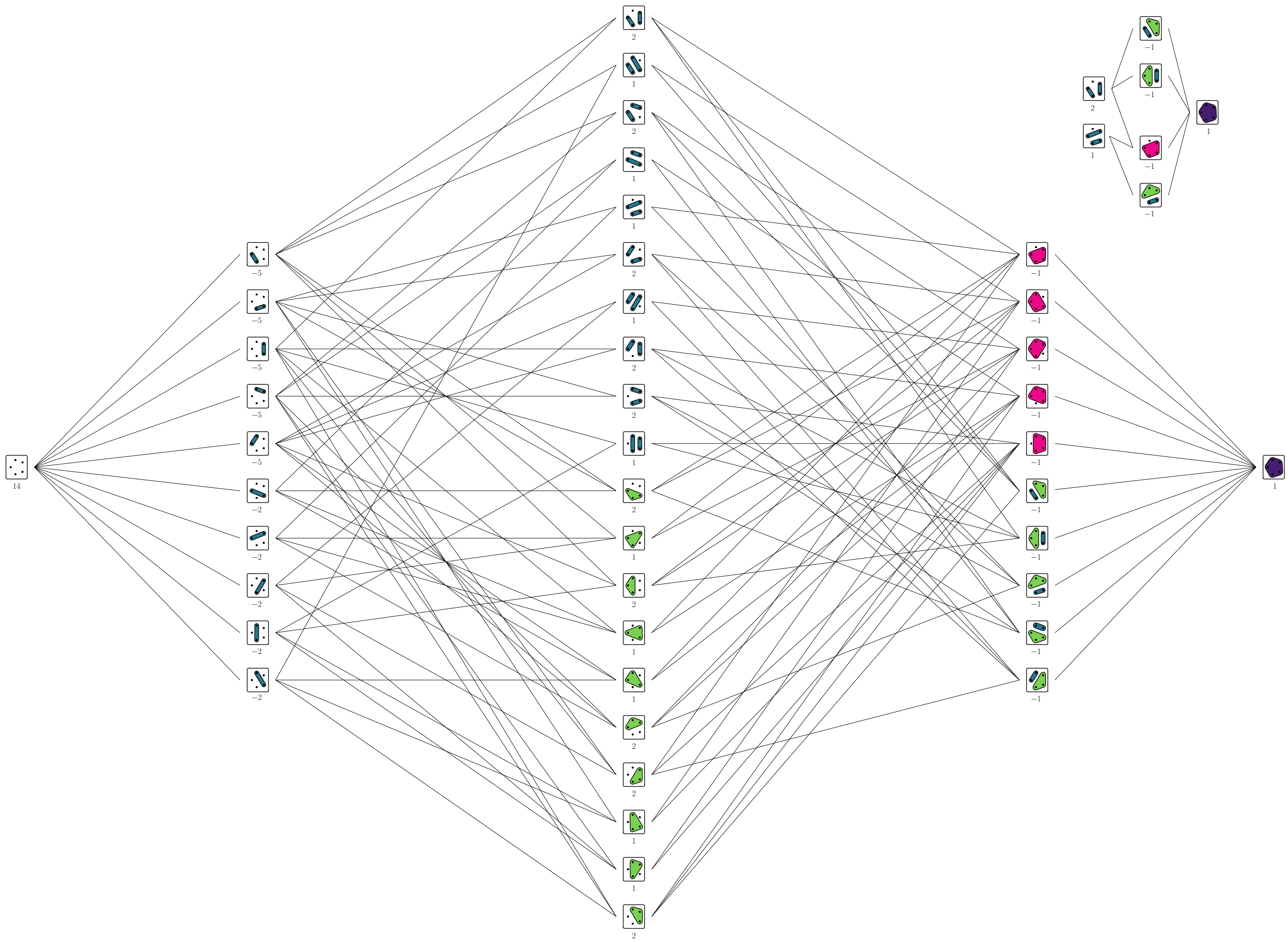

Figure 6. **Sub-permutation lattice for an $l = 5$ cycle $\sigma = (01234)$.** All $c_5 = 42$ sub-permutations $e \subseteq \xi \subseteq \sigma$ are shown as as non-crossing partitions of points, ordered counter-clockwise from the leftmost point on an $l = 5$-polygon. Nodes of the lattice depict sub-permutations $e \subseteq \xi \subseteq \sigma$ between minimal e and maximal σ, ordered in l columns, by size $0 \leq |\xi| \leq l-1$, from left to right. Edges of the lattice depict sub-permutation relationships, from left to right. At a given node ξ, edges and paths to rightward nodes ξ' represent larger sup-permutations $\xi \subset \xi' \subseteq \sigma$, such that $1 \leq \left|\xi^{-1}\xi'\right| \leq \left|\xi^{-1}\sigma\right| \leq l-1$. Such sup-permutations $\{\xi \subseteq \xi' \subseteq \sigma\}$ of ξ form of a sub-lattice with minimal ξ and maximal σ. Permutations ξ are labelled with their localized permutation change of basis coefficient $\phi_l(\sigma,\xi) = c(\xi^{-1}\sigma)$, such that $\sum_{\pi\subseteq\xi\subseteq\sigma} \phi_l(\sigma,\xi) = \delta_{\sigma\pi}$. Inset (top right) is the sub-lattices forming a set S_σ, with minimal permutations $\pi \in T_\sigma = \{\pi = (01)(23), (03)(12) \subset \sigma = (01234)\}$, with support $\Gamma = \Gamma_\pi = \{0,1,2,3\} \subset \Gamma_\sigma = \{0,1,2,3,4\}$, and a smallest common sup-permutation $\pi_\sigma = (0123) \subset \sigma = (01234)$, such that $\sum_{\xi\in S_\sigma} \phi_l(\sigma,\xi) = \sum^{\pi\in T_\sigma}_{\pi\subseteq\xi\subseteq\sigma} \phi_l(\sigma,\xi) - \sum^{\pi=\pi_\sigma}_{\pi\subseteq\xi\subseteq\sigma} \phi_l(\sigma,\xi) = 0$.

Proof. To prove properties of the localized permutation basis, we use properties of sub-permutations $\pi \subseteq \sigma$ such that $\left|\pi^{-1}\sigma\right| = |\sigma| - |\pi|$, which represent a partial ordering between permutations [118], and is transitive, such that if $\pi \subseteq \xi$, and $\xi \subseteq \sigma$, then $\pi \subseteq \sigma$, and if $\pi \subseteq \xi \subseteq \sigma$, then $\xi \subseteq \pi^{-1}\sigma$. Concurrently, the number of sub-permutations of an l-length cycle also corresponds to the number of non-crossing partitions of the set $[l]$ that respect the order of the permuted indices [118]. As depicted in Fig. 6, the set of sub-permutations $\{\pi \subseteq \xi \subseteq \sigma\} \cong \{e \subseteq \xi \subseteq \pi^{-1}\sigma\}$ forms a lattice containing all paths of sub-permutations between some minimum π and maximum σ permutations. We also note that any sub-permutations $\{\pi \subseteq \lambda\}$ of a cycle λ, with identical support $\Gamma_\pi = \Gamma \subseteq \Gamma_\lambda$ and locality $l_\pi = l \leq l_\lambda$, are within a set of permutations that is isomorphic to the l-order permutations $\mathcal{S}_l$. Such sub-permutations $\{\pi \subseteq \lambda\}$ are also always sub-permutations of a maximal Γ-support, l-length cycle $\gamma_\lambda \subseteq \lambda$. Similarly, many permutation-dependent quantities can be partitioned by the support $\Gamma_\sigma = \Gamma \subseteq [t]$ and locality $l_\sigma = l = |\Gamma| \leq t$ of permutations σ. Properties of sub-permutations also hold independently for disjoint cycles, given permutations $\pi = \prod_{\lambda\in\sigma}\pi_\lambda$ can be decomposed into disjoint sub-permutations $\{\pi_\lambda \subseteq \lambda\}_{\lambda\in\sigma}$ for any sup-permutation $\sigma = \prod_{\lambda\in\sigma}\lambda \supseteq \pi$. It follows that sup-permutation dependent quantities can often be factorized, such as $[\hat{\pi}] = \prod_{\lambda\in\sigma}[\hat{\pi}_\lambda]$, and $\phi_t(\sigma,\pi) = \prod_{\lambda\in\sigma}\phi_{l_\lambda}(\lambda,\pi_\lambda)$.

We will prove each property $i)$ to $iv)$ of the mapping between the permutation and localized permutation bases separately.

i) To prove the number of sub-permutations and localized permutations, we note that the number of sub-permutations, which equals the number of non-crossing partitions, equals the Catalan numbers, $c_l = \frac{1}{l+1}\binom{2l}{l}$ [119]. The number of sub-permutations between permutations $\pi \subseteq \sigma$ subsequently are,

$$c_{\sigma,\pi} = |\{\xi \in \mathcal{S}_t \ : \ \pi \subseteq \xi \subseteq \sigma\}| = c_{\pi^{-1}\sigma}\ \delta_{\sigma \supseteq \pi} \quad : \quad c_\sigma = |\{\xi \in \mathcal{S}_t \ : \ \xi \subseteq \sigma\}| = \prod_{\lambda \in \sigma} c_{l_\lambda}\ . \tag{A12}$$

The number of locality-l blocks of permutations is $\binom{t}{l}$, and the number of permutations non-trivially permuting their l-locality support Γ, is by definition, $!l = \left\lfloor \frac{l!}{e} + \frac{1}{2} \right\rfloor$, the number of length-$l$ derangements [120].

ii) To prove the existence of the localized permutation basis, we will show that an invertible relationship exists, also known as Möbius inversion [121], between the normalized permutations $\hat{\mathcal{S}}_t$, and a set of operators $[\hat{\mathcal{S}}_t]$. Let $\phi_t^{-1} \ : \ [\hat{\mathcal{S}}_t] \to \hat{\mathcal{S}}_t$ be such a change of basis, with associated elements of $\phi_t^{-1}(\sigma,\pi) = \delta_{\sigma\supseteq\pi}$. Such a matrix is lower-triangular, with a unit diagonal. Its sub-permutation dependent structure further impose that its strictly lower-triangular component is t-nilpotent, given the (σ,π) element of its k powers is the number of $k+1 \leq \left|\pi^{-1}\sigma\right|+1 \leq t$-length paths in the lattice. The change of basis matrix inverse thus exists, $\phi_t \ : \ \hat{\mathcal{S}}_t \to [\hat{\mathcal{S}}_t]$, in fact with a closed form Möbius inverse [121], with identical sub-permutation dependent structure $\phi_t(\sigma,\pi), \phi_t^{-1}(\sigma,\pi) \propto \delta_{\sigma\supseteq\pi}$, that only depends on the overlaps of permutations,

$$\phi_t = \sum_{k\in[t]} (I_{t!} - \phi_t^{-1})^k \quad , \quad \phi_t(\sigma,\pi) = \phi_t(\pi^{-1}\sigma)\ \delta_{\sigma\supseteq\pi} \quad : \quad \phi_t(\sigma) = c(\sigma) = \prod_{\lambda\in\sigma} c(\lambda) \quad , \quad \sum_{\pi\subseteq\sigma} \phi_t(\pi) = \delta_{\sigma e}\,. \tag{A13}$$

Such relationships are independent of the representation dimension d, and solely depend on the structure of permutations.

iii) From the invertible relationship between the bases, both the permutations and the localized permutations are linearly independent for $d \geq t$ [64], and the localized permutations are in one-to-one correspondence with the permutations.

iv) To prove that the localized permutations are localized, we consider the effect of the support of a basis operator string, $P = \otimes_{i\in\Gamma_P} P_i \in \mathcal{P}_d^t$, with support $\Gamma_P \subseteq \Gamma_\sigma \subseteq [t]$, on its overlap with a localized permutation $[\hat{\sigma}] \in [\hat{\mathcal{S}}_t]$, with potentially non-localized support $\Gamma_{[\hat{\sigma}]} = \cup_{\pi\subseteq\sigma}\Gamma_\pi = \Gamma_\sigma$. As depicted in Fig. 7, the overlap between a permutation and an operator string is the tracefullness of products of operators, in the order of the permuted indices within each disjoint cycle of the permutation. Such products of operators are invariant under any sup-permutations of the permutation, meaning non-zero overlaps of operator strings with all permutations with a common sup-permutation $\sigma \supseteq \pi, \pi'$, must be identical,

$$\chi_d^{(t)}(\hat{\pi}, P) \;=\; \chi_d^{(t)}(\hat{\pi}', P) \;=\; \chi_d^{(t)}(\hat{\sigma}, P) \;=\; \frac{1}{\chi_d^{(t)}(\sigma)} \prod_{\lambda\in\sigma} \chi_d^{(t)}\big(\prod_{i\in\Gamma_\lambda} P_i\big) \;\propto\; \delta_{\Gamma_\sigma\supseteq\Gamma_P}\ \delta_{\Gamma_\pi\supseteq\Gamma_P}\ \delta_{\Gamma_{\pi'}\supseteq\Gamma_P} \quad \forall\ \sigma \supseteq \pi, \pi'\ . \tag{A14}$$

Therefore overlaps of localized permutations with an operator string reduce to summations of coefficients ϕ_t over the subset of permutations, $S_\sigma = \{\xi \subseteq \sigma \ : \ \Gamma_\xi \supseteq \Gamma_P \ , \ \chi_d^{(t)}(\xi, P) \neq 0\}$ that have non-zero overlap with the operator string,

$$\frac{\chi_d^{(t)}([\hat{\sigma}], P)}{\chi_d^{(t)}(\hat{\sigma}, P)} \;=\; \sum_{\xi\subseteq\sigma} \phi_t(\sigma,\xi)\ \frac{\chi_d^{(t)}(\hat{\xi}, P)}{\chi_d^{(t)}(\hat{\sigma}, P)} \;=\; \sum_{\xi\in S_\sigma} \phi_t(\sigma,\xi)\ \frac{\chi_d^{(t)}(\hat{\xi}, P)}{\chi_d^{(t)}(\hat{\sigma}, P)} \;=\; \sum_{\xi\in S_\sigma} \phi_t(\sigma,\xi) \;\overset{?}{\propto}\; \delta_{\Gamma_P\Gamma_\sigma}\ . \tag{A15}$$

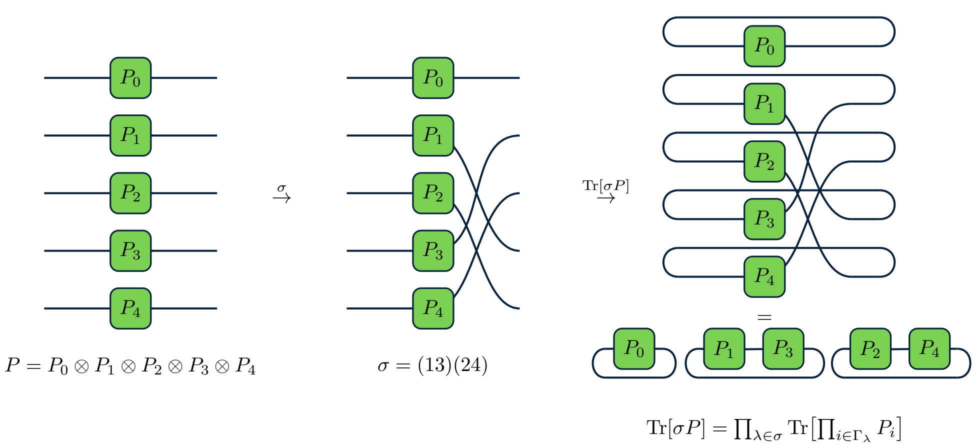

Figure 7. **Overlap between an operator string P, and a permutation σ.** Overlaps are equal to the tracefullness of the products of an operator's components $\otimes_{\lambda\in\sigma}\prod_{i\in\Gamma_\lambda} P_i$, in the order of indices $\cup_{\lambda\in\sigma}\Gamma_\lambda$ permuted by a permutation's disjoint cycles.

Given summations of the coefficients ϕ_t over lattices $\{\pi \subseteq \xi \subseteq \sigma\}$ are non-zero if $\pi = \sigma$, we seek to express summations over S_σ in terms of summations over lattices, potentially leading to cancellations. The set $S_\sigma = \cup_{\pi \in T_\sigma} S_{\sigma\pi}$, although not a lattice, can be partitioned into lattices, $S_{\sigma\pi} = \{\xi \in S_\sigma \ : \ \pi \subseteq \xi \subseteq \sigma\}$, each defined by a unique minimal permutation, $\pi \in T_\sigma = \{\pi \in S_\sigma \ : \ \pi \not\supset \xi \ \ \forall\, \xi \in S_\sigma\}$, as in the inset of Fig. 6. Crucially, all minimal $\pi \in T_\sigma$, given their identical support $\Gamma_\pi = \Gamma_P$, have a unique smallest common $\Gamma_{\pi_\sigma} = \Gamma_P$-support sup-permutation π_σ, such that $\pi \subseteq \pi_\sigma \subseteq \sigma$, and $\pi_\sigma = \prod_{\lambda \in \sigma} \pi_{\sigma\lambda}$, with components $\{\pi_{\sigma\lambda} \subseteq \lambda\}_{\lambda \in \sigma}$. This common sup-permutation π_σ is the minimum of the intersection of all the lattices $\cap_{\pi \in T_\sigma} S_{\sigma\pi} = \{\xi \in S_\sigma : \pi_\sigma \subseteq \xi \subseteq \sigma\}$, is itself a sub-permutation of the unique Γ_P-support permutation $\gamma_\sigma = \prod_{\lambda \in \sigma} \gamma_{\sigma\lambda} \supseteq \pi_\sigma$, with cycles $\{\gamma_{\sigma\lambda} : \pi_{\sigma\lambda} \subseteq \gamma_{\sigma\lambda} \subseteq \lambda\}_{\lambda \in \sigma}$, and is bounded in size, namely in the equal support case, $\Gamma_P = \Gamma_\sigma \ \rightarrow \ \pi_\sigma \subseteq \sigma$, whereas in the strict subset of support case, $\Gamma_P \subset \Gamma_\sigma \ \rightarrow \ \pi_\sigma \subset \sigma$.

Given this decomposition of a set of permutations into a union of lattices, we use the inclusion-exclusion principle [122], that relates counting of unions as counting of intersections, less multiply counted elements in multiple intersections,

$$\sum_{\xi \in \cup_{\pi \in T_\sigma} S_{\sigma\pi}} = \sum_{\{\} \subset T \subseteq T_\sigma} (-1)^{|T|-1} \sum_{\xi \in \cap_{\pi \in T} S_{\sigma\pi}} . \tag{A16}$$

For example, in the equal support case, if there is a single minimal permutation $\pi = \sigma$, forming a single lattice, $S_{\sigma\pi} = S_\sigma = \{\sigma\}$, then trivially $\sum_{\xi \in S_\sigma} \phi_t(\sigma, \xi) = 1$, whereas in the strict subset of support case, if there is a single minimal permutation $\pi \subset \sigma$, forming a single lattice, $S_{\sigma\pi} = S_\sigma = \{\pi \subseteq \xi \subseteq \sigma\}$, then trivially $\sum_{\xi \in S_\sigma} \phi_t(\sigma, \xi) = 0$. It is therefore essential to understand the non-trivial intersections of lattices in order to understand these summations.

Although intersections of lattices, $S'_\sigma = \cap_{\pi \in T} S_{\sigma\pi}$ for some $\{\} \subset T \subseteq T_\sigma \ : \ |T| > 1$, are not necessarily lattices [123, 124], given the sub-permutation relationships, any sup-permutation $\xi'' \supseteq \xi'$ of any permutation in the intersection $\xi' \in S'_\sigma$, is necessarily also in this intersection, $\xi'' \in S'_\sigma$. Therefore, intersections of lattices can be themselves decomposed into new lattices, $S'_\sigma = \cup_{\pi \in T'_\sigma} S'_{\sigma\pi}$, each defined by a unique minimal permutation, $\pi \in T'_\sigma = \{\pi \in S'_\sigma \ : \ \pi \not\supset \xi \ \ \forall\, \xi \in S'_\sigma\}$. Further, given the new lattices are subsets of the old lattice intersections, the new minimal permutations must have an identical smallest common sup-permutation of $\pi'_\sigma = \pi_\sigma$, and for any $\pi' \in T'_\sigma$, there exists $\pi \in T$, such that $\pi \subset \pi' \subseteq \pi_\sigma$, and thus $|T'_\sigma| < |T_\sigma|$. Each intersection of lattices can thus be recursively defined as unions of sub-lattices, $S_\sigma \rightarrow S'_\sigma \rightarrow \cdots \rightarrow S^*_\sigma$, as depicted in Fig. 8, where colored patches and nodes represent their non-intersecting components, with colored minimal elements $T_\sigma \rightarrow T'_\sigma \rightarrow \cdots \rightarrow T^*_\sigma$. Each union in the sequence is spanned by less and less sub-lattices, and terminates after at most $|\sigma|+1$ iterations, yielding a single lattice $S^*_\sigma = \{\pi_\sigma \subseteq \xi \subseteq \sigma\}$, with minimal element $T^*_\sigma = \{\pi_\sigma\}$.

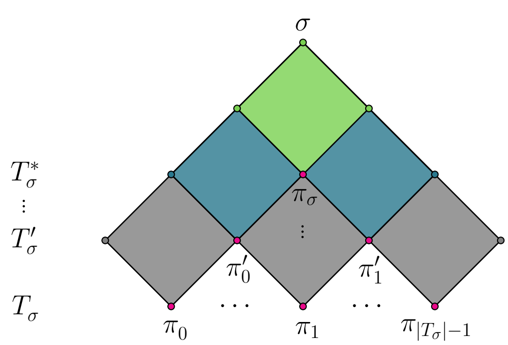

Figure 8. **Decomposition of $S_\sigma = \cup_{\pi \in \bar{T}_\sigma} S_{\sigma\pi}$ into intersecting sub-lattices $S_{\sigma\pi} = \{\pi \subseteq \xi \subseteq \sigma\}$.** Gray, blue, green patches and nodes depict disjoint subsets of each sub-lattice, with pink nodes for minimal permutations $\pi \in \bar{T}_\sigma = \cup\{T_\sigma, T'_\sigma, \ldots, T^*_\sigma\} : \pi \subseteq \pi_\sigma \subseteq \sigma$.

Summations over S_σ therefore take the form of summations over lattices, defined by some set of minimal elements $\bar{T}_\sigma = \{\xi \ : \ \exists\, \pi \in T_\sigma \ : \ \pi \subseteq \xi \subseteq \pi_\sigma\}$, weighted by some non-negative multiplicities and signs $\{n_\pi, k_\pi\}_{\pi \in \bar{T}_\sigma}$, such that,

$$\sum_{\xi \in S_\sigma} \phi_t(\sigma, \xi) \ = \ \sum_{\pi \in \bar{T}_\sigma} (-1)^{k_\pi} \ n_\pi \sum_{\pi \subseteq \xi \subseteq \sigma} \phi_t(\sigma, \xi) \ = \ (-1)^{k_\sigma} \ n_\sigma \ \delta_{\pi_\sigma \sigma} \ \propto \ \delta_{\Gamma_P \Gamma_\sigma} . \tag{A17}$$

Operator strings P with a strict subset of support $\Gamma_P \subset \Gamma_\sigma$ therefore have zero overlap with localized permutations $[\hat{\sigma}]$, meaning localized permutations are localized, with definitive support, $\Gamma_{[\hat{\sigma}]} = \Gamma_\sigma$, with overlaps of,

$$[\hat{\sigma}] \ = \ \sum_{P \in \mathcal{P}^t_d} \chi^{(t)}_d([\hat{\sigma}], P) \ P \quad : \quad \chi^{(t)}_d([\hat{\sigma}], P) \ = \ (-1)^{k_\sigma} \ n_\sigma \ \frac{\chi^{(t)}_d(\sigma, P)}{\chi^{(t)}_d(\sigma)} \delta_{\Gamma_P \Gamma_\sigma} \quad \rightarrow \quad \chi^{(t)}_d([\hat{\sigma}], [\hat{\pi}]) \ \propto \ \delta_{\Gamma_\sigma \Gamma_\pi} . \tag{A18}$$

Intriguingly, such arguments could potentially hold for the localization of other change of bases $\{\hat{X}\} \to \{[\hat{X}]\}$ that are in terms of other partial orderings $Y \subseteq X$ between operators, their Möbius inverses ϕ_t, and their summations $\sum_{Y\in S_X} \phi_t(X,Y)$, given symmetries of overlaps with sub and sup-operators $\chi_d^{(t)}(\hat{Y},P) = \chi_d^{(t)}(\hat{X},P)$ for $Y \subseteq X$, $P \in \mathcal{P}_d^t$.

To prove the bounds on the overlap between localized permutations, we recall their definition,

$$\chi_d^{(t)}([\hat{\sigma}],[\hat{\pi}]) \;=\; \sum_{\substack{\eta\subseteq\sigma\\ \kappa\subseteq\pi}} \phi_t(\sigma,\eta)\;\phi_t(\pi,\kappa)\frac{\chi_d^{(t)}(\eta,\kappa)}{\chi_d^{(t)}(\eta)\chi_d^{(t)}(\kappa)} \qquad\leftrightarrow\qquad \frac{\chi_d^{(t)}(\sigma,\pi)}{\chi_d^{(t)}(\sigma)\chi_d^{(t)}(\pi)} \;=\; \sum_{\substack{\eta\subseteq\sigma\\ \kappa\subseteq\pi}} \chi_d^{(t)}([\hat{\eta}],[\hat{\kappa}]) \;. \tag{A19}$$

Given bounds on permutation sizes, $|\sigma|+|\pi|-\left|\sigma^{-1}\pi\right| \le |\sigma|+|\pi|$, their overlaps are, $\chi_d^{(t)}(\sigma,\pi)/\chi_d^{(t)}(\sigma)\chi_d^{(t)}(\pi) \le d^{|\sigma|+|\pi|}$, given the monotonicity of the Catalan numbers $c_l \le c_{l+1}$, the Möbius functions are, $\left|c(\pi^{-1}\sigma)\right| \le |c(\sigma)| \le c_\sigma$, and given summations are bounded by the largest summand times the number of summands, $|\chi_d^{(t)}([\hat{\sigma}],[\hat{\pi}])| \le c_\sigma^2\, c_\pi^2\, d^{|\sigma|+|\pi|}\,\delta_{\Gamma_\sigma\Gamma_\pi}$.

Finally, we note there is a trade-off in this choice of localized basis. The mapping from the localized permutations to the permutations has a simple, attractive closed form, with known enumerations of the expansions in terms of sub-permutations, with unit expansion coefficients. Permutations can equally be expanded in terms of a basis of strings of orthogonal operators, grouped by their support. However such groupings are not enumerated by the sub-permutations, and the associated expansion coefficients have no closed form, requiring the algorithms used in this proof.

□

Appendix B: Properties of Moment Operators

In these appendices, we consider properties of t-th order moment operators for various ensembles, generalizing previous expressions [18], and adapt approaches for integrating over ensembles of operators [51, 62, 63].

1. Twirls and Moment Operators

Here, we denote the t-th order moment operator for an ensembles of channels $\mathcal{C} \subseteq \mathcal{B}[\mathcal{B}[\mathcal{H}]]$ as $\widehat{\mathcal{T}}_{\mathcal{C}}^{(t)} = \int_{\mathcal{C}} d\Lambda\; \widehat{\Lambda^{\otimes t}}$. Then, the moment operator for k-concatenated independent ensembles $\{\mathcal{C}_i\}_{i\in[k]}$ is the product of the moment operators,

$$\widehat{\mathcal{T}}_{\{\mathcal{C}_i\}}^{(t)} \;=\; \prod_{i\in[k]} \widehat{\mathcal{T}}_{\mathcal{C}_i}^{(t)} \;=\; \prod_{i\in[k]} \int_{\mathcal{C}_i} d\Lambda_i\; \widehat{\Lambda_i^{\otimes t}} \;=\; \widehat{\mathcal{D}}_d^{\otimes t} \;+\; \widehat{\Delta}_{\{\mathcal{C}_i\}}^{(t)} \;, \tag{B1}$$

where we defined

$$\widehat{\mathcal{D}}_d^{\otimes t} = \frac{1}{d^t} |I_d^{\otimes t}\rangle\!\rangle\langle\!\langle I_d^{\otimes t}| \quad\text{and}\quad \widehat{\Delta}_{\{\mathcal{C}_i\}}^{(t)} = \frac{1}{d^t} \sum_{P\in\mathcal{S}_{\mathcal{C}_0}^{(t)}\setminus\{I_d^{\otimes t}\},\; S\in\mathcal{S}_{\mathcal{C}_{k-1}}^{(t)}} \tau_{\{\mathcal{C}_i\}}^{(t)}(P,S) |P\rangle\!\rangle\langle\!\langle S| \;. \tag{B2}$$

Here, $\tau_{\{\mathcal{C}_i\}}^{(t)}$ denote the composite expansion coefficients, also known as the transfer matrix elements. These composite transfer matrices can be expanded in terms of the individual transfer matrices $\{\tau_{\Lambda_i}^{(t)}\}_{i\in[k]}$ for each moment operator, and overlaps $\chi_d^{(t)}$ between the elements of some ensemble-dependent bases $\{\mathcal{S}_{\mathcal{C}_i}^{(t)} \subseteq \mathcal{B}[\mathcal{H}^{\otimes t}]\}$ as,

$$\tau_{\{\mathcal{C}_i\}}^{(t)}(P,S) = \sum_{\substack{\{P_i,S_i\in\mathcal{S}_{\mathcal{C}_i}^{(t)}\}_{i\in[k]}\\ P_0=P,\; S_{k-1}=S}} \prod_{i\in[k]} \tau_{\mathcal{C}_i}^{(t)}(P_i,S_i) \prod_{i\in[k-1]} \chi_d^{(t)}(S_i,P_{i+1}) \;. \tag{B3}$$

In the case of concatenating k identical ensembles, i.e., when $\mathcal{C}_i = \mathcal{C}\ \forall i \in [k]$, our notation is replaced with k-powers, or k-superscripts, namely $\widehat{\mathcal{T}}_{\{\mathcal{C}_i\}}^{(t)} \to \widehat{\mathcal{T}}_{\mathcal{C}}^{(t)k}$, $\widehat{\Delta}_{\{\mathcal{C}_i\}}^{(t)} \to \widehat{\Delta}_{\mathcal{C}}^{(t,k)}$, and $\tau_{\{\mathcal{C}_i\}}^{(t)} \to \tau_{\mathcal{C}}^{(t,k)}$.

To compare two ensembles $\mathcal{C},\mathcal{C}'$, let us define the difference between their t-th order moment operators as

$$\widehat{\Delta}_{\mathcal{C}\mathcal{C}'}^{(t)} = \widehat{\mathcal{T}}_{\mathcal{C}}^{(t)} - \widehat{\mathcal{T}}_{\mathcal{C}'}^{(t)} \qquad\to\qquad \epsilon_{\mathcal{C}\mathcal{C}'}^{(t)} = \left\|\widehat{\Delta}_{\mathcal{C}\mathcal{C}'}^{(t)}\right\| \quad,\quad \epsilon_{\mathcal{C}}^{(t)} = \left\|\widehat{\Delta}_{\mathcal{C}}^{(t)}\right\| \;. \tag{B4}$$

Various norms $\|\cdot\|_*$ may be chosen, denoted with $\epsilon^{(t|*)}$ superscripts, and unless noted, we use the Hilbert-Schmidt norm.

To finish, we note that for the special cases when the ensembles contain strictly unital or unitary operators, we can readily prove the following result concerning how the associated moment operators preserve the support of their inputs.

Lemma 1. *Given an orthogonal with respect to support basis, unital ensemble twirls $\mathcal{C}_{\mathcal{U}}$ are support non-increasing and unitary ensemble twirls $\mathcal{U}$ are support preserving,*

$$\widehat{\mathcal{T}}^{(t)}_{\mathcal{C}_{\mathcal{U}}} = \frac{1}{d^t} \sum_{\substack{P,S\in\mathcal{S}^{(t)}_{\mathcal{C}_{\mathcal{U}}} \\ \Gamma_P\subseteq\Gamma_S}} \tau^{(|S|)}_{\mathcal{C}_{\mathcal{U}}}(P,S)\ |P\rangle\!\rangle\langle\!\langle S| \qquad \leftrightarrow \qquad \tau^{(t)}_{\mathcal{C}_{\mathcal{U}}}(P,S) = \tau^{(|S|)}_{\mathcal{C}_{\mathcal{U}}}(P,S)\ \delta_{\Gamma_P\subseteq\Gamma_S} \tag{B5}$$

$$\widehat{\mathcal{T}}^{(t)}_{\mathcal{U}} = \frac{1}{d^t} \sum_{\substack{P,S\in\mathcal{S}^{(t)}_{\mathcal{U}} \\ \Gamma_P=\Gamma_S}} \tau^{(|S|)}_{\mathcal{U}}(P,S)\ |P\rangle\!\rangle\langle\!\langle S| \qquad \leftrightarrow \qquad \tau^{(t)}_{\mathcal{U}}(P,S) = \tau^{(|S|)}_{\mathcal{U}}(P,S)\ \delta_{\Gamma_P=\Gamma_S}\ . \tag{B6}$$

Proof. Let us consider consider the action of unital ensemble $\mathcal{C}_{\mathcal{U}}$ twirls on $|X|$-local operators $X = \otimes_{i\in\Gamma_X} X_i \in \mathcal{B}[\mathcal{H}^{\otimes t}]$, given identity operators are invariant under unital operators,

$$\mathcal{T}^{(t)}_{\mathcal{C}_{\mathcal{U}}}(X) = \int_{\mathcal{C}_{\mathcal{U}}} d\Lambda\ \Lambda^{\otimes t}(X)\ = \int_{\mathcal{C}_{\mathcal{U}}} d\Lambda\ \otimes_{i\in\Gamma_X} \Lambda_i(X)\ = \sum_{\substack{P,S\in\mathcal{S}^{(|X|)}_{\mathcal{C}_{\mathcal{U}}} \\ \Gamma_P\subseteq\Gamma_S=\Gamma_X}} \tau^{(|X|)}_{\mathcal{C}_{\mathcal{U}}}(P,S)\ \chi^{(|X|)}_d(S,X)\ P\ = \mathcal{T}^{(|X|)}_{\mathcal{C}_{\mathcal{U}}}(X)\ . \tag{B7}$$

It follows that unital transfer matrices reduce to being non-zero $\tau^{(t)}_{\mathcal{C}_{\mathcal{U}}}(P,S) = \tau^{(|X|)}_{\mathcal{C}_{\mathcal{U}}}(P,S) \sim \delta_{\Gamma_P\subseteq\Gamma_S}$ for $P,S\in\mathcal{S}^{(t)}_{\mathcal{C}_{\mathcal{U}}}$ when mapping between basis operators of strictly common support. In the case of unitary ensembles $\mathcal{U}$, the associated moment operators are also self-adjoint, $\mathcal{T}^{(t)\dagger}_{\mathcal{U}} = \mathcal{T}^{(t)}_{\mathcal{U}}$, constraining such twirls to map between operators of identical support. □

2. Properties of Haar and cHaar Twirls

Here, we derive expressions and properties for moment operators for the Haar, cHaar, and Depolarize ensembles. Throughout this section, we will use the fact that the t-th moment operator over the representation of a group projects onto its t-th order commutant. We begin with the following definitions of the relevant ensembles.

Definition 1 (Haar ensemble)**.** *The t-th order twirl over the Haar ensemble $\mathcal{U}(d)$ is defined as the twirl over t copies of unitaries sampled from the group $\mathcal{U}(d)$ according to the Haar measure. The associated twirl and moment operator are*

$$\mathcal{T}^{(t)}_{\mathcal{U}(d)}(\cdot)\ =\ \int_{\mathcal{U}(d)} dU\ U^{\otimes t}\ \cdot\ U^{\otimes t\dagger} \qquad \to \qquad \widehat{\mathcal{T}}^{(t)}_{\mathcal{U}(d)}\ =\ \int_{\mathcal{U}(d)} dU\ U^{\otimes t}\otimes U^{\otimes t*}\ . \tag{B8}$$

Since the t-th order commutant of the unitary group corresponds to the system-permuting representation of the Symmetric group we can always express the moment operator as

$$\widehat{\mathcal{T}}^{(t)}_{\mathcal{U}(d)} = \frac{1}{d^t} \sum_{\sigma,\pi\in\mathcal{S}_t} \tau^{(t)}_{\mathcal{U}(d)}(\sigma,\pi)\ |\sigma\rangle\!\rangle\langle\!\langle\pi| \qquad \textit{with} \qquad \tau^{(t)}_{\mathcal{U}(d)}(\sigma,\pi) = W^{(t)-1}_d(\sigma,\pi)\ . \tag{B9}$$

Definition 2 (cHaar ensemble)**.** *The t-th order twirl over the channel-Haar (cHaar) ensemble $\mathcal{C}(d,d_{\mathcal{E}})$ is defined as a measure over t copies of random channels, and is a uniform, Lebesgue measure when $d_{\mathcal{E}} = d^2$. Using the Stinespring formalism, we can re-write this average as a twirl with respect to the Haar ensemble of unitaries acting on t copies of a composite space, of a system and environment $\mathcal{H}\otimes\mathcal{E}$ of dimensions $d, d_{\mathcal{E}}$, and a local pure environment state $\nu_{\mathcal{E}}\in\mathcal{B}[\mathcal{E}]$,*

$$\mathcal{T}^{(t)}_{\mathcal{C}(d,d_{\mathcal{E}})}(\cdot)\ =\ \mathrm{Tr}_{\mathcal{E}^{\otimes t}}\left[\mathcal{T}^{(t)}_{\mathcal{U}(dd_{\mathcal{E}})}(\cdot\ \otimes\ \nu^{\otimes t}_{\mathcal{E}})\right] \qquad \to \qquad \widehat{\mathcal{T}}^{(t)}_{\mathcal{C}(d,d_{\mathcal{E}})}\ =\ I^{\otimes t}_d\otimes\langle\!\langle I^{\otimes t}_{d_{\mathcal{E}}}|\ \ \widehat{\mathcal{T}}^{(t)}_{\mathcal{U}(dd_{\mathcal{E}})}\ \ I^{\otimes t}_d\otimes|\nu^{\otimes t}_{\mathcal{E}}\rangle\!\rangle\ . \tag{B10}$$

Since the cHaar ensemble is generated by the unitary Haar ensemble over $\mathcal{U}(dd_{\mathcal{E}})$, $\widehat{\mathcal{T}}^{(t)}_{\mathcal{C}(d,d_{\mathcal{E}})}$ can be obtained by projecting onto the commutant of $\mathcal{U}(dd_{\mathcal{E}})$, acting on the environment state $\nu_{\mathcal{E}}$, and then tracing out the environment. This leads to

$$\widehat{\mathcal{T}}^{(t)}_{\mathcal{C}(d,d_{\mathcal{E}})} = \frac{1}{d^t} \sum_{\sigma,\pi\in\mathcal{S}_t} \tau^{(t)}_{\mathcal{C}(d,d_{\mathcal{E}})}(\sigma,\pi)\ |\sigma\rangle\!\rangle\langle\!\langle\pi| \qquad \textit{with} \qquad \tau^{(t)}_{\mathcal{C}(d,d_{\mathcal{E}})}(\sigma,\pi) = \chi^{(t)}_{d_{\mathcal{E}}}(\sigma)\ \tau^{(t)}_{\mathcal{U}(dd_{\mathcal{E}})}(\sigma,\pi)\ . \tag{B11}$$

Definition 3 (Depolarize ensemble)**.** *The t-th order twirl over the depolarizing (Depolarize) ensemble $\mathcal{D}(d)$ is defined via a single maximally depolarizing channel $\mathcal{D}_d(\cdot)$, with unit probability. The associated twirl and moment operator are*

$$\mathcal{T}^{(t)}_{\mathcal{D}(d)}(\cdot) = \mathcal{D}^{\otimes t}_d(\cdot) = \frac{\mathrm{Tr}\,[\cdot]}{d^t} I^{\otimes t}_d \qquad \to \qquad \widehat{\mathcal{T}}^{(t)}_{\mathcal{D}(d)} = \widehat{\mathcal{D}}^{\otimes t}_d = \frac{1}{d^t}|I^{\otimes t}_d\rangle\!\rangle\langle\!\langle I^{\otimes t}_d|\ . \tag{B12}$$

The Depolarize moment operator projects onto an invariant, trace-preserving basis, of solely the identity operator $\{I^{\otimes t}_d\}$.

We now derive explicit forms of the Haar, cHaar, and Depolarize moment operators.

Theorem 6 (Haar Moment Operators are Support-Preserving and Block-Diagonal in Localized Basis)**.** *The Haar moment operator is block-diagonal in the localized permutation basis, as per its support,*

$$\widehat{\mathcal{T}}^{(t)}_{\mathcal{U}(d)} = \bigoplus_{\Gamma\subseteq[t]} \widehat{\mathcal{T}}^{(\Gamma)}_{\mathcal{U}(d)} = \begin{bmatrix} \widehat{\mathcal{T}}^{(\{\})}_{\mathcal{U}(d)} & & & \\ & \widehat{\mathcal{T}}^{(\{i,j\})}_{\mathcal{U}(d)} & & \\ & & \ddots & \\ & & & \widehat{\mathcal{T}}^{([t])}_{\mathcal{U}(d)} \end{bmatrix} \tag{B13}$$

where

$$\widehat{\mathcal{T}}^{(\Gamma)}_{\mathcal{U}(d)} = \frac{1}{d^t} \sum_{\substack{\sigma,\pi\in\mathcal{S}_t \\ \Gamma_\sigma=\Gamma_\pi=\Gamma \\ l=|\Gamma|}} \tau^{(l)}_{\mathcal{U}(d)}([\hat\sigma],[\hat\pi])\ |[\hat\sigma]\rangle\!\rangle\langle\!\langle[\hat\pi]|\ , \quad \textit{and} \quad \tau^{(l)}_{\mathcal{U}(d)}([\hat\sigma],[\hat\pi]) = \sum_{\substack{\sigma\subseteq\eta\in\mathcal{S}_l \\ \pi\subseteq\kappa\in\mathcal{S}_l}} \chi^{(l)}_d(\eta)\ W^{(l)-1}_d(\eta,\kappa)\ \chi^{(l)}_d(\kappa)\ , \tag{B14}$$

with orthogonal support-dependent block moment operators $\widehat{\mathcal{T}}^{(\Gamma)}_{\mathcal{U}(d)}$, spanned by localized permutations $[\hat\sigma],[\hat\pi]\in[\hat{\mathcal{S}}_t]$ with support $\Gamma=\Gamma_{[\hat\sigma]}=\Gamma_{[\hat\pi]}$, and exact identity and transposition coefficients of,

$$\tau^{(t)}_{\mathcal{U}(d)}([\hat\sigma],[\hat e]) = \delta_{\sigma e} \qquad , \qquad \tau^{(t)}_{\mathcal{U}(d)}([\hat\tau],[\hat\tau']) = \frac{1}{d^2-1}\delta_{\tau\tau'}\ . \tag{B15}$$

Theorem 7 (cHaar Moment Operators are Support-Non-Decreasing and Block-Lower-Triangular in Localized Basis)**.** *The cHaar moment operator is block-lower-triangular in the localized permutation basis, as per its support,*

$$\widehat{\mathcal{T}}^{(t)}_{\mathcal{C}(d,d_\mathcal{E})} = \begin{bmatrix} \widehat{\mathcal{T}}^{(\{\})}_{\mathcal{C}(d,d_\mathcal{E})} & & & & \\ \widehat{\mathcal{T}}^{(\{\},\{i,j\})}_{\mathcal{C}(d,d_\mathcal{E})} & \widehat{\mathcal{T}}^{(\{i,j\})}_{\mathcal{C}(d,d_\mathcal{E})} & & & \\ \widehat{\mathcal{T}}^{(\{\},\{i,j,k\})}_{\mathcal{C}(d,d_\mathcal{E})} & \cdots & \widehat{\mathcal{T}}^{(\{i,j,k\})}_{\mathcal{C}(d,d_\mathcal{E})} & & \\ \vdots & \ddots & \widehat{\mathcal{T}}^{(\Gamma,\Gamma')}_{\mathcal{C}(d,d_\mathcal{E})} & \ddots & \\ \widehat{\mathcal{T}}^{(\{\},[t])}_{\mathcal{C}(d,d_\mathcal{E})} & \cdots & \cdots & \cdots & \widehat{\mathcal{T}}^{([t])}_{\mathcal{C}(d,d_\mathcal{E})} \end{bmatrix} \tag{B16}$$

where

$$\widehat{\mathcal{T}}^{(\Gamma,\Gamma')}_{\mathcal{C}(d,d_\mathcal{E})} = \frac{1}{d^t} \sum_{\substack{\sigma,\pi\in\mathcal{S}_t \\ \Gamma_\sigma=\Gamma\supseteq\Gamma_\pi=\Gamma' \\ l=|\Gamma|\geq l'=|\Gamma'|}} \tau^{(l)}_{\mathcal{C}(d,d_\mathcal{E})}([\hat\sigma],[\hat\pi])\ |[\hat\sigma]\rangle\!\rangle\langle\!\langle[\hat\pi]|\ , \quad \textit{and} \quad \tau^{(l)}_{\mathcal{C}(d,d_\mathcal{E})}([\hat\sigma],[\hat\pi]) = \sum_{\substack{\sigma\subseteq\eta\in\mathcal{S}_l \\ \pi\subseteq\kappa\in\mathcal{S}_l}} \chi^{(l)}_{dd_\mathcal{E}}(\eta)\ W^{(l)-1}_{dd_\mathcal{E}}(\eta,\kappa)\ \chi^{(l)}_d(\kappa)\ , \tag{B17}$$

with orthogonal support-dependent block moment operators $\widehat{\mathcal{T}}^{(\Gamma,\Gamma')}_{\mathcal{C}(d,d_\mathcal{E})}$, spanned by localized permutations $[\hat\sigma],[\hat\pi]\in[\hat{\mathcal{S}}_t]$ with support $\Gamma_{[\hat\sigma]}=\Gamma\supseteq\Gamma'=\Gamma_{[\hat\pi]}$, and exact identity and transposition coefficients of,

$$\tau^{(t)}_{\mathcal{C}(d,d_\mathcal{E})}([\hat e],[\hat e]) = 1 \quad , \quad \tau^{(t)}_{\mathcal{C}(d,d_\mathcal{E})}([\hat\tau],[\hat e]) = \frac{d_\mathcal{E}-1}{d^2d_\mathcal{E}^2-1} \quad , \qquad \tau^{(t)}_{\mathcal{C}(d,d_\mathcal{E})}([\hat\tau],[\hat\tau']) = \frac{d_\mathcal{E}}{d^2d_\mathcal{E}^2-1}\delta_{\tau\tau'}\ . \tag{B18}$$

Proof. In what follows, the moment operators are expressed in terms of the orthogonal with respect to support localized permutation basis $[\hat{\mathcal{S}}_t]$. In particular, cHaar ensemble moment operators are generated by support-preserving Haar ensemble moment operators in the composite system and environment space,

$$\widehat{\mathcal{T}}^{(t)}_{\mathcal{U}(dd_\mathcal{E})} = \frac{1}{d^td^t_\mathcal{E}} \sum_{\sigma,\pi\in\mathcal{S}_t} \tau^{(t)}_{\mathcal{U}(dd_\mathcal{E})}(\sigma,\pi)\ |\sigma_\mathcal{H}\sigma_\mathcal{E}\rangle\!\rangle\langle\!\langle\pi_\mathcal{H}\pi_\mathcal{E}| \tag{B19}$$

$$= \frac{1}{d^td^t_\mathcal{E}} \sum_{\substack{\eta,\eta'\in\mathcal{S}_t \\ \kappa,\kappa'\in\mathcal{S}_t \\ \Gamma_\eta\cup\Gamma_{\eta'}=\Gamma_\kappa\cup\Gamma_{\kappa'} \\ \Gamma=\Gamma_\eta\cup\Gamma_{\eta'}\ ,\ l=|\Gamma|}} \sum_{\substack{\eta,\eta'\subseteq\sigma\in\mathcal{S}_l \\ \kappa,\kappa'\subseteq\pi\in\mathcal{S}_l}} \chi^{(l)}_{dd_\mathcal{E}}(\sigma)\ \tau^{(l)}_{\mathcal{U}(dd_\mathcal{E})}(\sigma,\pi)\ \chi^{(l)}_{dd_\mathcal{E}}(\pi)\ |[\hat\eta_\mathcal{H}][\hat\eta'_\mathcal{E}]\rangle\!\rangle\langle\!\langle[\hat\kappa_\mathcal{H}][\hat\kappa'_\mathcal{E}]|\ . \tag{B20}$$

In particular, we note that the support of the input permutations κ, κ' can be changed from $\Gamma_\eta \cup \Gamma_{\eta'} = \Gamma_\kappa \cup \Gamma_{\kappa'} \rightarrow \Gamma_\eta \cup \Gamma_{\eta'} \supseteq \Gamma_\kappa \cup \Gamma_{\kappa'}$, as any coefficients arising from summations over $\pi \supseteq \kappa, \kappa'$ will cancel to zero for $\Gamma_\kappa \cup \Gamma_{\kappa'} \subset \Gamma_\eta \cup \Gamma_{\eta'}$. Hence, we can write the Haar ensemble moment operator in the composite system and environment space as,

$$\widehat{\mathcal{T}}^{(t)}_{\mathcal{U}(dd_\mathcal{E})} = \frac{1}{d^t d^t_\mathcal{E}} \sum_{\substack{\eta,\eta' \in \mathcal{S}_t \\ \eta,\eta' \subseteq \sigma \in \mathcal{S}_l \\ \Gamma = \Gamma_\eta \cup \Gamma_{\eta'}\ ,\ l = |\Gamma|}} \ \sum_{\substack{\kappa,\kappa' \in \mathcal{S}_l \\ \kappa,\kappa' \subseteq \pi \in \mathcal{S}_l \\ ,\ \Gamma_\eta \cup \Gamma_{\eta'} \supseteq \Gamma_\kappa \cup \Gamma_{\kappa'}}} \chi^{(l)}_{dd_\mathcal{E}}(\sigma)\ \tau^{(l)}_{\mathcal{U}(dd_\mathcal{E})}(\sigma,\pi)\ \chi^{(l)}_{dd_\mathcal{E}}(\pi)\ |[\hat{\eta}_\mathcal{H}][\hat{\eta}'_\mathcal{E}]\rangle\!\rangle\langle\!\langle[\hat{\kappa}_\mathcal{H}][\hat{\kappa}'_\mathcal{E}]| \tag{B21}$$

$$= \frac{1}{d^t d^t_\mathcal{E}} \sum_{\substack{\eta,\eta' \in \mathcal{S}_t \\ \eta,\eta' \subseteq \sigma \in \mathcal{S}_l \\ \Gamma = \Gamma_\eta \cup \Gamma_{\eta'} = \Gamma_\sigma\ ,\ l = |\Gamma|}} \ \sum_{\substack{\pi \in \mathcal{S}_l \\ \Gamma_\pi \subseteq \Gamma}} \chi^{(l)}_{dd_\mathcal{E}}(\sigma)\ \tau^{(l)}_{\mathcal{U}(dd_\mathcal{E})}(\sigma,\pi)\ |[\hat{\eta}_\mathcal{H}][\hat{\eta}'_\mathcal{E}]\rangle\!\rangle\langle\!\langle \pi_\mathcal{H}\pi_\mathcal{E}|\ . \tag{B22}$$

From here, we can show that the cHaar ensemble moment is support-non-decreasing

$$\widehat{\mathcal{T}}^{(t)}_{\mathcal{C}(d,d_\mathcal{E})} = I_d^{\otimes t} \otimes \langle\!\langle I_{d_\mathcal{E}}^{\otimes t}|\ \ \widehat{\mathcal{T}}^{(t)}_{\mathcal{U}(dd_\mathcal{E})}\ \ I_d^{\otimes t} \otimes |\nu_\mathcal{E}^{\otimes t}\rangle\!\rangle \tag{B23}$$

$$= \frac{1}{d^t d^t_\mathcal{E}} \sum_{\substack{\eta,\eta' \in \mathcal{S}_t \\ \eta,\eta' \subseteq \sigma \in \mathcal{S}_l \\ \Gamma = \Gamma_\eta \cup \Gamma_{\eta'}\ ,\ l = |\Gamma|}} \ \sum_{\substack{\pi \in \mathcal{S}_l \\ \Gamma_\pi \subseteq \Gamma}} \chi^{(l)}_{dd_\mathcal{E}}(\sigma)\ \tau^{(l)}_{\mathcal{U}(dd_\mathcal{E})}(\sigma,\pi)\ \langle\!\langle I_{d_\mathcal{E}}^{\otimes t}|[\hat{\eta}'_\mathcal{E}]\rangle\!\rangle\ \langle\!\langle \pi_\mathcal{E}|\nu_\mathcal{E}^{\otimes t}\rangle\!\rangle\ |[\hat{\eta}_\mathcal{H}]\rangle\!\rangle\langle\!\langle \pi_\mathcal{H}| \tag{B24}$$

$$= \frac{1}{d^t} \sum_{\substack{\eta \in \mathcal{S}_t \\ \eta \subseteq \sigma \in \mathcal{S}_l \\ \Gamma = \Gamma_\eta\ ,\ l = |\Gamma|}} \ \sum_{\substack{\pi \in \mathcal{S}_l \\ \Gamma_\pi \subseteq \Gamma}} \chi^{(l)}_{dd_\mathcal{E}}(\sigma)\ \tau^{(l)}_{\mathcal{U}(dd_\mathcal{E})}(\sigma,\pi)\ |[\hat{\eta}_\mathcal{H}]\rangle\!\rangle\langle\!\langle \pi_\mathcal{H}| \tag{B25}$$

$$= \frac{1}{d^t} \sum_{\substack{\eta \in \mathcal{S}_t \\ \eta \subseteq \sigma \in \mathcal{S}_l \\ \Gamma = \Gamma_\eta\ ,\ l = |\Gamma|}} \ \sum_{\substack{\pi \in \mathcal{S}_l \\ \kappa \subseteq \pi \in \mathcal{S}_l \\ \Gamma_\eta \subseteq \Gamma}} \chi^{(l)}_{dd_\mathcal{E}}(\sigma)\ \tau^{(l)}_{\mathcal{U}(dd_\mathcal{E})}(\sigma,\pi)\ \chi^{(l)}_{d}(\pi)\ |[\hat{\eta}_\mathcal{H}]\rangle\!\rangle\langle\!\langle [\hat{\kappa}_\mathcal{H}]| \tag{B26}$$

$$= \frac{1}{d^t} \sum_{\substack{\sigma,\pi \in \mathcal{S}_t \\ \Gamma = \Gamma_\sigma \supseteq \Gamma_\pi\ ,\ l = |\Gamma|}} \tau^{(l)}_{\mathcal{C}(d,d_\mathcal{E})}([\hat{\sigma}],[\hat{\pi}])\ |[\hat{\sigma}]\rangle\!\rangle\langle\!\langle [\hat{\pi}]| \quad : \quad \tau^{(l)}_{\mathcal{C}(d,d_\mathcal{E})}([\hat{\sigma}],[\hat{\pi}]) = \sum_{\substack{\sigma \subseteq \eta \in \mathcal{S}_l \\ \pi \subseteq \kappa \in \mathcal{S}_l}} \chi^{(l)}_{dd_\mathcal{E}}(\eta)\ \tau^{(l)}_{\mathcal{U}(dd_\mathcal{E})}(\eta,\kappa)\ \chi^{(l)}_{d}(\kappa)\ . \tag{B27}$$

□

To illustrate these properties, we compute the exact $t = 3, 4, 5$-order Haar and cHaar transfer matrices, for $d_\mathcal{E} = d^2$ [91], as shown in Fig. 9. Here, we can see that a sparse, support-dependent block-diagonal or block-lower-triangular structure occurs. Within the transfer matrices, many of the $\binom{t}{l}$ locality-l blocks, of size $!l$, are identical, and correspond to $l \leq t$-order twirl blocks, due to isomorphisms between permutations. Such structure is primarily attributed to the localized permutation basis being orthogonal with respect to support. For the Haar ensemble, the Haar invariance enforces the support preserving behavior, and for the cHaar ensemble, the environment space partial trace enforces the support increasing behavior. From these results, and inspired by similar discussions in [109], we provide intriguing identities for permutations $\sigma, \pi \in \mathcal{S}_t$ of support $\Gamma = \Gamma_\sigma \supseteq \Gamma_\pi$ and locality $l = |\Gamma|$, and general dimensions $d, d_\mathcal{E}$,

$$\sum_{\substack{\eta,\kappa \in \mathcal{S}_t \\ \eta \supseteq \sigma,\ \kappa \supseteq \pi}} \chi^{(t)}_{dd_\mathcal{E}}(\eta)\ W^{(t)-1}_{dd_\mathcal{E}}(\eta,\kappa)\ \chi^{(t)}_{d}(\kappa) \ \ = \sum_{\substack{\eta,\kappa \in \mathcal{S}_l \\ \eta \supseteq \sigma,\ \kappa \supseteq \pi}} \chi^{(l)}_{dd_\mathcal{E}}(\eta)\ W^{(l)-1}_{dd_\mathcal{E}}(\eta,\kappa)\ \chi^{(l)}_{d}(\kappa)\ \left(\delta_{\Gamma_\sigma \supseteq \Gamma_\pi}\delta_{d_\mathcal{E}>1} + \delta_{\Gamma_\sigma = \Gamma_\pi}\delta_{d_\mathcal{E}=1}\right) . \tag{B28}$$

Next, we turn our attention to computing the exact $t = 1, 2$-order moment operator for the k-concatenated cHaar ensemble. In particular, $\widehat{\mathcal{T}}^{(t)}_{\mathcal{C}(d,d_\mathcal{E})}$ is not a projector (it is not obtained as the average of a group), then $\widehat{\mathcal{T}}^{(t)k}_{\mathcal{C}(d,d_\mathcal{E})} \neq \widehat{\mathcal{T}}^{(t)}_{\mathcal{C}(d,d_\mathcal{E})}$, and we are interested in its convergence properties with k. A straightforward calculation leads to the following lemma.

Lemma 2 (Exact $t = 1, 2$-order cHaar Moment Operators)**.** *k-concatenated $t = 1, 2$-order cHaar moment operators, for all dimensions $d, d_\mathcal{E}$, are,*

$$\widehat{\mathcal{T}}^{(1)k}_{\mathcal{C}(d,d_\mathcal{E})} = \frac{1}{d}|[\hat{e}]\rangle\!\rangle\langle\!\langle[\hat{e}]|$$

$$\widehat{\mathcal{T}}^{(2)k}_{\mathcal{C}(d,d_\mathcal{E})} = \frac{1}{d^2}|[\hat{e}]\rangle\!\rangle\langle\!\langle[\hat{e}]| + \frac{1}{d^2}\frac{d_\mathcal{E}-1}{d^2 d_\mathcal{E}^2 - 1}\left[1 + \sum_{s>0}^{k-1}\left[d_\mathcal{E}\frac{d^2-1}{d^2 d_\mathcal{E}^2-1}\right]^s\right]|[\hat{\tau}]\rangle\!\rangle\langle\!\langle[\hat{e}]| + \frac{1}{d^2}\frac{1}{d^2-1}\left[d_\mathcal{E}\frac{d^2-1}{d^2 d_\mathcal{E}^2-1}\right]^k|[\hat{\tau}]\rangle\!\rangle\langle\!\langle[\hat{\tau}]|\ , \tag{B29}$$

where $[\hat{e}] = I_d^{\otimes t}$ is the identity operator, and $[\hat{\tau}] = dS_d - I_d^{\otimes 2} = \sum_{P \in \mathcal{P}_d \setminus \{I_d\}} P \otimes P^\dagger$ is the localized transposition operator.

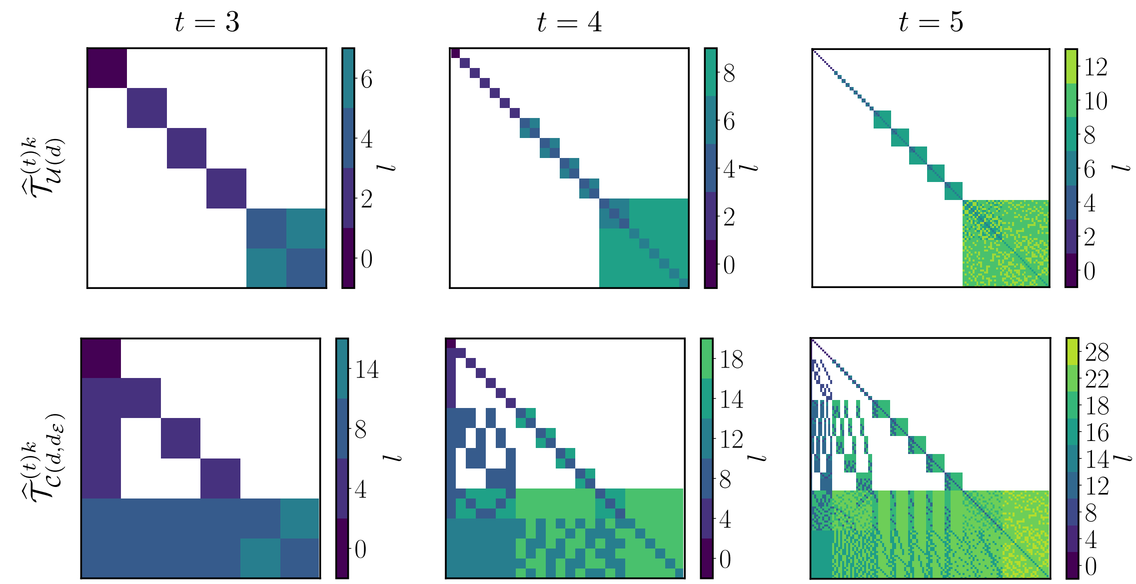

Figure 9. **Moment operators in a basis that block diagonalizes the Haar moment operator as a permutation transfer matrix.** Scaling of the t-th order moment operator transfer matrix elements for the Haar (top) and cHaar (bottom) ensembles in the localized permutation basis, for $t = 3, 4, 5$ and $d_\mathcal{E} = d^2$. Depicted are matrix elements corresponding to all pairs of $t!$ permutations $\sigma, \pi \in \mathcal{S}_t$, sorted by their support, from the smallest identity (top-left corner), to largest cycles (bottom-right corner), with leading-order scaling l (colored gradient), such that, $\tau^{(t)}_{\mathcal{U}(d)}([\hat\sigma],[\hat\pi]), \tau^{(t)}_{\mathcal{C}(d,d_\mathcal{E})}([\hat\sigma],[\hat\pi]) \sim \mathcal{O}(1/d^l)$. A sparse, support-dependent block-diagonal (Haar) or block-lower-triangular (cHaar) structure of coefficients occurs, and transposition coefficients (light-purple top-left diagonal) are uniquely diagonal $1/(d^2-1)$ (Haar), or diagonal and non-unital $(d_\mathcal{E} - \delta_{\pi e})/(d^2 d_\mathcal{E}^2 - 1)$ (cHaar).

From the previous results for individual transfer matrix elements, and overlaps between basis operators, we can bound the transfer matrix elements for general t-th order moment operators for the k-concatenated cHaar ensemble.

Theorem 8 (Asymptotic t-th order cHaar Moment Operators)**.** *k-concatenated t-th order cHaar moment operators, in the asymptotic limit $dd_\mathcal{E} \to \infty$, have transfer matrix elements upper bounded by,*

$$\left|\tau^{(l,k)}_{\mathcal{C}(d,d_\mathcal{E})}([\hat\sigma],[\hat\pi])\right| \;\leq\; \frac{1}{d^{|\sigma|+|\pi|}\, d_\mathcal{E}^{|\sigma|+(k-1)|\Gamma_\pi|/2}}\; C^{(l,k)}_\tau \;\delta_{\Gamma_\sigma \supseteq \Gamma_\pi}\;, \tag{B30}$$

for $\sigma, \pi \in \mathcal{S}_l$, $l \in [t+1]$, given t,k-dependent constants $C^{(l,k)}_\tau = \frac{1}{l!^2\; c_l^4}\; C^{(l)k}_\lambda$, and $C^{(l)}_\lambda = l!^2\; \chi_l\; c_l^7$.

Proof. By definition, for $\sigma, \pi \in \mathcal{S}_l$ we have bounds for the individual transfer matrix elements and overlaps of,

$$\tau^{(l)}_{\mathcal{C}(d,d_\mathcal{E})}([\hat\sigma],[\hat\pi]) \;=\; \sum_{\sigma\subseteq\eta\;,\;\pi\subseteq\kappa} \chi^{(l)}_{dd_\mathcal{E}}(\eta)\; W^{(l)-1}_{dd_\mathcal{E}}(\eta,\kappa)\; \chi^{(l)}_{d}(\kappa)\; \delta_{\Gamma_\sigma\supseteq\Gamma_\pi} \;\;\leq\;\; \frac{1}{d^{|\sigma|+|\pi|} d_\mathcal{E}^{|\sigma|}}\; C^{(l)}_\tau\; \delta_{\Gamma_\sigma\supseteq\Gamma_\pi} \tag{B31}$$

$$\chi^{(l)}_{d}([\hat\sigma],[\hat\pi]) \;=\; \sum_{\eta\subseteq\sigma\;,\;\kappa\subseteq\pi} c(\eta^{-1}\sigma)\; c(\kappa^{-1}\pi)\; \frac{\chi^{(l)}_{d}(\eta,\kappa)}{\chi^{(l)}_{d}(\eta)\chi^{(l)}_{d}(\kappa)}\; \delta_{\Gamma_\sigma\Gamma_\pi} \;\;\leq\;\; d^{|\sigma|+|\pi|}\; C^{(l)}_\chi\; \delta_{\Gamma_\sigma\Gamma_\pi} \tag{B32}$$

$$C^{(l)}_\tau \;=\; \chi_l\; c_l^3 \qquad , \qquad C^{(l)}_\chi \;=\; c_\sigma^2\; c_\pi^2\;\;, \tag{B33}$$

up to l-dependent constants $\chi_l > 0$ [109]. We can also bound the Catalan, Möbius, and Weingarten functions as,

$$c(\sigma) \;\leq\; |c(\sigma)| \;=\; \prod_{\lambda\in\sigma} c_{|\lambda|} \;\leq\; c_\sigma \;=\; \prod_{\lambda\in\sigma} c_{|\lambda|+1} \;\leq\; c_l \;\leq\; e^{2l} \quad:\quad \binom{t}{l} \leq e^{2l}\left(\frac{t}{2l}\right)^l \tag{B34}$$

$$W^{(l)-1}_{d}(\sigma) \;\leq\; \chi_l\; c(\sigma)\; \frac{1}{d^{|\sigma|}} \;\leq\; \chi_l\; c_\sigma\; \frac{1}{d^{|\sigma|}} \tag{B35}$$

$$|\sigma| + |\pi| \pm |\sigma^{-1}\pi| \;\gtreqless\; |\sigma| + |\pi| \;\;,\;\; |\sigma| + |\sigma^{-1}\pi| \;\geq\; |\sigma| \;\;,\;\; \lceil l/2 \rceil \;\leq\; |\sigma| \;\leq\; l-1 \tag{B36}$$

$$|\{\sigma \supseteq \pi \in \mathcal{S}_l\}| \;=\; c_\sigma \;\leq\; c_l \quad,\quad |\{\sigma \subseteq \pi \in \mathcal{S}_l\}| \;\leq\; c_{l-|\sigma|} \;\leq\; c_l\;. \tag{B37}$$

Then, k-concatenations of $l \leq t$-order cHaar moment operators have transfer matrix elements for $\sigma, \pi \in \mathcal{S}_l$ of the form,

$$\tau^{(l,k)}_{\mathcal{C}(d,d_\mathcal{E})}([\hat{\sigma}],[\hat{\pi}]) \;=\; \sum_{\substack{\{\sigma_i,\pi_i\in\mathcal{S}_{l_i}\}_{i\in[k]}\\ \Gamma_{\pi_{i-1}}=\Gamma_{\pi_i}\subseteq\Gamma_{\sigma_i}=\Gamma_{\sigma_{i+1}}\\ \sigma_{k-1}=\sigma\;,\;\pi_0=\pi\;,\;l_i=|\Gamma_{\sigma_i}|}} \prod_{i\in[k]} \tau^{(l_i)}_{\mathcal{C}(d,d_\mathcal{E})}([\hat{\sigma}_i],[\hat{\pi}_i]) \prod_{i\in[k-1]} W^{(l_i)-1}_{dd_\mathcal{E}}([\hat{\pi}_{i+1}],[\hat{\sigma}_i]) \;. \tag{B38}$$

Given the support constraints $\{\} \subseteq \Gamma_\pi \subseteq \cdots \subseteq \Gamma_{\pi_i} \subseteq \Gamma_{\sigma_i} \subseteq \cdots \subseteq \Gamma_\sigma \subseteq [l]$ of the propagated permutations, the telescoping of the system dimension scalings cancelling from identical scaling of the transfer matrices $\sim d^{|\sigma_i|+|\pi_i|}$ and overlaps $\sim 1/d^{|\sigma_i|+|\pi_{i+1}|}$, and the environment dimension scalings $\sim 1/d_\mathcal{E}^{|\sigma_i|}$ being bounded by the at least $|\Gamma_\pi|$-locality minimal permutations $\{\sigma_i \;:\; \Gamma_{\sigma_i} = \Gamma_{\pi_i} = \Gamma_\pi \;,\; |\sigma_i| \geq |\Gamma_\pi|/2 \geq |\Gamma_\pi|/2\}_{i\in[k-1]}$, propagating through until the $k-1$-th moment operator, the k-concatenated $l \leq t$-order transfer matrix elements are bounded by,

$$\left|\tau^{(l,k)}_{\mathcal{C}(d,d_\mathcal{E})}([\hat{\sigma}],[\hat{\pi}])\right| \;\leq\; \frac{1}{d^{|\sigma|+|\pi|}d_\mathcal{E}^{|\sigma|+(k-1)|\Gamma_\pi|/2}}\; C^{(l,k)}_\tau \;\delta_{\Gamma_\sigma\supseteq\Gamma_\pi} \tag{B39}$$

$$C^{(l,k)}_\tau \;=\; l!^{2(k-1)}\;\chi_l^k\;c_l^{3k+4(k-1)} \;=\; \frac{1}{l!^2\;c_l^4}\;\left(l!^2\;\chi_l\;c_l^7\right)^k \;=\; \frac{1}{l!^2\;c_l^4}\;C^{(l)k}_\lambda \qquad , \qquad C^{(l)}_\lambda \;:=\; l!^2\;\chi_l\;c_l^7 \;. \tag{B40}$$

Finally, we note that the simplest non-trivial transpositions are orthogonal, with exact transfer matrices and overlaps of,

$$\tau^{(l,k)}_{\mathcal{C}(d,d_\mathcal{E})}([\hat{\tau}],[\hat{e}]) = \frac{d_\mathcal{E}-1}{d^2d_\mathcal{E}^2-1}\left[1+\sum_{s>0}^{k-1}\left[d_\mathcal{E}\frac{d^2-1}{d^2d_\mathcal{E}^2-1}\right]^s\right] \qquad , \qquad \tau^{(l,k)}_{\mathcal{C}(d,d_\mathcal{E})}([\hat{\tau}],[\hat{\tau}']) = \frac{1}{d^2-1}\left[d_\mathcal{E}\frac{d^2-1}{d^2d_\mathcal{E}^2-1}\right]^k \;\delta_{\tau\tau'} \tag{B41}$$

$$\chi^{(l)}_d([\hat{\tau}],[\hat{\tau}']) = \;\left(d^2-1\right)\;\delta_{\tau\tau'} \;. \tag{B42}$$

□

3. Spectral Properties and Hierarchies of Moment Operators

In this section we derive spectral properties for the moment operators of the reference ensembles. These will allow us to derive a hierarchy between the Haar, cHaar, and Depolarize ensembles via their moment operator norms.

First, we investigate the leading eigenvectors of the cHaar moment operator.

Theorem 9 (Leading Eigenvectors of cHaar Moment Operators)**.** *The t-th order k-concatenated cHaar moment operator has a unit leading eigenvalue $\lambda = 1$, a leading right-eigenvector $|\psi\rangle\rangle$ of a uniform combination of permutations, weighted by their characters, and a leading left-eigenvector $|\varphi\rangle\rangle$ reflecting trace-preservation. That is, the cHaar moment operator*

$$\widehat{\mathcal{T}}^{(t)k}_{\mathcal{C}(d,d_\mathcal{E})} = \frac{|\psi\rangle\rangle\langle\langle\varphi|}{\langle\langle\varphi|\psi\rangle\rangle} \;+\;\cdots\;, \qquad : \qquad |\psi\rangle\rangle = \sum_{\sigma\in\mathcal{S}_t}\chi^{(t)}_{dd_\mathcal{E}}(\sigma)\;|\sigma\rangle\rangle \qquad , \qquad |\varphi\rangle\rangle = |e\rangle\rangle \tag{B43}$$

has a leading-order, concatenation-invariant, non-unital, and trace preserving contribution.

Proof. To confirm that such vectors $|\psi\rangle\rangle, |\varphi\rangle\rangle$ are eigenvectors of the cHaar moment operator, our goal is to show that,

$$\widehat{\mathcal{T}}^{(t)}_{\mathcal{C}(d,d_\mathcal{E})}|\psi\rangle\rangle = |\psi\rangle\rangle \qquad \text{and} \qquad \langle\langle\varphi|\widehat{\mathcal{T}}^{(t)}_{\mathcal{C}(d,d_\mathcal{E})} = \langle\langle\varphi| \;. \tag{B44}$$

Given the orthogonality of the characters and Weingarten functions, the invariance of the characters under inverses of permutations, $\chi^{(t)}_d(\pi,\varsigma) = \chi^{(t)}_d(\pi^{-1}\varsigma) = \chi^{(t)}_d(\varsigma^{-1}\pi)$, and letting $\varsigma \to \pi^{-1}\varsigma$, we find,

$$\begin{aligned}\widehat{\mathcal{T}}^{(t)}_{\mathcal{C}(d,d_\mathcal{E})}|\psi\rangle\rangle &= \sum_{\sigma,\pi,\varsigma\in\mathcal{S}_t}\chi^{(t)}_{d_\mathcal{E}}(\sigma)\;W^{(t)-1}_{dd_\mathcal{E}}(\sigma,\pi)\;\chi^{(t)}_d(\pi,\varsigma)\;\chi^{(t)}_{dd_\mathcal{E}}(\varsigma)\;|\sigma\rangle\rangle = \sum_{\sigma,\pi,\varsigma\in\mathcal{S}_t}\chi^{(t)}_{d_\mathcal{E}}(\sigma)\;W^{(t)-1}_{dd_\mathcal{E}}(\sigma,\pi)\;\chi^{(t)}_{dd_\mathcal{E}}(\pi,\varsigma^{-1})\;\chi^{(t)}_d(\varsigma)\;|\sigma\rangle\rangle = |\psi\rangle\rangle\\ \langle\langle\varphi|\widehat{\mathcal{T}}^{(t)}_{\mathcal{C}(d,d_\mathcal{E})} &= \sum_{\sigma,\pi\in\mathcal{S}_t}\chi^{(t)}_d(\sigma)\;\chi^{(t)}_{d_\mathcal{E}}(\sigma)\;W^{(t)-1}_{dd_\mathcal{E}}(\sigma,\pi)\;\langle\langle\pi| = \sum_{\sigma,\pi\in\mathcal{S}_t}\chi^{(t)}_{dd_\mathcal{E}}(\sigma)\;W^{(t)-1}_{dd_\mathcal{E}}(\sigma,\pi)\;\langle\langle\pi| = \sum_{\pi\in\mathcal{S}_t}\delta_{\pi e}\;\langle\langle\pi| = \langle\langle\varphi| \;.\end{aligned} \tag{B45}$$

Given the Jucys-Murphy character sums [125], $\sum_{\sigma\in\mathcal{S}_t}\chi^{(t)}_d(\sigma) = \binom{d+t-1}{t}\frac{t!}{d^t}$, the eigenvectors have overlaps of,

$$\chi^{(t)}_d(v) = \frac{1}{d^t}\langle\langle e|\psi\rangle\rangle = \sum_{\sigma\in\mathcal{S}_t}\chi^{(t)}_{d^2d_\mathcal{E}}(\sigma) = \binom{d^2d_\mathcal{E}+t-1}{t}\frac{t!}{d^{2t}d_\mathcal{E}^t} = 1+\binom{t}{2}\frac{1}{d^2d_\mathcal{E}}+\mathcal{O}\left(\frac{1}{d^4d_\mathcal{E}^2}\right) \;. \tag{B46}$$

□

In fact, we can bound all of the eigenvalues, the norm, and the trace of the Haar, cHaar, and Depolarize ensembles, by making use of the following lemma.

Lemma 3 (Spectrum with respect to Localized Permutation Basis). *The spectrum of the cHaar moment operator in the linearly-independent, but not strictly orthogonal localized permutation basis is,*

$$\text{Spectrum of } \widehat{\mathcal{T}}^{(t)k}_{\mathcal{C}(d,d_\mathcal{E})} = \text{Spectrum of } \left[\tilde{\tau}^{(t,k)}_{\mathcal{C}(d,d_\mathcal{E})}(\sigma,\pi)\right]_{\sigma,\pi\in\mathcal{S}_t} \quad : \quad \begin{aligned} \tilde{\tau}^{(t,k)}_{\mathcal{C}(d,d_\mathcal{E})}(\sigma,\pi) &= \sum_{\varsigma\in\mathcal{S}_t} \tau^{(t,k)}_{\mathcal{C}(d,d_\mathcal{E})}([\hat{\sigma}],[\hat{\varsigma}])\ \chi^{(t)}_d([\hat{\varsigma}],[\hat{\pi}]) \\ \tilde{\tau}^{(t,k)\prime}_{\mathcal{C}(d,d_\mathcal{E})}(\sigma,\pi) &= \sum_{\varsigma\in\mathcal{S}_t} \chi^{(t)}_d([\hat{\sigma}],[\hat{\varsigma}])\ \tau^{(t,k)}_{\mathcal{C}(d,d_\mathcal{E})}([\hat{\varsigma}],[\hat{\pi}]) \end{aligned} \ , \tag{B47}$$

expressed as the spectrum of the modified transfer matrix $\tilde{\tau}^{(t,k)}_{\mathcal{C}(d,d_\mathcal{E})}$ in the localized permutation basis, with bounded elements,

$$\tilde{\tau}^{(l,k)}_{\mathcal{C}(d,d_\mathcal{E})}(\sigma,\pi) \ \le\ \frac{1}{d^{|\sigma|-|\pi|}}\frac{1}{d_\mathcal{E}^{|\sigma|+(k-1)l'/2}}\ C^{(l,l',k)}_{\tilde{\tau}}\ \delta_{\Gamma_\sigma\supseteq\Gamma_\pi} \tag{B48}$$

$$\tilde{\tau}^{(l,k)\prime}_{\mathcal{C}(d,d_\mathcal{E})}(\sigma,\pi) \ \le\ d^{|\sigma|-|\pi|}\frac{1}{d_\mathcal{E}^{l/2+(k-1)l'/2}}\ C^{(l,l',k)}_{\tilde{\tau}'}\ \delta_{\Gamma_\sigma\supseteq\Gamma_\pi}\ , \tag{B49}$$

given l,k-dependent constants $C^{(l,l',k)}_{\tilde{\tau}} = l!^{2(k-1)}\ l'!\ \chi_l^k\ c_l^{3k+4(k-1)}\ c_{l'}^4$ and $C^{(l,l',k)}_{\tilde{\tau}'} = l!^{2(k-1)}\ l!\ \chi_l^k\ c_l^{3k+4(k-1)}\ c_l^4$.

Proof. Let $\widehat{\mathcal{T}} = \sum_{\alpha,\beta\in\mathcal{S}}\tau(\alpha,\beta)|\alpha\rangle\!\rangle\langle\!\langle\beta|$ be a matrix, expressed in terms of a linearly independent basis $\mathcal{S}\ni\alpha,\beta$, with overlaps $\chi(\alpha,\beta)=\langle\!\langle\alpha|\beta\rangle\!\rangle$. Its eigenvalue equation, given an eigenvector $|\nu\rangle\!\rangle=\sum_{\alpha\in\mathcal{S}}\nu(\alpha)\ |\alpha\rangle\!\rangle$, and eigenvalue λ, is,

$$\widehat{\mathcal{T}}|\nu\rangle\!\rangle = \lambda\ |\nu\rangle\!\rangle \quad\to\quad \sum_{\xi,\beta\in\mathcal{S}}\tau(\alpha,\xi)\ \chi(\xi,\beta)\ \nu(\beta) = \lambda\ \nu(\alpha)\quad \forall\ \alpha \quad\to\quad \tilde{\tau}\ \nu = \lambda\ \nu \quad : \quad \tilde{\tau}(\alpha,\beta) = \sum_{\xi\in\mathcal{S}}\tau(\alpha,\xi)\ \chi(\xi,\beta)\ . \tag{B50}$$

Therefore the eigenvectors and eigenvalues of $\widehat{\mathcal{T}}$ are equivalent to the eigenvectors of components $\{\nu(\alpha)\}_{\alpha\in\mathcal{S}}$ and eigenvalues λ of the modified matrix $\tilde{\tau}$, with respect to the basis $\mathcal{S}$.

To bound the modified cHaar transfer matrix elements, we rely on the block-lower-triangular structure of the cHaar moment operator in this localized basis, The cHaar moment operator can be partitioned into blocks labeled by supports Γ,Γ', with associated localities $l=|\Gamma|, l'=|\Gamma'|$, and spanned by permutations $\sigma\in\mathcal{S}_l$ and $\pi\in\mathcal{S}_{l'}$ such that $\Gamma_\sigma=\Gamma\supseteq\Gamma'=\Gamma_\pi$. We also recall that the transfer matrix and overlaps for these blocks have bounds of,

$$\left|\tau^{(l,k)}_{\mathcal{C}(d,d_\mathcal{E})}([\hat{\sigma}],[\hat{\pi}])\right| \ \le\ \frac{1}{d^{|\sigma|+|\pi|}d_\mathcal{E}^{|\sigma|+(k-1)l/2}}\ C^{(l,k)}_\tau\ \delta_{\Gamma_\sigma\supseteq\Gamma_\pi} \quad , \quad \left|\chi^{(l)}_d([\hat{\sigma}],[\hat{\pi}])\right| \ \le\ d^{|\sigma|+|\pi|}\ C^{(l)}_\chi\ \delta_{\Gamma_\sigma\Gamma_\pi}, \tag{B51}$$

$$C^{(l,k)}_\tau \ =\ l!^{2(k-1)}\ \chi_l^k\ c_l^{3k+4(k-1)} \quad , \quad C^{(l)}_\chi \ =\ c_l^4\ . \tag{B52}$$

The modified transfer matrix elements, as sums over permutations $\varsigma\in\mathcal{S}_l$ can thus be bounded by,

$$\tilde{\tau}^{(l,k)}_{\mathcal{C}(d,d_\mathcal{E})}(\sigma,\pi) = \sum_{\substack{\varsigma\in\mathcal{S}_l \\ \Gamma_\sigma=\Gamma\supseteq\Gamma'=\Gamma_\varsigma=\Gamma_\pi}} \tau^{(l,k)}_{\mathcal{C}(d,d_\mathcal{E})}([\hat{\sigma}],[\hat{\varsigma}])\ \chi^{(l')}_d([\hat{\varsigma}],[\hat{\pi}]) \ \le\ \max_{\substack{\sigma,\varsigma,\pi\in\mathcal{S}_l \\ \Gamma_\sigma=\Gamma\supseteq\Gamma'=\Gamma_\varsigma=\Gamma_\pi}} !l'\ \left|\tau^{(l,k)}_{\mathcal{C}(d,d_\mathcal{E})}([\hat{\sigma}],[\hat{\varsigma}])\right|\ \left|\chi^{(l')}_d([\hat{\varsigma}],[\hat{\pi}])\right| \tag{B53}$$

$$\le \frac{1}{d^{|\sigma|-|\pi|}}\frac{1}{d_\mathcal{E}^{|\sigma|+(k-1)l'/2}}\ C^{(l,l',k)}_{\tilde{\tau}}\ \delta_{\Gamma_\sigma\supseteq\Gamma_\pi} \ \le\ C^{(l,l',k)}_{\tilde{\tau}}\ \delta_{\Gamma_\sigma\supseteq\Gamma_\pi} \begin{cases} \left(\frac{1}{d d_\mathcal{E}}\right)^{l/2} & {\scriptstyle l>0 \atop l'=0} \\ \frac{1}{d_\mathcal{E}^{k+(k+1)(l-l')/2}}\left(\frac{d}{d_\mathcal{E}^k}\right)^{((l'-2)-(l-l'))/2} & {\scriptstyle l>0 \atop l'>0} \end{cases}$$

$$C^{(l,l',k)}_{\tilde{\tau}} \ = l'!\ C^{(l,k)}_\tau\ C^{(l')}_\chi = l!^{2(k-1)}\ l'!\ \chi_l^k\ c_l^{3k+4(k-1)}\ c_{l'}^4\ , \tag{B54}$$

$$\tilde{\tau}^{(l,k)\prime}_{\mathcal{C}(d,d_\mathcal{E})}(\sigma,\pi) = \sum_{\substack{\varsigma\in\mathcal{S}_l \\ \Gamma_\sigma=\Gamma_\varsigma=\Gamma\supseteq\Gamma'=\Gamma_\pi}} \chi^{(l)}_d([\hat{\sigma}],[\hat{\varsigma}])\ \tau^{(l,k)}_{\mathcal{C}(d,d_\mathcal{E})}([\hat{\varsigma}],[\hat{\pi}]) \ \le\ \max_{\substack{\sigma,\varsigma,\pi\in\mathcal{S}_l \\ \Gamma_\sigma=\Gamma_\varsigma=\Gamma\supseteq\Gamma'=\Gamma_\pi}} !l\ \left|\chi^{(l)}_d([\hat{\sigma}],[\hat{\varsigma}])\right|\ \left|\tau^{(l,k)}_{\mathcal{C}(d,d_\mathcal{E})}([\hat{\varsigma}],[\hat{\pi}])\right| \tag{B55}$$

$$\le d^{|\sigma|-|\pi|}\frac{1}{d_\mathcal{E}^{l/2+(k-1)l'/2}}\ C^{(l,l',k)}_{\tilde{\tau}'}\ \delta_{\Gamma_\sigma\supseteq\Gamma_\pi} \ \le\ C^{(l,l',k)}_{\tilde{\tau}'}\ \delta_{\Gamma_\sigma\supseteq\Gamma_\pi} \begin{cases} d_\mathcal{E}^{l/2-1}\left(\frac{d}{d_\mathcal{E}}\right)^{l-1} & {\scriptstyle l>0 \atop l'=0} \\ \frac{1}{d_\mathcal{E}^{k+(1-2k)(l-l')/2}}\left(\frac{d}{d_\mathcal{E}^k}\right)^{((l-2)+(l-l'))/2} & {\scriptstyle l>0 \atop l'>0} \end{cases}$$

$$C^{(l,l',k)}_{\tilde{\tau}'} \ = l!\ C^{(l,k)}_\tau\ C^{(l)}_\chi = l!^{2(k-1)}\ l!\ \chi_l^k\ c_l^{3k+4(k-1)}\ c_l^4\ , \tag{B56}$$

where the final bounds follow from considering scaling separately for each case of $l'=0, l'>0$, and from the fact that $l>0$-locality permutations $\sigma\in\mathcal{S}_l$ have sizes bounded by $l/2\le\lceil l/2\rceil\le|\sigma|\le l-1$. $\square$

We now will make use of the derived bounds for the modified k-concatenated, t-order cHaar moment operator transfer matrix elements to bound its spectra, norm, and trace, and assess its overall spectral properties.

Theorem 10 (Spectrum of cHaar Moment Operator)**.** *The t-th order cHaar moment operator has a single leading eigenvalue $\lambda = 1$, and in the asymptotic limit $dd_{\mathcal{E}} \to \infty$, for fixed t, all non-leading eigenvalues are upper bounded by,*

$$0 \;\leq\; |\lambda| \;\leq\; \frac{1}{d_{\mathcal{E}}^{l/2}}\, C_{\lambda}^{(l)} \;<\; 1 \;, \tag{B57}$$

and $|\lambda| < 1$ for sufficiently large $d_{\mathcal{E}}$, given t-dependent constants $C_{\lambda}^{(l)} = l!^2\ \chi_l\ c_l^7$, and some $2 \leq l \leq t$.

Proof. Here, we use Gershgorin's circle theorem [126], in the non-orthogonal, block matrix case. Eigenvalues λ of the Γ-support, $l \geq 2$-locality, $!l$-sized diagonal blocks of the cHaar moment operator, satisfy

$$|\lambda| \;=\; \left|\lambda^k\right|^{\frac{1}{k}} \;\leq\; \left[!l \max_{\substack{\sigma,\pi\in\mathcal{S}_l\\ \Gamma_\sigma=\Gamma_\pi=\Gamma}} \left|\tilde{\tau}^{(l,k)}_{\mathcal{C}(d,d_{\mathcal{E}})}(\sigma,\pi)\right|\right]^{\frac{1}{k}} \;\leq\; \left[\frac{1}{d_{\mathcal{E}}^k}\left(\frac{d}{d_{\mathcal{E}}^k}\right)^{\frac{l-2}{2}} C_{\lambda}^{(l,k)}\right]^{\frac{1}{k}} \;\overset{k\to\infty}{\leq}\; \frac{1}{d_{\mathcal{E}}^{l/2}}\, C_{\lambda}^{(l)} \tag{B58}$$

$$C_{\lambda}^{(l,k)} \;=\; l!\; C_{\tilde{\tau}}^{(l,l,k)} \;=\; l!^{2k}\ \chi_l^k\ c_l^{7k} \;=\; C_{\lambda}^{(l)k} \qquad , \qquad C_{\lambda}^{(l)} \;=\; l!^2\ \chi_l\ c_l^7 \;, \tag{B59}$$

given eigenvalues of k-concatenations of a matrix are equal to k powers of the eigenvalues, given the bounds on the modified k-concatenated transfer matrix elements $\tilde{\tau}^{(l,k)}_{\mathcal{C}(d,d_{\mathcal{E}})}$, and given we can take the $k \to \infty$ limit for fixed $t, d, d_{\mathcal{E}}$. □

Theorem 11 (Norm and Trace of Moment Operators)**.** *The k-concatenated Haar, cHaar, and Depolarize moment operators have norms and traces, in the asymptotic limit $dd_{\mathcal{E}} \to \infty$ for fixed t, k, upper bounded by,*

$$\left\|\widehat{\mathcal{T}}^{(t)k}_{\mathcal{U}(d)}\right\|^2 = \mathrm{Tr}\left[\widehat{\mathcal{T}}^{(t)k}_{\mathcal{U}(d)}\right] = t! \qquad , \qquad \left\|\widehat{\mathcal{T}}^{(t)k}_{\mathcal{D}(d)}\right\|^2 = \mathrm{Tr}\left[\widehat{\mathcal{T}}^{(t)k}_{\mathcal{D}(d)}\right] = 1 \tag{B60}$$

$$\left\|\widehat{\mathcal{T}}^{(t)k}_{\mathcal{C}(d,d_{\mathcal{E}})}\right\|^2 \leq 1 \;+\; \frac{1}{d_{\mathcal{E}}^2}\, C^{(t,k)}_{\|\tau\|} \qquad , \qquad \mathrm{Tr}\left[\widehat{\mathcal{T}}^{(t)k}_{\mathcal{C}(d,d_{\mathcal{E}})}\right] \leq 1 \;+\; \frac{1}{d_{\mathcal{E}}^k}\, C^{(t,k)}_{\mathrm{Tr}[\tau]} \;, \tag{B61}$$

given t, k-dependent constants $C^{(t,k)}_{\|\tau\|} = \frac{t!^2-1}{t!^2}\ C_{\lambda}^{(t)2k}$, $C^{(t,k)}_{\mathrm{Tr}[\tau]} = (t!^2-1)\ C_{\lambda}^{(t)k}$, and $C_{\lambda}^{(t)} = t!^2\ \chi_t\ c_t^7$.

Proof. The exact norm and trace of the k-concatenated t-th order Haar and Depolarize moment operators are given by

$$\left\|\widehat{\mathcal{T}}^{(t)k}_{\mathcal{U}(d)}\right\|^2 = \left\|\widehat{\mathcal{T}}^{(t)}_{\mathcal{U}(d)}\right\|^2 = \mathrm{Tr}\left[\widehat{\mathcal{T}}^{(t)}_{\mathcal{U}(d)}\right] = \sum_{\sigma,\pi\in\mathcal{S}_t} W_d^{(t)-1}(\sigma,\pi)\ \chi_d^{(t)}(\pi,\sigma) = |\mathcal{S}_t| \tag{B62}$$

$$\left\|\widehat{\mathcal{T}}^{(t)k}_{\mathcal{D}(d)}\right\|^2 = \left\|\widehat{\mathcal{T}}^{(t)}_{\mathcal{D}(d)}\right\|^2 = \mathrm{Tr}\left[\widehat{\mathcal{T}}^{(t)}_{\mathcal{D}(d)}\right] = 1\;. \tag{B63}$$

The exact norm and trace of t-th order cHaar moment operator do not appear to simplify, however can be bounded as,

$$\left\|\widehat{\mathcal{T}}^{(t)k}_{\mathcal{C}(d,d_{\mathcal{E}})}\right\|^2 = \sum_{\substack{\Gamma'\subseteq\Gamma\subseteq[t]\\ |\Gamma|=l\geq l'=|\Gamma'|}} \sum_{\substack{\sigma,\pi\in\mathcal{S}_l\\ \Gamma_\sigma=\Gamma\supseteq\Gamma'=\Gamma_\pi}} \tilde{\tau}^{(l,k)\prime}_{\mathcal{C}(d,d_{\mathcal{E}})}(\sigma,\pi)\ \tilde{\tau}^{(l,k)}_{\mathcal{C}(d,d_{\mathcal{E}})}(\sigma,\pi) \tag{B64}$$

$$\leq 1 \;+\; (t!^2-1) \max_{\substack{l'\leq l\leq t\\ l\leq 2}} \frac{1}{d_{\mathcal{E}}^{l+(k-1)l'}}\ C_{\tilde{\tau}}^{(l,l',k)}\ C_{\tilde{\tau}'}^{(l,l',k)} \;\leq\; 1 \;+\; \frac{1}{d_{\mathcal{E}}^2} C^{(t,k)}_{\|\tau\|} \tag{B65}$$

$$\mathrm{Tr}\left[\widehat{\mathcal{T}}^{(t)k}_{\mathcal{C}(d,d_{\mathcal{E}})}\right] = \sum_{s\in[t!]} \lambda_s^k \;\leq\; 1 \;+\; (t!^2-1) \max_{2\leq l\leq t} \frac{1}{d_{\mathcal{E}}^{kl/2}}\ C_{\lambda}^{(l)k} \;\leq\; 1 \;+\; \frac{1}{d_{\mathcal{E}}^k} C^{(t,k)}_{\mathrm{Tr}[\tau]} \tag{B66}$$

$$C^{(t,k)}_{\|\tau\|} = (t!^2-1)\ \ C_{\tilde{\tau}}^{(t,t,k)2} = \frac{t!^2-1}{t!^2}\ C_{\lambda}^{(t)2k} \quad , \quad C^{(t,k)}_{\mathrm{Tr}[\tau]} = (t!^2-1)\ C_{\lambda}^{(t)k} \quad , \quad C_{\lambda}^{(t)} = t!^2\ \chi_t\ c_t^7 \;, \tag{B67}$$

given the bounds on the modified k-concatenated transfer matrix elements $\tilde{\tau}^{(l,k)}_{\mathcal{C}(d,d_{\mathcal{E}})}, \tilde{\tau}^{(l,k)\prime}_{\mathcal{C}(d,d_{\mathcal{E}})}$. □

For sufficiently large $d_{\mathcal{E}}$, all non-leading cHaar moment operator eigenvalues are $|\lambda| < 1$, which agrees with the fact that all channels have $\lambda \leq 1$ [127]. For $l = 2$, we have exactly $\lambda = d_{\mathcal{E}}(d^2-1)/(d^2 d_{\mathcal{E}}^2 - 1)$.

Next, we can find that all trace-preserving moment operators over ensembles $\mathcal{C}$ can be expressed as the Depolarize moment operator $\widehat{\mathcal{T}}^{(t)}_{\mathcal{D}(d)}$, plus deviations $\widehat{\Delta}^{(t)}_{\mathcal{C}}$, and the cHaar $\mathcal{C}(d,d_{\mathcal{E}})$ ensemble is invariant under the Haar $\mathcal{U}(d)$ ensemble,

$$\widehat{\mathcal{T}}^{(t)}_{\mathcal{C}} = \widehat{\mathcal{T}}^{(t)}_{\mathcal{D}(d)} \ + \ \widehat{\Delta}^{(t)}_{\mathcal{C}} \tag{B68}$$

$$\widehat{\mathcal{T}}^{(t)}_{\mathcal{C}(d,d_{\mathcal{E}})} = I_d^{\otimes 2t} \otimes \langle\!\langle I_d^{\otimes t}| \ \widehat{\mathcal{T}}^{(t)}_{\mathcal{U}(dd_{\mathcal{E}})} \ I_d^{\otimes 2t} \otimes |\nu_{\mathcal{E}}^{\otimes t}\rangle\!\rangle \tag{B69}$$

$$\widehat{\mathcal{T}}^{(t)}_{\mathcal{D}(d)}\widehat{\mathcal{T}}^{(t)}_{\mathcal{C}} = \widehat{\mathcal{T}}^{(t)\dagger}_{\mathcal{C}}\widehat{\mathcal{T}}^{(t)}_{\mathcal{D}(d)} = \widehat{\mathcal{T}}^{(t)}_{\mathcal{D}(d)} \quad , \quad \widehat{\mathcal{T}}^{(t)}_{\mathcal{U}(d)}\widehat{\mathcal{T}}^{(t)}_{\mathcal{C}(d,d_{\mathcal{E}})} = \widehat{\mathcal{T}}^{(t)}_{\mathcal{C}(d,d_{\mathcal{E}})} \ \widehat{\mathcal{T}}^{(t)}_{\mathcal{U}(d)} = \widehat{\mathcal{T}}^{(t)}_{\mathcal{C}(d,d_{\mathcal{E}})} \quad , \quad \widehat{\mathcal{T}}^{(t)k}_{\mathcal{U}(d)} = \widehat{\mathcal{T}}^{(t)}_{\mathcal{U}(d)} \quad , \quad \widehat{\mathcal{T}}^{(t)\dagger}_{\mathcal{U}(d)} = \widehat{\mathcal{T}}^{(t)}_{\mathcal{U}(d)} \ , \tag{B70}$$

From these relationships, and the exact forms of the moment operators obtained above, the following relationships hold:

Corollary 1 (cHaar Moment Operator Limits)**.** *The k-concatenated t-th order cHaar moment operator satisfies,*

$$\lim_{dd_{\mathcal{E}}\to d} \widehat{\mathcal{T}}^{(t)k}_{\mathcal{C}(d,d_{\mathcal{E}})} = \widehat{\mathcal{T}}^{(t)}_{\mathcal{U}(d)} \qquad , \qquad \lim_{dd_{\mathcal{E}}\to \infty} \widehat{\mathcal{T}}^{(t)k}_{\mathcal{C}(d,d_{\mathcal{E}})} = \widehat{\mathcal{T}}^{(t)}_{\mathcal{D}(d)} \ . \tag{B71}$$

Corollary 2 (cHaar Moment Operator Converges to Depolarizing and Non-Unital Channel)**.** *The k-concatenated t-th order cHaar moment operator is inherently depolarizing, with next-leading-order concatenation-independent non-unital behavior, and concatenation-dependent unital terms. For all dimensions, given traceless operators $T_d^{(t,k)}, T_d^{(t,k)\prime} \in \mathcal{B}[\mathcal{H}^{\otimes t}]$,*

$$\lim_{dd_{\mathcal{E}}\to\infty} \widehat{\mathcal{T}}^{(t)k}_{\mathcal{C}(d,d_{\mathcal{E}})} = \frac{1}{d^t}|I_d^{\otimes t}\rangle\!\rangle\langle\!\langle I_d^{\otimes t}| \ + \ \frac{1}{d^t}\mathcal{O}\!\left(\frac{1}{dd_{\mathcal{E}}}\right) \ |T_d^{(t,k)}\rangle\!\rangle\langle\!\langle I_d^{\otimes t}| \ + \ \frac{1}{d^t}\mathcal{O}\!\left(\frac{1}{d^2 d_{\mathcal{E}}^k}\right)|T_d^{(t,k)}\rangle\!\rangle\langle\!\langle T_d^{(t,k)\prime}| \ . \tag{B72}$$

Indeed, in the localized permutation basis, the projector Haar moment operators have an invariant block-diagonal structure, and the non-projector cHaar moment operators have a block-lower-triangular structure, whose elements decay with dimension and concatenations. The average behavior of random quantum channels is thus a useful pedagogical tool, given it interpolates with dimension, between trivial environments and unitarity, and large dimensions and depolarization. In Fig. 10, we numerically illustrate this bounded behavior for moment operator norms, for different t, k, and $d_{\mathcal{E}}$.

The previous results indicate that as we concatenate cHaar random channels k times, the resulting composite transfer matrix elements are suppressed, which intuitively means the moment operator norm should also decrease until it converges to the norm of a single projector with a unit eigenvalue. In the large $dd_{\mathcal{E}} \to \infty$ limit, this intuition is verified, but given that the cHaar moment operator has no closed form, we instead postulate that the following conjecture should hold.

Conjecture 1 (Hierarchy of Haar, cHaar, Depolarize Ensembles)**.** *Given there exists an exact hierarchy between the k-concatenated, t-th order Haar, cHaar, and Depolarize moment operators in the asymptotic limit $dd_{\mathcal{E}} \to \infty$,*

$$1 = \left\|\widehat{\mathcal{T}}^{(t)}_{\mathcal{D}(d)}\right\|^2 \leq \left\|\widehat{\mathcal{T}}^{(t)k}_{\mathcal{C}(d,d_{\mathcal{E}})}\right\|^2 \leq \left\|\widehat{\mathcal{T}}^{(t)}_{\mathcal{U}(d)}\right\|^2 = t! \ , \tag{B73}$$

it is conjectured that such a hierarchy holds for all concatenations k, orders t, and dimensions $d, d_{\mathcal{E}}$. Further, it is conjectured that the cHaar norm is monotonically non-increasing with environment dimension and concatenations,

$$\left\|\widehat{\mathcal{T}}^{(t)k}_{\mathcal{C}(d,d'_{\mathcal{E}})}\right\|^2 \overset{?}{\geq} \left\|\widehat{\mathcal{T}}^{(t)k}_{\mathcal{C}(d,d_{\mathcal{E}})}\right\|^2 \qquad \forall \ d'_{\mathcal{E}} \geq d_{\mathcal{E}} \qquad , \qquad \left\|\widehat{\mathcal{T}}^{(t)k'}_{\mathcal{C}(d,d_{\mathcal{E}})}\right\|^2 \overset{?}{\leq} \left\|\widehat{\mathcal{T}}^{(t)k}_{\mathcal{C}(d,d_{\mathcal{E}})}\right\|^2 \qquad \forall \ k' \geq k \ . \tag{B74}$$

Given the relationship between trace-preserving moment operators and the invariant Depolarize moment operator,

$$\left\|\widehat{\mathcal{T}}^{(t)}_{\mathcal{C}} - \widehat{\mathcal{T}}^{(t)}_{\mathcal{D}(d)}\right\|^2 = \left\|\widehat{\mathcal{T}}^{(t)}_{\mathcal{C}}\right\|^2 \ - \ \left\|\widehat{\mathcal{T}}^{(t)}_{\mathcal{D}(d)}\right\|^2 \geq 0 \qquad \to \qquad \left\|\widehat{\mathcal{T}}^{(t)}_{\mathcal{C}}\right\|^2 \geq \left\|\widehat{\mathcal{T}}^{(t)}_{\mathcal{D}(d)}\right\|^2 \ , \tag{B75}$$

and given the partial invariance of the Haar and cHaar ensembles, the conjecture is true if

$$\left\|\widehat{\mathcal{T}}^{(t)k}_{\mathcal{C}(d,d_{\mathcal{E}})} - \widehat{\mathcal{T}}^{(t)}_{\mathcal{U}(d)}\right\|^2 \geq 0 \quad \to \quad \left\|\widehat{\mathcal{T}}^{(t)}_{\mathcal{U}(d)}\right\|^2 - \left\|\widehat{\mathcal{T}}^{(t)k}_{\mathcal{C}(d,d_{\mathcal{E}})}\right\|^2 \geq 2\left(\mathrm{Tr}\left[\widehat{\mathcal{T}}^{(t)k}_{\mathcal{C}(d,d_{\mathcal{E}})}\right] - \left\|\widehat{\mathcal{T}}^{(t)k}_{\mathcal{C}(d,d_{\mathcal{E}})}\right\|^2\right) \overset{?}{\geq} 0 \ . \tag{B76}$$

This conjecture, and this bound on the difference of Haar and cHaar moment operator norms, holds in the asymptotic limit $dd_{\mathcal{E}} \to \infty$, where both norms converge to 1. However, our analysis is limited in part by the lack of closed forms for the Weingarten functions, and from our approaches of (possibly loosely) bounding quantities in terms of maximum cHaar transfer matrix elements. Different, possibly representation-theoretic based approaches are to necessary to prove this non-trivial, but important hierarchy conjecture for general $d, d_{\mathcal{E}}$.

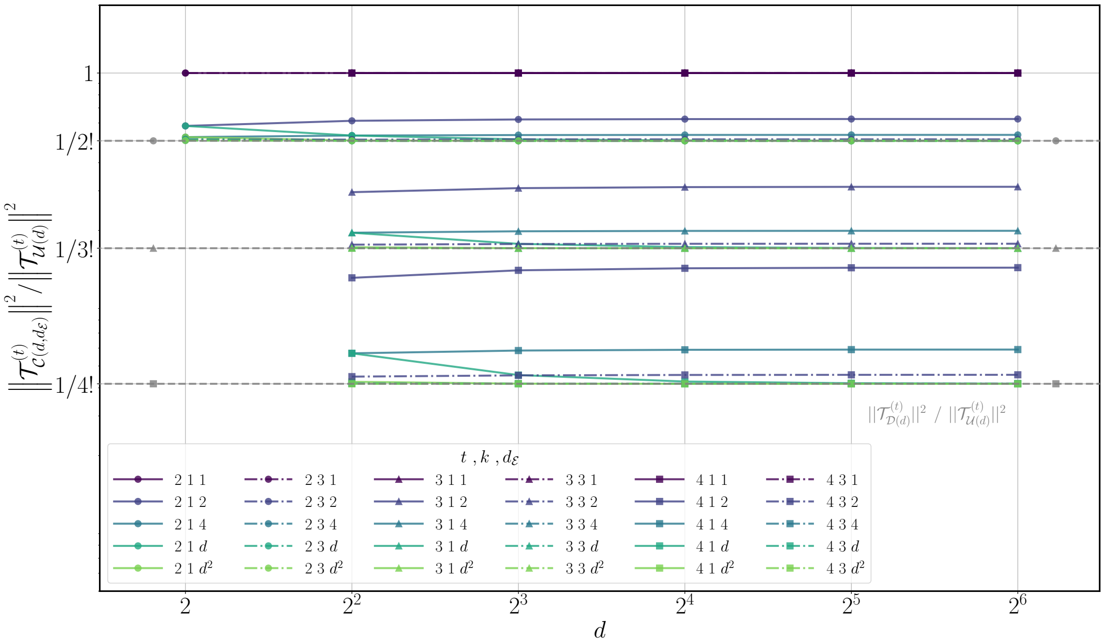

Figure 10. **cHaar moment operator norm scaling.** Norm of the k-concatenated t-th order cHaar moment operator $\|\widehat{\mathcal{T}}^{(t,k)}_{\mathcal{C}(d,d_\mathcal{E})}\|^2$ as a function of the system dimension d, for environment dimensions $d_\mathcal{E}$, concatenations k, and orders t. Norms are computed exactly using the derived closed form expressions for $\widehat{\mathcal{T}}^{(t,k)}_{\mathcal{C}(d,d_\mathcal{E})}$, and symbolic methods [91]. Norms converge to their leading-order, system dimension d-independent term, and decay monotonically with environment dimension $d_\mathcal{E}$ or with concatenation k, for all t, from $\|\widehat{\mathcal{T}}^{(t,k)}_{\mathcal{U}(d)}\|^2 = t!$ (1 on scaled axis) to $\|\widehat{\mathcal{T}}^{(t,k)}_{\mathcal{D}(d)}\|^2 = 1$ ($1/t!$ grey dashed lines on scaled axis). The depicted scalings confirm the bounded behavior, even for constant environment dimension $d_\mathcal{E} < d$, of the cHaar moment operator as it interpolates between unitary and depolarizing behavior. Further, such results suggest that the conjecture that $\|\widehat{\mathcal{T}}^{(t,k)}_{\mathcal{C}(d,d_\mathcal{E})}\|^2 \leq \|\widehat{\mathcal{T}}^{(t,k)}_{\mathcal{U}(d)}\|^2$ holds absolutely.

Appendix C: Moment Operators of Noisy Ensembles

In these appendices, we calculate moment operators for ensembles of noisy unitary channels, and investigate their scaling with respect to concatenations, dimension, and noise parameters.

1. Composite Noise Channels

Here, we seek to calculate moment operators for composite ensembles of channels $\mathcal{N} \circ U$, where the unitary component is sampled as $U \sim \mathcal{U}$, and where the noise component is a fixed noise channel $\mathcal{N}$. The t-th order composite moment operator, $\widehat{\mathcal{T}}^{(t)}_{\mathcal{UN}} = \mathcal{N} \circ \widehat{\mathcal{T}}^{(t)}_{\mathcal{U}}$, then may be concatenated k times,

$$\widehat{\mathcal{T}}^{(t)k}_{\mathcal{UN}} = \prod_{i\in[k]} \widehat{\mathcal{N}^{\otimes t}} \circ \widehat{\mathcal{T}}^{(t)}_{\mathcal{U}} = \widehat{\mathcal{D}}^{\otimes t}_{d} + \widehat{\Delta}^{(t,k)}_{\mathcal{UN}} \tag{C1}$$

with trace-preserving depolarization $\widehat{\mathcal{D}}^{\otimes t}_{d}$ and deviation $\widehat{\Delta}^{(t,k)}_{\mathcal{UN}}$ terms of

$$\widehat{\mathcal{D}}^{\otimes t}_{d} = \frac{1}{d^t} |I^{\otimes t}_d\rangle\!\rangle\langle\!\langle I^{\otimes t}_d| \quad \text{and} \quad \widehat{\Delta}^{(t,k)}_{\mathcal{UN}} = \frac{1}{d^t} \sum_{P\in\mathcal{S}^{(t)}_{\mathcal{N}}\setminus\{I^{\otimes t}_d\},\ S\in\mathcal{S}^{(t)}_{\mathcal{U}}} \tau^{(t,k)}_{\mathcal{UN}}(P,S)\ |P\rangle\!\rangle\langle\!\langle S|\ . \tag{C2}$$

The concatenated composite transfer matrices $\tau^{(t,k)}_{\mathcal{UN}}$ can be expressed in terms of the component transfer matrices $\tau^{(t)}_{\mathcal{N}}, \tau^{(t)}_{\mathcal{U}}$, with respective bases $\mathcal{S}^{(t)}_{\mathcal{U}}, \mathcal{S}^{(t)}_{\mathcal{N}}$ for each component

$$\tau^{(t)}_{\mathcal{UN}}(P,S) = \sum_{Q\in\mathcal{S}^{(t)}_{\mathcal{N}},\ T\in\mathcal{S}^{(t)}_{\mathcal{U}}} \tau^{(t)}_{\mathcal{N}}(P,Q)\ \chi^{(t)}_d(Q,T)\ \tau^{(t)}_{\mathcal{U}}(T,S) \quad \rightarrow \quad \tau^{(t,k)}_{\mathcal{UN}}(P,S) = \sum_{Q\in\mathcal{S}^{(t)}_{\mathcal{U}},\ T\in\mathcal{S}^{(t)}_{\mathcal{N}}} \tau^{(t)}_{\mathcal{UN}}(P,Q)\ \chi^{(t)}_d(Q,T)\ \tau^{(t,k-1)}_{\mathcal{UN}}(T,S)\ . \tag{C3}$$

The most appropriate basis for each component, and whether $\mathcal{S}^{(t)}_{\mathcal{U}} \neq \mathcal{S}^{(t)}_{\mathcal{N}}$, greatly affects their interplay, and the ease of analysis of the scaling of moment operator norms $\epsilon^{(t,k)2}_{\mathcal{UN}} = \left\|\widehat{\Delta}^{(t,k)}_{\mathcal{UN}}\right\|^2$, with variables t, k, d, γ, η.

2. Unital and Non-unital Noise and Haar Random Unitary Channels

Here, we consider Haar random unitary channels $\mathcal{U} = \mathcal{U}(d)$ followed by fixed diagonal unital or non-unital noise $\mathcal{N} = \mathcal{N}_\gamma, \mathcal{N}_{\gamma\eta}$, and derive the scaling of their k-concatenated t-th order composite moment operators with t, k, d, γ, η.

For the unitary component, t-th order moment operators over $\mathcal{U}(d)$ are, in an orthogonal basis $\mathcal{S}^{(t)}_{\mathcal{U}(d)} = \mathcal{P}^t_d$,

$$\widehat{\mathcal{T}}^{(t)}_{\mathcal{U}(d)} = \int_{\mathcal{U}(d)} dU\ U^{\otimes t} \otimes U^{\otimes t\,*} = \frac{1}{d^t} \sum_{P,S\in\mathcal{P}^t_d} \tau^{(t)}_{\mathcal{U}(d)}(P,S)\ |P\rangle\!\rangle\langle\!\langle S| \quad : \quad \tau^{(t)}_{\mathcal{U}(d)}(P,S)\ \propto\ \delta_{\Gamma_P\Gamma_S}\ . \tag{C4}$$

This expansion in a strictly orthogonal basis can be derived from the localized permutation basis $[\hat{\sigma}] \in [\hat{\mathcal{S}}_t]$, given $[\hat{\sigma}] = \sum^{P\in\mathcal{P}^t_d}_{\Gamma_P=\Gamma_\sigma} \chi^{(t)}_d([\hat{\sigma}],P)\ P$, with associated transfer matrix coefficients $\tau^{(t)}_{\mathcal{U}(d)}([\hat{\sigma}],[\hat{\pi}]) \propto \delta_{\Gamma_\sigma\Gamma_\pi}$. The orthogonal basis transfer matrix coefficients $\tau^{(t)}_{\mathcal{U}(d)}(P,S)$ thus also have support-dependent and concatenation-invariant block structure,

$$\tau^{(t)}_{\mathcal{U}(d)}(P,S) = \sum_{\substack{\sigma,\pi\in\mathcal{S}_t\\ \Gamma_\sigma=\Gamma_P=\Gamma_S=\Gamma_\pi}} \chi^{(t)}_d(P,[\hat{\sigma}])\ \tau^{(t)}_{\mathcal{U}(d)}([\hat{\sigma}],[\hat{\pi}])\ \chi^{(t)}_d([\hat{\pi}],S)\ \delta_{\Gamma_P\Gamma_S} \to \sum_{Q\in\mathcal{S}^{(t)}_{\mathcal{U}(d)}} \tau^{(t)}_{\mathcal{U}(d)}(P,Q)\ \tau^{(t)}_{\mathcal{U}(d)}(Q,S)\ =\ \tau^{(t)}_{\mathcal{U}(d)}(P,S)\ . \tag{C5}$$

For the noise component, the noise model $\mathcal{N} = \mathcal{N}_{\gamma\eta}$ moment operator has the special form of diagonal unital noise and non-unital noise components $\gamma_d(P) \geq \gamma\ ,\ \eta_d(P) \leq \eta$ for $P \in \mathcal{S}_{\mathcal{N}_{\gamma\eta}}$, in an orthogonal basis $\mathcal{S}_{\mathcal{N}_{\gamma\eta}} = \mathcal{P}_d$,

$$\widehat{\mathcal{N}}_{\gamma\eta} = \frac{1}{d} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ \eta_d(S) & 1-\gamma_d(S) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \eta_d(P) & 0 & \cdots & 1-\gamma_d(P) \end{bmatrix}_{P,S\in\mathcal{P}_d\setminus\{I_d\}} \quad \to \quad \widehat{\mathcal{N}^{\otimes t}_{\gamma\eta}} = \frac{1}{d^t} \sum_{P,S\in\mathcal{S}^{(t)}_{\mathcal{N}_{\gamma\eta}}} \tau^{(t)}_{\mathcal{N}_{\gamma\eta}}(P,S)\ |P\rangle\!\rangle\langle\!\langle S|\ , \tag{C6}$$

and appropriate additional constraints on γ_d, η_d such that $\mathcal{N}$ is completely positive [92]. The t-th order noise component transfer matrix $\tau^{(t)}_{\mathcal{N}_{\gamma\eta}}$, for basis operators $P, S \in \mathcal{S}^{(t)}_{\mathcal{N}_{\gamma\eta}} = \mathcal{P}^t_d$, is

$$\tau^{(t)}_{\mathcal{N}_{\gamma\eta}}(P,S) = \prod_{i\in\Gamma_S} (1-\gamma_d(P_i)) \prod_{i\in\Gamma_P\setminus\Gamma_S} \eta_d(P_i)\ \delta_{P_{\Gamma_S}S_{\Gamma_S}}\ \in\ \mathcal{O}\Big((1-\gamma)^{|S|}\ \eta^{|P|-|S|}\Big)\ \delta_{P_{\Gamma_S}S_{\Gamma_S}}\ . \tag{C7}$$

We restrict to this special form (a diagonal component arising from Pauli-like-noise, and a dense first column arising from a non-unital component) as we seek to understand exact theoretical properties. More general noise models with other non-diagonal elements, even if initially sparse, become dense and intractable to analyze upon concatenations.

Finally, let us assess the leading-order behavior of the composite and concatenated transfer matrices,

$$\widehat{\mathcal{T}}^{(t)}_{\mathcal{U}(d)\mathcal{N}_{\gamma\eta}} = \widehat{\mathcal{N}^{\otimes t}_{\gamma\eta}} \circ \widehat{\mathcal{T}}^{(t)}_{\mathcal{U}(d)} = \frac{1}{d^t} \sum_{\substack{P\in\mathcal{S}^{(t)}_{\mathcal{N}_{\gamma\eta}}\\ S\in\mathcal{S}^{(t)}_{\mathcal{U}(d)}}} \tau^{(t)}_{\mathcal{U}(d)\mathcal{N}_{\gamma\eta}}(P,S)\ |P\rangle\!\rangle\langle\!\langle S| \quad \to \quad \widehat{\mathcal{T}}^{(t)k}_{\mathcal{U}(d)\mathcal{N}_{\gamma\eta}} = \frac{1}{d^t} \sum_{\substack{P\in\mathcal{S}^{(t)}_{\mathcal{N}_{\gamma\eta}}\\ S\in\mathcal{S}^{(t)}_{\mathcal{U}(d)}}} \tau^{(t,k)}_{\mathcal{U}(d)\mathcal{N}_{\gamma\eta}}(P,S)\ |P\rangle\!\rangle\langle\!\langle S|\ . \tag{C8}$$

In particular, given the forms of the individual component transfer matrices, we seek to understand the competition that arises between the unitary components preserving locality, and non-unital noise components increasing locality.

Lemma 4 (Unital and Non-unital Noise and Haar Random Unitary Channel Moment Operators)**.** *Composite ensembles, consisting of a Haar random unitary component with a basis $S \in \mathcal{S}^{(t)}_{\mathcal{U}(d)} = \mathcal{P}^t_d$, and a noise component with a basis $P \in \mathcal{S}^{(t)}_{\mathcal{N}_{\gamma\eta}} = \mathcal{P}^t_d$, and with diagonal unital, and general non-unital noise coefficients, $\gamma_d(P) \geq \gamma\ ,\ \eta_d(P) \leq \eta$ for $P \in \mathcal{P}^t_d\setminus\{I^{\otimes t}_d\}$, have k-concatenated t-th order composite transfer matrices of, for $P,S \in \mathcal{P}^t_d$*

$$\tau^{(t,k)}_{\mathcal{U}(d)\mathcal{N}_{\gamma\eta}}(P,S) = \mathcal{O}\Big((1-\gamma)^{k|S|}\eta^{|P|-|S|}\Big)\ \tau^{(t)}_{\mathcal{U}(d)}(P_{\Gamma_S},S)\ \delta_{\Gamma_P\supseteq\Gamma_S}\ , \tag{C9}$$

where the leading-order $\mathcal{O}(\gamma,\eta)$ scaling is strictly due to the unspecified noise component, with no explicit dimension-dependence from the unitary component.

Proof. To begin, we note that the unitary component $\mathcal{U}(d)$ is support-preserving, whereas the non-unital noise component $\mathcal{N}_{\gamma\eta}$ is support-non-decreasing, with individual component transfer matrices, for $P,S\in\mathcal{P}_d^t=\mathcal{S}^{(t)}_{\mathcal{U}(d)}=\mathcal{S}^{(t)}_{\mathcal{N}_{\gamma\eta}}$, of

$$\tau^{(t)}_{\mathcal{U}(d)\mathcal{N}_{\gamma\eta}}(P,S)=\prod_{i\in\Gamma_S}(1-\gamma_d(P_i))\prod_{i\in\Gamma_P\setminus\Gamma_S}\eta_d(P_i)\ \tau^{(t)}_{\mathcal{U}(d)}(P_{\Gamma_S},S)\ \delta_{\Gamma_P\supseteq\Gamma_S}\in\mathcal{O}\Big((1-\gamma)^{|S|}\eta^{|P|-|S|}\Big)\,\tau^{(t)}_{\mathcal{U}(d)}(P_{\Gamma_S},S)\ \delta_{\Gamma_P\supseteq\Gamma_S}\,. \tag{C10}$$

From locality preservation of the unitary component, diagonal unital noise scaling always occurs, for all $\Gamma_P\supseteq\Gamma_S$, whereas non-unital noise scaling only occurs when strictly $\Gamma_P\supset\Gamma_S$, and concatenated operators are constrained to be support-increasing $\Gamma_P\subseteq\cdots\subseteq\Gamma_T\subseteq\cdots\subseteq\Gamma_S$. For instance, the $k=2$ transfer matrix elements are

$$\begin{aligned}
\tau^{(t,2)}_{\mathcal{U}(d)\mathcal{N}_{\gamma\eta}}(P,S) &= \sum_{\substack{T\in\mathcal{P}_d^t\\ \Gamma_S\subseteq\Gamma_T\subseteq\Gamma_P}}\tau^{(t)}_{\mathcal{U}(d)\mathcal{N}_{\gamma\eta}}(P,T)\ \tau^{(t)}_{\mathcal{U}(d)\mathcal{N}_{\gamma\eta}}(T,S) && \text{(C11)}\\
&= \sum_{\substack{T\in\mathcal{P}_d^t\\ \Gamma_S\subseteq\Gamma_T\subseteq\Gamma_P}}\prod_{i\in\Gamma_T}(1-\gamma_d(P_i))\prod_{i\in\Gamma_P\setminus\Gamma_T}\eta_d(P_i)\prod_{i\in\Gamma_S}(1-\gamma_d(T_i))\prod_{i\in\Gamma_T\setminus\Gamma_S}\eta_d(T_i)\ \tau^{(t)}_{\mathcal{U}(d)}(P_{\Gamma_T},T)\ \tau^{(t)}_{\mathcal{U}(d)}(T_{\Gamma_S},S) && \text{(C12)}\\
&= \mathcal{O}\Big((1-\gamma)^{2|S|}\ \eta^{|P|-|S|}\Big)\left[\sum_{\substack{T\in\mathcal{P}_d^t\\ \Gamma_T=\Gamma_S\subseteq\Gamma_P}}\tau^{(t)}_{\mathcal{U}(d)}(P_{\Gamma_S},T)\ \tau^{(t)}_{\mathcal{U}(d)}(T,S)\ +\ \mathcal{O}\Big((1-\gamma)\Big)\right] && \text{(C13)}\\
&= \mathcal{O}\Big((1-\gamma)^{2|S|}\ \eta^{|P|-|S|}\Big)\ \tau^{(t)}_{\mathcal{U}(d)}(P_{\Gamma_S},S)\ \delta_{\Gamma_P\supseteq\Gamma_S}\ , && \text{(C14)}
\end{aligned}$$

and k-concatenated composite transfer matrices are thus by induction,

$$\tau^{(t,k)}_{\mathcal{U}(d)\mathcal{N}_{\gamma\eta}}(P,S)=\mathcal{O}\Big((1-\gamma)^{k|S|}\ \eta^{|P|-|S|}\Big)\ \tau^{(t)}_{\mathcal{U}(d)}(P_{\Gamma_S},S)\ \delta_{\Gamma_P\supseteq\Gamma_S}\ . \tag{C15}$$

□

Therefore, the leading-order noise scaling is dictated by the concatenation-invariant unitary components being scaled by strictly unital noise until the last concatenation, at which point non-unital noise increases the support from $\Gamma_S\to\Gamma_P$. Intermediate non-unital noise at concatenations $0<s<k$ also increase the support $\Gamma_S\to\Gamma_T\to\Gamma_P$, however such intermediate noise contribute next-leading-order unital noise scaling $\mathcal{O}((1-\gamma)^{(k-s)|T|})$. Intermediate non-unital noise, in the case it maps unitary component basis operators to larger-support unitary component basis operators, also contribute dimension-dependent scaling when such concatenations occur between unitary components of differing support.

We now assess the leading-order scaling of moment operator norms for concatenations of composite ensembles.

Theorem 12 (Unital and Non-unital Noise and Haar Random Unitary Channel Moment Operator Norms)**.** *k-concatenations of t-th order ensembles of Haar random unitaries, and diagonal unital and non-unital noise, with noise coefficients $\gamma_d(P)\geq\gamma$, $\eta_d(P)\geq\eta$ for $P\in\mathcal{P}_d\setminus\{I_d\}$, have moment operator norms of,*

$$\epsilon^{(t,k)2}_{\mathcal{U}(d)\mathcal{N}_\gamma}=\binom{t}{2}\ \mathcal{O}\big((1-\gamma)^{4k}\big)\qquad,\qquad\epsilon^{(t,k)2}_{\mathcal{U}(d)\mathcal{N}_{\gamma\eta}}=t\ |\mathcal{P}_d\setminus\{I_d\}|\ \mathcal{O}\big(\eta^2\big)\ . \tag{C16}$$

From the previous theorem it follows that as $k\to\infty$, unital noise decreases moment operator norms for $\gamma>0$, and approaches a Depolarize t-design, whereas non-unital noise increases moment operator norms for $\eta>0$, and never approaches a Depolarize t-design.

Proof. Moment operator norms follow directly from their expansion in an orthogonal basis. Composite unital noise and unitary component moment operators may be decomposed into support $\Gamma\subseteq[t]$-dependent unitary component projectors $\widehat{\mathcal{T}}^{(\Gamma)}_{\mathcal{U}(d)}=(1/d^t)\sum^{P,S\in\mathcal{S}^{(t)}_{\mathcal{U}(d)}}_{\Gamma_P=\Gamma_S=\Gamma}\tau^{(t)}_{\mathcal{U}(d)}(P,S)\ |P\rangle\!\rangle\langle\!\langle S|$, scaled by unital noise $\gamma_d\geq\gamma$. Such moment operators are therefore have deviations from the Depolarize term $\mathcal{D}_d^{\otimes t}=(1/d^t)|I_d^{\otimes t}\rangle\!\rangle\langle\!\langle I_d^{\otimes t}|$, scaled by non-unital noise $\eta_d\geq\eta$,

$$\widehat{\mathcal{T}}^{(t)k}_{\mathcal{U}(d)\mathcal{N}_\gamma}\ =\ \mathcal{D}_d^{\otimes t}\ +\sum_{\Gamma\subseteq[t],l=|\Gamma|>1}\mathcal{O}\big((1-\gamma)^{kl}\big)\ \widehat{\mathcal{T}}^{(\Gamma)}_{\mathcal{U}(d)}\qquad,\qquad\widehat{\mathcal{T}}^{(t)k}_{\mathcal{U}(d)\mathcal{N}_\eta}\ =\ \mathcal{D}_d^{\otimes t}\ +\sum_{\substack{\Gamma\subseteq[t],l=|\Gamma|>0\\ P\in\mathcal{P}_d,\Gamma_P=\Gamma}}\mathcal{O}\big(\eta^l\big)\ |P\rangle\!\rangle\langle\!\langle I_d^{\otimes t}|\ +\ \cdots \tag{C17}$$

$$\Big\|\widehat{\mathcal{T}}^{(t)k}_{\mathcal{U}(d)\mathcal{N}_\gamma}\Big\|^2\ =\ 1\ +\sum_{2\leq l\leq t}\binom{t}{l}\ \mathcal{O}\big((1-\gamma)^{2kl}\big)\qquad,\qquad\Big\|\widehat{\mathcal{T}}^{(t)k}_{\mathcal{U}(d)\mathcal{N}_\eta}\Big\|^2\ =\ 1\ +\sum_{1\leq l\leq t}\binom{t}{l}\ |\mathcal{P}_d\setminus\{I_d\}|^l\ \mathcal{O}\big(\eta^{2l}\big)\ . \tag{C18}$$

There are $t|\mathcal{P}_d\setminus\{I_d\}|$ leading-order 1-locality non-unital terms, and $\binom{t}{2}$ leading-order 2-locality unital terms. □

3. Unital Noise and parametrized Random Unitary Channels

Here, we consider parametrized random unitary channels $\mathcal{U} = \mathcal{G}_\Theta$, and fixed diagonal unital noise $\mathcal{N} = \mathcal{N}_\gamma$, and derive the scaling of their k-concatenated t-th order composite moment operators with t, k, d, γ.

For the unitary component $\mathcal{U} = \mathcal{G}_\Theta$, we assume parametrized unitary channels

$$\mathcal{G}_\Theta = \{U_\theta^G = e^{-i\theta G} \ , \ \theta \sim \Theta\} \ , \tag{C19}$$

generated by single hermitian, involutory operators $G \in \mathcal{P}_d$, with a distribution of parameters Θ, and a basis for the generator's commutant of $\mathcal{S}_G \subseteq \mathcal{P}_d$. t copies of the parametrized unitaries for involutory generators G have the form,

$$U_\theta^{G\otimes t} \quad = \sum_{l\in[t+1]} u_\theta^{(l)} \ G^{(l)} \quad : \quad G^{(l)} = \sum_{\substack{\Gamma\subseteq[t] \\ |\Gamma|=l}} \otimes_{i\in\Gamma} G_i \quad , \quad u_\theta^{(l)} = (-i)^l \ \cos^t\theta \ \tan^l\theta \ , \tag{C20}$$

and the associated t-th order moment operators are functions of trigonometric moments of the parameter distribution,

$$\widehat{\mathcal{T}}^{(t)}_{G\mathcal{G}_\Theta} \quad = \quad \langle U_\theta^{G\,\otimes t} \otimes U_\theta^{G\,\otimes t\,*}\rangle_{\theta\sim\Theta} \quad = \sum_{l,l'\in[t+1]} \left\langle u_\theta^{(l)} u_\theta^{(l')*} \right\rangle_{\theta\sim\Theta} \ G^{(l)} \otimes G^{(l')*} \ . \tag{C21}$$

Crucial to this analysis is an understanding of the t-th order commutant of parametrized random unitaries $\mathcal{S}^{(t)}_{\mathcal{G}_\Theta}$, which depends non-trivially on the generator commutant $\mathcal{S}_G$, and can be described by the locality $l \in [t+1]$ of its operators,

$$\mathcal{S}^{(t)}_{\mathcal{G}_\Theta} = \bigcup_{l\in[t+1]} \mathcal{S}^{(t|l)}_{\mathcal{G}_\Theta} \ : \ \mathcal{S}^{(t|l)}_{\mathcal{G}_\Theta} = \{X \in \mathcal{S}^{(t)}_{\mathcal{G}_\Theta} \ : \ |X| = l\} \ . \tag{C22}$$

A complete understanding of the entire commutant is non-trivial, and is left for future work. However, given the form of the moment operator, and that higher t-th order commutants include lower $l \le t$-locality commutants, therefore the $l = 1$-locality local component of commutants is the generator commutant at each of the t copies of the space,

$$\mathcal{S}^{(t|1)}_{\mathcal{G}_\Theta} \ = \ \bigcup_{i\in[t]} \mathcal{S}_{G_i} \ . \tag{C23}$$

We can also evaluate the trigonometric moments involved in this moment operator exactly for $t = 1, 2$.

Lemma 5 ($t = 1, 2$ Twirls of parametrized Random Unitaries)**.** *Twirls over $t = 1, 2$-order parametrized random unitary ensembles $e^{-i\theta G} \sim \mathcal{G}_\Theta$, given Hermitian involutory generators G, uniform parameter distributions Θ, and $X \in \mathcal{B}[\mathcal{H}]$, are*

$$\mathcal{T}^{(1)}_{\mathcal{G}_\Theta}(X) = \frac{1}{2}\left[X + GXG\right] \tag{C24}$$

$$\mathcal{T}^{(2)}_{\mathcal{G}_\Theta}(X) = \frac{1}{8}\left[3\left(X + (G\otimes G)\,X\,(G\otimes G)\right) - \{G\otimes G, X\} + (I\otimes G + G\otimes I)\,X\,(I\otimes G + G\otimes I)\right] \ . \tag{C25}$$

Given these insights into the local component of the commutant, the t-th order $\mathcal{G}_\Theta$ ensemble moment operator is

$$\widehat{\mathcal{T}}^{(t)}_{\mathcal{G}_\Theta} = \frac{1}{d^t} \sum_{P\in\mathcal{S}^{(t|1)}_{\mathcal{G}_\Theta}} |P\rangle\!\rangle\langle\!\langle P| \ + \ \frac{1}{d^t} \sum_{P,S\in\mathcal{S}^{(t)}_{\mathcal{G}_\Theta}\setminus\mathcal{S}^{(t|1)}_{\mathcal{G}_\Theta}} \tau^{(t,k)}_{\mathcal{G}_\Theta}(P,S) \ |P\rangle\!\rangle\langle\!\langle S| \ . \tag{C26}$$

For the noise component, the noise model $\mathcal{N} = \mathcal{N}_\gamma$ moment operator has the special form of diagonal unital noise components $\gamma_d(P) \ge \gamma$ for $P \in \mathcal{S}_{\mathcal{N}_\gamma}$, in an orthogonal basis $\mathcal{S}_{\mathcal{N}_\gamma} = \mathcal{P}_d$,

$$\widehat{\mathcal{N}}_\gamma = \frac{1}{d}\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1-\gamma_d(S) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1-\gamma_d(P) \end{bmatrix}_{P,S\in\mathcal{P}_d\setminus\{I_d\}} \quad \to \quad \widehat{\mathcal{N}_\gamma^{\otimes t}} = \frac{1}{d^t}\sum_{P,S\in\mathcal{S}^{(t)}_{\mathcal{N}_\gamma}} \tau^{(t)}_{\mathcal{N}_\gamma}(P,S) \ |P\rangle\!\rangle\langle\!\langle S| \ , \tag{C27}$$

The t-th order noise component transfer matrix $\tau^{(t)}_{\mathcal{N}_\gamma}$, for basis operators $P, S \in \mathcal{S}^{(t)}_{\mathcal{N}_\gamma} = \mathcal{P}_d^t$, is

$$\tau^{(t)}_{\mathcal{N}_\gamma}(P,S) = \prod_{i\in\Gamma_S}\left(1 - \gamma_d(S_i)\right) \delta_{PS} \ \in \ \mathcal{O}\Big((1-\gamma)^{|S|}\Big) \ \delta_{PS} \ . \tag{C28}$$

The t-th order noise component basis can also be described by the locality $l \in [t+1]$ of its operators,

$$\mathcal{S}^{(t)}_{\mathcal{N}_\gamma} = \bigcup_{l\in[t+1]} \mathcal{S}^{(t|l)}_{\mathcal{N}_\gamma} \;:\; \mathcal{S}^{(t|l)}_{\mathcal{N}_\gamma} = \{X \in \mathcal{S}^{(t)}_{\mathcal{N}_\gamma} \;:\; |X| = l\} \;. \tag{C29}$$

and the local component is the noise component basis at each of the t copies of the space,

$$\mathcal{S}^{(t|1)}_{\mathcal{N}_\gamma} = \cup_{i\in[t]} \mathcal{S}_{\mathcal{N}_\gamma i}. \tag{C30}$$

Given the transfer matrices for the unitary and noise components, and that their respective bases satisfy $\mathcal{S}_G \subseteq \mathcal{S}_{\mathcal{N}_\gamma} = \mathcal{P}_d \;\rightarrow\; \mathcal{S}^{(t|1)}_{\mathcal{G}_\Theta} \subseteq \mathcal{S}^{(t|1)}_{\mathcal{N}_\gamma} = \cup_{i\in[t]}\mathcal{P}_{d_i}$, the k-concatenated t-th order composite moment operator is

$$\widehat{\mathcal{T}}^{(t)k}_{\mathcal{G}_\Theta\mathcal{N}_\gamma} = \frac{1}{d^t} \sum_{P,S\in\mathcal{S}^{(t|1)}_{\mathcal{G}_\Theta}} \tau^{(t,k)}_{\mathcal{G}_\Theta\mathcal{N}_\gamma}(P,S)\; |P\rangle\!\rangle\langle\!\langle S| \;+\; \frac{1}{d^t} \sum_{\substack{P\in\mathcal{S}^{(t)}_{\mathcal{N}_\gamma}\backslash\mathcal{S}^{(t|1)}_{\mathcal{N}_\gamma}\\ S\in\mathcal{S}^{(t)}_{\mathcal{G}_\Theta}\backslash\mathcal{S}^{(t|1)}_{\mathcal{G}_\Theta}}} \tau^{(t,k)}_{\mathcal{G}_\Theta\mathcal{N}_\gamma}(P,S)\; |P\rangle\!\rangle\langle\!\langle S| \;, \tag{C31}$$

with diagonal k-concatenated t-th order composite transfer matrices, for the local operators $P,S \in \mathcal{S}^{(t|1)}_{\mathcal{G}_\Theta}$, of

$$\tau^{(t,k)}_{\mathcal{G}_\Theta\mathcal{N}_\gamma}(P,S) \;=\; \prod_{i\in\Gamma_S} (1-\gamma_d(S_i))^k\; \delta_{PS} \;\in\; \mathcal{O}\Big((1-\gamma)^{k|S|}\Big)\; \delta_{PS} \quad:\quad P,S \in \mathcal{S}^{(t|1)}_{\mathcal{G}_\Theta} \;. \tag{C32}$$

Finally, we can assess the leading-order scaling of moment operator norms for k-concatenations of composite ensembles.

Theorem 13 (Unital Noise and parametrized Random Unitary Channel Moment Operator Norms)**.** *k-concatenations of composite ensembles, consisting of parametrized random unitaries, generated by involutory operators G with commutants $\mathcal{S}_G$, and diagonal unital noise $\mathcal{N}_\gamma$ with noise coefficients $\gamma_d(P) \geq \gamma$ for $P \in \mathcal{S}_{\mathcal{N}_\gamma}\backslash\{I_d\}$, have moment operator norms of,*

$$\epsilon^{(t,k)2}_{\mathcal{G}_\Theta\mathcal{N}_\gamma} = t\; |\mathcal{S}_G\backslash\{I_d\}|\; \mathcal{O}\big((1-\gamma)^{2k}\big) \;. \tag{C33}$$

Therefore unital noise decreases moment operator norms for $\gamma > 0$, and as $k \rightarrow \infty$, approaches a Depolarize t-design.

Proof. Leading-order behavior of the composite moment operators is associated with local operators in the commutant, $\mathcal{S}^{(t|1)}_{\mathcal{G}_\Theta} = \cup_i \mathcal{S}_{G_i}$. Under the assumption of the intersection of the unitary and noise component bases of $\mathcal{S}^{(t|1)}_{\mathcal{G}_\Theta} \subseteq \mathcal{S}^{(t|1)}_{\mathcal{N}_\gamma}$, the local operator components of the associated moment operator transfer matrices are invariant under k-concatenations, up to scaling by diagonal noise coefficients. The moment operator norm is thus,

$$\epsilon^{(t,k)2}_{\mathcal{G}_\Theta\mathcal{N}_\gamma} = \sum_{P\in\mathcal{S}^{(t|1)}_{\mathcal{G}_\Theta}\backslash\{I_d^{\otimes t}\}} \mathcal{O}\big((1-\gamma)^{2k}\big) \;+\; \sum_{\substack{P,T\in\mathcal{S}^{(t)}_{\mathcal{N}_\gamma}\backslash\mathcal{S}^{(t|1)}_{\mathcal{N}_\gamma}\\ S,Q\in\mathcal{S}^{(t)}_{\mathcal{G}_\Theta}\backslash\mathcal{S}^{(t|1)}_{\mathcal{G}_\Theta}}} \chi^{(t)}_d(P,T)\chi^{(t)}_d(S,Q)\tau^{(t,k)}_{\mathcal{G}_\Theta\mathcal{N}_\gamma}(P,Q)\tau^{(t,k)}_{\mathcal{G}_\Theta\mathcal{N}_\gamma}(T,S) \tag{C34}$$

$$= t\; |\mathcal{S}_G\backslash\{I_d\}|\; \mathcal{O}\big((1-\gamma)^{2k}\big) \;+\; \mathcal{O}\big((1-\gamma)^{2k+2}\big) \;. \tag{C35}$$

□

Given both Haar random and parametrized unitary ensembles with unital noise approach a Depolarize t-design, this suggests that unital noise may universally impose depolarization on arbitrary unitary ensembles.

Appendix D: Moments and Expectation Values

In these appendices, we generalize previous results [18], and show how to use channel moment operators to bound biases and variances of expectation values of functions over channel ensembles. Such bounds allow us to study concentration properties of expectation values, leading to notions of channel-design-induced concentration phenomena.

1. Statistics of Ensembles

In what follows, given an operator $X \in \mathcal{B}[\mathcal{H}^{\otimes t}]$, we will denote as $X_\Lambda = \Lambda^{\otimes t}(X)$ the action of the t-th fold tensor product action of a channel Λ. Further, for $\rho, O \in \mathcal{B}[\mathcal{H}]$, the following property holds $\mathrm{Tr}\left[\rho_\Lambda O\right] = \mathrm{Tr}\left[\rho\ O_{\Lambda^\dagger}\right]$.

Given a channel-dependent function $\mathcal{F}(\Lambda)$, for $\Lambda \sim \mathcal{C}$, its t-th order moments with respect to an ensemble $\mathcal{C}$ are

$$\left\langle \mathcal{F}(\Lambda)^t \right\rangle_{\mathcal{C}} = \int_{\mathcal{C}} d\Lambda\ \mathcal{F}(\Lambda)^t \ , \tag{D1}$$

and we denote its mean as $\mu_{\mathcal{C}} = \langle \mathcal{F}(\Lambda) \rangle_{\mathcal{C}}$, and its variance as $\sigma^2_{\mathcal{C}} = \left\langle \mathcal{F}(\Lambda)^2 \right\rangle_{\mathcal{C}} - \langle \mathcal{F}(\Lambda) \rangle^2_{\mathcal{C}}$.

Given that $\mathcal{F}(\Lambda)$ is a random variable over the ensemble $\mathcal{C}$, Chebyshev's inequality allows us to upper bound the probability that $\mathcal{F}(\Lambda)$ deviates from its mean, in terms of its variance, as

$$p\left(|\mathcal{F}(\Lambda) - \mu_{\mathcal{C}}| \geq \epsilon\right) \leq \frac{\sigma^2_{\mathcal{C}}}{\epsilon^2} \ . \tag{D2}$$

Here, we study the mean and variance of channel-dependent functions, which generally depend on so-called inherent terms, which are due to taking expectation values over large spaces, and depend on so-called design terms, which are due to design properties of ensembles with respect to reference ensembles. Such expressions are with respect to quantum state inputs $\rho \in \mathcal{B}[\mathcal{H}]$, and hermitian observables $O \in \mathcal{B}[\mathcal{H}]$ that act non-trivially in a subspace $\mathcal{W} \subseteq \mathcal{H}$ of dimension $d_{\mathcal{W}} \leq d$, with complement space $\overline{\mathcal{W}} \subseteq \mathcal{H}$ of dimension $d_{\overline{\mathcal{W}}} = d/d_{\mathcal{W}}$. Such expressions are also with respect to reference ensembles $\mathcal{C}'$, such as the cHaar $\mathcal{C}' = \mathcal{C}(d, d_{\mathcal{E}})$ or Depolarize $\mathcal{C}' = \mathcal{D}(d)$ ensembles. To bound various quantities, we will use Schatten norms $\|A\|_p^p = \mathrm{Tr}\left[|A|^p\right]\ ,\ p \in [1, \infty]$ for $A, B \in \mathcal{B}[\mathcal{H}^{\otimes t}]$, which satisfy monotonicity $\|A\|_p \leq \|A\|_q\ ,\ q \leq p$, sub-additivity $\|A + B\|_p \leq \|A\|_p + \|B\|_p$, sub-multiplicativity $\|AB\|_p \leq \|A\|_p \|B\|_p$, and Holder's $\|AB\|_r \leq \|A\|_p \|B\|_q\ ,\ 1/p + 1/q = 1/r$, and von-Neumann $|\mathrm{Tr}\left[AB\right]| \leq \|A\|_p \|B\|_q\ ,\ 1/p + 1/q = 1$ inequalities [128].

Let us also recall the $t = 1, 2$-order twirls $\left\langle \rho_\Lambda^{\otimes t} \right\rangle_{\mathcal{C}'}$ of the specific reference ensembles $\mathcal{C}'$ considered, with respect to input states $\rho \in \mathcal{B}[\mathcal{H}]$, given the identity I_d, and swap operators S_d acting on $\mathcal{H}$.
For the cHaar reference ensemble $\mathcal{C}' = \mathcal{C}(d, d_{\mathcal{E}})$,

$$\langle \rho_\Lambda \rangle_{\mathcal{C}(d,d_{\mathcal{E}})} = \frac{1}{d}\ \mathrm{Tr}\left[\rho\right]\ I_d \tag{D3}$$

$$\left\langle \rho_\Lambda^{\otimes 2} \right\rangle_{\mathcal{C}(d,d_{\mathcal{E}})} = \frac{1}{d^2} \frac{1}{1 - \frac{1}{d^2 d_{\mathcal{E}}^2}} \left[\left[\mathrm{Tr}^2\left[\rho\right] - \frac{1}{d d_{\mathcal{E}}} \mathrm{Tr}\left[\rho^2\right] \right] I_d^{\otimes 2}\ +\ \frac{1}{d_{\mathcal{E}}} \left[-\frac{1}{d d_{\mathcal{E}}} \mathrm{Tr}^2\left[\rho\right] + \mathrm{Tr}\left[\rho^2\right] \right] S_d \right]\ , \tag{D4}$$

and for the Depolarize reference ensemble $\mathcal{C}' = \mathcal{D}(d)$,

$$\langle \rho_\Lambda \rangle_{\mathcal{D}(d)} = \frac{1}{d}\ \mathrm{Tr}\left[\rho\right]\ I_d \tag{D5}$$

$$\left\langle \rho_\Lambda^{\otimes 2} \right\rangle_{\mathcal{D}(d)} = \frac{1}{d^2}\ \mathrm{Tr}^2\left[\rho\right]\ I_d^{\otimes 2}\ . \tag{D6}$$

Finally, we recall that t-th order twirls of ensembles $\mathcal{C}$, with respect to input states $\rho \in \mathcal{B}[\mathcal{H}]$, can be written in terms of reference ensembles $\mathcal{C}'$, as $\left\langle \rho_\Lambda^{\otimes t} \right\rangle_{\mathcal{C}} = \left\langle \rho_\Lambda^{\otimes t} \right\rangle_{\mathcal{C}'} + \Delta^{(t)}_{\mathcal{C}\mathcal{C}'}(\rho^{\otimes t})$, where the deviation has a p-norm of $\epsilon^{(t|p)}_{\mathcal{C}\mathcal{C}'}(\rho^{\otimes t}) = \|\Delta^{(t)}_{\mathcal{C}\mathcal{C}'}(\rho^{\otimes t})\|_p$.

2. Expectation Value Bias

Let us consider the channel-dependent function $\mathcal{F}(\Lambda) = \mathcal{L}_\Lambda$ of an expectation value, with respect to input states $\rho \in \mathcal{B}[\mathcal{H}]$ and hermitian observables $O \in \mathcal{B}[\mathcal{H}]$,

$$\mathcal{L}_\Lambda(\rho, O) = \mathrm{Tr}\left[\rho_\Lambda O\right]\ . \tag{D7}$$

First, we can study how much expectation value means $\langle \mathcal{L}_\Lambda \rangle_{\mathcal{C}}$ and $\langle \mathcal{L}_\Lambda \rangle_{\mathcal{C}'}$ differ. Concretely we bound the bias $\langle |\mathcal{L}_\Lambda - \mu_{\mathcal{C}'}| \rangle_{\mathcal{C}}$ where $\mu_{\mathcal{C}'} = \langle \mathcal{L}_\Lambda \rangle_{\mathcal{C}'}$.

Theorem 14 (Expectation Value Bias)**.** *Let $\mathcal{C},\mathcal{C}'$ be ensembles of quantum channels, let ρ be a quantum state, and let O be an observable that acts non-trivially on a subspace $\mathcal{W}\subseteq\mathcal{H}$ of dimension $d_{\mathcal{W}}$, with complement space $\overline{\mathcal{W}}$. The following bound on the expectation value $\mathcal{L}_\Lambda$ bias holds*

$$\langle|\mathcal{L}_\Lambda(\rho,O)-\mu_{\mathcal{C}'}(\rho,O)|\rangle_{\mathcal{C}} \;\leq\; \nu_{\mathcal{C}'}(\rho,O) \;+\; \alpha_{\mathcal{C}\mathcal{C}'}(\rho,O)\;, \tag{D8}$$

where, given $S_{d_{\mathcal{W}}}$ is the swap operator acting on $\mathcal{W}$, and $I_{d_{\overline{\mathcal{W}}}}$ is the identity operator acting on $\overline{\mathcal{W}}$, the inherent $\nu_{\mathcal{C}'}$ and design $\alpha_{\mathcal{C}\mathcal{C}'}$ expectation value bias terms are respectively,

$$\nu_{\mathcal{C}'}(\rho,O)=\sqrt{\mathrm{Tr}\left[\left(\langle\rho_\Lambda^{\otimes 2}\rangle_{\mathcal{C}'}-\langle\rho_\Lambda\rangle_{\mathcal{C}'}^{\otimes 2}\right)\left(S_{d_{\mathcal{W}}}\otimes I_{d_{\overline{\mathcal{W}}}}^{\otimes 2}\right)\right]}\;\sqrt{d_{\mathcal{W}}}\;\|O\|_\infty \tag{D9}$$

$$\alpha_{\mathcal{C}\mathcal{C}'}(\rho,O)=\sqrt{\mathrm{Tr}\left[\left(\Delta^{(2)}_{\mathcal{C}\mathcal{C}'}\,(\rho^{\otimes 2})-\Delta_{\mathcal{C}\mathcal{C}'}(\rho)\;\otimes\;\langle\rho_\Lambda\rangle_{\mathcal{C}'}-\langle\rho_\Lambda\rangle_{\mathcal{C}'}\otimes\;\Delta_{\mathcal{C}\mathcal{C}'}(\rho)\right)\left(S_{d_{\mathcal{W}}}\otimes I_{d_{\overline{\mathcal{W}}}}^{\otimes 2}\right)\right]}\;\sqrt{d_{\mathcal{W}}}\;\|O\|_\infty\;. \tag{D10}$$

Proof. To simplify our analysis of the bias of expectation values, we will denote the deviation of a transformed input $\rho\to\rho_\Lambda\in\mathcal{B}[\mathcal{H}]$, where $\Lambda\sim\mathcal{C}$, from its mean with respect to a reference ensemble $\mathcal{C}'$, as

$$\tilde{\rho}_{\Lambda|\mathcal{C}'}=\rho_\Lambda-\langle\rho_\Lambda\rangle_{\mathcal{C}'}\;\;, \tag{D11}$$

and the $t=1,2$-order moments of these deviations, with respect to an ensemble $\mathcal{C}$, are

$$\left\langle\tilde{\rho}_{\Lambda|\mathcal{C}'}\right\rangle_{\mathcal{C}}=\Delta_{\mathcal{C}\mathcal{C}'}(\rho) \tag{D12}$$

$$\left\langle\tilde{\rho}^{\otimes 2}_{\Lambda|\mathcal{C}'}\right\rangle_{\mathcal{C}}=\;\langle\rho_\Lambda^{\otimes 2}\rangle_{\mathcal{C}'}-\langle\rho_\Lambda\rangle_{\mathcal{C}'}^{\otimes 2}+\Delta^{(2)}_{\mathcal{C}\mathcal{C}'}(\rho^{\otimes 2})-\Delta_{\mathcal{C}\mathcal{C}'}(\rho)\otimes\langle\rho_\Lambda\rangle_{\mathcal{C}'}-\langle\rho_\Lambda\rangle_{\mathcal{C}'}\otimes\Delta_{\mathcal{C}\mathcal{C}'}(\rho)\;. \tag{D13}$$

The average expectation value bias can then be bounded using the following norm inequalities,

$$\langle|\mathcal{L}_\Lambda(\rho,O)-\mu_{\mathcal{C}'}(\rho,O)|\rangle_{\mathcal{C}}=\langle|\mathrm{Tr}\left[\rho_\Lambda O\otimes I\right]-\langle\mathrm{Tr}\left[\rho_\Lambda O\otimes I\right]\rangle_{\mathcal{C}'}|\rangle_{\mathcal{C}} \tag{D14}$$

$$=\left\langle\left|\mathrm{Tr}\left[\tilde{\rho}_{\Lambda|\mathcal{C}'}O\otimes I\right]\right|\right\rangle_{\mathcal{C}} \tag{D15}$$

$$=\left\langle\left|\mathrm{Tr}_{\mathcal{W}}\left[\mathrm{Tr}_{\overline{\mathcal{W}}}\left[\tilde{\rho}_{\Lambda|\mathcal{C}'}\right]O\right]\right|\right\rangle_{\mathcal{C}} \tag{D16}$$

$$\leq\left\langle\left\|\mathrm{Tr}_{\overline{\mathcal{W}}}\left[\tilde{\rho}_{\Lambda|\mathcal{C}'}\right]\right\|_1\;\|O\|_\infty\right\rangle_{\mathcal{C}} \tag{D17}$$

$$\leq\left\langle\|I_{d_{\mathcal{W}}}\|_2\;\left\|\mathrm{Tr}_{\overline{\mathcal{W}}}\left[\tilde{\rho}_{\Lambda|\mathcal{C}'}\right]\right\|_2\;\|O\|_\infty\right\rangle_{\mathcal{C}} \tag{D18}$$

$$=\left\langle\sqrt{\left\|\mathrm{Tr}_{\overline{\mathcal{W}}}\left[\tilde{\rho}_{\Lambda|\mathcal{C}'}\right]\right\|_2^2}\right\rangle_{\mathcal{C}}\;\|I_{d_{\mathcal{W}}}\|_2\;\|O\|_\infty \tag{D19}$$

$$\leq\sqrt{\left\langle\left\|\mathrm{Tr}_{\overline{\mathcal{W}}}\left[\tilde{\rho}_{\Lambda|\mathcal{C}'}\right]\right\|_2^2\right\rangle_{\mathcal{C}}}\;\|I_{d_{\mathcal{W}}}\|_2\;\|O\|_\infty \tag{D20}$$

$$=\sqrt{\left\langle\mathrm{Tr}_{\mathcal{W}}\left[\mathrm{Tr}^2_{\overline{\mathcal{W}}}\left[\tilde{\rho}_{\Lambda|\mathcal{C}'}\right]\right]\right\rangle_{\mathcal{C}}}\;\|I_{d_{\mathcal{W}}}\|_2\;\|O\|_\infty \tag{D21}$$

$$=\sqrt{\left\langle\mathrm{Tr}_{\mathcal{W}^{\otimes 2}}\left[\mathrm{Tr}_{\overline{\mathcal{W}}}\left[\tilde{\rho}_{\Lambda|\mathcal{C}'}\right]^{\otimes 2}S_{d_{\mathcal{W}}}\right]\right\rangle_{\mathcal{C}}}\;\|I_{d_{\mathcal{W}}}\|_2\;\|O\|_\infty \tag{D22}$$

$$=\sqrt{\left\langle\mathrm{Tr}_{(\mathcal{W}\otimes\overline{\mathcal{W}})^{\otimes 2}}\left[\tilde{\rho}^{\otimes 2}_{\Lambda|\mathcal{C}'}\;S_{d_{\mathcal{W}}}\otimes I^{\otimes 2}_{d_{\overline{\mathcal{W}}}}\right]\right\rangle_{\mathcal{C}}}\;\|I_{d_{\mathcal{W}}}\|_2\;\|O\|_\infty \tag{D23}$$

$$=\sqrt{\mathrm{Tr}\left[\left\langle\tilde{\rho}^{\otimes 2}_{\Lambda|\mathcal{C}'}\right\rangle_{\mathcal{C}}\left(S_{d_{\mathcal{W}}}\otimes I^{\otimes 2}_{d_{\overline{\mathcal{W}}}}\right)\right]}\;\|I_{d_{\mathcal{W}}}\|_2\;\|O\|_\infty \tag{D24}$$

$$=\sqrt{\mathrm{Tr}\left[\left(\langle\rho_\Lambda^{\otimes 2}\rangle_{\mathcal{C}'}-\langle\rho_\Lambda\rangle_{\mathcal{C}'}^{\otimes 2}+\Delta^{(2)}_{\mathcal{C}\mathcal{C}'}(\rho^{\otimes 2})-\Delta_{\mathcal{C}\mathcal{C}'}(\rho)\otimes\langle\rho_\Lambda\rangle_{\mathcal{C}'}-\langle\rho_\Lambda\rangle_{\mathcal{C}'}\otimes\Delta_{\mathcal{C}\mathcal{C}'}(\rho)\right)\left(S_{d_{\mathcal{W}}}\otimes I^{\otimes 2}_{d_{\overline{\mathcal{W}}}}\right)\right]}\;\|I_{d_{\mathcal{W}}}\|_2\;\|O\|_\infty \tag{D25}$$

$$\leq\left(\sqrt{\mathrm{Tr}\left[\left(\langle\rho_\Lambda^{\otimes 2}\rangle_{\mathcal{C}'}-\langle\rho_\Lambda\rangle_{\mathcal{C}'}^{\otimes 2}\right)\left(S_{d_{\mathcal{W}}}\otimes I^{\otimes 2}_{d_{\overline{\mathcal{W}}}}\right)\right]}\right. \tag{D26}$$

$$\left.+\sqrt{\mathrm{Tr}\left[\left(\Delta^{(2)}_{\mathcal{C}\mathcal{C}'}(\rho^{\otimes 2})-\Delta_{\mathcal{C}\mathcal{C}'}(\rho)\otimes\langle\rho_\Lambda\rangle_{\mathcal{C}'}-\langle\rho_\Lambda\rangle_{\mathcal{C}'}\otimes\Delta_{\mathcal{C}\mathcal{C}'}(\rho)\right)\left(S_{d_{\mathcal{W}}}\otimes I^{\otimes 2}_{d_{\overline{\mathcal{W}}}}\right)\right]}\;\right)\;\|I_{d_{\mathcal{W}}}\|_2\;\|O\|_\infty$$

$$=\nu_{\mathcal{C}'}(\rho,O)\;+\;\alpha_{\mathcal{C}\mathcal{C}'}(\rho,O)\;. \tag{D27}$$

□

Corollary 3 (Expectation Value Bias for Reference Ensembles)**.** *For specific observables O and reference ensembles $\mathcal{C}'$, the expectation value mean is*

$$\mu_{\mathcal{C}'}(\rho,O)=\frac{\operatorname{Tr}[\rho]\operatorname{Tr}[O]}{d}=\begin{cases}0 & \operatorname{Tr}[O]=0\ ,\ \mathcal{C}'=\mathcal{C}(d,d_{\mathcal{E}})\\ \frac{1}{d} & \operatorname{Tr}[O]=1\ ,\ \mathcal{C}'=\mathcal{C}(d,d_{\mathcal{E}})\\ 0 & \operatorname{Tr}[O]=0\ ,\ \mathcal{C}'=\mathcal{D}(d)\\ \frac{1}{d} & \operatorname{Tr}[O]=1\ ,\ \mathcal{C}'=\mathcal{D}(d)\end{cases}\ , \tag{D28}$$

and the inherent $\nu_{\mathcal{C}'}$ and design $\alpha_{\mathcal{C}\mathcal{C}'}$ expectation value bias terms are respectively,

$$\nu_{\mathcal{C}'}(\rho,O)=\begin{cases}\mathcal{O}\left(\sqrt{\frac{d_{\mathcal{W}}^2}{dd_{\mathcal{E}}}}\right)\ \|\rho\|_2\|O\|_\infty & \mathcal{C}'=\mathcal{C}(d,d_{\mathcal{E}})\\ 0 & \mathcal{C}'=\mathcal{D}(d)\end{cases} \tag{D29}$$

$$\alpha_{\mathcal{C}\mathcal{C}'}(\rho,O)\leq\sqrt{\epsilon_{\mathcal{C}\mathcal{C}'}^{(2|\diamond)}\|\rho\|_1^2+2\ \epsilon_{\mathcal{C}\mathcal{C}'}^{(1|\diamond)}\|\rho\|_1}\ \|O\|_\infty\ , \tag{D30}$$

under the assumption that the $t=1$-order twirl for a reference ensemble $\mathcal{C}'$ is proportional to the identity, $\langle\rho_\Lambda\rangle_{\mathcal{C}'}=\frac{\operatorname{Tr}[\rho]}{d}I$.

Proof. The expectation value mean follows directly from its definition $\mu_{\mathcal{C}'}(\rho,O)=\langle\mathcal{L}_\Lambda(\rho,O)\rangle_{\mathcal{C}'}=\operatorname{Tr}[\langle\rho_\Lambda\rangle_{\mathcal{C}'}O]$, and substituting in the assumption that the $t=1$-order twirl for the reference ensemble $\mathcal{C}'$ is $\langle\rho_\Lambda\rangle_{\mathcal{C}'}=\frac{\operatorname{Tr}[\rho]}{d}I$.

The expectation value bias inherent term follows from the deviation $\langle\rho_\Lambda^{\otimes 2}\rangle_{\mathcal{C}'}-\langle\rho_\Lambda\rangle_{\mathcal{C}'}^{\otimes 2}$ of the $t=2$ moment for specific reference ensembles $\mathcal{C}'$, given the forms of their respective $t=1,2$-order twirls $\langle\rho_\Lambda^{\otimes 2}\rangle_{\mathcal{C}'}, \langle\rho_\Lambda\rangle_{\mathcal{C}'}^{\otimes 2}$.

For the cHaar reference ensemble $\mathcal{C}'=\mathcal{C}(d,d_{\mathcal{E}})$,

$$\langle\rho_\Lambda^{\otimes 2}\rangle_{\mathcal{C}(d,d_{\mathcal{E}})}-\langle\rho_\Lambda\rangle_{\mathcal{C}(d,d_{\mathcal{E}})}^{\otimes 2}=\frac{1}{d^2}\frac{1}{1-\frac{1}{d^2d_{\mathcal{E}}^2}}\left[\left[\frac{1}{d^2d_{\mathcal{E}}^2}\operatorname{Tr}^2[\rho]-\frac{1}{dd_{\mathcal{E}}}\operatorname{Tr}\left[\rho^2\right]\right]I_d^{\otimes 2}\ +\ \frac{1}{d_{\mathcal{E}}}\left[-\frac{1}{dd_{\mathcal{E}}}\operatorname{Tr}^2[\rho]+\operatorname{Tr}\left[\rho^2\right]\right]S_d\right] \tag{D31}$$

$$=\frac{1}{d^2}\ \|\rho\|_2^2\ \left[\mathcal{O}\left(\frac{1}{dd_{\mathcal{E}}}\right)\ I_d^{\otimes 2}\ +\ \mathcal{O}\left(\frac{1}{d_{\mathcal{E}}}\right)\ S_d\right]\ , \tag{D32}$$

and for the Depolarize reference ensemble $\mathcal{C}'=\mathcal{D}(d)$,

$$\langle\rho_\Lambda^{\otimes 2}\rangle_{\mathcal{D}(d)}-\langle\rho_\Lambda\rangle_{\mathcal{D}(d)}^{\otimes 2}=\frac{1}{d^2}\ \operatorname{Tr}^2[\rho]\ I_d^{\otimes 2}\ -\ \left(\frac{1}{d}\ \operatorname{Tr}[\rho]\ I_d\right)^{\otimes 2}\ =\ 0\ . \tag{D33}$$

It follows that the expectation value bias inherent term is,

$$\nu_{\mathcal{C}'}(\rho,O)=\sqrt{\operatorname{Tr}\left[\left(\langle\rho_\Lambda^{\otimes 2}\rangle_{\mathcal{C}'}-\langle\rho_\Lambda\rangle_{\mathcal{C}'}^{\otimes 2}\right)\left(S_{d_{\mathcal{W}}}\otimes I_{d_{\overline{\mathcal{W}}}}^{\otimes 2}\right)\right]}\ \|I_{d_{\mathcal{W}}}\|_2\ \|O\|_\infty \tag{D34}$$

$$=\begin{cases}\sqrt{\frac{1}{d^2}\ \left[\mathcal{O}\left(\frac{1}{dd_{\mathcal{E}}}\right)\ \operatorname{Tr}\left[S_{d_{\mathcal{W}}}\otimes I_{d_{\overline{\mathcal{W}}}}^{\otimes 2}\right]\ +\ \mathcal{O}\left(\frac{1}{d_{\mathcal{E}}}\right)\ \operatorname{Tr}\left[I_{d_{\mathcal{W}}}^{\otimes 2}\otimes S_{d_{\overline{\mathcal{W}}}}\right]\right]}\ \sqrt{d_{\mathcal{W}}}\ \|\rho\|_2\ \|O\|_\infty & \mathcal{C}'=\mathcal{C}(d,d_{\mathcal{E}})\\ 0 & \mathcal{C}'=\mathcal{D}(d)\end{cases} \tag{D35}$$

$$=\begin{cases}\mathcal{O}\left(\sqrt{\frac{d_{\mathcal{W}}^2}{dd_{\mathcal{E}}}}\right)\ \|\rho\|_2\|O\|_\infty & \mathcal{C}'=\mathcal{C}(d,d_{\mathcal{E}})\\ 0 & \mathcal{C}'=\mathcal{D}(d)\end{cases}\ . \tag{D36}$$

The expectation value bias design term follows from the following norm inequalities,

$$\alpha_{\mathcal{C}\mathcal{C}'}(\rho,O)=\sqrt{\operatorname{Tr}\left[\left(\Delta_{\mathcal{C}\mathcal{C}'}^{(2)}(\rho^{\otimes 2})-\Delta_{\mathcal{C}\mathcal{C}'}(\rho)\otimes\langle\rho_\Lambda\rangle_{\mathcal{C}'}-\langle\rho_\Lambda\rangle_{\mathcal{C}'}\otimes\Delta_{\mathcal{C}\mathcal{C}'}(\rho)\right)\left(S_{d_{\mathcal{W}}}\otimes I_{d_{\overline{\mathcal{W}}}}^{\otimes 2}\right)\right]}\ \|I_{d_{\mathcal{W}}}\|_2\ \|O\|_\infty \tag{D37}$$

$$\leq\min_{\frac{1}{p}+\frac{1}{q}=1}\sqrt{\epsilon_{\mathcal{C}\mathcal{C}'}^{(2|p)}\ \|\rho\|_p^2\ +\ 2\ \epsilon_{\mathcal{C}\mathcal{C}'}^{(1|p)}\ \|\langle\rho_\Lambda\rangle_{\mathcal{C}'}\|_p\ \|\rho\|_p}\ \|I_{d_{\mathcal{W}}}\|_q\ \|I_d\|_q\ \|O\|_\infty \tag{D38}$$

$$\leq\sqrt{\epsilon_{\mathcal{C}\mathcal{C}'}^{(2|\diamond)}\|\rho\|_1^2\ +\ 2\ \epsilon_{\mathcal{C}\mathcal{C}'}^{(1|\diamond)}\|\rho\|_1}\ \|O\|_\infty\ . \tag{D39}$$

□

3. Expectation Value Variance

Second, we can study the variance of expectation values, which can be bounded as,

$$\sigma^2_{\mathcal{C}}(\rho, O) \;=\; \left\langle \mathcal{L}_\Lambda(\rho,O)^2 \right\rangle_{\mathcal{C}} - \left\langle \mathcal{L}_\Lambda(\rho,O) \right\rangle^2_{\mathcal{C}} \;\leq\; \left\langle \mathcal{L}_\Lambda(\rho,O)^2 \right\rangle_{\mathcal{C}} \;. \tag{D40}$$

Theorem 15 (Expectation Value Variance)**.** *Let $\mathcal{C},\mathcal{C}'$ be ensembles of quantum channels, let ρ be a quantum state, and let O be an observable. The following bound on the expectation value $\mathcal{L}_\Lambda$ variance holds*

$$\sigma^2_{\mathcal{C}}(\rho, O) \;\leq\; \varsigma_{\mathcal{C}'}(\rho, O) \;+\; \beta_{\mathcal{C}\mathcal{C}'}(\rho, O) \;, \tag{D41}$$

where, the inherent $\varsigma_{\mathcal{C}'}$ and design $\beta_{\mathcal{C}\mathcal{C}'}$ expectation value variance terms are respectively,

$$\varsigma_{\mathcal{C}'}(\rho, O) = \mathrm{Tr}\left[\left\langle \rho_\Lambda^{\otimes 2} \right\rangle_{\mathcal{C}'} O^{\otimes 2}\right] \tag{D42}$$

$$\beta_{\mathcal{C}\mathcal{C}'}(\rho, O) = \left|\mathrm{Tr}\left[\Delta^{(2)}_{\mathcal{C}\mathcal{C}'}(\rho^{\otimes 2})\; O^{\otimes 2}\right]\right| \;. \tag{D43}$$

Proof. The expectation value variance can be bounded, given its definition $\sigma^2_{\mathcal{C}}(\rho,O) \leq \left\langle \mathcal{L}_\Lambda(\rho,O)^2 \right\rangle_{\mathcal{C}} = \mathrm{Tr}\left[\left\langle \rho_\Lambda^{\otimes 2} \right\rangle_{\mathcal{C}} O^{\otimes 2}\right]$, and given twirls over ensembles $\mathcal{C}$ can be written in terms of reference ensembles $\mathcal{C}'$ as, $\left\langle \rho_\Lambda^{\otimes 2} \right\rangle_{\mathcal{C}} = \left\langle \rho_\Lambda^{\otimes 2} \right\rangle_{\mathcal{C}'} + \Delta^{(2)}_{\mathcal{C}\mathcal{C}'}(\rho^{\otimes 2})$. □

Corollary 4 (Expectation Value Variance for Reference Ensembles)**.** *For specific observables O and reference ensembles $\mathcal{C}'$, the expectation value variance is*

$$\sigma^2_{\mathcal{C}'}(\rho, O) = \begin{cases} \frac{1}{d^2}\frac{1}{1-\frac{1}{d^2 d_{\mathcal{E}}^2}}\left[\left[\frac{1}{d^2 d_{\mathcal{E}}^2}\mathrm{Tr}^2\left[\rho\right] - \frac{1}{d d_{\mathcal{E}}}\mathrm{Tr}\left[\rho^2\right]\right]\mathrm{Tr}^2\left[O\right] + \frac{1}{d_{\mathcal{E}}}\left[-\frac{1}{d d_{\mathcal{E}}}\mathrm{Tr}^2\left[\rho\right] + \mathrm{Tr}\left[\rho^2\right]\right]\mathrm{Tr}\left[O^2\right]\right] & \mathcal{C}' = \mathcal{C}(d, d_{\mathcal{E}}) \\ 0 & \mathcal{C}' = \mathcal{D}(d) \end{cases}, \tag{D44}$$

$$= \begin{cases} \mathcal{O}\left(\frac{1}{d^2 d_{\mathcal{E}}}\right)\; \|\rho\|_2^2\; \|O\|_2^2 & \mathcal{C}' = \mathcal{C}(d, d_{\mathcal{E}}) \\ 0 & \mathcal{C}' = \mathcal{D}(d) \end{cases} \;. \tag{D45}$$

and the inherent $\varsigma_{\mathcal{C}'}$ and design $\beta_{\mathcal{C}\mathcal{C}'}$ expectation value variance terms are respectively,

$$\varsigma_{\mathcal{C}'}(\rho, O) = \mathrm{Tr}\left[\left\langle \rho_\Lambda^{\otimes 2} \right\rangle_{\mathcal{C}'} O^{\otimes 2}\right] = \begin{cases} \mathcal{O}\left(\frac{1}{d^2 d_{\mathcal{E}}}\right)\; \|\rho\|_2^2\; \|O\|_2^2 & \mathrm{Tr}[O] = 0 \;,\; \mathcal{C}' = \mathcal{C}(d, d_{\mathcal{E}}) \\ \mathcal{O}\left(\frac{1}{d^2}\right) & \mathrm{Tr}[O] = 1 \;,\; \mathcal{C}' = \mathcal{C}(d, d_{\mathcal{E}}) \\ 0 & \mathrm{Tr}[O] = 0 \;,\; \mathcal{C}' = \mathcal{D}(d) \\ \frac{1}{d^2} & \mathrm{Tr}[O] = 1 \;,\; \mathcal{C}' = \mathcal{D}(d) \end{cases} \tag{D46}$$

$$\beta_{\mathcal{C}\mathcal{C}'}(\rho, O) = \left|\mathrm{Tr}\left[\Delta^{(2)}_{\mathcal{C}\mathcal{C}'}(\rho^{\otimes 2}) O^{\otimes 2}\right]\right| \leq \; \min\left\{\|\rho\|_2^2\; \|O\|_2^2\; \epsilon^{(2|\mathrm{op})}_{\mathcal{C}\mathcal{C}'} \quad,\quad \|\rho\|_1^2\; \|O\|_\infty^2\; \epsilon^{(2|\diamond)}_{\mathcal{C}\mathcal{C}'}\right\} \;. \tag{D47}$$

Proof. The expectation value variance inherent term follows from the following norm inequalities, for specific reference ensembles $\mathcal{C}'$, given the forms of the $t=2$-order twirls $\left\langle \rho_\Lambda^{\otimes 2} \right\rangle_{\mathcal{C}'}$,

$$\varsigma_{\mathcal{C}'}(\rho, O) = \mathrm{Tr}\left[\left\langle \rho_\Lambda^{\otimes 2} \right\rangle_{\mathcal{C}'} O^{\otimes 2}\right] \tag{D48}$$

$$= \begin{cases} \frac{1}{d^2}\frac{1}{1-\frac{1}{d^2 d_{\mathcal{E}}^2}}\left[\left[\mathrm{Tr}^2\left[\rho\right] - \frac{1}{d d_{\mathcal{E}}}\mathrm{Tr}\left[\rho^2\right]\right]\mathrm{Tr}^2\left[O\right] \;+\; \frac{1}{d_{\mathcal{E}}}\left[-\frac{1}{d d_{\mathcal{E}}}\mathrm{Tr}^2\left[\rho\right] + \mathrm{Tr}\left[\rho^2\right]\right]\mathrm{Tr}\left[O^2\right]\right] & \mathcal{C}' = \mathcal{C}(d, d_{\mathcal{E}}) \\ \frac{\mathrm{Tr}^2[\rho]\mathrm{Tr}^2[O]}{d^2} & \mathcal{C}' = \mathcal{D}(d) \end{cases} \tag{D49}$$

$$= \begin{cases} \mathcal{O}\left(\frac{1}{d^2 d_{\mathcal{E}}}\right)\; \|\rho\|_2^2\; \|O\|_2^2 & \mathrm{Tr}[O] = 0 \;,\; \mathcal{C}' = \mathcal{C}(d, d_{\mathcal{E}}) \\ \mathcal{O}\left(\frac{1}{d^2}\right) & \mathrm{Tr}[O] = 1 \;,\; \mathcal{C}' = \mathcal{C}(d, d_{\mathcal{E}}) \\ 0 & \mathrm{Tr}[O] = 0 \;,\; \mathcal{C}' = \mathcal{D}(d) \\ \frac{1}{d^2} & \mathrm{Tr}[O] = 1 \;,\; \mathcal{C}' = \mathcal{D}(d) \end{cases} \;. \tag{D50}$$

The expectation value variance design term follows from the following norm inequalities,

$$\beta_{\mathcal{C}\mathcal{C}'}(\rho, O) = \left|\mathrm{Tr}\left[\Delta^{(2)}_{\mathcal{C}\mathcal{C}'}(\rho^{\otimes 2}) O^{\otimes 2}\right]\right| \leq \min_{\frac{1}{p}+\frac{1}{q}=1}\; \epsilon^{(2|p)}_{\mathcal{C}\mathcal{C}'}(\rho^{\otimes 2})\; \|O\|_q^2 \tag{D51}$$

$$\leq \; \min_{\frac{1}{p}+\frac{1}{q}=1}\; \|\rho\|_p^2\; \|O\|_q^2\; \epsilon^{(2|p)}_{\mathcal{C}\mathcal{C}'} \tag{D52}$$

$$\leq \; \min\left\{\|\rho\|_2^2\; \|O\|_2^2\; \epsilon^{(2|\mathrm{op})}_{\mathcal{C}\mathcal{C}'} \quad,\quad \|\rho\|_1^2\; \|O\|_\infty^2\; \epsilon^{(2|\diamond)}_{\mathcal{C}\mathcal{C}'}\right\} \;. \tag{D53}$$

□