# Typical Correlation Length of Sequentially Generated Tensor Network States

Daniel Haag ,[1,2,3,*] Flavio Baccari ,[1,3] and Georgios Styliaris [1,3]

[1] *Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Str. 1, 85748 Garching, Germany*

[2] *Physik-Department, Technische Universität München, James-Franck-Str. 1, 85748 Garching, Germany*

[3] *Munich Center for Quantum Science and Technology (MCQST), Schellingstr. 4, 80799 München, Germany*

The complexity of quantum many-body systems is manifested in the vast diversity of their correlations, making it challenging to distinguish the generic from the atypical features. This can be addressed by analyzing correlations through ensembles of random states, chosen to faithfully embody the relevant physical properties. Here, we focus on spins with local interactions, whose correlations are extremely well captured by tensor network states. Adopting an operational perspective, we define ensembles of random tensor network states in one and two spatial dimensions that admit a sequential generation. As such, they directly correspond to outputs of quantum circuits with a sequential architecture and random gates. In one spatial dimension, the ensemble explores the entire family of matrix product states, while in two spatial dimensions, it corresponds to random isometric tensor network states. We extract the scaling behavior of the average correlations between two subsystems as a function of their distance. Using elementary concentration results, we then deduce the typical case for measures of correlation such as the von Neumann mutual information and a measure arising from the Hilbert-Schmidt norm. We find for all considered cases that the typical behavior is an exponential decay (for both one and two spatial dimensions). We observe the consistent emergence of a correlation length that depends only on the underlying spatial dimension and not the considered measure. Remarkably, increasing the bond dimension leads to a higher correlation length in one spatial dimension but has the opposite effect in two spatial dimensions.

## I. INTRODUCTION

The behavior of correlations in quantum many-body systems is an inherently difficult problem to characterize. Specifying a generic *n*-particle state requires exponentially many parameters, a fact which reflects the enormous variety of correlations possible in the quantum realm. Nonetheless, significant insights can be gained about the nature of correlations by utilizing random ensembles of states. A celebrated result along this direction shows that random states sampled uniformly from the full Hilbert space of an *n*-particle system typically exhibit strong correlations, as manifested by a volume-law behavior of the entanglement entropy [1–5]. However, there is by now clear evidence that the set of physically relevant states constitutes an exponentially small subset of the full Hilbert space of an *n*-particle system [6], bringing into question

the relevance and utility of conclusions obtained under the assumption of uniform sampling from the full Hilbert space.

For quantum spin systems with local interactions, tensor network states have been exceedingly successful at capturing relevant properties [7]. They exhibit an area law for the entanglement entropy by construction and are, therefore, good candidates to represent many physically relevant many-body states. Their pre-eminent one-dimensional representatives, matrix product states (MPS), have been shown to represent faithfully ground states of gapped local Hamiltonians [8–10] and have given rise to the complete classification of topological phases of matter in one dimension [11,12]. MPS have been generalized to their counterparts in two (or more) spatial dimensions, projected entangled pair states (PEPS). While only a weaker link between local Hamiltonians and PEPS has been proven rigorously, two-dimensional PEPS are known to efficiently represent a wide class of strongly correlated states [7,13], including states with power-law [14] and topological correlations [15–17].

The importance of defining ensembles of random tensor network states for the purpose of exploring typical properties of physically relevant states has been recognized more

---

*daniel.haag@tum.de

than a decade ago [18]. MPS ensembles have been utilized to gain insights into, among other things, the typicality of expectation values of local observables [18,19], equilibration under Hamiltonian time evolution [20], the entropy of subsystems [21], nonstabilizerness [22], and, most relevant for this work, the behavior of correlations [23–27]. In particular, correlation functions of random inhomogeneous MPS (that is, MPS whose local tensors can be different) were shown to exhibit almost surely an exponential decay [24,28]. A qualitatively similar behavior was observed also for correlation functions of translation-invariant MPS and PEPS with random Gaussian entries [25]. Instead of incorporating the randomness directly at the level of states, one can also consider random local Hamiltonians and examine their ground states. The typical behavior of correlations for this case was found to depend on the nature of randomness, allowing for both long and short-range correlated states [23,29,30].

Here, we approach the problem of typical correlations in random MPS and PEPS from a more operational point of view. We introduce families of inhomogeneous random tensor network states that arise from a sequential generation in a quantum computer. Such ensembles are, by definition, directly connected to the study of quantum circuits with a sequential architecture and random gates, where each unitary gate is independently chosen randomly from the uniform (Haar) measure. In the one-dimensional case, every MPS admits such a preparation [31], where the bond dimension dictates the number of overlapping qudits between any two successive gates. In the two-dimensional setting, our ensemble can be understood as being uniform over the space of so-called isometric tensor network states (isoTNS) [32], which are PEPS with a given bond dimension. In this case, the resulting family of random circuits is composed of two-dimensional circuits with local overlapping gates, each resembling a tile acting on a neighborhood of qudits [33]. Although isoTNS are only a subfamily of PEPS, they are known to contain a rich variety of strongly correlated states such as topological models [34].

For the above ensembles, we study the scaling behavior of the average correlations between two subsystems as a function of their distance. We then utilize this average behavior of correlations to deduce the typical case via concentration inequalities. Instead of using well-known correlation functions, we perform the analysis using a measure of correlation arising from the Hilbert-Schmidt norm. Although, in a generic many-body setting, such a measure might have undesirable properties, we show that it is particularly suited in the context of tensor network states because it bounds the trace distance as well as all connected correlation functions. For MPS, we also consider the Rényi-$\alpha$ mutual information. Given a technical conjecture, we compute the average correlations for all integer values of $\alpha \geq 1$. We then use those results to retrieve

the von Neumann mutual information [35] via analytic continuation.

We confirm analytically the common intuition that inhomogeneous MPS typically exhibit exponentially decaying correlations. We show that a single common correlation length $\xi_{1D}$ persists among different measures of correlation. We obtain a similar quantitative conclusion for two-dimensional isoTNS, where we observe the emergence of a different correlation length $\xi_{2D}$ that is also consistent among different measures of correlation. $\xi_{1D}$ and $\xi_{2D}$ have a surprisingly weak dependence on the underlying bond dimension $D$ of the tensor network, and both MPS and isoTNS remain short-range correlated for all values of $D$. However, $\xi_{1D}$ and $\xi_{2D}$ have exactly opposite behaviors when the bond dimension is varied; $\xi_{1D}$ monotonically increases, while $\xi_{2D}$ monotonically decreases. Our findings also establish that exponentially decaying correlations are typical for the family of (inhomogeneous) isoTNS and consequently for the random states produced by the corresponding quantum circuit architecture.

While an exponential decay of correlations is not an unexpected result for MPS, the weak dependence of the average correlation length on the bond dimension is particularly surprising. Intuitively, one might expect that a higher bond dimension allows for a more effective spreading of correlations. Our results show that that is not the case for most MPS. The starkly different behavior of $\xi_{1D}$ and $\xi_{2D}$ is also remarkable. While increasing the bond dimension affects the typical spreading of correlations only modestly in one dimension, it is detrimental to the average correlation length in two dimensions.

The paper is structured as follows. In Sec. II, we introduce our families of sequentially generated tensor network states and the main technical tools required to compute their average properties. In Sec. III, we summarize our results for both MPS and isoTNS. In Secs. IV and V, we, respectively, discuss our findings for MPS and isoTNS in detail. Lastly, we devote Sec. VI to final observations and potential followups to our work.

## II. PRELIMINARIES

In this section, we introduce the main technical concepts that will be needed throughout the paper. In Sec. II A, we review the relevant families of sequentially generated tensor network states in one and two dimensions. Section II B is devoted to the measures of correlation we are interested to estimate. In Sec. II C, we explain how to compute averages with respect to the Haar measure. Lastly, Sec. II D introduces the graphical notation we will use to present and prove our results.

### A. Tensor network states

In one dimension, the pre-eminent tensor network structure are matrix product states (MPS) [7]. An *n*-particle

MPS with open boundary conditions and local (physical) dimension $d$ is given by

$$|\psi\rangle = \sum_{i_1,\ldots,i_n} \langle L|A^{(1)}_{i_1}\cdots A^{(n)}_{i_n}|R\rangle |i_1\cdots i_n\rangle, \qquad (1)$$

where $A^{(j)} \in \mathbb{C}^{D\times D}$, $|L\rangle \in \mathbb{C}^D$ is the left boundary condition, and $|R\rangle \in \mathbb{C}^D$ is the right boundary condition. $D$ is called the bond dimension of the MPS. A commonly used graphical notation for Eq. (1) is
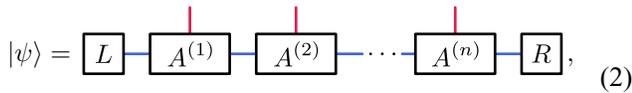
$$|\psi\rangle = \boxed{L}-\boxed{A^{(1)}}-\boxed{A^{(2)}}-\cdots-\boxed{A^{(n)}}-\boxed{R}, \qquad (2)$$

where vertical (red) legs represent physical space indices $(\mathbb{C}^d)$, and horizontal (blue) legs represent bond space indices $(\mathbb{C}^D)$.

From its definition, it might not be evident how an MPS can be generated because each tensor $A$ does not necessarily correspond to a physical process. However, the representation of an MPS in terms of tensors is not unique. This can be resolved by imposing a convenient canonical form [36]. Any MPS in such a canonical form can be seen as a state generated sequentially by applying unitary matrices $U^{(1)}, \ldots, U^{(n)} \in U(dD)$ to a product state initialized in $|0\rangle^{\otimes n}$ for the physical space and $|0\rangle$ for the bond space [31]. The resulting state is given by
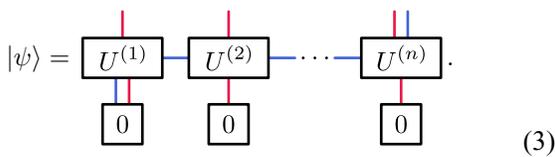
$$|\psi\rangle = \boxed{U^{(1)}}-\boxed{U^{(2)}}-\cdots-\boxed{U^{(n)}}. \qquad (3)$$

Note that the final site has dimension $dD$, while all other sites have dimension $d$. As we will see later, the final site will not play a significant role in our analysis, making its different dimension not an issue. In Fig. 1(a1), we sketch an equivalent representation of sequential generation in terms of isometries instead of unitary matrices.

The family of MPS is thus equivalent to states resulting from quantum circuits that have a sequential architecture and act on input product states. The architecture is a consequence of the connectivity of the MPS network [see Fig. 1(a1)]. In this picture, larger bond dimensions translate to wider gates, each acting on $1 + \lceil \log_d(D)\rceil$ qudits.

For example, for $D = d^2$, one has

$$|\psi\rangle = \quad (4)$$
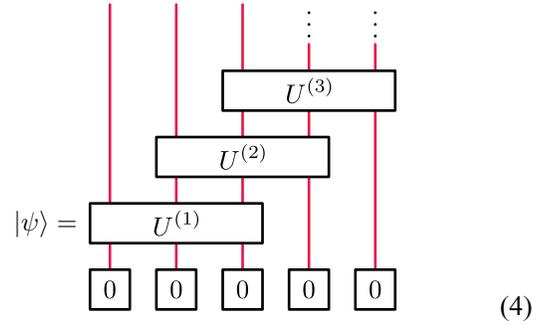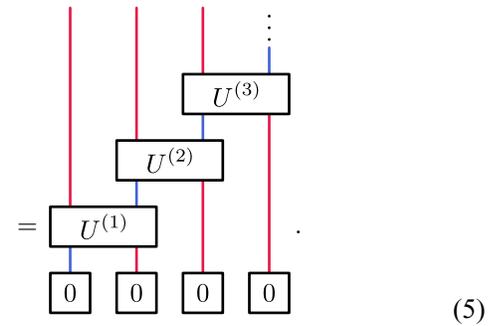
$$= \quad . \quad (5)$$

Naturally, using this correspondence, all of our results can be translated to the language of quantum circuits with the described architecture.

Projected entangled-pair states (PEPS) are the generalization of MPS to two (or more) dimensions [7]. Because no simple generalization of the sequential generation of MPS to arbitrary PEPS is known, we restrict ourselves to the rich family of two-dimensional isometric tensor network states (isoTNS), which were first defined in Ref. [32] (see also Ref. [37]).

Much like MPS, isoTNS can be generated sequentially by applying unitary matrices to a product state initialized in $|0\rangle^{\otimes mn}$ for the physical space [33], where $m$ denotes the number of rows and $n$ the number of columns of the underlying rectangular lattice. We will use the sequential generation sketched in Fig. 1(b1), which is a generalization of the one proposed in Ref. [33]. Each box corresponds to an isometry that arises from a unitary matrix $U^{(i,j)} \in U(dD^2)$. In particular, isometries in the bulk can be drawn as
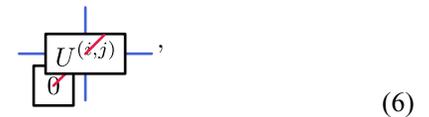
$$\boxed{U^{(i,j)}}, \qquad (6)$$

as indicated in Fig. 1(b2). The diamond in Fig. 1(b1) indicates the so-called orthogonality center of the isoTNS. Its row and column constitute the orthogonality hypersurface, which can be treated like an MPS. That is, if an operator is supported only on the orthogonality hypersurface,

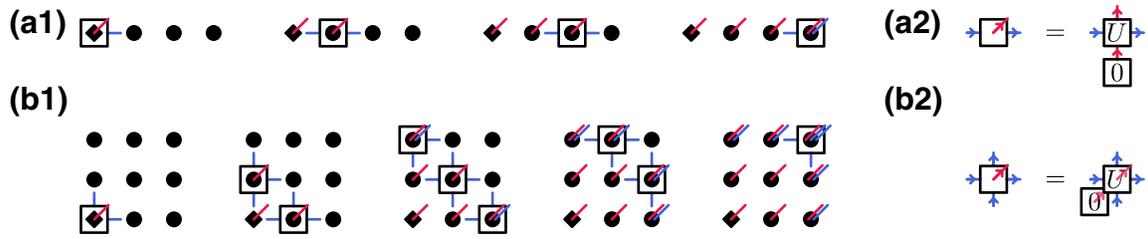FIG. 1. Sequential generation of MPS and isoTNS. Each circle represents a site of the finalized state and boxes represent the isometries of the sequential generation. (a1) Sequential generation of an MPS with physical dimension $d$ and bond dimension $D$. The diamond indicates the origin of the process. (a2) Each isometry arises from a unitary matrix of $U(dD)$, with input and output as shown. The (blue) ancillary system is initialized at the first step of the sequential generation, transferred along the process, and accumulated at the final step. (b1) Sequential generation of an isoTNS with physical dimension $d$ and bond dimension $D$. In addition to indicating the origin of the process, the diamond also indicates the orthogonality center of the isoTNS. (b2) Each isometry arises from a unitary matrix of $U(dD^2)$. Ancillary systems are initialized and eventually accumulated at the boundary of the isoTNS at different steps of the sequential generation.

its expectation value with respect to the isoTNS reduces to that of the underlying MPS [32]. Although isoTNS of a given bond dimension form by definition only a subclass of PEPS, they are known to contain states with a rich structure of correlations, such as topological models [34]. On top, their properties make isoTNS a suitable candidate for studying correlations analytically, which is otherwise a generally challenging task in more than one dimension.

IsoTNS correspond to quantum circuits on a two-dimensional grid with local overlapping gates, which now resemble tiles. Increasing the bond dimension translates to larger tile sizes and overlaps, as in the MPS case. The corresponding architecture is dictated by the connectivity of the isoTNS network [see Fig. 1(b1)], and it is tedious (although straightforward), which is why we refer the reader to Ref. [33] for details.

### B. Quantifying correlations

Correlations express that knowledge about one subsystem can convey information about another. They are quantified by different measures that frequently arise from an information-theoretic perspective and are based on operationally motivated tasks. A prime example is the von Neumann mutual information [35]

$$I(A : B) = S(\varrho_A) + S(\varrho_B) - S(\varrho_{AB}), \qquad (7)$$

where

$$S(\varrho) = -\operatorname{tr}[\varrho \log(\varrho)] \qquad (8)$$

is the von Neumann entropy. $A$ and $B$ are two disjoint subsystems of a larger system, and $\varrho_A$ and $\varrho_B$ denote the marginals of $\varrho_{AB}$. The von Neumann mutual information captures the total (classical and quantum) amount of correlations between $A$ and $B$ as it is equal to the minimum rate of randomness required to asymptotically turn $\varrho_{AB}$

into a product state [38]. It is also a non negative quantity and nonincreasing under local operations [35], both of which are desirable properties for a measure of correlation. The latter means that a quantum channel [35] acting on $A$ or $B$ alone (for example, by discarding part of a subsystem) cannot increase $I(A : B)$. Unfortunately, the analytical treatment of the von Neumann mutual information is impractical because computing the logarithm of $\varrho$ generally requires the knowledge of its full spectrum.

To overcome this issue, an alternative is to consider a particular Rényi-$\alpha$ generalization of the mutual information

$$I_\alpha(A : B) = S_\alpha(\varrho_A) + S_\alpha(\varrho_B) - S_\alpha(\varrho_{AB}), \qquad (9)$$

where

$$S_\alpha(\varrho) = \frac{1}{1 - \alpha} \log[\operatorname{tr}(\varrho^\alpha)] \qquad (10)$$

is the Rényi-$\alpha$ entropy. As is apparent from the definition, for integer values of $\alpha$, its evaluation is considerably simpler. The Rényi-$\alpha$ mutual information has been investigated in the context of conformal field theories [39–41], free fermions [42], and quantum dynamics [43,44]. We will use later that the limit of $\alpha \to 1$ of $I_\alpha(A : B)$ recovers the von Neumann mutual information. The mentioned positive aspects notwithstanding, unlike the von Neumann mutual information, Eq. (9) does not arise from a (generalized) divergence [45], and the Rényi-$\alpha$ mutual information can be negative [46,47] and increasing under local operations. It is thus hard to interpret it as a proper measure of correlation in general. Nevertheless, for certain families of initial states (see, for example, Ref. [44]), monotonicity and non negativity can be restored. Henceforth, we will mostly focus on the case of $\alpha = 2$, but we will also consider an analytic continuation on positive integer values of $\alpha$. As we will show, in the present context of

tensor network states, the case of $\alpha = 2$ appropriately captures the decay of correlations at large distances between subsystems $A$ and $B$ with little effort.

In addition to the previous quantities, we would also like to probe the trace distance

$$T(A : B) = \frac{1}{2} \|\varrho_{AB} - \varrho_A \otimes \varrho_B\|_1, \qquad (11)$$

where $\|\cdot\|_p$ denotes the Schatten $p$-norm [48]. For an operator $X$, the Schatten $p$-norm is given by

$$\|X\|_p = \mathrm{tr}\left[\left(X^\dagger X\right)^{p/2}\right]^{1/p}. \qquad (12)$$

$T(A : B)$ has a well-known operational interpretation as it quantifies the optimal distinguishability between $\varrho_{AB}$ and the product of its marginals $\varrho_A \otimes \varrho_B$ by a two-element generalized global measurement [49]. Moreover, the trace distance upper bounds the (properly normalized) connected correlation function [45]:

$$T(A : B) \geq 2 \frac{|\langle M_A \otimes M_B \rangle - \langle M_A \rangle \langle M_B \rangle|}{\|M_A\|_\infty \|M_B\|_\infty} \qquad (13)$$

Although the bound can be tight, the two quantities are different whenever product measurements are ineffective in distinguishing $\rho_{AB}$ from $\rho_A \otimes \rho_B$, a fact used in quantum data hiding [50].

As one expects from its operational interpretation, the trace distance satisfies the monotonicity property under local operations [49]. However, $T(A : B)$ is usually hard to compute exactly. We will now argue that investigating

$$N(A : B) = \|\varrho_{AB} - \varrho_A \otimes \varrho_B\|_2^2 \qquad (14)$$

meaningfully probes $T(A : B)$ for tensor network states with constant (that is, size-independent) bond dimension, all while being much simpler to treat.

In general, for mixed many-body states, the two measures can have vast disagreement because it holds [48] that

$$\|X\|_2 \leq \|X\|_1 \leq \sqrt{\mathrm{rank}(X)}\|X\|_2. \qquad (15)$$

Both bounds are tight, and the upper bound is saturated for $X \propto I$. As such, for arbitrary mixed states of an exponentially large Hilbert space, the factor $\mathrm{rank}(X)$ can render the upper bound useless. Crucially, in this work, we investigate (random) tensor network states with fixed bond dimension $D$. Let $\partial R$ denote the boundary of a system $R$ and $|\partial R|$ its size (number of sites). The ranks of $\varrho_A$ and $\varrho_B$ are, respectively, upper bounded by $D^{|\partial A|}$ and $D^{|\partial B|}$, and that of $\varrho_{AB}$ is

upper bounded by $D^{|\partial A| + |\partial B|}$ [7]. Thus,

$$\mathrm{rank}(\varrho_{AB} - \varrho_A \otimes \varrho_B) \leq 2D^{|\partial A| + |\partial B|}, \qquad (16)$$

yielding the bound

$$\frac{1}{2}\sqrt{N(A : B)} \leq T(A : B) \leq \sqrt{\frac{D^{|\partial A| + |\partial B|}}{2}}\sqrt{N(A : B)}. \quad (17)$$

For MPS (one dimension) with connected subsystems $A$ and $B$, the bound reads

$$\frac{1}{2}\sqrt{N(A : B)} \leq T(A : B) \leq \frac{D^2}{\sqrt{2}}\sqrt{N(A : B)}. \qquad (18)$$

That is, the bound is independent of the sizes of $A$ and $B$, unlike in the generic case of Eq. (15). This suggests that, for reasonably small bond dimension, $N(A : B)$ is a reliable probe of correlations [as quantified by $T(A : B)$] between subsystems $A$ and $B$. We will expand on this point later.

### C. $k$-fold twirl

Let $X$ be an operator acting on $(\mathbb{C}^q)^{\otimes k}$. The $k$-fold twirl of $X$ with respect to the Haar measure on the unitary group $U(q)$ is defined [51–53] as

$$\mathcal{T}_U^{(k)}(X) = \int dU\, U^{\otimes k} X \left(U^\dagger\right)^{\otimes k}. \qquad (19)$$

One can employ the Schur-Weyl duality for unitary groups to show [51,54] that

$$\mathcal{T}_U^{(k)}(X) = \sum_{\sigma, \tau \in S_k} \mathrm{Wg}\left(\sigma \tau^{-1}, q\right) P_\sigma^{(q)} \mathrm{tr}\left[X \left(P_\tau^{(q)}\right)^T\right], \quad (20)$$

where

$$P_\pi^{(q)} : v_1 \otimes \cdots \otimes v_k \mapsto v_{\sigma^{-1}(1)} \otimes \cdots \otimes v_{\sigma^{-1}(k)} \qquad (21)$$

is the representation of $\pi \in S_k$ on $(\mathbb{C}^q)^{\otimes k}$, where $S_k$ is the symmetric group. $\mathrm{Wg}\left(\sigma \tau^{-1}, q\right) = \left(G^{-1}\right)_{\sigma \tau}$ [55] is the Weingarten function, where $G \in \mathbb{R}^{k! \times k!}$ denotes the Gram matrix whose entries are given by

$$G_{\sigma \tau} = \mathrm{tr}\left[P_\sigma^{(q)} \left(P_\tau^{(q)}\right)^T\right] = q^{\#\left(\sigma \tau^{-1}\right)}. \qquad (22)$$

Here, $\#(\pi)$ counts the number of cycles in the decomposition of $\pi \in S_k$ into disjoint cycles. Thus, $\mathrm{Wg}(\pi, q)$ depends only on the conjugacy class of $\pi$ [54]. In Appendix A, we show how to obtain Eq. (20) from Eq. (19) by using a result of Ref. [54].

### D. Graphical notation

In this section, we introduce the graphical notation used throughout this paper. To keep the images compact, we employ the operator-vector correspondence. Let $\{|i\rangle\}$ denote the standard basis of $\mathbb{C}^q$. Then, the operator-vector correspondence [48] is defined by

$$\mathrm{vec}(|i\rangle\langle j|) = |i\rangle \otimes |j\rangle \qquad (23)$$

and extended linearly to the vector space at large.

Because we consider the standard (product) basis to be fixed, we do not distinguish between tensors (as multidimensional arrays) and their basis-independent counterparts (such as vectors and operators). Let $X$ be an operator acting on $(\mathbb{C}^q)^{\otimes k}$. Using the operator-vector correspondence, we denote it by

$$\overset{\parallel}{\boxed{X}} = \mathrm{vec}(X). \qquad (24)$$

Note that the orientation of the legs does not have any meaning in our images. That is,

$$\overset{\parallel}{\boxed{X}} = \boxed{X}\!\!= \;=\boxed{X}\overset{\parallel}{} = =\!\boxed{X}. \qquad (25)$$

When we need the transpose of an operator, we will explicitly use

$$\overset{\parallel}{\boxed{X^T}} = \mathrm{vec}(X^T). \qquad (26)$$

As such, when we contract two operators $X$ and $Y$, we mean the trace of their product:

$$\boxed{X}\!\!=\!\!\boxed{Y} = \mathrm{tr}(XY) \qquad (27)$$

Let us state the two most prominent operators we will encounter. We will see

$$\overset{\parallel}{\boxed{P}}\!\!-\!\sigma = \mathrm{vec}\left(P_\sigma^{(q)}\right), \qquad (28)$$

where the horizontal (green) leg is permutation valued, and

$$\overset{\parallel}{\boxed{0}} = \mathrm{vec}\left(|0\rangle\langle 0|^{\otimes k}\right). \qquad (29)$$

Their relevant contractions are

$$\sigma\!-\!\boxed{P}\!\!=\!\!\boxed{P}\!-\!\tau = \mathrm{tr}\left(P_\sigma^{(q)} P_\tau^{(q)}\right) = q^{\#(\sigma\tau)} \qquad (30)$$

and

$$\boxed{0}\!\!=\!\!\boxed{P}\!-\!\sigma = \mathrm{tr}\left(|0\rangle\langle 0|^{\otimes k} P_\sigma^{(q)}\right) = 1. \qquad (31)$$

Moving forward, we will not explicitly write the operator vec, as it shall be clear from the context.

With the definition of the Weingarten matrix,

$$\sigma\!-\!\boxed{W}\!-\!\tau = \mathrm{Wg}\left(\sigma\tau^{-1}, q\right), \qquad (32)$$

we can then write the $k$-fold twirl [see Eq. (20)] as

$$\mathcal{T}_U^{(k)}(X) = \boxed{P^T}\!-\!\boxed{W}\!-\!\boxed{P}, \\ \overset{\boxed{X}}{} \qquad (33)$$

where the contraction of two green legs corresponds to a summation over the permutations of $S_k$.

## III. SUMMARY OF RESULTS

In this paper, we analyze the average behavior of correlations in random tensor network states. Through the average, we obtain conclusions about the generic case. Our work focuses on the disordered case, that is, the case where each local tensor is independent. Our setting can be equivalently understood as an investigation of correlations in states resulting from quantum circuits with a sequential architecture and random gates.

In one dimension, generic MPS are known to exhibit exponentially decaying correlations in the translation-invariant case [56]. This is due to the fact that injectivity is a generic property [7]. On the other hand, injectivity alone is not enough to guarantee an exponential decay of correlations for an inhomogeneous sequence of tensors. Nevertheless, the exponential decay of correlations is widely expected to persist without translational invariance, but it has never been rigorously studied so far in this setting.

In two (or more) dimensions, the landscape of correlations is much richer. For instance, already in two dimensions, certain PEPS corresponding to thermal states

of classical models are known to exhibit power-law correlations [14]. Moreover, prominent topological states, such as quantum double models [57] (which include the toric code) and string-net models [16,17], admit a description in terms of PEPS. On the other hand, for translation-invariant PEPS whose tensors' entries are drawn from a Gaussian measure, it is known that correlations typically decay exponentially [25].

Computing correlations in higher-dimensional systems usually poses a significant challenge because they can be mediated through different paths connecting the two subsystems of interest. Here, we restrict our analysis to two-dimensional isoTNS. This rich class of tensor network states is relevant in both the analytical and the numerical context [58–62], all while admitting a simple physical interpretation through sequential generation [see Fig. 1(b1)]. Moreover, its mathematical properties make the analytical study of correlations in two dimensions tractable.

For isoTNS, it is expected that correlations between two subsystems decay exponentially if they are both on the orthogonality hypersurface because the calculation reduces to the contraction of an MPS [32]. Nonetheless, isoTNS can represent a rich variety of topological models, as all string-net models admit an exact and explicit description in terms of isoTNS [34] (on the appropriate underlying lattice). This motivates us to study the typical behavior of correlations in isoTNS, particularly between subsystems that extend beyond the orthogonality hypersurface.

To investigate the decay of correlations in our two families of random tensor network states, one must specify the ensembles to draw from. Here we adopt an operational perspective and relate our measures of randomness directly to the sequential generation process. Because that is defined with respect to isometries, one can incorporate randomness at the level of the underlying unitary matrices. A natural choice is to draw each unitary matrix from the Haar measure on the appropriate unitary group. This approach was introduced for MPS in Ref. [18] (see also Ref. [19]), and it can directly be applied to higher-dimensional tensor network states that admit a sequential generation, such as isoTNS, yielding normalized states by construction. Although one can sample random translation-invariant states with this method, we investigate the disordered case by drawing each unitary matrix independently from the Haar measure.

Because we are interested in the decay of correlations, we focus on computing average correlations between two subsystems $A$ and $B$ as a function of their distance $r$. For random MPS and isoTNS, we consider subsystems $A$ and $B$ as sketched in Figs. 2(a) and 2(b), respectively. In both cases, $A$ and $B$ stretch across $a$ and $b$ consecutive (horizontal) sites. In Fig. 2(b), $A$ and $B$ touch the orthogonality hypersurface and stretch across $h$ vertical sites. We will relax this condition later.
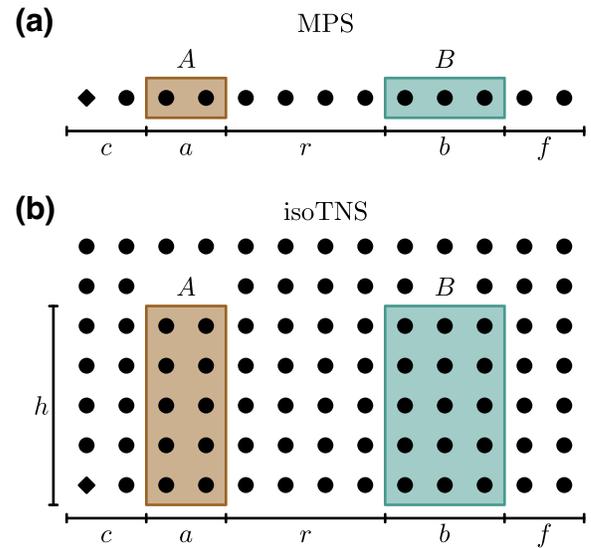


FIG. 2. We investigate average correlations between two subsystems $A$ and $B$ as a function of their (horizontal) distance $r$. $A$ and $B$, respectively, stretch across $a$ and $b$ consecutive (horizontal) sites. (a) The diamond indicates the origin of the sequential generation of the MPS. (b) In addition to indicating the origin of the sequential generation, the diamond also indicates the orthogonality center of the isoTNS. For now, we restrict ourselves to $A$ and $B$ that touch the orthogonality hypersurface and stretch across $h$ consecutive vertical sites.

For all of the measures of correlation we study, we find that the average with respect to the considered ensemble of states decays exponentially. We formalize this type of behavior in Definition 1.

*Definition 1.*—Let $\mathcal{C}(A : B)$ denote a measure of correlation. We say that the average of $\mathcal{C}(A : B)$ with respect to a given ensemble of random states decays exponentially if

$$\mathbb{E}\mathcal{C}(A : B) = K \exp\left(-\frac{r}{\xi}\right) + O\left[\exp\left(-\frac{r}{\chi}\right)\right], \quad (34)$$

where $K$ is constant with respect to $r$, and $\xi > \chi$ is the average correlation length for $\mathcal{C}(A : B)$.

Remarkably, we find that a single average correlation length persists throughout the different families of measures of correlation and that it depends only on the underlying spatial dimension. We later pinpoint the origin of this behavior to the invariance of a spectral gap of a family of transfer matrices. For MPS,

$$\xi = \left[\log\left(\frac{d^2D^2 - 1}{dD^2 - d}\right)\right]^{-1} \equiv \xi_{\text{1D}}, \quad (35a)$$

and for isoTNS,

$$\xi = \left[\log\left(\frac{d^2D^4 - 1}{dD^3 - dD}\right)\right]^{-1} \equiv \xi_{\text{2D}}. \quad (35b)$$

Note that the average correlation length for MPS coincides with that for isoTNS for $d \to dD$. This seemingly small modification will prove to change the qualitative behavior substantially.

Before moving to the detailed presentation of our methods and results, we briefly comment on the considered measures of correlation and the implications of our findings, first for MPS and then for isoTNS.

### A. Results in one dimension (MPS)

In one dimension, we compute the averages of the Rényi-2 mutual information $I_2(A : B)$, the 2-norm expression $N(A : B)$, and the von Neumann mutual information $I(A : B)$ (see Sec. II B for the definitions of the measures of correlation), with subsystems $A$ and $B$ as sketched in Fig. 2(a).

We find that the averages decay exponentially as specified in Definition 1 with the same correlation length $\xi_{1D}$ (see Results 1, 2, and 3). The derivation for $I(A : B)$ relies on a technical conjecture (see Conjecture 1), which we will discuss in detail later. In addition, we show that the same conjecture is enough to assert that $\xi_{1D}$ is also the average correlation length for $I_\alpha(A : B)$ for any integer value of $\alpha \geq 1$ (see Corollary 4). In short, the same average correlation length $\xi_{1D}$ persists across different measures of correlation.

Interestingly, $\xi_{1D}$ depends only weakly on the bond dimension $D$. In particular, for $d, D \geq 2$,

$$\xi_{1D} = \left\{ \log \left[ \frac{d}{\zeta_{1D}(d,D)} \right] \right\}^{-1} \tag{36}$$

with

$$\frac{4}{5} \leq \zeta_{1D}(d,D) < 1 \tag{37}$$

is monotonically increasing with $D$.

Because we are concerned with random tensor network states, $\xi_{1D}$ is obtained after averaging over realizations. It is then natural to ask if exponentially decaying correlations are typical and, if so, what is the typical correlation length for an individual realization. This motivates the investigation of the concentration of the distribution around its average. To that end, we will show that it is exponentially unlikely in $r$ that $N(A : B)$ and $I(A : B)$ decay slower than with $\xi_{1D}$ (see Corollaries 1 and 3). Our result for $N(A : B)$ allows us to deduce that the average of the trace distance $T(A : B)$ decays at least exponentially with correlation length $\xi \leq 2\xi_{1D}$, and it leads to a concentration result for $T(A : B)$ (see Corollary 2).

As stated before, MPS are known to exhibit exponentially decaying correlations [56]. However, our finding that the average correlation length is almost independent of the bond dimension is novel and implies that long-range correlated states are very atypical members of the ensemble.

### B. Results in two dimensions (isoTNS)

In two dimensions, we compute the averages of the Rényi-2 mutual information $I_2(A : B)$ and the 2-norm expression $N(A : B)$, where subsystems $A$ and $B$ are sketched in Fig. 2(b).

As in one dimension, we find that the averages decay exponentially as specified in Definition 1 with the same average correlation length $\xi_{2D}$ (see Results 4 and 5).

The correlation length $\xi_{2D}$ displays a qualitatively different dependence on the bond dimension $D$. In particular, for $d, D \geq 2$,

$$\xi_{2D} = \left\{ \log \left[ \frac{d}{\zeta_{2D}(d,D)} \right] \right\}^{-1} \tag{38}$$

with

$$0 < \zeta_{2D}(d,D) \leq \frac{8}{21} \tag{39}$$

is monotonically decreasing with $D$, in contrast to its one-dimensional counterpart. As such, the largest correlation length is achieved for $D = 2$.

For $N(A : B)$, we can extend the applicability of our results to any size and shape of subsystems $A$ and $B$; we find the decay to be at least exponential with correlation length $\xi = \xi_{2D}$ (see Corollary 5). We furthermore prove a concentration result for $N(A : B)$ expressing that it is highly unlikely that $N(A : B)$ decays slower than with $\xi_{2D}$ (see Corollary 6). This also allows us to draw a similar conclusion about the behavior of $T(A : B)$ (see Corollary 7).

While it is expected for isoTNS that correlations decay exponentially if they are both on the orthogonality hypersurface [32], our finding that the exponential decay persists for subsystems in the bulk is novel. Moreover, the different behavior of $\xi_{2D}$ as compared to $\xi_{1D}$ is particularly surprising; for typical isoTNS, increasing the bond dimension reduces the average correlation length.

### C. Further discussion

Here, we discuss three aspects of our results on a more qualitative level: the absence of large average correlation lengths, the opposite dependence of said correlation lengths on the bond dimension in one and two dimensions, and the effect of blocking sites together.

From the analytical expressions for $\xi_{1D}$ and $\xi_{2D}$ [see Eq. (35)], it follows that, for a fixed physical dimension $d$, both average correlation lengths are upper bounded by a constant, irrespective of the bond dimension $D$. In one dimension, the sequential-generation scheme corresponds to actions of unitary matrices on two neighboring $d$-dimensional sites together with a $D$-dimensional system. The unitary matrices act as an information carrier, conveying information from one physical qudit to another through

interaction. Our results shows that, when these unitary matrices are random and independent from each other, the resulting correlations are, on average, weak.

The average correlation lengths for isoTNS can be obtained from the one for MPS by replacing $d \to dD$. This simple fact, however, leads to $\xi_{2D}$ being a decreasing function of $D$, in contrast to $\xi_{1D}$. To explain this feature, we use the fact (which we will prove later) that $\xi_{2D}$ is independent of the vertical extent $h$ of subsystems $A$ and $B$. As a result, we can gain sufficient insight from the limiting case of a $1 \times n$ isoTNS, which corresponds to a certain MPS graphically given by

$$|\psi\rangle = \text{[diagram]} \qquad (40)$$

Importantly, this MPS has local dimension $dD$, explaining why the average correlation lengths coincide for $d \to dD$. An increase of the bond dimension of the isoTNS thus corresponds to an increase of the local dimension of the corresponding MPS, which dominates the behavior of $\xi_{1D}$.

Finally, we comment on the effect of blocking sites together. It holds that

$$\xi_{1D} = \frac{1}{\log(d)} + O\left(\frac{1}{D^2}\right), \qquad (41)$$

implying that $\xi_{1D} \approx 1/\log(d)$ for $D \gg 1$. This suggests the following scale invariance property, valid in the context of this approximation. Consider two families $|\psi\rangle$ and $|\widetilde{\psi}\rangle$ of random MPS with $\widetilde{d} = d^q$ and $\widetilde{D} = D$. Graphically, for $q = 2$,

$$|\psi\rangle = \text{[diagram]} \qquad (42)$$

and

$$|\widetilde{\psi}\rangle = \text{[diagram]} \qquad (43)$$

Our simple observation is that, although it holds that $\widetilde{\xi} = \xi/q$, the effective distances (that is, the number of unitary matrices separating any two physical sites) are also rescaled as $\widetilde{r} = r/q$. Consequently, the circuit architectures generating $|\psi\rangle$ and $|\widetilde{\psi}\rangle$ have scale-invariant (average) correlations within the approximation $\xi_{1D} \approx 1/\log(d)$.

## IV. CORRELATIONS IN ONE DIMENSION

In this section, we state and discuss the results for random MPS summarized in Sec. III A in more detail. Before doing that, we develop the tools behind our proofs in Secs. IV A and IV B. In Sec. IV C, we compute the average of $I_2(A : B)$, and in Sec. IV D, we investigate the decays of $N(A : B)$ and $T(A : B)$. Finally, we discuss the behavior of $I(A : B)$ in Sec. IV E.

When computing the averages of measures of correlation for random MPS, we will exploit a simplification with respect to the scenario depicted in Fig. 2(a). Instead of allowing for an arbitrary number of sites before subsystem $A$, we prove our statements in the limit of $c \to \infty$. As we show in Appendix J, this does not constitute a limitation because the $c$ initial sites do not affect the decay of correlations and, therefore, neither the average correlation length $\xi_{1D}$. Furthermore, we will see that the $f$ sites after subsystem $B$ do not play a role in the computation of average correlations, as it is expected for any sequentially generated state.

### A. Transfer matrices

The key challenge for computing the average of each measure of correlation will be evaluating multiple expressions of the form

$$\text{tr}\left(P\mathbb{E}|\psi\rangle\langle\psi|^{\otimes k}\right), \qquad (44)$$

where

$$P = \left(P_e^{(d)}\right)^{\otimes c} \otimes \left(P_\alpha^{(d)}\right)^{\otimes a} \otimes \left(P_e^{(d)}\right)^{\otimes r}$$
$$\otimes \left(P_\beta^{(d)}\right)^{\otimes b} \otimes \left(P_e^{(d)}\right)^{\otimes(f-1)} \otimes P_e^{(dD)}. \qquad (45)$$

The permutation $\alpha \in S_k$ acts on the sites comprising subsystem $A$, while $\beta \in S_k$ acts on the sites comprising $B$. The exact forms of $\alpha$ and $\beta$ as well as the number of required replicas $k$ depend on the considered measure of correlation and will be specified later. It shall also become clear why sites belonging to neither $A$ nor $B$ are acted upon by the trivial permutation $e \in S_k$. In the following, we show that Eq. (44), for random MPS, reduces to multiplying matrices $T_\rho \in \mathbb{R}^{k! \times k!}$ with $\rho \in S_k$ whose definition will be natural. Because their role is analogous to the known transfer matrices mediating correlations, we will also adopt this term here.

Before introducing the transfer matrices, we must analyze $\mathbb{E}|\psi\rangle\langle\psi|^{\otimes k}$. To that end, let us define $V^{(j)} = U^{(j)} \otimes \overline{U^{(j)}}$. Then, by Eq. (3),


$$(46)$$

and


$$(47)$$

By computing the $k$-fold twirl [see Eq. (33)], we obtain the building block


$$(48)$$


$$(49)$$

where the (green) dot represents a Kronecker $\delta$ on three permutation indices. Note that we have not drawn the contraction of a permutation matrix with $|0\rangle\langle 0|^{\otimes k}$ because it is trivial by Eq. (31). The average of a random MPS is then given by


$$(50)$$

We could, in principle, work with the building block above. However, it is not convenient to have dangling bond (blue) legs whose dimension grows with $D$. By cutting permutation-valued (green) legs instead, we obtain a building block with fixed dimension for fixed $k$. With that building block, the average of a random MPS is given by


$$(51)$$

The entries of the initial vector $\langle I_k| \in \mathbb{R}^{k!}$ are given by


$$(52)$$

where we have used Eq. (31). The tensors in the bulk are given via


$$(53)$$

and the final tensor is given via


$$(54)$$

Computing an expression of the form of Eq. (44) corresponds to contracting each tensor $S$ with $P_\rho^{(d)}$, which leads us to the promised definition of a transfer matrix $T_\rho \in \mathbb{R}^{k! \times k!}$. Using Eq. (30), its entries are given by


$$(55)$$

$$= \sum_{\sigma \in S_k} Wg(\sigma\tau^{-1}, dD)\, d^{\#(\sigma,p)} D^{\#(\sigma\theta^{-1})} \qquad (56)$$

We define $T_\rho$ with respect to the basis defined by the map $s_i \mapsto e_i$, where $s_i$ is the $i$th element of $S_k = \{s_1, \dots, s_{k!}\}$, and $\{e_i\}$ is the standard basis of $\mathbb{R}^{k!}$. In Appendix D, we

find that $T_\rho$ is block triangular if the elements of $S_k$ are ordered in a certain way.

As alluded to in Eq. (45), the final tensor $S'$ will be contracted with the trivial permutation $e \in S_k$ in our computations. The final vector $|F_k\rangle \in \mathbb{R}^{k!}$ is thus defined via

$$\tau - \boxed{F} = \tau - \boxed{S'} = \tau - \boxed{W} - \boxed{P} \quad (57)$$

$$= \sum_{\sigma \in S_k} Wg(\sigma\tau^{-1}, dD)(dD)^{\#(\sigma\epsilon)} = \delta_{e\tau}, \quad (58)$$

where we have used the definition of the Weingarten function.

Using the definitions of $T_e$ and $|F_k\rangle$, it is easy to confirm that $T_e|F_k\rangle = |F_k\rangle$. Graphically, this implies the simplification

$$\boxed{I} - \boxed{T_\rho} - \boxed{T_e} - \boxed{T_e} - \boxed{F} = \boxed{I} - \boxed{T_\rho} - \boxed{F}. \quad (59)$$

From this, it also follows that $\mathbb{E}|\psi\rangle\langle\psi|^{\otimes k}$ is properly normalized:

$$\mathrm{tr}\left(\mathbb{E}|\psi\rangle\langle\psi|^{\otimes k}\right) = \langle I_k|T_e^{n-1}|F_k\rangle = \langle I_k|F_k\rangle = 1 \quad (60)$$

With what we have laid out above, we can write Eq. (44) in terms of transfer matrices:

$$\mathrm{tr}\left(P\mathbb{E}|\psi\rangle\langle\psi|^{\otimes k}\right) = \langle I_k|T_e^c T_\alpha^a T_e^r T_\beta^b|F_k\rangle \quad (61)$$

We provide a simple Mathematica package [63] that defines $T_\rho$ with $\rho \in S_k$ for $k \in \{1, \ldots, 20\}$ according to Eq. (56). The package relies on the package provided by the authors of Ref. [64] for evaluating the Weingarten function.

### B. Estimating the decay of correlations

The decay of the average of each measure of correlation is necessarily determined by the $r$ sites separating subsystems $A$ and $B$. As we will see in the following sections, this will, for each measure, translate to a simple statement in terms of the just-defined transfer matrices. In particular, we will find that the decay of correlations is determined by $T_e^r$ with $e \in S_k$. Taking this as a fact for now, we connect the decay of correlations with the spectrum of $T_e$.

The spectrum of $T_e$ depends on $k$ because $k$ determines its dimension and entries. Still, we can make general statements about $T_e$ for any $k \geq 2$. In particular, we will prove the following statements in Appendices B and C.

*Proposition 1.*—The eigenvalues of $T_e$ with $e \in S_k$ are non negative for any $k \geq 2$.

*Proposition 2.*—$T_e$ with $e \in S_k$ is diagonalizable for any $k \geq 2$.

*Proposition 3.*—Let $\lambda_1 > \lambda_2 > \cdots \geq 0$ denote the distinct eigenvalues of $T_e$ with $e \in S_k$. Then, $\lambda_1 = 1$, and it is nondegenerate for any $k \geq 2$ if $d \geq 2$.

Given the statements above, we can expand $T_e^r$ as

$$T_e^r = |R_1\rangle\langle L_1| + \lambda_2^r \sum_{\mu=1}^{w_2} |R_2^{(\mu)}\rangle\langle L_2^{(\mu)}| + O\left(\lambda_3^r\right), \quad (62)$$

where $|R_i^{(\mu)}\rangle$ denotes the $\mu$th right eigenvector corresponding to $\lambda_i$, $\langle L_i^{(\mu)}|$ denotes the $\mu$th left eigenvector corresponding to $\lambda_i$, and $w_i$ denotes the degeneracy of $\lambda_i$.

The asymptotic decay of correlations is thus determined by the subleading eigenvalue $\lambda_2$ of $T_e$, and the average correlation length is given by

$$\xi = -\frac{1}{\log(|\lambda_2|)}. \quad (63)$$

The argument behind this is similar to one known from the analysis of correlations in translation-invariant MPS [56,65], where the decay is determined by the subleading eigenvalue of the relevant transfer matrix.

### C. Rényi-2 mutual information

We start our analysis of correlations in random MPS with the simplest case, namely the computation of the average of the Rényi-2 mutual information

$$I_2(A:B) = \log\left[\mathrm{tr}\left(\varrho_{AB}^2\right)\right]$$
$$- \log\left[\mathrm{tr}\left(\varrho_A^2\right)\right]$$
$$- \log\left[\mathrm{tr}\left(\varrho_B^2\right)\right]. \quad (64)$$

The analytical treatment turns out to be comparatively simple if one assumes that $\mathbb{E}\log(X) = \log(\mathbb{E}X)$, as is frequently done in this context [66–68]. The assumption can furthermore be justified with the fact that the purities of $\varrho_A$, $\varrho_B$, and $\varrho_{AB}$ concentrate around their averages, as we discuss towards the end of this section. That said, we will not make the assumption further below when we study the von Neumann mutual information. The analysis there will require transfer matrices $T_\rho$ with $\rho \in S_k$ for all $k \geq 2$, while $\rho \in S_2$ will suffice here because only the averages of purities are needed. Our first result is summarized below.

*Result 1.*—The average of $I_2(A:B)$ with respect to the random MPS ensemble and subsystems $A$ and $B$ as sketched in Fig. 2(a) decays exponentially as specified in Definition 1 with the average correlation length $\xi_{1D}$ defined in Eq. (35a).

*Proof.*—We split the proof into four steps. The exact same structure will also appear in the proofs for the other

measures of correlation. Thus, this proof serves as the simplest example and a point of reference for later proofs. Step 1: We rewrite $\mathbb{E}I_2(A:B)$ in terms of expressions of the form of Eq. (44). To that end, we make the assumption that $\mathbb{E}\log(X) = \log(\mathbb{E}X)$. Then,

$$\begin{aligned}
\mathbb{E}I_2(A:B) &= \log\left[\mathbb{E}\operatorname{tr}\left(\varrho_{AB}^2\right)\right] \\
&\quad - \log\left[\mathbb{E}\operatorname{tr}\left(\varrho_A^2\right)\right] \\
&\quad - \log\left[\mathbb{E}\operatorname{tr}\left(\varrho_B^2\right)\right].
\end{aligned} \tag{65}$$

$\mathbb{E}\operatorname{tr}\left(\varrho_A^2\right)$, $\mathbb{E}\operatorname{tr}\left(\varrho_B^2\right)$, and $\mathbb{E}\operatorname{tr}\left(\varrho_{AB}^2\right)$ can already be written in the desired form. For example,

$$\mathbb{E}\operatorname{tr}\left(\varrho_{AB}^2\right) = \operatorname{tr}\left(P\mathbb{E}|\psi\rangle\langle\psi|^{\otimes 2}\right) \tag{66}$$

with

$$\begin{aligned}
P &= \left(P_e^{(d)}\right)^{\otimes c} \otimes \left(P_{(12)}^{(d)}\right)^{\otimes a} \otimes \left(P_e^{(d)}\right)^{\otimes r} \\
&\quad \otimes \left(P_{(12)}^{(d)}\right)^{\otimes b} \otimes \left(P_e^{(d)}\right)^{\otimes(f-1)} \otimes P_e^{(dD)}.
\end{aligned} \tag{67}$$

Step 2: We express $\mathbb{E}I_2(A:B)$ in terms of the transfer matrices defined in Sec. IV A. Given the previous step, it is easy to confirm that

$$\begin{aligned}
\mathbb{E}I_2(A:B) &= \log\left(\langle I_2|T_e^c T_{(12)}^a T_e^r T_{(12)}^b|F_2\rangle\right) \\
&\quad - \log\left(\langle I_2|T_e^c T_{(12)}^a|F_2\rangle\right) \\
&\quad - \log\left(\langle I_2|T_e^{c+a+r} T_{(12)}^b|F_2\rangle\right).
\end{aligned} \tag{68}$$

Step 3: We expand $\mathbb{E}I_2(A:B)$ in terms of the spectrum of $T_e$ with $e \in S_2$ [see Eq. (62)]. Using the relevant expressions for the Weingarten function, it is evident [69] that

$$T_e = \begin{pmatrix} 1 & \dfrac{d^2D - D}{d^2D^2 - 1} \\ 0 & \dfrac{dD^2 - d}{d^2D^2 - 1} \end{pmatrix} \tag{69}$$

is diagonalizable with

$$\lambda_1 = 1 \quad \text{and} \quad \lambda_2 = \frac{dD^2 - d}{d^2D^2 - 1}. \tag{70}$$

Expanding $T_e^c$ and taking the limit of $c \to \infty$ yields

$$\begin{aligned}
\mathbb{E}I_2(A:B) &= \log\left(\langle L_1|T_{(12)}^a T_e^r T_{(12)}^b|F_2\rangle\right) \\
&\quad - \log\left(\langle L_1|T_{(12)}^a|F_2\rangle\right) \\
&\quad - \log\left(\langle L_1|T_{(12)}^b|F_2\rangle\right),
\end{aligned} \tag{71}$$

where we have used that $\langle I_2|R_1\rangle = 1$. After expanding also $T_e^r$ and using that $|F_2\rangle = |R_1\rangle$, we have

$$\begin{aligned}
\mathbb{E}I_2(A:B) &= \log\Big(\langle L_1|T_{(12)}^a|R_1\rangle\langle L_1|T_{(12)}^b|R_1\rangle \\
&\quad + \lambda_2^r \langle L_1|T_{(12)}^a|R_2\rangle\langle L_2|T_{(12)}^b|R_1\rangle\Big) \\
&\quad - \log\left(\langle L_1|T_{(12)}^a|R_1\rangle\right) \\
&\quad - \log\left(\langle L_1|T_{(12)}^b|R_1\rangle\right).
\end{aligned} \tag{72}$$

Step 4: Finally, we can write $\mathbb{E}I_2(A:B)$ in the form of Definition 1. That is,

$$\begin{aligned}
&\mathbb{E}I_2(A:B) \\
&= \log\left(1 + \lambda_2^r \frac{\langle L_1|T_{(12)}^a|R_2\rangle\langle L_2|T_{(12)}^b|R_1\rangle}{\langle L_1|T_{(12)}^a|R_1\rangle\langle L_1|T_{(12)}^b|R_1\rangle}\right) \quad (73) \\
&= \lambda_2^r \frac{\langle L_1|T_{(12)}^a|R_2\rangle\langle L_2|T_{(12)}^b|R_1\rangle}{\langle L_1|T_{(12)}^a|R_1\rangle\langle L_1|T_{(12)}^b|R_1\rangle} + O\left(\lambda_2^{2r}\right) \quad (74) \\
&\equiv K\exp\left(-\frac{r}{\xi}\right) + O\left[\exp\left(-\frac{2r}{\xi}\right)\right], \quad (75)
\end{aligned}$$

where

$$K = \frac{\langle L_1|T_{(12)}^a|R_2\rangle\langle L_2|T_{(12)}^b|R_1\rangle}{\langle L_1|T_{(12)}^a|R_1\rangle\langle L_1|T_{(12)}^b|R_1\rangle} \tag{76}$$

and

$$\xi = -\frac{1}{\log(\lambda_2)} = -\left[\log\left(\frac{dD^2 - d}{d^2D^2 - 1}\right)\right]^{-1} = \xi_{1D}. \tag{77}$$

This concludes the proof. ∎

We are now in a position to discuss the concentration of the purities of $\varrho_A$, $\varrho_B$, and $\varrho_{AB}$ around their averages, which justifies $\mathbb{E}\log(X) = \log(\mathbb{E}X)$. By expanding $\mathbb{E}\operatorname{tr}\left(\varrho_A^2\right) = \langle L_1|T_{(12)}^a|R_1\rangle$, one can confirm that

$$\mathbb{E}\operatorname{tr}\left(\varrho_A^2\right) = \left(\frac{d+1}{d}\right)^2 \frac{1}{D^2} + O\left(\frac{1}{d^a} + \frac{1}{D^4}\right). \tag{78}$$

Similarly, it holds that

$$\mathbb{E}\operatorname{tr}\left(\varrho_{AB}^2\right) = \left(\frac{d+1}{d}\right)^4 \frac{1}{D^4} + O\left(\frac{1}{d^{a+b}} + \frac{1}{D^6}\right). \tag{79}$$

The purities of $\varrho_A$, $\varrho_B$, and $\varrho_{AB}$ are thus close to their minimum values, implying concentration by Markov's inequality and justifying $\mathbb{E}\log(X) = \log(\mathbb{E}X)$.

As discussed earlier, the Rényi-2 mutual information is lacking many of the desirable properties that a sound measure of correlation ought to fulfill. On top, our computation

simplifies considerably because we are using the assumption that $\mathbb{E} \log(X) = \log(\mathbb{E}X)$, which amounts to ignoring statistical fluctuations in the different realizations. In the following section, we will see that $N(A : B)$ decays exponentially with the same average correlation length $\xi_{1D}$. We will furthermore show that $N(A : B)$ concentrates around its average, providing evidence that fluctuations can be safely ignored in our context.

### D. Trace distance and 2-norm

In this section, we investigate average correlations as quantified by the trace distance $T(A : B)$. As anticipated in Sec. II B, this is a challenging task. However, as laid out there, the 2-norm expression $N(A : B)$ reliably estimates $T(A : B)$ for the case of random MPS. Hence, we compute the average of

$$N(A : B) = \|\varrho_{AB} - \varrho_A \otimes \varrho_B\|_2^2 \tag{80}$$

$$= \mathrm{tr}\left(\varrho_{AB}^2\right)$$

$$+ \mathrm{tr}\left(\varrho_A^2\right)\mathrm{tr}\left(\varrho_B^2\right)$$

$$- 2\,\mathrm{tr}\left[\varrho_{AB}\left(\varrho_A \otimes \varrho_B\right)\right]. \tag{81}$$

Because of its connection to the Hilbert-Schmidt inner product, the average of $N(A : B)$ can be computed without any simplifying assumptions. Making use of the transfer-matrix techniques introduced above, we prove the following result.

*Result 2.*—The average of $N(A : B)$ with respect to the random MPS ensemble and subsystems $A$ and $B$ as sketched in Fig. 2(a) decays exponentially as specified in Definition 1 with the average correlation length $\xi_{1D}$ defined in Eq. (35a).

*Sketch of proof.*—The proof follows the same procedure as that of Result 1. Here, we sketch the main steps and refer to Appendix H for more details.

In Step 1, we write $\mathbb{E}N(A : B)$ in terms of expressions of the form of Eq. (44). The second summand in Eq. (81) requires permutations of $S_4$ because

$$\mathbb{E}\,\mathrm{tr}\left(\varrho_A^2\right)\mathrm{tr}\left(\varrho_B^2\right) = \mathrm{tr}\left(P\mathbb{E}|\psi\rangle\langle\psi|^{\otimes 4}\right) \tag{82}$$

with

$$P = \left(P_e^{(d)}\right)^{\otimes c} \otimes \left(P_{(12)}^{(d)}\right)^{\otimes a} \otimes \left(P_e^{(d)}\right)^{\otimes r}$$

$$\otimes \left(P_{(34)}^{(d)}\right)^{\otimes b} \otimes \left(P_e^{(d)}\right)^{\otimes (f-1)} \otimes P_e^{(dD)}. \tag{83}$$

The first summand and the third summand, respectively, require only permutations of $S_2$ and $S_3$. In Step 2, we thus write $\mathbb{E}N(A : B)$ in terms of transfer matrices $T_\rho$ with $\rho \in S_4$.

This means that the average correlation length is determined by the subleading eigenvalue of $T_e$ with $e \in S_4$. Let

$\lambda_1 > \lambda_2 > \cdots \geq 0$ denote the distinct eigenvalues of $T_e$. In Step 3, we find that

$$\lambda_1 = 1 \qquad \text{and} \qquad \lambda_2 = \frac{dD^2 - d}{d^2D^2 - 1}, \tag{84}$$

just like for $T_e$ with $e \in S_2$. The former is nondegenerate, while the degeneracy of the latter is given by the number of transposition in $S_4$,

$$w_2 = \binom{4}{2} = 6. \tag{85}$$

Thus, the average correlation length for $N(A : B)$ coincides with that for $I_2(A : B)$, as we conclude in Step 4. ∎

The above result establishes the exponential decay of the average. However, one is usually interested in knowing if typical instances are expected to have the same exponential decay. This can be easily established by Markov's inequality because $N(A : B)$ is non negative and its average decays to zero as a function of the distance $r$.

*Corollary 1.*—For subsystems $A$ and $B$ as sketched in Fig. 2(a), sufficiently large $r$, and all $0 < \varepsilon < 1$, the random MPS ensemble satisfies

$$\Pr\left\{N(A : B) \geq K \exp\left[-\frac{(1-\varepsilon)r}{\xi_{1D}}\right]\right\} \leq \exp\left(-\frac{\varepsilon r}{\xi_{1D}}\right), \tag{86}$$

where $K$ is constant with respect to $r$.

*Proof.*—By Result 2, for sufficiently large $r$, we can bound $\mathbb{E}N(A : B) \leq K \exp(-r/\xi_{1D})$. Because $N(A : B)$ is non negative, by Markov's inequality, we have, for $\eta > 0$,

$$\Pr\left[N(A : B) \geq \eta K \exp\left(-\frac{r}{\xi_{1D}}\right)\right]$$

$$\leq \Pr\left[N(A : B) \geq \eta\mathbb{E}N(A : B)\right] \tag{87}$$

$$\leq \frac{1}{\eta}. \tag{88}$$

The result follows with $\eta = \exp(\varepsilon r/\xi_{1D})$. ∎

The above corollary reflects that it is exponentially unlikely in $r$ that $N(A : B)$ decays slower than with the average correlation length $\xi_{1D}$. Because we have already established that the average case exhibits an exponential decay with correlation length $\xi_{1D}$, the average case is also typical.

By combining the above result with Eq. (18), we can now also bound the correlation length for $T(A : B)$.

*Corollary 2.*—For subsystems $A$ and $B$ as sketched in Fig. 2(a), sufficiently large $r$, and all $0 < \varepsilon < 1$, the
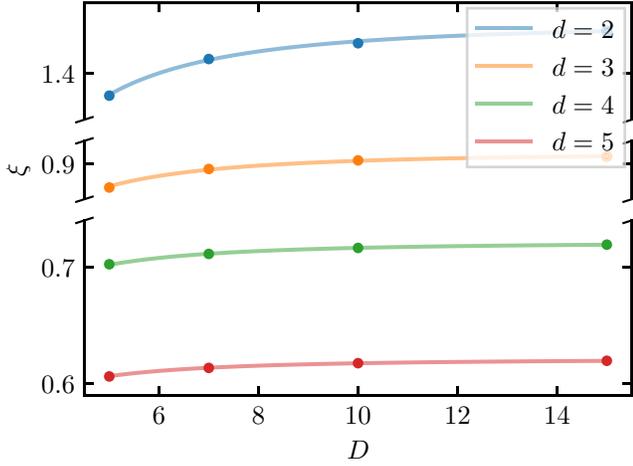
FIG. 3. Numerically obtained average correlation length $\xi$ for different $d$ and $D$. The data points are obtained by fitting the average value of $I(A : B)$ against $r \in \{5, 7, 9, 11, 13, 15\}$ for $a = b = 1$. The sample size of 10 000 suffices for the error bars to lie within the plot points. The opaque curves correspond to $\xi_{1D}$ [see Eq. (35a)].

random MPS ensemble satisfies

$$\Pr\left\{T(A : B) \geq K \exp\left[-\frac{(1-\varepsilon)r}{2\xi_{1D}}\right]\right\} \leq \exp\left(-\frac{\varepsilon r}{\xi_{1D}}\right), \tag{89}$$

where $K$ is constant with respect to $r$.

*Proof.*—It holds that

$$\mathbb{E}T(A : B) \leq \sqrt{\mathbb{E}[T(A : B)]^2} \tag{90}$$

$$\leq \frac{D^2}{2}\sqrt{\mathbb{E}N(A : B)} \tag{91}$$

$$\leq K \exp\left(-\frac{r}{2\xi_{1D}}\right), \tag{92}$$

where, in the last line, we have assumed $r$ to be sufficiently large. The result follows as in the proof of Corollary 1. ∎

Thus, with overwhelming probability, correlations as quantified by $T(A : B)$ decay exponentially with $\xi \leq 2\xi_{1D}$.

### E. Von Neumann mutual information

The fact that $I_2(A : B)$ and $N(A : B)$ have the same average correlation length $\xi_{1D}$ motivates the question whether other measures of correlation behave similarly. In this section, we will provide compelling evidence that $\xi_{1D}$ is indeed the average correlation length also for the von Neumann mutual information $I(A : B)$.

We start by numerically investigating the behavior of $I(A : B)$ for random MPS. We have generated MPS according to our measure, computed the average of $I(A : B)$, and extracted the average correlation length from fits. As Fig. 3

shows, the numerically obtained average correlation length coincides well with $\xi_{1D}$. It should be noted that we have set $c = 0$ for our numerical analysis. We discuss in Appendix J why this does not affect the average correlation length. For more details on our numerical analysis, see Appendix K.

We now turn to the analytical computation of the average of

$$I(A : B) = \mathrm{tr}\left[\rho_{AB} \log\left(\rho_{AB}\right)\right]$$
$$- \mathrm{tr}\left[\rho_A \log\left(\rho_A\right)\right]$$
$$- \mathrm{tr}\left[\rho_B \log\left(\rho_B\right)\right]. \tag{93}$$

To be able to make use of the transfer-matrix techniques introduced above, we employ two replica tricks to write $\mathbb{E}I(A : B)$ in terms of expressions of the form of Eq. (44). First, we write $S(\rho)$ as the limit of $\alpha \to 1$ of $S_\alpha(\rho)$:

$$S(\rho) = \lim_{\alpha \to 1} \frac{1}{1 - \alpha} \log[\mathrm{tr}(\varrho^\alpha)] \tag{94}$$

Second, instead of assuming again that $\mathbb{E}\log(X) = \log(\mathbb{E}X)$, we use

$$\mathbb{E}\log(X) = \lim_{v \to 0} \frac{1}{v} \log\left(\mathbb{E}X^v\right). \tag{95}$$

We are thus dealing with expressions of the form $\mathrm{tr}\left(P\mathbb{E}|\psi\rangle\langle\psi|^{\otimes v\alpha}\right)$, which require transfer matrices $T_\rho$ with $\rho \in S_{v\alpha}$ (see Appendix I). This means that knowing the spectrum of $T_e$ with $e \in S_{v\alpha}$ for all $v\alpha \geq 2$ allows us to draw conclusions about the decay of the average of $I(A : B)$.

While this is in principle a daunting task, we are able to prove several properties of the transfer matrix $T_e$ with $e \in S_k$ for any $k \geq 2$ (see Propositions 1, 2, and 3).

In Appendix D, we furthermore show that $T_e$ with $e \in S_k$ has eigenvalue

$$\mu_2 = \frac{dD^2 - d}{d^2D^2 - 1} \tag{96}$$

for any $k \geq 2$. Its degeneracy is at least

$$v_2 = \binom{k}{2}, \tag{97}$$

the number of transpositions in $S_k$. We conjecture that $\mu_2$ is the subleading eigenvalue of $T_e$ with $e \in S_k$ for any $k \geq 2$ and that it has degeneracy $v_2$. We know this conjecture to hold for $k \in \{2, 3, 4\}$, and we have numerical evidence suggesting so for $k \in \{5, 6, 7\}$ [63]. We were not able to prove the statement outright, but in the following we argue that it is the only missing step to show that $\xi_{1D}$ is the average correlation length for $I(A : B)$. Let us state the conjecture formally below.

*Conjecture 1.*—Let $\lambda_1 > \lambda_2 > \cdots \geq 0$ denote the distinct eigenvalues of $T_e$ with $e \in S_k$. Then, $\lambda_2 = \mu_2$ with degeneracy $w_2 = v_2$ for any $k \geq 2$.

If Conjecture 1 holds, the properties of $T_e$ with $e \in S_{v\alpha}$ that are relevant for determining the decay of correlations are independent of $v\alpha$. In Appendix I, we argue that the replica limit does not affect this and prove the following result.

*Result 3.*—If Conjecture 1 holds, the average of $I(A : B)$ with respect to the random MPS ensemble and subsystems $A$ and $B$ as sketched in Fig. 2(a) decays exponentially as specified in Definition 1 with the average correlation length $\xi_{1D}$ defined in Eq. (35a).

We provide a concentration result also for $I(A : B)$. The statement and its proof are identical to that for $N(A : B)$. It is exponentially unlikely in $r$ that $I(A : B)$ decays slower than with the average correlation length $\xi_{1D}$, and the average case is also typical.

*Corollary 3.*—If Conjecture 1 holds, for subsystems $A$ and $B$ as sketched in Fig. 2(a), sufficiently large $r$, and all $0 < \varepsilon < 1$, the random MPS ensemble satisfies

$$\Pr\left\{I(A : B) \geq K \exp\left[-\frac{(1-\varepsilon)r}{\xi}\right]\right\} \leq \exp\left(-\frac{\varepsilon r}{\xi}\right), \tag{98}$$

where $K$ is constant with respect to $r$.

Another corollary of Result 3 is that $\xi_{1D}$ is also the average correlation length for $I_\alpha(A : B)$ for any integer value of $\alpha \geq 1$. We prove also this statement in Appendix I.

*Corollary 4.*—If Conjecture 1 holds, for any integer value of $\alpha \geq 1$, the average of $I_\alpha(A : B)$ with respect to the random MPS ensemble and subsystems $A$ and $B$ as sketched in Fig. 2(a) decays exponentially as specified in Definition 1 with the average correlation length $\xi_{1D}$ defined in Eq. (35a).

Finally, let us summarize the reason behind the persistent appearance of the average correlation length $\xi_{1D}$. In all of the examined cases, the asymptotic behavior of the correlations was determined by the asymptotic decay of $T_e^r$. Although the transfer matrix $T_e$ with $e \in S_k$ does depend on the number of replicas $k$, its asymptotic decay does not (given Conjecture 1), resulting in a common average correlation length $\xi_{1D}$ across measures of correlation with different complexity.

## V. CORRELATIONS IN TWO DIMENSIONS

In this section, we state and discuss in more detail the results for random isoTNS summarized in Sec. III B. In Sec. V A, we develop the two-dimensional analogue to the transfer matrices introduced in Sec. IV A, the tool behind our proofs. In Sec. V C, we compute the average of $I_2(A : B)$, and in Sec. V D, we investigate the decays of $N(A : B)$ and $T(A : B)$.

We show the exponential decay of the average for each considered measure of correlation and subsystems $A$ and $B$ as sketched in Fig. 4(a) as a function of the horizontal distance $r$ between $A$ and $B$. In particular, we prove that $I_2(A : B)$ and $N(A : B)$ have a common average correlation length that is independent of the sizes of $A$ and $B$. Moreover, thanks to the more amenable properties of $N(A : B)$, we are able to show that average correlations decay exponentially also for subsystems $A$ and $B$ that do not need to touch the orthogonality hypersurface [see Fig. 4(b)]. As in one dimension, we prove our statements in the limit of $c \to \infty$.

### A. Transfer matrices

As in one dimension, computing the average of each measure of correlation will involve computing multiple terms of the form

$$\mathrm{tr}\left(P\mathbb{E}|\psi\rangle\langle\psi|^{\otimes k}\right), \tag{99}$$

where $P$ has a similar tensor product structure as Eq. (45), adapted to the two-dimensional setting considered here.
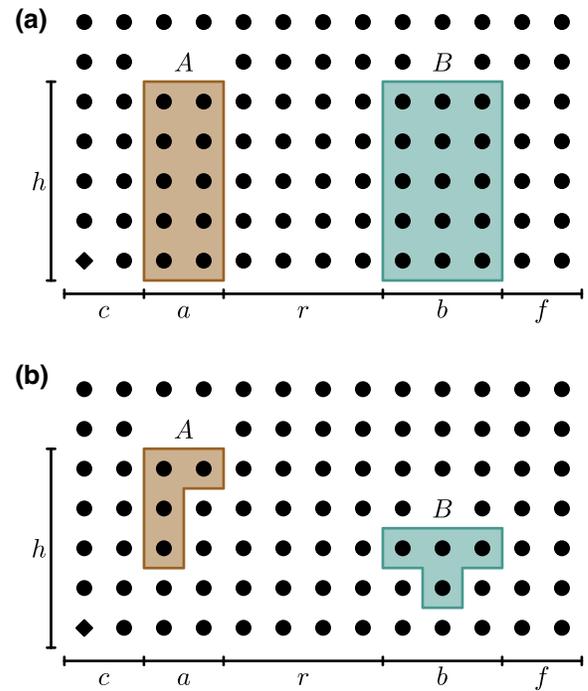


FIG. 4. We investigate average correlations in random isoTNS between two subsystems $A$ and $B$ as a function of their horizontal distance $r$. $A$ and $B$, respectively, stretch across $a$ and $b$ consecutive horizontal sites. In addition to indicating the origin of the sequential generation, the diamond also indicates the orthogonality center of the isoTNS. (a) For Results 4 and 5, we consider $A$ and $B$ that touch the orthogonality hypersurface and stretch across $h$ consecutive vertical sites. (b) We will provide an additional result for the 2-norm expression $N(A : B)$ for arbitrary (but fixed) $A$ and $B$ that do not need to touch the orthogonality hypersurface.

The number of required replicas $k$ again depends on the considered measure of correlation.

We will thus need a two-dimensional analogue to the transfer matrices introduced in Sec. IV A. In contrast to the one-dimensional case, the size of the resulting transfer matrices will also depend on the geometry of the considered subsystems, making their analysis much more challenging. However, the procedure of defining the tensors is similar to that in one dimension. We provide an overview here and refer to Appendix L for more details.

We define $V^{(i,j)} = U^{(i,j)} \otimes \overline{U^{(i,j)}}$, where $U^{(i,j)} \in U(dD^2)$ is the unitary matrix depicted in Fig. 1(b2). By computing the $k$-fold twirl [see Eq. (33)], we obtain the building block

$$\text{(100)}$$

$$\text{(101)}$$

As in one dimension, it is convenient to define a building block for which the contraction of bond (blue) legs is implicit. This results in a tensor with only permutation-valued (green) legs. As we show in Appendix L, the resulting bulk tensor is given via

$$\text{(102)}$$

For the sake of brevity, the expressions for the boundary tensors are stated in Appendix L.

Computing expressions of the form of Eq. (99) corresponds to contracting each tensor $S$ with $P_\rho^{(d)}$. The entries of the resulting bulk tensor $T_\rho \in \mathbb{R}^{k! \times k! \times k! \times k!}$ are given by

$$\text{(103)}$$

$$= \sum_{\sigma \in S_k} W g(\sigma \tau^{-1}, dD^2) d^{\#(\sigma,p)} D^{\#(\sigma v^{-1})} \tag{104}$$

In our computations, the site in the top-right corner belongs to neither $A$ nor $B$. It is thus acted upon by the trivial permutation $e \in S_k$. As we show in Appendix L, the corresponding tensor is given by $T_e = |F_k\rangle\langle F_k|$.

In Appendix L, we furthermore show that a property similar to Eq. (59) also holds for random isoTNS. That is, tensors corresponding to the trivial permutation $e \in S_k$ on the boundary simplify. For example,

$$\text{(105)}$$

As for the one-dimensional case, the proper normalization of $\mathbb{E}|\psi\rangle\langle\psi|^{\otimes k}$ follows directly from this property.

Instead of thinking in terms of contractions of two-dimensional tensor networks, it will later prove beneficial to think again in terms of multiplications of matrices. To that end, for any height $h$ of subsystems $A$ and $B$ [see. Fig. 4(a)], we define

$$\text{(106)}$$

as well as



$$(107)$$

For subsystems $A$ and $B$ as sketched in Fig. 4(a), we can now write Eq. (99) in terms of transfer matrices:

$$\mathrm{tr}\left(P\mathbb{E}|\psi\rangle\langle\psi|^{\otimes k}\right) = \langle\mathcal{I}_k|\mathcal{T}_e^c\mathcal{T}_\alpha^a\mathcal{T}_e^r\mathcal{T}_\beta^b|\mathcal{F}_k\rangle \qquad (108)$$

We provide an additional Mathematica package [63] that defines $T_\rho$ with $\rho \in S_k$ for $k \in \{1,\ldots,20\}$ according to Eq. (104). Once again, the package relies on the one provided by the authors of Ref. [64] for evaluating the Weingarten function.

### B. Estimating the decay of correlations

Equation (108) implies that, for subsystems $A$ and $B$ as sketched in Fig. 4(a), the decay of the average of each measure of correlation will again reduce to a statement in terms of transfer matrices. In particular, the decay will be determined by the spectrum of $\mathcal{T}_e$ with $e \in S_k$. Notice that, in addition to a dependence on $k$, the form and properties of $\mathcal{T}_e$ now depend also on the height $h$ of subsystems $A$ and $B$ [see. Fig. 4(a)]. However, as we prove in the following sections, its two leading eigenvalues are independent of $h$ for at least $k = 2$ and $k = 4$. Crucially, this will allow us to make statements about the decay of the averages of $I_2(A : B)$ and $N(A : B)$ for arbitrary $h$.

Note that we could, in principle, investigate vertical separation instead of horizontal separation because



$$(109)$$

The underlying exchange of indices does not affect the spectrum of the relevant identity transfer matrix and thus neither the average correlation length. This reflects the fact that the sequential generation procedure is symmetric in the horizontal and vertical spatial directions.

### C. Rényi-2 mutual information

In this section, we compute the average of the Rényi-2 mutual information $I_2(A : B)$ for random isoTNS that are

generated as sketched in Fig. 1(b1). Subsystems $A$ and $B$ are defined in Fig. 4(a).

As in one dimension, we will make the assumption that $\mathbb{E}\log(X) = \log(\mathbb{E}X)$ (see Sec. IV C). Step 1 of the proof of Result 4 (see Appendix O) is thus almost identical to Step 1 of the proof of Result 1.

Also Step 2 is largely analogous. To see this, let us take $\mathbb{E}\,\mathrm{tr}\left(\varrho_A^2\right)$ as an example. With $A$ and $B$ as defined in Fig. 4(a) and $x = (12)$,



$$(110)$$



$$(111)$$

$$= \langle\mathcal{I}_2|\mathcal{T}_e^c\mathcal{T}_x^a|\mathcal{F}_2\rangle. \qquad (112)$$

The resulting expression for $\mathbb{E}I_2(A : B)$,

$$\begin{aligned}
\mathbb{E}I_2(A : B) = &\log\left(\langle\mathcal{I}_2|\mathcal{T}_e^c\mathcal{T}_x^a\mathcal{T}_e^r\mathcal{T}_x^b|\mathcal{F}_2\rangle\right) \\
&- \log\left(\langle\mathcal{I}_2|\mathcal{T}_e^c\mathcal{T}_x^a|\mathcal{F}_2\rangle\right) \\
&- \log\left(\langle\mathcal{I}_2|\mathcal{T}_e^{c+a+r}\mathcal{T}_x^b|\mathcal{F}_2\rangle\right),
\end{aligned} \qquad (113)$$

resembles Eq. (68) closely.

The remaining technical challenge is the analysis of the spectrum of $\mathcal{T}_e$ with $e \in S_2$. Let $\lambda_1 > \lambda_2 > \cdots \geq 0$ denote its distinct eigenvalues. We show in Appendix M that

$$\lambda_1 = 1 \qquad \text{and} \qquad \lambda_2 = \frac{dD^3 - dD}{d^2D^4 - 1} \qquad (114)$$

and that both eigenvalues are nondegenerate. In the proof, that holds for any $h$, we map the contraction of tensors defining $\mathcal{T}_e$ with $e \in S_2$ to a multiplication of matrices. Using the Weingarten calculus, we show that $\mathcal{T}_e$ is upper block triangular, a property that simplifies the analysis of its spectrum. The main difficulty is then to prove that the specified $\lambda_2$ is indeed the subleading eigenvalue for all $h$. We do this by exploiting substochasticity.

Following the same reasoning as before, we find that the average of $I_2(A:B)$ decays exponentially as specified in Definition 1. We prove this result in Appendix O.

*Result 4.*—The average of $I_2(A:B)$ with respect to the random isoTNS ensemble and subsystems $A$ and $B$ as sketched in Fig. 4(a) decays exponentially as specified in Definition 1 with the average correlation length $\xi_{2D}$ defined in Eq. (35b).

### D. Trace distance and 2-norm

In this section, we show the exponential decay of the average of $N(A:B)$ for random isoTNS. As for the one-dimensional case, we do that to eventually make conclusions about the behavior of the trace distance $T(A:B)$.

While it is not trivial to compute the average of $N(A:B)$ with respect to the random isoTNS ensemble and subsystems $A$ and $B$ as sketched in Fig. 4(a), the computation follows along the lines of what we have laid out in Sec. IV D. In particular, we find that the decay of the average of $N(A:B)$ is determined by the spectrum of the transfer matrix $\mathcal{T}_e$ with $e \in S_4$.

Let $\lambda_1 > \lambda_2 > \cdots \geq 0$ denote the distinct eigenvalues of $\mathcal{T}_e$ with $e \in S_4$. As we show in Appendix N, for any $h$,

$$\lambda_1 = 1 \quad \text{and} \quad \lambda_2 = \frac{dD^3 - dD}{d^2 D^4 - 1}. \qquad (115)$$

As in one dimension, the former is nondegenerate, while the degeneracy of the latter is given by the number of transpositions in $S_4$,

$$w_2 = \binom{4}{2} = 6. \qquad (116)$$

While the analysis of the spectrum is, in principle, similar to that of the spectrum of $\mathcal{T}_e$ with $e \in S_2$, it is considerably more technical. This is largely due to the fact that the matrices whose multiplication defines $\mathcal{T}_e$ with $e \in S_4$ are significantly more complex. Still, we find also $\mathcal{T}_e$ with $e \in S_4$ to be upper block triangular, allowing us to show that the specified $\lambda_2$ is indeed the subleading eigenvalue.

Given $\lambda_2$ of $\mathcal{T}_e$ with $e \in S_4$ and the arguments developed in Sec. IV D, we can state the first result of this section, which we prove in Appendix P.

*Result 5.*—The average of $N(A:B)$ with respect to the random isoTNS ensemble and subsystems $A$ and $B$ as sketched in Fig. 4(a) decays exponentially as specified in Definition 1 with the average correlation length $\xi_{2D}$ defined in Eq. (35b).

We now turn to the case of correlations in isoTNS with arbitrary (but fixed) subsystems $A$ and $B$ [see Fig. 4(b)]. The decay of the average of $N(A:B)$ for arbitrary $A$ and $B$ can be bounded by employing the fact that the Schatten

2-norm satisfies [70]

$$\|\mathrm{tr}_B (X_{AB})\|_2 \leq \sqrt{\dim(B)}\, \|X_{AB}\|_2 , \qquad (117)$$

where $X_{AB}$ is any bipartite operator and $\dim(B)$ is the dimension of the Hilbert space that is traced out. This means that

$$N(A:B) \leq d^{|A_C| + |B_C|} N(A':B'), \qquad (118)$$

where $A$ is now an arbitrary subsystem, $A'$ is its (minimal) enclosing rectangle that touches the hypersurface, and $A_C = A' - A$. $B'$ and $B_C$ are defined analogously for $B$.

Using the statement above, we can bound the decay of the average of $N(A:B)$ for arbitrary $A$ and $B$ as a straightforward corollary of Result 5. We stress that we consider the regime in which the distance $r$ between $A$ and $B$ grows.

*Corollary 5.*—For arbitrary subsystems $A$ and $B$, the average of $N(A:B)$ with respect to the random isoTNS ensemble decays as

$$N(A:B) = O\left[\exp\left(-\frac{r}{\xi_{2D}}\right)\right], \qquad (119)$$

where the average correlation length $\xi_{2D}$ is defined in Eq. (35b).

Finally, we state a concentration result for $N(A:B)$, which, in combination with Eq. (17), also allows us to draw conclusions about the typical behavior of $T(A:B)$. As before, we consider arbitrary (but fixed) subsystems $A$ and $B$ [see Fig. 4(b)].

*Corollary 6.*—For arbitrary subsystems $A$ and $B$, sufficiently large $r$, and all $0 < \varepsilon < 1$, the random isoTNS ensemble satisfies

$$\mathrm{Pr}\left\{N(A:B) \geq K \exp\left[-\frac{(1-\varepsilon)r}{\xi_{2D}}\right]\right\} \leq \exp\left(-\frac{\varepsilon r}{\xi_{2D}}\right), \qquad (120)$$

where $K$ is constant with respect to $r$ and the average correlation length $\xi_{2D}$ is defined in Eq. (35b).

*Corollary 7.*—For arbitrary subsystems $A$ and $B$, sufficiently large $r$, and all $0 < \varepsilon < 1$, the random isoTNS ensemble satisfies

$$\mathrm{Pr}\left\{T(A:B) \geq K \exp\left[-\frac{(1-\varepsilon)r}{2\xi_{2D}}\right]\right\} \leq \exp\left(-\frac{\varepsilon r}{\xi_{2D}}\right), \qquad (121)$$

where $K$ is constant with respect to $r$ and the average correlation length $\xi_{2D}$ is defined in Eq. (35b).

The proofs and discussions of these results are identical to their one-dimensional counterparts.

The tools we have developed in this and the previous sections should, in principle, allow us to make statements

also about the decay of the average of the von Neumann mutual information $I(A : B)$ with respect to the random isoTNS ensemble. As in one dimension, we would need to investigate the spectrum of $\mathcal{T}_e$ with $e \in S_k$ for all $k \geq 2$. However, as the analysis of the spectrum of $\mathcal{T}_e$ is already quite technical for $e \in S_4$ with our methods, we refrain from tackling the spectrum for $k \geq 5$ here.

## VI. CONCLUSION

We have investigated the average behavior of correlations between two distant subsystems $A$ and $B$ for ensembles of random MPS and isoTNS. As measures of correlation, we have considered the Rényi-$\alpha$ mutual information, a measure arising from the Hilbert-Schmidt norm, the trace distance, and the von Neumann mutual information. We have shown that the average of each considered measure exhibits an exponential decay. Our results can equivalently be seen as describing states resulting from quantum circuits with a sequential architecture and Haar random gates.

By leveraging the Weingarten calculus, we have developed a mathematical framework that allows to infer the average correlation length from the subleading eigenvalue of an appropriately defined transfer matrix. We have computed the averages of the Rényi-$\alpha$ mutual information and the measure arising from the Hilbert-Schmidt norm to show the emergence of an average correlation length that depends only on the underlying spatial dimension but not the considered measure. In particular, the average correlation length for random MPS increases weakly with the bond dimension $D$ and converges rapidly (as $D$ grows) to $1/\log(d)$, where $d$ is the physical dimension. In contrast, the average correlation length for random isoTNS decreases with $D$. The highest average correlation length

for random isoTNS is achieved with the lowest nontrivial bond dimension ($D = 2$).

Using elementary concentration results, we have furthermore deduced the typical behavior of the measure arising from the Hilbdert-Schmidt norm, which has in turn allowed us to make similar statements about the trace distance. For MPS, we have been able to give strong indications that the universal correlation length applies also to the von Neumann mutual information, and also any Rényi-$\alpha$ mutual information for integer values of $\alpha \geq 1$. It would be interesting to prove this behavior rigorously and also investigate its validity for isoTNS. Another possible future direction would be to study average correlations in more general random PEPS as well as other types of quantum circuit architectures. To extend our results from the class of isoTNS to generic PEPS, a protocol to sequentially generate the latter would be needed.

It could also be interesting to investigate the implications of our results on state-preparation schemes and variational algorithms for MPS and particularly isoTNS. Regarding the former, it would be interesting to see if typical sequentially generated states, being short-range correlated, can be prepared by shallower circuits. On the numerical side, we expect our findings to be relevant for the performance of variational algorithms that approximate long-range correlated states with MPS and isoTNS. Our results imply that (even more for isoTNS than for MPS) polynomially decaying states are very special examples of the two variational families.

## ACKNOWLEDGMENTS

## APPENDIX A: $K$-FOLD TWIRL

In this Appendix, we present some additional details on the $k$-fold twirl. In particular, we go from Eq. (19) to Eq. (20). To do this, we will need a result of Ref. [54], which appears as Corollary 2.4 in Ref. [51]. We state it without proof as Lemma 1.

*Lemma 1.*—Let $k$ be a positive integer, and $(i_1, \ldots, i_k)$, $(j_1, \ldots, j_k)$, $(\ell_1, \ldots, \ell_k)$, and $(m_1, \ldots, m_k)$ be $k$-tuples of positive integers. Then,

$$\int dU\, U_{i_1 j_i} \cdots U_{i_k j_k} \overline{U_{m_1 \ell_1}} \cdots \overline{U_{m_k \ell_k}} = \sum_{\sigma, \tau \in S_k} \mathrm{Wg}\left(\sigma^{-1}\tau, q\right) \delta_{i_1 m_{\sigma(1)}} \cdots \delta_{i_k m_{\sigma(k)}} \delta_{j_1 \ell_{\tau(1)}} \cdots \delta_{j_k l_{\tau(k)}}, \qquad (A1)$$

where the integration is with respect to the Haar measure on the unitary group $U(q)$.

By Lemma 1,

$$\left(\mathcal{T}_U^{(k)}(X)\right)_{i_1\cdots i_k m_1\cdots m_k} = \sum_{\substack{j_1,\ldots,j_k \\ \ell_1,\ldots,\ell_k}} \int dU \, U_{i_1 j_i}\cdots U_{i_k j_k} X_{j_1\cdots j_k \ell_1\cdots \ell_k} \overline{U_{m_1\ell_1}}\cdots \overline{U_{m_k\ell_k}} \tag{A2}$$

$$= \sum_{\substack{j_1,\ldots,j_k \\ \ell_1,\ldots,\ell_k}} \sum_{\sigma,\tau\in S_k} \mathrm{Wg}\left(\sigma^{-1}\tau,q\right) \delta_{i_1 m_{\sigma(1)}}\cdots\delta_{i_k m_{\sigma(k)}} X_{j_1\cdots j_k \ell_1\cdots \ell_k} \delta_{j_1 \ell_{\tau(1)}}\cdots\delta_{j_k l_{\tau(k)}} \tag{A3}$$

$$= \sum_{j_1,\ldots,j_k} \sum_{\sigma,\tau\in S_k} \mathrm{Wg}\left(\sigma^{-1}\tau,q\right) \delta_{i_1 m_{\sigma(1)}}\cdots\delta_{i_k m_{\sigma(k)}} X_{j_1\cdots j_k j_{\tau^{-1}(1)}\cdots j_{\tau^{-1}(k)}} \tag{A4}$$

$$= \sum_{\sigma,\tau\in S_k} \mathrm{Wg}\left(\sigma^{-1}\tau,q\right) \left(P_{\sigma^{-1}}^{(q)}\right)_{i_1\cdots i_k m_1\cdots m_k} \mathrm{tr}\left(XP_\tau^{(q)}\right), \tag{A5}$$

where, in the final line, we have used that

$$\left(P_{\sigma^{-1}}^{(q)}\right)_{i_1\cdots i_k m_1\cdots m_k} = \langle i_1\cdots i_k | P_{\sigma^{-1}}^{(q)} | m_1\cdots m_k\rangle \tag{A6}$$

$$= \langle i_1\cdots i_k | m_{\sigma(1)}\cdots m_{\sigma(k)}\rangle \tag{A7}$$

$$= \delta_{i_1 m_{\sigma(1)}}\cdots\delta_{i_k m_{\sigma(k)}} \tag{A8}$$

and that

$$\mathrm{tr}\left(XP_\tau^{(q)}\right) = \sum_{j_1,\ldots,j_k} \langle j_1\cdots j_k | XP_\tau^{(q)} | j_1\cdots j_k\rangle \tag{A9}$$

$$= \sum_{j_1,\ldots,j_k} \langle j_1\cdots j_k | X | j_{\tau^{-1}(1)}\cdots j_{\tau^{-1}(k)}\rangle \tag{A10}$$

$$= X_{j_1\cdots j_k j_{\tau^{-1}(1)}\cdots j_{\tau^{-1}(k)}}. \tag{A11}$$

Thus,

$$\mathcal{T}_U^{(k)}(X) = \sum_{\sigma,\tau\in S_k} \mathrm{Wg}\left(\sigma^{-1}\tau,q\right) P_{\sigma^{-1}}^{(q)} \mathrm{tr}\left(XP_\tau^{(q)}\right) \tag{A12}$$

$$= \sum_{\sigma,\tau\in S_k} \mathrm{Wg}\left(\sigma\tau^{-1},q\right) P_\sigma^{(q)} \mathrm{tr}\left(XP_{\tau^{-1}}^{(q)}\right) \tag{A13}$$

$$= \sum_{\sigma,\tau\in S_k} \mathrm{Wg}\left(\sigma\tau^{-1},q\right) P_\sigma^{(q)} \mathrm{tr}\left[X\left(P_\tau^{(q)}\right)^T\right], \tag{A14}$$

which coincides with Eq. (20).

In addition, let us confirm that $\mathcal{T}_U^{(k)}\left(P_\rho^{(q)}\right) = P_\rho^{(q)}$ [52,53]. Indeed,

$$\mathcal{T}_U^{(k)}\left(P_\rho^{(q)}\right) = \sum_{\sigma,\tau\in S_k} \mathrm{Wg}\left(\sigma\tau^{-1},q\right) P_\sigma^{(q)} \mathrm{tr}\left[P_\rho^{(q)}\left(P_\tau^{(q)}\right)^T\right] \tag{A15}$$

$$= \sum_{\sigma,\tau\in S_k} \mathrm{Wg}\left(\sigma\tau^{-1},q\right) P_\sigma^{(q)} q^{\#\left(\rho\tau^{-1}\right)} \tag{A16}$$

$$= \sum_{\sigma\in S_k} \delta_{\rho\sigma} P_\sigma^{(q)} \tag{A17}$$

$$= P_\rho^{(q)}, \tag{A18}$$

where, in the third line, we have used the definition of the Weingarten function.

## APPENDIX B: PROOFS OF PROPOSITIONS 1 AND 2

*Proposition 1.*—The eigenvalues of $T_e$ with $e \in S_k$ are non negative for any $k \geq 2$.

*Proposition 2.*—$T_e$ with $e \in S_k$ is diagonalizable for any $k \geq 2$.

To prove Propositions 1 and 2, we will define matrices $X \in \mathbb{R}^{k! \times k!}$ and $Y \in \mathbb{R}^{k! \times k!}$ so that $C_k = WXY$. We will then discuss some properties of those three matrices. The proofs themselves will boil down to similarity.

We define the diagonal matrix $X \in \mathbb{R}^{k! \times k!}$ via

$$X_{\sigma\sigma} = d^{\#(\sigma)}. \tag{B1}$$

It is evident that $X$ is positive definite.

We define the Gram matrix $Y \in \mathbb{R}^{k! \times k!}$ via

$$Y_{\sigma\theta} = \mathrm{tr}\left[ P_\sigma^{(D)} \left( P_\theta^{(D)} \right)^T \right] = D^{\#\left(\sigma\theta^{-1}\right)}. \tag{B2}$$

$Y$ is positive semidefinite because it is a Gram matrix.

The Weingarten matrix $\mathrm{Wg}\left(\sigma\tau^{-1}, q\right) = \left(G^{-1}\right)_{\sigma\tau}$ is positive definite because the Gram matrix $G$ [see Eq. (22)] is positive definite [71].

We will need the fact that $YW$ is positive semidefinite. $YW$ has non negative eigenvalues because it is a product of a positive semidefinite ($Y$) and a positive definite matrix ($W$) (see Corollary 7.6.2 of Ref. [72]). Furthermore, $YW$ is symmetric:

$$\langle\tau|YW|\theta\rangle = \sum_{\sigma \in S_k} \mathrm{Wg}\left(\sigma\tau^{-1}, dD\right) D^{\#\left(\sigma\theta^{-1}\right)} \tag{B3}$$

$$= \sum_{\pi \in S_k} \mathrm{Wg}\left(\pi\theta\tau^{-1}, dD\right) D^{\#(\pi)} \tag{B4}$$

$$= \sum_{\pi \in S_k} \mathrm{Wg}\left(\tau\theta^{-1}\pi^{-1}, dD\right) D^{\#(\pi)} \tag{B5}$$

$$= \sum_{\pi \in S_k} \mathrm{Wg}\left(\theta^{-1}\pi^{-1}\tau, dD\right) D^{\#(\pi)} \tag{B6}$$

$$= \sum_{\varphi \in S_k} \mathrm{Wg}\left(\theta^{-1}\varphi, dD\right) D^{\#\left(\tau\varphi^{-1}\right)} \tag{B7}$$

$$= \sum_{\varphi \in S_k} \mathrm{Wg}\left(\varphi\theta^{-1}, dD\right) D^{\#\left(\varphi\tau^{-1}\right)} \tag{B8}$$

$$= \langle\theta|YW|\tau\rangle. \tag{B9}$$

In the third line, we have used that $\mathrm{Wg}(\alpha, q) = \mathrm{Wg}\left(\alpha^{-1}, q\right)$, and, in the fourth line, we have used that $\mathrm{Wg}(\alpha, q) = \mathrm{Wg}\left(\beta\alpha\beta^{-1}, q\right)$. Both identities are a result of the Weingarten function being sensitive only to the conjugacy class of a given permutation.

*Proof of Proposition 1.*—$C_k = WXY$ is similar to $XYW$. A product of a positive definite ($X$) and a positive semidefinite matrix ($YW$), $XYW$ has non negative eigenvalues. The statement follows by similarity. ∎

*Proof of Proposition 2.*—$C_k = WXY$ is similar to $XYW$, which is similar to $X^{1/2}YWX^{1/2}$. Because $YW$ is symmetric, so is $X^{1/2}YWX^{1/2}$, which makes the latter diagonalizable. The statement follows. ∎

## APPENDIX C: PROOF OF PROPOSITION 3

*Proposition 3.*—Let $\lambda_1 > \lambda_2 > \cdots \geq 0$ denote the distinct eigenvalues of $T_e$ with $e \in S_k$. Then, $\lambda_1 = 1$, and it is non-degenerate for any $k \geq 2$ if $d \geq 2$.

*Proof.*—For what follows, it is convenient to define the transfer matrix $\Sigma_e$ with $e \in S_k$ via

$$
=\boxed{\Sigma_\rho}= \;=\; =\boxed{R}= \;=\; =\boxed{P^T}-\boxed{W}-\boxed{P}= \;.
\tag{C1}
$$

Our strategy will be to prove the claimed spectral property for $\Sigma_e$. This is enough because, for any two operators $X$ and $Y$ for which $XY$ and $YX$ are well defined, the sets of eigenvalues of $XY$ and $YX$ are the same (up to zeros and the multiplicity of the nonzero eigenvalues). By Eqs. (55) and (C1), $T_\rho$ and $\Sigma_\rho$ are related in exactly this way.

As $\Sigma_e$ arises from the contraction of the quantum channel underlying $R$ [see Eq. (48)] with the identity permutation, it can be understood as a generalization of the $k$-fold twirling operator. In particular, it involves an "environment" $E$ of dimension $\left(\mathbb{C}^d\right)^{\otimes k}$ that is eventually traced out. As a superoperator (that is, without using the operator-vector correspondence), it reads [see Eq. (19)]

$$
\Sigma_e(\cdot) = \int dU \; \mathrm{tr}_E \left\{ U^{\otimes k} \left[ \bigotimes_{l=1}^{k} |0\rangle\langle 0|_{E_l} \otimes (\cdot)_{S_l} \right] (U^\dagger)^{\otimes k} \right\},
\tag{C2}
$$

where the integration is with respect to the Haar measure on the unitary group $U(dD)$, $E = \bigotimes_{l=1}^{k} E_l$ corresponds to $\left(\mathbb{C}^d\right)^{\otimes k}$, and $S = \bigotimes_{l=1}^{k} S_l$ corresponds to $\left(\mathbb{C}^D\right)^{\otimes k}$. It is apparent that Eq. (C2) represents a convex combination of quantum channels (note the Stinespring dilation form), and thus the resulting operator is also a valid quantum channel. This implies that 1 is an eigenvalue of $\Sigma_e$ and that there is no eigenvalue of greater modulus [73].

We now prove that no other eigenvalue of the same modulus exists. To that end, we will show that $\Sigma_e$ is a primitive channel [73], which implies said property. A quantum channel is primitive if and only if the spanning space formed by products of its Kraus operators,

$$
\mathcal{K}_m = \mathrm{span}\left( \left\{ \prod_{k=1}^{m} K_{i_k} \right\} \right),
\tag{C3}
$$

is equal to the full matrix algebra for some integer $m$, that is, $\mathcal{K}_m = \mathcal{M}_{D^k}(\mathbb{C})$ [74]. We show that this condition is satisfied for $\Sigma_e$ if $d \geq 2$.

Indeed, the Haar integral in Eq. (C2), together with the partial trace, can be understood as a (redundant) Kraus decomposition. Precisely, we can take

$$
\{K_i\}_i = \left\{ \mathrm{tr}_E \left[ \left( \bigotimes_{l=1}^{k} |+\rangle\langle \psi_l|_{E_l} \otimes I_{S_l} \right) U^{\otimes k} \right] \right\}_{U,\psi},
\tag{C4}
$$

where $U \in U(dD)$, $|\psi_l\rangle \in \{|0\rangle, \ldots, |d-1\rangle\}$ is the computational basis of the $l$th replica of the environment, and $|+\rangle = \sum_{j=0}^{d-1} |j\rangle / \sqrt{d}$. The latter is a choice (instead of $|0\rangle$) made for later convenience. It remains to show that there exists an integer $m = m(k,d,D)$ such that $\mathcal{K}_m = \mathcal{M}_{D^k}(\mathbb{C})$ if $d \geq 2$. First of all, note that the above fails for $d = 1$ (that is, if the environment $E$ is trivial). This is because span $\left( \left\{ U^{\otimes k} \right\}_U \right)$ coincides with the symmetric subspace over the $k$ subsystems [48]. However, $d = 2$ is already enough to span the full matrix algebra.

An explicit construction to show this fact amounts to taking $U$ to be a controlled unitary gate, where the control system is $E$. In particular, consider

$$
|\psi_l\rangle = |\delta_{lr}\rangle \qquad \text{and} \qquad U = |0\rangle\langle 0|_E \otimes I_S + \sum_{j=1}^{d-1} |j\rangle\langle j|_E \otimes V_S,
\tag{C5}
$$

where $r \in \{1, \ldots, k\}$, and $V_S$ is unitary. This results in Kraus operators [see Eq. (C4)] of the form

$$
I_{S_1} \otimes \cdots \otimes I_{S_{r-1}} \otimes V \otimes I_{S_{r+1}} \cdots \otimes I_{S_k}.
\tag{C6}
$$

Taking (finite) products, as Eq. (C3) dictates, is enough to build a basis of the vector space $\mathcal{M}_{D^k}(\mathbb{C})$. Note that this construction requires two control levels (that is, $d \geq 2$).

This concludes the proof. ∎

## APPENDIX D: FURTHER PROPERTIES OF THE TRANSFER MATRIX $T_e$

In this Appendix, we state and prove statements about the structure of the transfer matrix $T_e$ with $e \in S_k$ that will motivate Conjecture 1. We prove the statements in Appendices E, F, and G.

Moving forward, we denote $T_e$ with $e \in S_k$ by $C_k$. Each entry

$$\langle \tau | C_k | \theta \rangle = \sum_{\sigma \in S_k} \mathrm{Wg}\left(\sigma \tau^{-1}, dD\right) d^{\#(\sigma)} D^{\#\left(\sigma \theta^{-1}\right)} \tag{D1}$$

of $C_k \in \mathbb{R}^{k! \times k!}$ is a sum of $k!$ terms. While this may sound daunting, $C_k$ exhibits a structure that reduces its complexity.

As formalized by Proposition 4, the entries $\langle \tau | C_k | \theta \rangle$ of $C_k$ exhibit a dependence on the conjugacy class of their indices. In particular, if we know the entries of the column given by $\theta \in S_k$, we also know the entries of the columns given by permutations in the same conjugacy class as $\theta$.

*Proposition 4.*—For any $\pi \in S_k$,

$$\langle \tau | C_k | \theta \rangle = \langle \pi \tau \pi^{-1} | C_k | \pi \theta \pi^{-1} \rangle. \tag{D2}$$

Certain entries of $C_k$ vanish, while others are given by the entries of $C_{k-1}$. Proposition 5 captures these two statements.

*Proposition 5.*—For all $\theta \in S_k$ with $\theta(k) = k$,

$$\langle \tau | C_k | \theta \rangle = \delta_{k, \tau(k)} \langle \tau^{\downarrow} | C_{k-1} | \theta^{\downarrow} \rangle, \tag{D3}$$

where $\rho^{\downarrow} \in S_{k-1}$ is the restriction of $\rho \in S_k$ with $\rho(k) = k$ to the permutation on $\{1, \ldots, k-1\}$.

To understand the strength of Propositions 4 and 5, let us have a look at $C_2$ and $C_3$, the two most simple transfer matrices. With bases $S_2 = \{e, (12)\}$ and $S_3 = \{e, (12), (13), (23), (123), (132)\}$, one finds that

$$C_2 = \begin{pmatrix} 1 & \alpha \\ 0 & \beta \end{pmatrix} \quad \text{and} \quad C_3 = \begin{pmatrix} 1 & \alpha & \alpha & \alpha & \gamma & \gamma \\ 0 & \beta & 0 & 0 & \delta & \delta \\ 0 & 0 & \beta & 0 & \delta & \delta \\ 0 & 0 & 0 & \beta & \delta & \delta \\ 0 & 0 & 0 & 0 & \varepsilon & \zeta \\ 0 & 0 & 0 & 0 & \zeta & \varepsilon \end{pmatrix}, \tag{D4}$$

where each Greek letter corresponds to some function of $d$ and $D$ whose exact form is not relevant here. The entries in the first four columns of $C_3$ are fully determined by those of $C_2$. The entries in the last two columns of $C_3$ do not arise from those of $C_2$, but the sixth column is a permutation of the fifth.

Note that we are deliberately choosing a basis $S_k = \{s_1, \ldots, s_{k!}\}$ that makes the special structure of $C_k$ more apparent. In particular, we sort permutations so that those with $i$ fixed points come before those with $i - 1$ fixed points. We group permutations that have common fixed points and then those that are in the the same conjugacy class. Given this basis, Propositions 4 and 5 imply that $C_k$ is block triangular with $k$ diagonal blocks.

We denote by $C_k^{(1)}$ the diagonal block with $\tau = \theta = e \in S_k$ and by $C_k^{(i)}$ with $2 \leq i \leq k$ the diagonal block corresponding to $\tau, \theta \in S_k$ with $k - i$ fixed points. The spectrum of $C_k$ is then given by

$$\lambda(C_k) = \lambda\left(C_k^{(1)}\right) \cup \lambda\left(C_k^{(2)}\right) \cup \cdots \cup \lambda\left(C_k^{(k)}\right). \tag{D5}$$

It is apparent that $C_k^{(1)}$ has a single, nondegenerate eigenvalue

$$\mu_1 = \langle e | C_k | e \rangle = 1. \tag{D6}$$

$C_k^{(2)}$ has a single, degenerate eigenvalue

$$\mu_2 = \langle (12)|C_k|(12)\rangle = \frac{dD^2 - d}{d^2 D^2 - 1},$$ (D7)

which corresponds to the expression of $\beta$ in Eq. (D4). The degeneracy of $\mu_2$ is given by the size of the block, which is in turn given by the number of transpositions in $S_k$,

$$v_2 = \binom{k}{2}.$$ (D8)

By Proposition 3, 1 is the leading eigenvalue of $C_k$ and nondegenerate. We conjecture that $\mu_2$ is the subleading eigenvalue of $C_k$ and that it has degeneracy $v_2$. The statement of Conjecture 1 holds for $k \in \{2, 3, 4\}$, and we have numerical evidence suggesting so for $k \in \{5, 6, 7\}$ [63].

*Conjecture 1.*—Let $\lambda_1 > \lambda_2 > \cdots \geq 0$ denote the distinct eigenvalues of $C_k$. Then, $\lambda_2 = \mu_2$ with degeneracy $w_2 = v_2$ for any $k \geq 2$.

As formalized by Proposition 6, the statements above hold for any transfer matrix $T_\rho$ with $\rho \in S_k$ because $T_\rho$ is similar to $C_k$.

*Proposition 6.*—For any $\rho \in S_k$,

$$T_\rho = Q_\rho^T C_k Q_\rho \qquad \text{with} \qquad Q_\rho = \sum_{\pi \in S_k} |\rho\pi\rangle\langle\pi|.$$ (D9)

## APPENDIX E: PROOF OF PROPOSITION 4

*Proposition 4.*—For any $\pi \in S_k$,

$$\langle\tau|C_k|\theta\rangle = \langle\pi\tau\pi^{-1}|C_k|\pi\theta\pi^{-1}\rangle.$$ (D2)

*Proof.*—It holds that

$$\langle\tau|C_k|\theta\rangle$$ (E1)

$$= \sum_{\sigma \in S_k} \text{Wg}\left(\sigma\tau^{-1}, dD\right) d^{\#(\sigma)} D^{\#\left(\sigma\theta^{-1}\right)}$$ (E2)

$$= \sum_{\varphi \in S_k} \text{Wg}\left(\pi^{-1}\varphi\pi\tau^{-1}, dD\right) d^{\#\left(\pi^{-1}\varphi\pi\right)} D^{\#\left(\pi^{-1}\varphi\pi\theta^{-1}\right)}$$ (E3)

$$= \sum_{\varphi \in S_k} \text{Wg}\left(\varphi\pi\tau^{-1}\pi^{-1}, dD\right) d^{\#\left(\pi^{-1}\varphi\pi\right)} D^{\#\left(\varphi\pi\theta^{-1}\pi^{-1}\right)}$$ (E4)

$$= \sum_{\varphi \in S_k} \text{Wg}\left[\varphi\left(\pi\tau\pi^{-1}\right)^{-1}, dD\right] d^{\#(\varphi)} D^{\#\left[\varphi\left(\pi\theta\pi^{-1}\right)^{-1}\right]}$$ (E5)

$$= \langle\pi\tau\pi^{-1}|C_k|\pi\theta\pi^{-1}\rangle,$$ (E6)

where, in the second line, we have used that the conjugation map is an isomorphism. In the third line, we have used that $\text{Wg}(\alpha, q) = \text{Wg}\left(\beta\alpha\beta^{-1}, q\right)$ and that $\#(\alpha) = \#\left(\beta\alpha\beta^{-1}\right)$. Both identities are a result of the two functions being sensitive only to the conjugacy class of a given permutation. ∎

## APPENDIX F: PROOF OF PROPOSITION 5

*Proposition 5.*—For all $\theta \in S_k$ with $\theta(k) = k$,

$$\langle\tau|C_k|\theta\rangle = \delta_{k,\tau(k)}\langle\tau^\downarrow|C_{k-1}|\theta^\downarrow\rangle,$$ (D3)

where $\rho^\downarrow \in S_{k-1}$ is the restriction of $\rho \in S_k$ with $\rho(k) = k$ to the permutation on $\{1, \ldots, k-1\}$.

To prove Proposition 5, we will need a number of ingredients. The main one will be Proposition 2.2 of Ref. [75], which we state without proof as Lemma 2.

*Lemma 2.*—For any $\rho \in S_k$,

$$\sum_{i=1}^{k-1} \mathrm{Wg}\big((ik)\pi, q\big) + q\, \mathrm{Wg}(\pi, q) = \delta_{k,\pi(k)}\, \mathrm{Wg}\left(\pi^{\downarrow}, q\right), \tag{F1}$$

where $(ik)$ denotes the transposition of elements $i$ and $k$.

To use Lemma 2, we must do two things. First, we must split the sum in Eq. (D1) into certain sums of $k$ terms. We employ Lemma 3 to achieve this.

*Lemma 3.*—For any $\alpha \in S_k$, there exists a $\beta \in S_k$ with $\beta(k) = k$ such that either $\alpha = \beta$ or $\alpha = (ik)\beta$ with $i \in \{1, \ldots, k-1\}$.

*Proof.*—If $\alpha(k) = k$, then $\alpha = \beta$. Otherwise, $k$ is in exactly one of the disjoint cycles of $\alpha$. Without loss of generality, say $\alpha = (kia_3a_4\cdots)(b_1b_2\cdots)\cdots$. Then, $\alpha = (ik)(ia_3a_4\cdots)(b_1b_2\cdots)\cdots \equiv (ik)\beta$. ∎

Second, we must ensure that summands in Eq. (D1) with $\mathrm{Wg}(\pi, dD)$ have a factor $dD$ while those with $\mathrm{Wg}\big((ik)\pi, dD\big)$ do not. We achieve this with Lemma 5 whose proof employs Lemma 4 of Ref. [76], which we state without proof as Lemma 4. The lemma also appears as Lemma 1 in Ref. [77].

*Lemma 4.*—Let $\alpha \in S_k$ be a transposition, and $\beta \in S_k$ such that $\#(\beta) = u$. If the elements exchanged by $\alpha$ are not in the same cycle of $\beta$, then $\#(\alpha\beta) = u - 1$.

*Lemma 5.*—Let $\varphi, \theta \in S_k$ with $\varphi(k) = k$ and $\theta(k) = k$. Then,

1. $\#\left(\varphi\theta^{-1}\right) = \#\left((ik)\varphi\theta^{-1}\right) + 1$ for all $i \in \{1, \ldots, k-1\}$.
2. $\#\left(\varphi\theta^{-1}\right) - \#(\varphi) = \#\left((ik)\varphi\theta^{-1}\right) - \#\left((ik)\varphi\right)$ for all $i \in \{1, \ldots, k-1\}$.

*Proof of 1.*—Let $\varphi, \theta \in S_k$ with $\varphi(k) = k$ and $\theta(k) = k$. Then, $\left(\varphi\theta^{-1}\right)(k) = k$. That is, $k$ is in a cycle by itself and thus not in the same cycle of $\varphi\theta^{-1}$ as $i$. The statement follows with Lemma 4. ∎

*Proof of 2.*—Let $\varphi, \theta \in S_k$ with $\varphi(k) = k$ and $\theta(k) = k$. Then, $i$ and $k$ are not in the same cycle of neither $\varphi\theta^{-1}$ nor $\varphi$. By Lemma 4, $\#\left(\varphi\theta^{-1}\right) = \#\left((ik)\varphi\theta^{-1}\right) + 1$ and $\#(\varphi) = \#\left((ik)\varphi\right) + 1$. Thus, $\#\left(\varphi\theta^{-1}\right) - \#(\varphi) = \#\left((ik)\varphi\theta^{-1}\right) + 1 - \#\left((ik)\varphi\right) - 1 = \left((ik)\varphi\theta^{-1}\right) - \#\left((ik)\varphi\right)$. ∎

With that, we can prove Proposition 5.

*Proof of Proposition 5.*—By Lemma 3,

$$\langle\tau|C_k|\theta\rangle = \sum_{\sigma\in S_k} \mathrm{Wg}\left(\sigma\tau^{-1}, dD\right) d^{\#(\sigma)} D^{\#\left(\sigma\theta^{-1}\right)} \tag{F2}$$

$$= \sum_{\substack{\varphi\in S_k \\ \varphi(k)=k}} \left\{ \sum_{i=1}^{k-1} \mathrm{Wg}\left((ik)\varphi\tau^{-1}, dD\right) d^{\#\left((ik)\varphi\right)} D^{\#\left((ik)\varphi\theta^{-1}\right)} + \mathrm{Wg}\left(\varphi\tau^{-1}, dD\right) d^{\#(\varphi)} D^{\#\left(\varphi\theta^{-1}\right)} \right\}. \tag{F3}$$

By Lemma 5, for $\theta \in S_k$ with $\theta(k) = k$,

$$\langle\tau|C_k|\theta\rangle = \sum_{\substack{\varphi\in S_k \\ \varphi(k)=k}} \frac{1}{d^{\#\left(\varphi\theta^{-1}\right)-\#(\varphi)}} \left\{ \sum_{i=1}^{k-1} \mathrm{Wg}\left((ik)\varphi\tau^{-1}, dD\right) (dD)^{\#\left((ik)\varphi\theta^{-1}\right)} + \mathrm{Wg}\left(\varphi\tau^{-1}, dD\right) (dD)^{\#\left(\varphi\theta^{-1}\right)} \right\} \tag{F4}$$

$$= \sum_{\substack{\varphi\in S_k \\ \varphi(k)=k}} \frac{(dD)^{\#\left(\varphi\theta^{-1}\right)-1}}{d^{\#\left(\varphi\theta^{-1}\right)-\#(\varphi)}} \left\{ \sum_{i=1}^{k-1} \mathrm{Wg}\left((ik)\varphi\tau^{-1}, dD\right) + dD\, \mathrm{Wg}\left(\varphi\tau^{-1}, dD\right) \right\} \tag{F5}$$

$$= \sum_{\substack{\varphi\in S_k \\ \varphi(k)=k}} d^{\#(\varphi)-1} D^{\#\left(\varphi\theta^{-1}\right)-1} \left\{ \sum_{i=1}^{k-1} \mathrm{Wg}\left((ik)\varphi\tau^{-1}, dD\right) + dD\, \mathrm{Wg}\left(\varphi\tau^{-1}, dD\right) \right\} \tag{F6}$$

$$= \sum_{\substack{\varphi \in S_k \\ \varphi(k)=k}} d^{\#(\varphi^{\downarrow})} D^{\#\left[\left(\varphi\theta^{-1}\right)^{\downarrow}\right]} \left\{ \sum_{i=1}^{k-1} \mathrm{Wg}\left((ik)\varphi\tau^{-1}, dD\right) + dD\,\mathrm{Wg}\left(\varphi\tau^{-1}, dD\right) \right\}, \tag{F7}$$

where, in the final line, we have used that $\#\left(\alpha^{\downarrow}\right) = \#(\alpha) - 1$.

By Lemma 2,

$$\langle \tau | C_k | \theta \rangle = \sum_{\substack{\varphi \in S_k \\ \varphi(k)=k}} \delta_{k,(\varphi\tau^{-1})(k)}\, \mathrm{Wg}\left[\left(\varphi\tau^{-1}\right)^{\downarrow}, dD\right] d^{\#(\varphi^{\downarrow})} D^{\#\left[\left(\varphi\theta^{-1}\right)^{\downarrow}\right]} \tag{F8}$$

$$= \delta_{k,\tau(k)} \sum_{\substack{\varphi \in S_k \\ \varphi(k)=k}} \mathrm{Wg}\left[\left(\varphi\tau^{-1}\right)^{\downarrow}, dD\right] d^{\#(\varphi^{\downarrow})} D^{\#\left[\left(\varphi\theta^{-1}\right)^{\downarrow}\right]} \tag{F9}$$

$$= \delta_{k,\tau(k)} \langle \tau^{\downarrow} | C_{k-1} | \theta^{\downarrow} \rangle, \tag{F10}$$

where, in the second line, we have used that $\left(\varphi\tau^{-1}\right)(k) = k$ if and only if $\tau(k) = k$ because $\varphi(k) = k$.

This concludes the proof. ∎

## APPENDIX G: PROOF OF PROPOSITION 6

*Proposition 6.*—For any $\rho \in S_k$,

$$T_\rho = Q_\rho^T C_k Q_\rho \qquad \text{with} \qquad Q_\rho = \sum_{\pi \in S_k} |\rho\pi\rangle\langle\pi|. \tag{D9}$$

*Proof.*—It holds that

$$\langle \tau | T_\rho | \theta \rangle = \sum_{\sigma \in S_k} \mathrm{Wg}\left(\sigma\tau^{-1}, dD\right) d^{\#(\sigma\rho)} D^{\#\left(\sigma\theta^{-1}\right)} \tag{G1}$$

$$= \sum_{\sigma \in S_k} \mathrm{Wg}\left(\rho\sigma\tau^{-1}\rho^{-1}, dD\right) d^{\#(\rho\sigma)} D^{\#\left(\rho\sigma\theta^{-1}\rho^{-1}\right)} \tag{G2}$$

$$= \sum_{\sigma \in S_k} \mathrm{Wg}\left[\rho\sigma(\rho\tau)^{-1}, dD\right] d^{\#(\rho\sigma)} D^{\#\left[\rho\sigma(\rho\theta)^{-1}\right]} \tag{G3}$$

$$= \sum_{\pi \in S_k} \mathrm{Wg}\left[\pi(\rho\tau)^{-1}, dD\right] d^{\#(\pi)} D^{\#\left[\pi(\rho\theta)^{-1}\right]} \tag{G4}$$

$$= \langle \rho\tau | C_k | \rho\theta \rangle, \tag{G5}$$

where, in the second line, we have used that $\mathrm{Wg}(\alpha, q) = \mathrm{Wg}\left(\beta\alpha\beta^{-1}, q\right)$ and that $\#(\alpha) = \#\left(\beta\alpha\beta^{-1}\right)$. Both identities are a result of the two functions being sensitive only to the conjugacy class of a given permutation. In the fourth line, we have used that the left-multiplication map is an isomorphism. ∎

## APPENDIX H: PROOF OF RESULT 2

*Result 2.*—The average of $N(A:B)$ with respect to the random MPS ensemble and subsystems $A$ and $B$ as sketched in Fig. 2(a) decays exponentially as specified in Definition 1 with the average correlation length $\xi_{1D}$ defined in Eq. (35a).

*Proof.*—We split the proof into four steps, following the structure of the proof of Result 1.

Step 1: We rewrite $\mathbb{E}N(A:B)$ in terms of expressions of the form of Eq. (44). With the Hilbert-Schmidt inner product,

$$\mathbb{E}N(A:B) = \mathbb{E}\operatorname{tr}\left(\varrho_{AB}^2\right) + \mathbb{E}\operatorname{tr}\left(\varrho_A^2\right)\operatorname{tr}\left(\varrho_B^2\right) - 2\mathbb{E}\operatorname{tr}\left[\varrho_{AB}\left(\varrho_A \otimes \varrho_B\right)\right]. \tag{H1}$$

It is then easy to confirm that

$$\mathbb{E}N(A:B) = \operatorname{tr}\left\{\left[\left(P_e^{(d)}\right)^{\otimes c} \otimes \left(P_{(12)}^{(d)}\right)^{\otimes a} \otimes \left(P_e^{(d)}\right)^{\otimes r} \otimes \left(P_{(12)}^{(d)}\right)^{\otimes b} \otimes \left(P_e^{(d)}\right)^{\otimes(f-1)} \otimes P_e^{(dD)}\right]\mathbb{E}|\psi\rangle\langle\psi|^{\otimes 4}\right\}$$

$$+ \operatorname{tr}\left\{\left[\left(P_e^{(d)}\right)^{\otimes c} \otimes \left(P_{(12)}^{(d)}\right)^{\otimes a} \otimes \left(P_e^{(d)}\right)^{\otimes r} \otimes \left(P_{(34)}^{(d)}\right)^{\otimes b} \otimes \left(P_e^{(d)}\right)^{\otimes(f-1)} \otimes P_e^{(dD)}\right]\mathbb{E}|\psi\rangle\langle\psi|^{\otimes 4}\right\}$$

$$- 2\operatorname{tr}\left\{\left[\left(P_e^{(d)}\right)^{\otimes c} \otimes \left(P_{(12)}^{(d)}\right)^{\otimes a} \otimes \left(P_e^{(d)}\right)^{\otimes r} \otimes \left(P_{(13)}^{(d)}\right)^{\otimes b} \otimes \left(P_e^{(d)}\right)^{\otimes(f-1)} \otimes P_e^{(dD)}\right]\mathbb{E}|\psi\rangle\langle\psi|^{\otimes 4}\right\}. \tag{H2}$$

Step 2: We express $\mathbb{E}N(A:B)$ in terms of the transfer matrices defined in Sec. IV A. Given the previous step, it is easy to confirm that

$$\mathbb{E}N(A:B) = \langle I_4|T_e^c T_{(12)}^a T_e^r T_{(12)}^b|F_4\rangle + \langle I_4|T_e^c T_{(34)}^a T_e^r T_{(12)}^b|F_4\rangle - 2\langle I_4|T_e^c T_{(12)}^a T_e^r T_{(13)}^b|F_4\rangle \tag{H3}$$

$$= \langle I_4|T_e^c T_{(12)}^a T_e^r \left(T_{(12)}^b + T_{(34)}^b - 2T_{(13)}^b\right)|F_4\rangle \tag{H4}$$

$$\equiv \langle I_4|T_e^c \mathcal{A} T_e^r \mathcal{B}|F_4\rangle, \tag{H5}$$

where we have defined

$$\mathcal{A} = T_{(12)}^a \qquad \text{and} \qquad \mathcal{B} = T_{(12)}^b + T_{(34)}^b - 2T_{(13)}^b. \tag{H6}$$

Step 3: We expand $\mathbb{E}N(A:B)$ in terms of the spectrum of $T_e$ with $e \in S_4$ [see Eq. (62)]. Let $\lambda_1 > \lambda_2 > \cdots \geq 0$ denote the distinct eigenvalues of $T_e$. It holds [63] that

$$\lambda_1 = 1 \qquad \text{and} \qquad \lambda_2 = \frac{dD^2 - d}{d^2 D^2 - 1}. \tag{H7}$$

The former is nondegenerate, while the degeneracy of the latter is given by the number of transposition in $S_4$ [63],

$$w_2 = \binom{4}{2} = 6. \tag{H8}$$

Expanding $T_e^c$ and taking the limit of $c \to \infty$ yields

$$\mathbb{E}N(A:B) = \langle L_1|\mathcal{A} T_e^r \mathcal{B}|F_4\rangle, \tag{H9}$$

where we have used that $\langle I_4|R_1\rangle = 1$. After expanding $T_e^r$ and using that $|F_4\rangle = |R_1\rangle$, we have

$$\mathbb{E}N(A:B) = \langle L_1|\mathcal{A}|R_1\rangle\langle L_1|\mathcal{B}|R_1\rangle + \lambda_2^r \sum_{\mu=1}^{w_2} \langle L_1|\mathcal{A}|R_2^{(\mu)}\rangle\langle L_2^{(\mu)}|\mathcal{B}|R_1\rangle + O\left(\lambda_3^r\right) \tag{H10}$$

$$= \lambda_2^r \sum_{\mu=1}^{w_2} \langle L_1|\mathcal{A}|R_2^{(\mu)}\rangle\langle L_2^{(\mu)}|\mathcal{B}|R_1\rangle + O\left(\lambda_3^r\right), \tag{H11}$$

where, in the second line, we have used that

$$\langle L_1|\mathcal{B}|R_1\rangle = 0, \tag{H12}$$

which we prove in the following.

In particular, we prove that $\langle L_1 | T_t^b | R_1 \rangle$ does not depend on the two elements the transposition $t \in S_4$ acts upon. We need some understanding of $T_t^b$ as well as the eigenvectors $\langle L_1 |$ and $| R_1 \rangle$ of $T_e$. For the former, we make use of the fact that $T_\rho^b$ with $\rho \in S_4$ is similar to to $T_e^b$ with $e \in S_4$,

$$T_\rho^b = \sum_{\pi, \varphi \in S_4} |\pi\rangle \langle \rho\pi | C_4^b | \rho\varphi \rangle \langle \varphi|. \tag{H13}$$

With

$$\alpha = \frac{d^2 D - D}{d^2 D^2 - 1}, \qquad \beta = \frac{dD^2 - d}{d^2 D^2 - 1}, \qquad f(u) = \sum_{i=0}^{u-1} \alpha\beta^i, \qquad \text{and} \qquad g(u) = \beta^u, \tag{H14}$$

we have

$$\left(T_e|s_i\rangle\right)_{1 \le i \le 7} = \begin{pmatrix} 1 & \alpha & \alpha & \alpha & \alpha & \alpha & \alpha \\ 0 & \beta & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \beta & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \beta & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \beta & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \beta & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \beta \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \left(T_e^b|s_i\rangle\right)_{1 \le i \le 7} = \begin{pmatrix} 1 & f(b) & f(b) & f(b) & f(b) & f(b) & f(b) \\ 0 & g(b) & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & g(b) & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & g(b) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & g(b) & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & g(b) & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & g(b) \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{H15}$$

For getting some understanding of $\langle L_1 |$, we make use of the fact that $\langle L_i^{(\mu)} | R_j^{(\nu)} \rangle = \delta_{ij} \delta_{\mu\nu}$. It is easy to confirm that

$$|R_1\rangle = |s_1\rangle \qquad \text{and} \qquad |R_2^{(\mu)}\rangle = \frac{\alpha}{\beta - 1}|s_1\rangle + |s_{\mu+1}\rangle, \tag{H16}$$

which implies that

$$\left(\langle L_1 | s_i \rangle\right)_{1 \le i \le 7} = \begin{pmatrix} 1 & \dfrac{\alpha}{\beta - 1} & \dfrac{\alpha}{\beta - 1} & \dfrac{\alpha}{\beta - 1} & \dfrac{\alpha}{\beta - 1} & \dfrac{\alpha}{\beta - 1} & \dfrac{\alpha}{\beta - 1} \end{pmatrix}. \tag{H17}$$

For any transposition $t \in S_4$, it thus holds that

$$\langle L_1 | T_t^b | R_1 \rangle = \sum_{\pi, \varphi \in S_4} \langle L_1 | \pi \rangle \langle t\pi | C_4^b | t\varphi \rangle \langle \varphi | R_1 \rangle \tag{H18}$$

$$= \sum_{\pi,\varphi \in S_4} \langle L_1|\pi\rangle\langle t\pi|C_4^b|t\varphi\rangle\delta_{s_1\varphi} \tag{H19}$$

$$= \sum_{\pi \in S_4} \langle L_1|\pi\rangle\langle t\pi|C_4^b|t\rangle \tag{H20}$$

$$= \sum_{\pi \in S_4} f(b)\langle L_1|\pi\rangle\langle t\pi|s_1\rangle + \sum_{\pi \in S_4} g(b)\langle L_1|\pi\rangle\langle t\pi|t\rangle \tag{H21}$$

$$= \sum_{\pi \in S_4} f(b)\langle L_1|\pi\rangle\delta_{t\pi} + \sum_{\pi \in S_4} g(b)\langle L_1|\pi\rangle\delta_{s_1\pi} \tag{H22}$$

$$= f(b)\langle L_1|t\rangle + g(b)\langle L_1|s_1\rangle \tag{H23}$$

$$= \frac{\alpha}{\beta-1}f(b) + g(b), \tag{H24}$$

which is independent of the two elements the transposition $t \in S_4$ acts upon. Thus,

$$\langle L_1|\mathcal{B}|R_1\rangle = \langle L_1|\left(T_{(12)}^b + T_{(34)}^b - 2T_{(13)}^b\right)|R_1\rangle = 0. \tag{H25}$$

Step 4:   Finally, we can write $\mathbb{E}N(A:B)$ in the form of Definition 1. That is,

$$\mathbb{E}N(A:B) \equiv K\exp\left(-\frac{r}{\xi}\right) + O\left[\exp\left(-\frac{r}{\chi}\right)\right], \tag{H26}$$

where

$$K = \sum_{\mu=1}^{w_2} \langle L_1|\mathcal{A}|R_2^{(\mu)}\rangle\langle L_2^{(\mu)}|\mathcal{B}|R_1\rangle \tag{H27}$$

and

$$\xi = -\frac{1}{\log(\lambda_2)} = -\left[\log\left(\frac{dD^2-d}{d^2D^2-1}\right)\right]^{-1} = \xi_{1D} > \chi. \tag{H28}$$

This concludes the proof.         ■

## APPENDIX I: PROOF OF RESULT 3 AND COROLLARY 4

*Result 3.*—If Conjecture 1 holds, the average of $I(A:B)$ with respect to the random MPS ensemble and subsystems $A$ and $B$ as sketched in Fig. 2(a) decays exponentially as specified in Definition 1 with the average correlation length $\xi_{1D}$ defined in Eq. (35a).

*Corollary 4.*—If Conjecture 1 holds, for any integer value of $\alpha \geq 1$, the average of $I_\alpha(A:B)$ with respect to the random MPS ensemble and subsystems $A$ and $B$ as sketched in Fig. 2(a) decays exponentially as specified in Definition 1 with the average correlation length $\xi_{1D}$ defined in Eq. (35a).

In this Appendix, we prove Result 3 and Corollary 4. The proof of the latter will follow directly from the proof of the former.

*Proof of Result 3.*—We split the proof into four steps, following the structure of the proof of Result 1. Because $I(A:B)$ and $I_2(A:B)$ are related, the steps are overall very similar. As is usual in the context of the replica trick [78], we will interchange the order of some limits without rigorous justification.

Step 1: We rewrite $\mathbb{E}I(A:B)$ in terms of expressions of the form of Eq. (44). To that end, we make use of Eqs. (94) and (95). With those,

$$\mathbb{E}I(A:B) = \lim_{\alpha\to 1}\lim_{v\to 0}\frac{1}{v\alpha - v}\left\{\log\left[\mathbb{E}\,\text{tr}\left(\varrho_{AB}^\alpha\right)^v\right] - \log\left[\mathbb{E}\,\text{tr}\left(\varrho_A^\alpha\right)^v\right] - \log\left[\mathbb{E}\,\text{tr}\left(\varrho_B^\alpha\right)^v\right]\right\}. \tag{I1}$$

$\mathbb{E}\,\text{tr}\left(\varrho_A^\alpha\right)^v$, $\mathbb{E}\,\text{tr}\left(\varrho_A^\alpha\right)^v$, and $\mathbb{E}\,\text{tr}\left(\varrho_{AB}^\alpha\right)^v$ can be written in the desired form. Let us define $x \in S_\alpha$ to be the cyclic permutation so that $x(i) = i + 1$ modulo $\alpha$ and

$$x_w = \left(\alpha(w-1)+1, \alpha(w-1)+2, \ldots, \alpha w\right) \in S_{v\alpha} \tag{I2}$$

with $w \in \{1, \ldots, v\}$. Then, for example,

$$\mathbb{E}\,\text{tr}\left(\varrho_{AB}^\alpha\right)^v = \text{tr}\left\{\left[\left(P_e^{(d)}\right)^{\otimes c} \otimes \left(P_x^{(d)}\right)^{\otimes a} \otimes \left(P_e^{(d)}\right)^{\otimes r} \otimes \left(P_x^{(d)}\right)^{\otimes b} \otimes \left(P_e^{(d)}\right)^{\otimes(f-1)} \otimes P_e^{(dD)}\right]\mathbb{E}|\psi\rangle\langle\psi|^{\otimes\alpha}\right\}^v \tag{I3}$$

$$= \text{tr}\left\{\left[\left(P_e^{(d)}\right)^{\otimes c} \otimes \left(P_{x_1\cdots x_v}^{(d)}\right)^{\otimes a} \otimes \left(P_e^{(d)}\right)^{\otimes r} \otimes \left(P_{x_1\cdots x_v}^{(d)}\right)^{\otimes b} \otimes \left(P_e^{(d)}\right)^{\otimes(f-1)} \otimes P_e^{(dD)}\right]\mathbb{E}|\psi\rangle\langle\psi|^{\otimes v\alpha}\right\}. \tag{I4}$$

Step 2: We express $\mathbb{E}I(A:B)$ in terms of the transfer matrices defined in Sec. IV A. Given the previous step, it is easy to confirm that

$$\mathbb{E}I(A:B) = \lim_{\alpha\to 1}\lim_{v\to 0}\frac{1}{v\alpha - v}\left[\log\left(\langle I_{v\alpha}|T_e^c T_{x_1\cdots x_v}^a T_e^r T_{x_1\cdots x_v}^b|F_{v\alpha}\rangle\right) - \log\left(\langle I_{v\alpha}|T_e^c T_{x_1\cdots x_v}^a|F_{v\alpha}\rangle\right)\right.$$
$$\left. - \log\left(\langle I_{v\alpha}|T_e^{c+a+r} T_{x_1\cdots x_v}^b|F_{v\alpha}\rangle\right)\right] \tag{I5}$$

$$\equiv \lim_{\alpha\to 1}\lim_{v\to 0}\frac{1}{v\alpha - v}\left[\log\left(\langle I_{v\alpha}|T_e^c\mathcal{A}T_e^r\mathcal{B}|F_{v\alpha}\rangle\right) - \log\left(\langle I_{v\alpha}|T_e^c\mathcal{A}|F_{v\alpha}\rangle\right) - \log\left(\langle I_{v\alpha}|T_e^{c+a+r}\mathcal{B}|F_{v\alpha}\rangle\right)\right], \tag{I6}$$

where we have defined

$$\mathcal{A} = T_{x_1\cdots x_v}^a \qquad \text{and} \qquad \mathcal{B} = T_{x_1\cdots x_v}^b. \tag{I7}$$

Step 3: We expand $\mathbb{E}I(A:B)$ in terms of the spectrum of $T_e$ with $e \in S_{v\alpha}$ [see Eq. (62)]. At this point, we assume Conjecture 1 to hold. That is, for any $v\alpha \geq 2$, we assume that

$$\lambda_2 = \frac{dD^2 - d}{d^2D^2 - 1} \tag{I8}$$

with degeneracy

$$w_2 = \binom{k}{2}. \tag{I9}$$

Expanding $T_e^c$ and taking the limit of $c \to \infty$ yields

$$\mathbb{E}I(A:B) = \lim_{\alpha\to 1}\lim_{v\to 0}\frac{1}{v\alpha - v}\left[\log\left(\langle L_1|\mathcal{A}T_e^r\mathcal{B}|F_{v\alpha}\rangle\right) - \log\left(\langle L_1|\mathcal{A}|F_{v\alpha}\rangle\right) - \log\left(\langle L_1|\mathcal{B}|F_{v\alpha}\rangle\right)\right], \tag{I10}$$

where we have used that $\langle I_{v\alpha}|R_1\rangle = 1$. After expanding $T_e^r$ and using that $|F_{v\alpha}\rangle = |R_1\rangle$, we have

$$\mathbb{E}I(A:B) = \lim_{\alpha\to 1}\lim_{v\to 0}\frac{1}{v\alpha - v}\left\{\log\left[\langle L_1|\mathcal{A}|R_1\rangle\langle L_1|\mathcal{B}|R_1\rangle + \lambda_2^r\sum_{\mu=1}^{w_2}\langle L_1|\mathcal{A}|R_2^{(\mu)}\rangle\langle L_2^{(\mu)}|\mathcal{B}|R_1\rangle + O\left(\lambda_3^r\right)\right]\right.$$
$$\left. - \log\left(\langle L_1|\mathcal{A}|R_1\rangle\right) - \log\left(\langle L_1|\mathcal{B}|R_1\rangle\right)\right\}. \tag{I11}$$

Step 4:   Finally, we can write $\mathbb{E}I(A:B)$ in the form of Definition 1. With $\Lambda = \max\left(\{\lambda_2^{2r}, \lambda_3^r\}\right)$,

$$
\mathbb{E}I(A:B) = \lim_{\alpha \to 1} \lim_{v \to 0} \frac{1}{v\alpha - v} \log\left[ 1 + \lambda_2^r \sum_{\mu=1}^{w_2} \frac{\langle L_1|\mathcal{A}|R_2^{(\mu)}\rangle\langle L_2^{(\mu)}|\mathcal{B}|R_1\rangle}{\langle L_1|\mathcal{A}|R_1\rangle\langle L_1|\mathcal{B}|R_1\rangle} + O\left(\lambda_3^r\right) \right] \tag{I12}
$$

$$
= \lim_{\alpha \to 1} \lim_{v \to 0} \frac{1}{v\alpha - v} \left[ \lambda_2^r \sum_{\mu=1}^{w_2} \frac{\langle L_1|\mathcal{A}|R_2^{(\mu)}\rangle\langle L_2^{(\mu)}|\mathcal{B}|R_1\rangle}{\langle L_1|\mathcal{A}|R_1\rangle\langle L_1|\mathcal{B}|R_1\rangle} + O(\Lambda) \right] \tag{I13}
$$

$$
\equiv \lim_{\alpha \to 1} \lim_{v \to 0} \frac{1}{v\alpha - v} \left\{ \widetilde{K}(v\alpha) \exp\left(-\frac{r}{\xi}\right) + O\left[ \exp\left(-\frac{r}{\chi}\right) \right] \right\}, \tag{I14}
$$

where

$$
\widetilde{K}(v\alpha) = \sum_{\mu=1}^{w_2} \frac{\langle L_1|\mathcal{A}|R_2^{(\mu)}\rangle\langle L_2^{(\mu)}|\mathcal{B}|R_1\rangle}{\langle L_1|\mathcal{A}|R_1\rangle\langle L_1|\mathcal{B}|R_1\rangle} \tag{I15}
$$

and

$$
\xi = -\frac{1}{\log(\lambda_2)} = -\left[\log\left(\frac{dD^2 - d}{d^2 D^2 - 1}\right)\right]^{-1} = \xi_{1D} > \chi. \tag{I16}
$$

As $\xi$ is independent of $v\alpha$, it cannot be affected by the replica limits. $\widetilde{K}(v\alpha)$ will converge to some $K$ that is guaranteed to be independent of $r$.

This concludes the proof.      ■

*Proof of Corollary 4.*—Using Eq. (95), it holds that

$$
\mathbb{E}I_\alpha(A:B) = \lim_{v \to 0} \frac{1}{v\alpha - v} \left\{ \log\left[\mathbb{E}\operatorname{tr}\left(\varrho_{AB}^\alpha\right)^v\right] - \log\left[\mathbb{E}\operatorname{tr}\left(\varrho_A^\alpha\right)^v\right] - \log\left[\mathbb{E}\operatorname{tr}\left(\varrho_B^\alpha\right)^v\right] \right\}. \tag{I17}
$$

Thus, the proof is identical to that of Result 3 without the limit of $\alpha \to 1$. The statement follows.      ■

## APPENDIX J: RESULT 3 WITH $c = 0$

In this Appendix, we prove a version of Result 3 with $c = 0$. Steps 1 and 2 of this proof are identical to Steps 1 and 2 of the proof of Result 3 with $c = 0$. That is, at the end of Step 2, we have

$$
\mathbb{E}I(A:B) = \lim_{\alpha \to 1} \lim_{v \to 0} \frac{1}{v\alpha - v} \left[ \log\left(\langle I_{v\alpha}|\mathcal{A}C_{v\alpha}^r\mathcal{B}|F_{v\alpha}\rangle\right) - \log\left(\langle I_{v\alpha}|\mathcal{A}|F_{v\alpha}\rangle\right) - \log\left(\langle I_{v\alpha}|C_{v\alpha}^{a+r}\mathcal{B}|F_{v\alpha}\rangle\right) \right]. \tag{J1}
$$

We start the proof at Step 3.
Step 3:   We expand $\mathbb{E}I(A:B)$ in terms of the spectrum of $T_e$ with $e \in S_{v\alpha}$ [see Eq. (62)]. We assume Conjecture 1 to hold. Expanding $T_e^r$ and using that $\langle I_{v\alpha}|R_1\rangle = 1$ yields

$$
\begin{aligned}
\mathbb{E}I(A:B) = \lim_{\alpha \to 1} \lim_{v \to 0} \frac{1}{v\alpha - v} \Bigg\{ &\log\left[ \langle I_{v\alpha}|\mathcal{A}|R_1\rangle\langle L_1|\mathcal{B}|F_{v\alpha}\rangle + \lambda_2^r \sum_{\mu=1}^{w_2} \langle I_{v\alpha}|\mathcal{A}|R_2^{(\mu)}\rangle\langle L_2^{(\mu)}|\mathcal{B}|F_{v\alpha}\rangle + O\left(\lambda_3^r\right) \right] \\
&- \log\left(\langle I_{v\alpha}|\mathcal{A}|F_{v\alpha}\rangle\right) \\
&- \log\left[ \langle L_1|\mathcal{B}|F_{v\alpha}\rangle + \lambda_2^{a+r} \sum_{\mu=1}^{w_2} \langle I_{v\alpha}|R_2^{(\mu)}\rangle\langle L_2^{(\mu)}|\mathcal{B}|F_{v\alpha}\rangle + O\left(\lambda_3^r\right) \right] \Bigg\}. 
\end{aligned} \tag{J2}
$$

Step 4:  We write $\mathbb{E}I(A:B)$ in the form of Definition 1. With $\Lambda = \max\left(\{\lambda_2^{2r}, \lambda_3^r\}\right)$,

$$\mathbb{E}I(A:B) = \lim_{\alpha\to 1}\lim_{v\to 0}\frac{1}{v\alpha - v}\left\{\log\left[1 + \lambda_2^r\sum_{\mu=1}^{w_2}\frac{\langle I_{v\alpha}|\mathcal{A}|R_2^{(\mu)}\rangle\langle L_2^{(\mu)}|\mathcal{B}|F_{v\alpha}\rangle}{\langle I_{v\alpha}|\mathcal{A}|R_1\rangle\langle L_1|\mathcal{B}|F_{v\alpha}\rangle} + O\left(\lambda_3^r\right)\right]\right.$$
$$\left. - \log\left[1 + \lambda_2^{a+r}\sum_{\mu=1}^{w_2}\langle I_{v\alpha}|R_2^{(\mu)}\rangle\frac{\langle L_2^{(\mu)}|\mathcal{B}|F_{v\alpha}\rangle}{\langle L_1|\mathcal{B}|F_{v\alpha}\rangle} + O\left(\lambda_3^r\right)\right]\right\} \tag{J3}$$

$$= \lim_{\alpha\to 1}\lim_{v\to 0}\frac{1}{v\alpha - v}\left[\lambda_2^r\sum_{\mu=1}^{w_2}\frac{\langle I_{v\alpha}|\mathcal{A}|R_2^{(\mu)}\rangle\langle L_2^{(\mu)}|\mathcal{B}|F_{v\alpha}\rangle}{\langle I_{v\alpha}|\mathcal{A}|R_1\rangle\langle L_1|\mathcal{B}|F_{v\alpha}\rangle} + \lambda_2^{a+r}\sum_{\mu=1}^{w_2}\langle I_{v\alpha}|R_2^{(\mu)}\rangle\frac{\langle L_2^{(\mu)}|\mathcal{B}|F_{v\alpha}\rangle}{\langle L_1|\mathcal{B}|F_{v\alpha}\rangle} + O(\Lambda)\right] \tag{J4}$$

$$= \lim_{\alpha\to 1}\lim_{v\to 0}\frac{1}{v\alpha - v}\left\{\lambda_2^r\sum_{\mu=1}^{w_2}\left[\frac{\langle I_{v\alpha}|\mathcal{A}|R_2^{(\mu)}\rangle\langle L_2^{(\mu)}|\mathcal{B}|F_{v\alpha}\rangle}{\langle I_{v\alpha}|\mathcal{A}|R_1\rangle\langle L_1|\mathcal{B}|F_{v\alpha}\rangle} + \lambda_2^a\langle I_{v\alpha}|R_2^{(\mu)}\rangle\frac{\langle L_2^{(\mu)}|\mathcal{B}|F_{v\alpha}\rangle}{\langle L_1|\mathcal{B}|F_{v\alpha}\rangle}\right] + O(\Lambda)\right\} \tag{J5}$$

$$\equiv \lim_{\alpha\to 1}\lim_{v\to 0}\frac{1}{v\alpha - v}\left\{\widetilde{K}'(v\alpha)\exp\left(-\frac{r}{\xi}\right) + O\left[\exp\left(-\frac{r}{\chi}\right)\right]\right\}, \tag{J6}$$

where

$$\widetilde{K}'(v\alpha) = \sum_{\mu=1}^{w_2}\left[\frac{\langle I_{v\alpha}|\mathcal{A}|R_2^{(\mu)}\rangle\langle L_2^{(\mu)}|\mathcal{B}|F_{v\alpha}\rangle}{\langle I_{v\alpha}|\mathcal{A}|R_1\rangle\langle L_1|\mathcal{B}|F_{v\alpha}\rangle} + \lambda_2^a\langle I_{v\alpha}|R_2^{(\mu)}\rangle\frac{\langle L_2^{(\mu)}|\mathcal{B}|F_{v\alpha}\rangle}{\langle L_1|\mathcal{B}|F_{v\alpha}\rangle}\right] \tag{J7}$$

and

$$\xi = -\frac{1}{\log(\lambda_2)} = -\left[\log\left(\frac{dD^2 - d}{d^2D^2 - 1}\right)\right]^{-1}\xi_{1D} > \chi. \tag{J8}$$

Again, $\widetilde{K}'(v\alpha)$ will converge to some $K'$ that is guaranteed to be independent of $r$. While $K'$ is different from $K$ in general, the correlation length $\xi$ is independent of $c$.

## APPENDIX K: NUMERICAL ANALYSIS

In this Appendix, we briefly review our numerical analysis of the von Neumann mutual information $I(A:B)$ in Sec. IV E. We fix $d$ and $D$, and we set $a = b = 1$ and $r = 5$. (i) We generate $a + r + b + 1$ Haar-random unitary matrices of $U(dD)$ to define $|\psi\rangle$. This definition makes the assumption that there are no sites before subsystem $A$ (that is, $c = 0$). We discuss in Appendix J why this does not affect the average correlation length. By setting $f = 1$, we furthermore use the fact that the sites after subsystem $B$ do not play a role as a result of the sequential generation [see Eq. (59)]. (ii) We compute $I(A:B)$ with respect to $|\psi\rangle$. (iii) We repeat steps (i) and (ii) 10 000 times to compute the average of $I(A:B)$. (iv) We repeat steps (i) through (iii) for $r \in \{7, 9, 11, 13, 15\}$, plot the averages of $I(A:B)$ against $r$, and fit the data to extract the average correlation length. (v) To obtain Fig. 3, we repeat steps (i) through (iv) for different $d$ and $D$.

## APPENDIX L: TRANSFER MATRICES IN TWO DIMENSIONS

This Appendix expands on Sec. V A. We will state the definitions of the boundary tensors and prove Eq. (105).

From the main text, recall that we define $V^{(i,j)} = U^{(i,j)} \otimes \overline{U^{(i,j)}}$. By computing the $k$-fold twirl, we obtain the building block



$$\tag{L1}$$

With that, we have



$$\mathbb{E}|\psi\rangle\langle\psi|^{\otimes k} = \qquad = \qquad , \tag{L2}$$

where we have cut permutation-valued (green) legs instead of bond (blues) ones in the second step. The tensor $S$ is stated in Eq. (102). $S'$ and $S''$ reflect the different boundary conditions.

After contracting $S'$ with $P_\rho^{(d)}$, we obtain



$$\tau - T_\rho - \theta = \tau - S'' - \theta = \tau - W - P - P^T - \theta \tag{L3}$$

$$= \sum_{\sigma \in S_k} Wg(\sigma\tau^{-1}, dD^2) d^{\#(\sigma,p)} D^{\#(\sigma\theta^{-1})} \tag{L4}$$

at the top boundary and



$$\tau - T_\rho = \tau - S'' = \tau - W - P \tag{L5}$$

$$= \sum_{\sigma \in S_k} Wg(\sigma\tau^{-1}, dD^2) d^{\#(\sigma,p)} D^{\#(\sigma v^{-1})} \tag{L6}$$

at the right boundary.

We will always contract $S''$ with $P_e^{(d)}$. The tensor in the top-right corner thus plays the same role as the final vector $|F_k\rangle = e_1 \in \mathbb{R}^{k!}$ does in one dimension. In fact, $T_e = |F_k\rangle\langle F_k|$,



$$\tau - T_e = \tau - S'' = \tau - W - P \equiv \tau - F \tag{L7}$$

$$= \sum_{\sigma \in S_k} Wg\left(\sigma\tau^{-1}, dD^2\right) \left(dD^2\right)^{\#(\sigma e)} = \delta_{e\tau}, \tag{L8}$$

where we have defined



$$\tag{L9}$$

If a tensor corresponding to $e \in S_k$ is contracted with $|F_k\rangle$, it factorizes. For tensors at the top boundary, we have



$$\tag{L10}$$

$$= \sum_{\sigma \in S_k} \mathrm{Wg}\left(\sigma\tau^{-1}, dD^2\right)\left(dD^2\right)^{\#(\sigma e)} = \delta_{e\tau}, \tag{L11}$$

for those at the right boundary, we have



$$\tag{L12}$$

$$= \sum_{\sigma \in S_k} \mathrm{Wg}\left(\sigma\tau^{-1}, dD^2\right)\left(dD^2\right)^{\#(\sigma e)} = \delta_{e\tau}, \tag{L13}$$

and for those in the bulk, we have



$$\tag{L14}$$

$$= \sum_{\sigma \in S_k} \mathrm{Wg}\left(\sigma\tau^{-1}, dD^2\right)\left(dD^2\right)^{\#(\sigma e)} = \delta_{e\tau}. \tag{L15}$$

The identities above lead to Eq. (105):



$$(\text{L16})$$



$$(\text{L17})$$

## APPENDIX M: SPECTRUM OF THE TRANSFER MATRIX $T_e$ WITH $e \in S_2$

In this Appendix, we state and prove two lemmas concerning the spectrum of $\mathcal{T}_e$ with $e \in S_2$.
Let us start with some preliminaries. We define

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad \text{and} \qquad |+\rangle = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \tag{M1}$$

and map the contraction of tensors defining $\mathcal{T}_e$ with $e \in S_2$ to a multiplication of matrices:



$$(\text{M2})$$

where



$$(\text{M3})$$

with

$$\alpha = \frac{d^2 D^3 - D}{d^2 D^4 - 1}, \qquad \beta = \frac{dD^3 - dD}{d^2 D^4 - 1}, \qquad \gamma = \frac{d^2 D^2 - D^2}{d^2 D^4 - 1}, \qquad \text{and} \qquad \delta = \frac{dD^2 - d}{d^2 D^4 - 1}. \tag{M4}$$

Note that

$$\langle i| \boxed{M} = 0 \quad \text{if} \quad i \neq j,$$
$$\langle j|$$

(M5)

and

$$N = \boxed{M} \begin{array}{c} |0\rangle \\ \langle +| \end{array} = \begin{pmatrix} 1 & \alpha \\ 0 & \beta \end{pmatrix}$$

(M6)

is equal to $T_e$ with $e \in S_2$ for $d \to dD$. With that, it easy to check that

$$\langle i| \boxed{M} \begin{array}{c} |0\rangle \\ |o\rangle \end{array} = \frac{|i\rangle\langle i|}{\langle +|} \boxed{M} \begin{array}{c} |0\rangle \\ |o\rangle \end{array} = \langle o|N|i\rangle\langle i|.$$

(M7)

We also introduce an analytical notation. We define

$$M_j = I^{\otimes(j-1)} \otimes M \otimes I^{\otimes(h-j)}.$$

(M8)

$\mathcal{T}_e$ with $e \in S_2$ is then given by

$$\mathcal{T}_e = \left( I^{\otimes h} \otimes \langle 0| \right) \left( M_h \cdots M_1 \right) \left( |+\rangle \otimes I^{\otimes h} \right).$$

(M9)

With $i_1, \ldots, i_h \in \{0, 1\}$ and $o_1, \ldots, o_h \in \{0, 1\}$, the entries of $\mathcal{T}_e$ with $e \in S_2$ are given by

$$\left( \langle o_1, \ldots, o_h| \otimes \langle 0| \right) \left( M_h \cdots M_1 \right) \left( |+\rangle \otimes |i_1, \ldots, i_h\rangle \right).$$

(M10)

Our first lemma states that $\mathcal{T}_e$ with $e \in S_2$ is block triangular, where our definition of blocks arises from the indexing of rows and columns in base 2. In particular, with $2 \leq j \leq h$, the $j$th diagonal block of $\mathcal{T}_e$, which we denote by $\mathcal{T}_e^{(j)} \in \mathbb{R}^{2^{j-1} \times 2^{j-1}}$, has fixed indices $i_h = \cdots i_{j+1} = o_h = \cdots = o_{j+1} = 0$ and $i_j = o_j = 1$. Using Eq. (M14), it is given by



(M11)

The first diagonal block, which we denote by $\mathcal{T}_e^{(1)} \in \mathbb{R}^{2\times2}$, is given by



(M12)

In particular, we will prove that a block of $\mathcal{T}_e$ is zero if its defining row digit $o_j$ is higher than its defining column digit $i_j$, which implies that $\mathcal{T}_e$ is upper block triangular. As the proof relies exclusively on Eq. (M7), $\mathcal{T}_e$ inherits its upper block triangularity from the upper triangularity of $N$.

*Lemma 6.*—$\mathcal{T}_e$ with $e \in S_2$ is upper block triangular because

$$\left(I^{\otimes(j-1)} \otimes \langle 0| \otimes \langle 0|^{\otimes(h-j)} \otimes \langle 0|\right) \left(M_h \cdots M_1\right) \left(|+\rangle \otimes I^{\otimes(j-1)} \otimes |1\rangle \otimes |0\rangle^{\otimes(h-j)}\right) = 0. \tag{M13}$$

*Proof.*—From
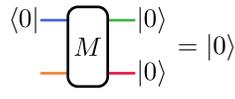


$$\tag{M14}$$

and
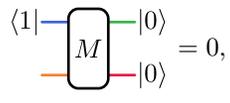


$$\tag{M15}$$

it follows that
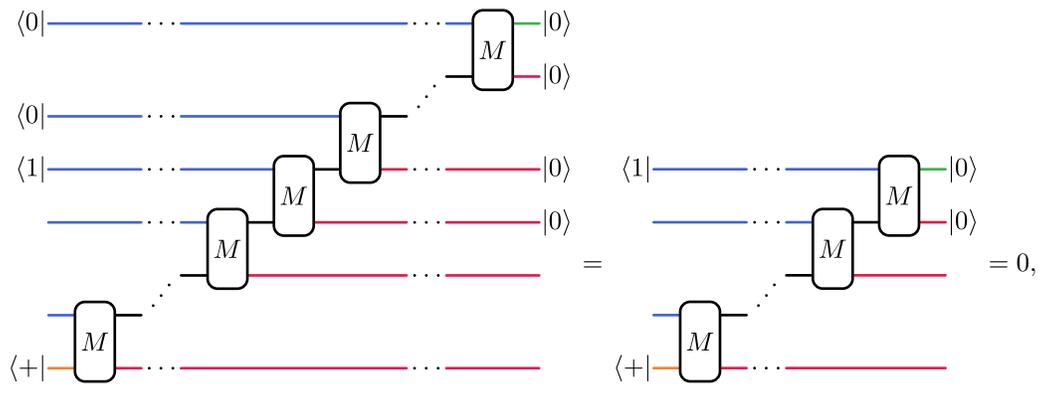


$$\tag{M16}$$

which concludes the proof. ∎

In our second lemma, we utilize the block triangularity of $\mathcal{T}_e$ with $e \in S_2$ to make a direct statement about its two leading eigenvalues.

*Lemma 7.*—Let $|\lambda_1| > |\lambda_2| > \cdots \geq 0$ denote the distinct eigenvalues of $\mathcal{T}_e$ with $e \in S_2$. Then, for any $h$, $\lambda_1 = 1$ and $\lambda_2 = \beta$. Furthermore, $\lambda_1$ and $\lambda_2$ are nondegenerate.

*Proof.*—The spectrum of $\mathcal{T}_e$ with $e \in S_2$ is given by the union of the spectra of its diagonal blocks. It is evident that the first block $\mathcal{T}_e^{(1)}$ has eigenvalues 1 and $\beta$. In the following, we show that any other diagonal block $\mathcal{T}_e^{(j)}$ with $2 \leq j \leq h$ can be written as a product of $\beta$ and a strictly substochastic matrix, implying that its eigenvalues are strictly smaller than $\beta$. We structure the proof in steps.

Step 1: We show that any diagonal block $\mathcal{T}_e^{(j)}$ with $2 \leq j \leq h$ can be written as $\beta$ times a matrix. From



$$\tag{M17}$$

it follows that

$$\mathcal{T}_e^{(j)} = \quad\quad = \beta \quad\quad .$$

(M18)

Step 2: We argue that the matrix

(M19)

is strictly column substochastic. It holds that

$$\boxed{M} = \begin{pmatrix} 1 & \alpha & \alpha & \gamma \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & \beta & \beta & \delta \end{pmatrix}$$

(M20)

is column substochastic. While the first column of $M$ evidently sums to 1, the sums of the other columns are strictly bounded by 1. As a result, $M_{j-1} \cdots M_1$ is column substochastic. The boundary condition $|1\rangle$ does not affect this because it specifies a subset of columns of the matrix $M_{j-1} \cdots M_1$. In fact, it imposes strict substochasticity because this subset does not include the only column summing to 1. Also the boundary condition $\langle +|$ does not affect the substochasticity. The boundary condition means that Eq. (M19) is a sum of matrices. Each of these matrices comprises a disjoint subset of rows of the matrix $M_{j-1} \cdots M_1$. Because $M_{j-1} \cdots M_1$, the matrix comprising the whole set of rows, is column substochastic, so is the sum of the matrices comprising the disjoint subsets.

Step 3: 1 and $\beta$ are the only eigenvalues of $\mathcal{T}_e^{(1)}$. They are nondegenerate. Because the matrix in Eq. (M20) is strictly column substochastic, the eigenvalues of any diagonal block $\mathcal{T}_e^{(j)}$ with $2 \leq j \leq h$ are strictly smaller than $\beta$.

The statement follows. ∎

As a preparation for Appendix N, we provide the proofs of Lemmas 6 and 7 in analytical notation.

*Proof of Lemma 6 in analytical notation.*—From

$$\big(\langle 0| \otimes \langle 0|\big) M \big(I \otimes |0\rangle\big) = |0\rangle$$

(M21)

and

$$\big(\langle 0| \otimes \langle 0|\big) M \big(I \otimes |1\rangle\big) = 0,$$

(M22)

it follows that

$$\big(I^{\otimes(j-1)} \otimes \langle 0| \otimes \langle 0|^{\otimes(h-j)} \otimes \langle 0|\big) \big(M_h \cdots M_1\big) \big(|+\rangle \otimes I^{\otimes(j-1)} \otimes |1\rangle \otimes |0\rangle^{\otimes(h-j)}\big)$$

$$= \big(I^{\otimes(j-1)} \otimes \langle 0| \otimes \langle 0|\big) \big(M_j \cdots M_1\big) \big(|+\rangle \otimes I^{\otimes(j-1)} \otimes |1\rangle\big)$$

(M23)

$$= 0,$$

(M24)

which concludes the proof. ∎

*Proof of Lemma 7 in analytical notation.*—As in the version with graphical notation, we structure the proof in steps, without repeating the details.

Step 1: From

$$\big(\langle 1|\otimes\langle 0|\big)M\big(I\otimes|0\rangle\big)=\beta|1\rangle, \tag{M25}$$

it follows that

$$\mathcal{T}_e^{(j)}=\big(I^{\otimes(j-1)}\otimes\langle 1|\otimes\langle 0|\big)\big(M_j\cdots M_1\big)\big(|+\rangle\otimes I^{\otimes(j-1)}\otimes|1\rangle\big) \tag{M26}$$

$$=\beta\big(I^{\otimes(j-1)}\otimes\langle 1|\big)\big(M_{j-1}\cdots M_1\big)\big(|+\rangle\otimes I^{\otimes(j-1)}\big). \tag{M27}$$

Step 2: It holds that

$$\big(I^{\otimes(j-1)}\otimes\langle 1|\big)\big(M_{j-1}\cdots M_1\big)\big(|+\rangle\otimes I^{\otimes(j-1)}\big) \tag{M28}$$

is strictly column substochastic.

Step 3: 1 and $\beta$ are the only eigenvalues of $\mathcal{T}_e^{(1)}$. They are nondegenerate. Because the matrix in Eq. (M28) is strictly column substochastic, the eigenvalues of any diagonal block $\mathcal{T}_e^{(j)}$ with $2\le j\le h$ are strictly smaller than $\beta$.

The statement follows. ■

## APPENDIX N: SPECTRUM OF THE TRANSFER MATRIX $T_e$ WITH $e\in S_4$

In this Appendix, we state and prove two lemmas concerning the spectrum of $\mathcal{T}_e$ with $e\in S_4$. Using the same notation as in Appendix M, we will draw on results from that Appendix.

$\mathcal{T}_e$ with $e\in S_4$ is defined by Eq. (M2), now with $M\in\mathbb{R}^{576\times 576}$. As in Appendix M, it holds that

$$\big(\langle o|\otimes\langle 0|\big)M\big(I\otimes|i\rangle\big)=\langle o|N|i\rangle\langle i|, \tag{N1}$$

where

$$N=\big(\langle +|\otimes I\big)M\big(I\otimes|0\rangle\big)$$

$$=\begin{pmatrix}
1 & \alpha & \alpha & \alpha & \alpha & \alpha & \alpha & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \gamma & \eta & \eta & \eta & \eta & \eta & \eta & \rho & \rho & \rho \\
0 & \beta & 0 & 0 & 0 & 0 & 0 & \delta & \delta & \delta & \delta & 0 & 0 & 0 & 0 & \theta & \theta & \iota & \theta & \iota & \theta & \sigma & \tau & \tau \\
0 & 0 & \beta & 0 & 0 & 0 & 0 & \delta & \delta & 0 & 0 & \delta & \delta & 0 & 0 & \iota & \theta & \theta & \theta & \theta & \iota & \tau & \sigma & \tau \\
0 & 0 & 0 & \beta & 0 & 0 & 0 & 0 & 0 & \delta & \delta & \delta & \delta & 0 & 0 & \theta & \iota & \theta & \iota & \theta & \theta & \tau & \tau & \sigma \\
0 & 0 & 0 & 0 & \beta & 0 & 0 & \delta & \delta & 0 & 0 & 0 & 0 & \delta & \delta & \theta & \iota & \theta & \iota & \theta & \theta & \tau & \tau & \sigma \\
0 & 0 & 0 & 0 & 0 & \beta & 0 & 0 & 0 & \delta & \delta & 0 & 0 & \delta & \delta & \iota & \theta & \theta & \theta & \theta & \iota & \tau & \sigma & \tau \\
0 & 0 & 0 & 0 & 0 & 0 & \beta & 0 & 0 & 0 & 0 & \delta & \delta & \delta & \delta & \theta & \theta & \iota & \theta & \iota & \theta & \sigma & \tau & \tau \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \varepsilon & \zeta & 0 & 0 & 0 & 0 & 0 & 0 & \kappa & \kappa & \lambda & \lambda & \kappa & \lambda & \upsilon & \upsilon & \upsilon \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \zeta & \varepsilon & 0 & 0 & 0 & 0 & 0 & 0 & \lambda & \lambda & \kappa & \kappa & \lambda & \kappa & \upsilon & \upsilon & \upsilon \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \varepsilon & \zeta & 0 & 0 & 0 & 0 & \kappa & \kappa & \kappa & \lambda & \lambda & \lambda & \upsilon & \upsilon & \upsilon \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \zeta & \varepsilon & 0 & 0 & 0 & 0 & \lambda & \lambda & \lambda & \kappa & \kappa & \kappa & \upsilon & \upsilon & \upsilon \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \varepsilon & \zeta & 0 & 0 & \kappa & \lambda & \kappa & \kappa & \lambda & \lambda & \upsilon & \upsilon & \upsilon \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \zeta & \varepsilon & 0 & 0 & \lambda & \kappa & \lambda & \lambda & \kappa & \kappa & \upsilon & \upsilon & \upsilon \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \varepsilon & \zeta & \kappa & \lambda & \lambda & \kappa & \kappa & \lambda & \upsilon & \upsilon & \upsilon \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \zeta & \varepsilon & \lambda & \kappa & \kappa & \lambda & \lambda & \kappa & \upsilon & \upsilon & \upsilon \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mu & \nu & \nu & \nu & \nu & \xi & \tau & \varphi & \tau \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \nu & \mu & \nu & \xi & \nu & \nu & \tau & \tau & \varphi \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \nu & \nu & \mu & \nu & \xi & \nu & \varphi & \tau & \tau \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \nu & \xi & \nu & \mu & \nu & \nu & \tau & \tau & \varphi \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \nu & \nu & \xi & \nu & \mu & \nu & \varphi & \tau & \tau \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \xi & \nu & \nu & \nu & \nu & \mu & \tau & \varphi & \tau \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & o & o & \pi & o & \pi & o & \chi & \psi & \psi \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \pi & o & o & o & o & \pi & \psi & \chi & \psi \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & o & \pi & o & \pi & o & o & \psi & \psi & \chi
\end{pmatrix} \tag{N2}$$

is equal to $T_e$ with $e\in S_4$ for $d\to dD$.

As in Appendix M, our first lemma states that $\mathcal{T}_e$ with $e \in S_4$ is block triangular. The definition of blocks now arises from the indexing of rows and columns in base 24. In particular, any diagonal block (but the first) is defined by $i_h = \cdots = i_{j+1} = o_h = \cdots = o_{j+1} = 0$ and $i_j = o_j \neq 0$. There are four classes of diagonal blocks:

(a) The first diagonal block is in its own class. It is given by

$$\left(I \otimes \langle 0|^{\otimes(h-1)} \otimes \langle 0|\right) \left(M_h \cdots M_1\right) \left(|+\rangle \otimes I \otimes |0\rangle^{\otimes(h-1)}\right) = \left(I \otimes \langle 0|\right) M \left(|+\rangle \otimes I\right) = N. \tag{N3}$$

(b) The second class of diagonal blocks corresponds to transpositions. $i_j$ and $o_j$ correspond to the same transposition. There are six sub-blocks in this class because there are six different transpositions in $S_4$.

(c) The third class of diagonal blocks corresponds to permutations with a single fixed point. $i_j$ and $o_j$ correspond to any of the two permutations with the same single fixed point. There are four sub-blocks in this class because there are four different choices of a single fixed point.

(d) The fourth class of diagonal blocks corresponds to permutations with no fixed point. $i_j$ and $o_j$ correspond to any of the nine permutations with no fixed point.

In particular, we will prove that a block of $\mathcal{T}_e$ is zero if its defining row digit $o_j$ is higher than its defining column digit $i_j$ stand in a certain relation to each other. As the proof relies exclusively on Eq. (N1), $\mathcal{T}_e$ again inherits its upper block triangularity from the upper triangularity of $N$.

*Lemma 8.*—$\mathcal{T}_e$ with $e \in S_4$ is upper block triangular.

*Proof.*—From Eqs. (N1), it follows that

$$\left(I^{\otimes(j-1)} \otimes \langle o_j| \otimes \langle 0|^{\otimes(h-j)} \otimes \langle 0|\right) \left(M_h \cdots M_1\right) \left(|+\rangle \otimes I^{\otimes(j-1)} \otimes |i_j\rangle \otimes |0\rangle^{\otimes(h-j)}\right) = 0 \tag{N4}$$

if

(a) $i_j$ corresponds to the trivial permutation and $o_j$ does not,

(b) $i_j$ corresponds to a transposition and $o_j$ corresponds to a different transposition or a permutation with one or no fixed point,

(c) $i_j$ corresponds to a permutation with a single fixed point and $o_j$ corresponds to a permutation with a different single fixed point or no fixed point,

which concludes the proof. ∎

In our second lemma, we utilize the block triangularity of $\mathcal{T}_e$ with $e \in S_4$ to make direct a statement about its two leading eigenvalues.

*Lemma 9.*—Let $|\lambda_1| > |\lambda_2| > \cdots \geq 0$ denote the distinct eigenvalues of $\mathcal{T}_e$ with $e \in S_4$. Then, for any $h$, $\lambda_1 = 1$ and $\lambda_2 = \beta$. Furthermore, $\lambda_1$ is nondegenerate, and $\lambda_2$ has a degeneracy of six.

*Proof.*—As is the case for $\mathcal{T}_e$ with $e \in S_2$, the spectrum of $\mathcal{T}_e$ with $e \in S_4$ is given by the union of the spectra of its diagonal blocks. The two leading eigenvalues of the first diagonal block $N$ are 1 and $\beta$. In the following, we show that any other diagonal block can be written as a product of $\beta$ and a matrix whose spectral radius is strictly bounded by 1, implying that its eigenvalues are strictly smaller than $\beta$. We again structure the proof in steps.

Step 1: We show that any diagonal block but the first can be written as a product of $\beta$ and a matrix. From Eq. (N1), it follows that,

(a) if $i_j$ and $o_j$ correspond to the same transposition,

$$\left(I^{\otimes(j-1)} \otimes \langle o_j| \otimes \langle 0|^{\otimes(h-j)} \otimes \langle 0|\right) \left(M_h \cdots M_1\right) \left(|+\rangle \otimes I^{\otimes(j-1)} \otimes |i_j\rangle \otimes |0\rangle^{\otimes(h-j)}\right)$$

$$= \left(I^{\otimes(j-1)} \otimes \langle o_j| \otimes \langle 0|\right) \left(M_j \cdots M_1\right) \left(|+\rangle \otimes I^{\otimes(j-1)} \otimes |i_j\rangle\right) \tag{N5}$$

$$= \beta \left(I^{\otimes(j-1)} \otimes \langle o_j|\right) \left(M_{j-1} \cdots M_1\right) \left(|+\rangle \otimes I^{\otimes(j-1)}\right), \tag{N6}$$

(b) if $i_j$ and $o_j$ correspond to any two permutations with the same single fixed point,

$$\left(I^{\otimes(j-1)} \otimes \langle o_j| \otimes \langle 0|^{\otimes(h-j)} \otimes \langle 0|\right) \left(M_h \cdots M_1\right) \left(|+\rangle \otimes I^{\otimes(j-1)} \otimes |i_j\rangle \otimes |0\rangle^{\otimes(h-j)}\right)$$

$$= \left(I^{\otimes(j-1)} \otimes \langle o_j| \otimes \langle 0|\right) \left(M_j \cdots M_1\right) \left(|+\rangle \otimes I^{\otimes(j-1)} \otimes |i_j\rangle\right) \tag{N7}$$

$$= p \left(I^{\otimes(j-1)} \otimes \langle o_j|\right) \left(M_{j-1} \cdots M_1\right) \left(|+\rangle \otimes I^{\otimes(j-1)}\right) \tag{N8}$$

$$< \frac{\beta}{2} \left(I^{\otimes(j-1)} \otimes \langle o_j|\right) \left(M_{j-1} \cdots M_1\right) \left(|+\rangle \otimes I^{\otimes(j-1)}\right), \tag{N9}$$

where $\{\varepsilon, \zeta\} \ni p < \beta/2$ [63] depends on $i_j$ and $o_j$,

(c) if $i_j$ and $o_j$ correspond to any permutation with no fixed point,

$$\left(I^{\otimes(j-1)} \otimes \langle o_j| \otimes \langle 0|^{\otimes(h-j)} \otimes \langle 0|\right) \left(M_h \cdots M_1\right) \left(|+\rangle \otimes I^{\otimes(j-1)} \otimes |i_j\rangle \otimes |0\rangle^{\otimes(h-j)}\right)$$

$$= \left(I^{\otimes(j-1)} \otimes \langle o_j| \otimes \langle 0|\right) \left(M_j \cdots M_1\right) \left(|+\rangle \otimes I^{\otimes(j-1)} \otimes |i_j\rangle\right) \tag{N10}$$

$$= p \left(I^{\otimes(j-1)} \otimes \langle o_j|\right) \left(M_{j-1} \cdots M_1\right) \left(|+\rangle \otimes I^{\otimes(j-1)}\right) \tag{N11}$$

$$< \frac{\beta}{9} \left(I^{\otimes(j-1)} \otimes \langle o_j|\right) \left(M_{j-1} \cdots M_1\right) \left(|+\rangle \otimes I^{\otimes(j-1)}\right), \tag{N12}$$

where $\{\mu, \nu, \xi, o, \pi, \tau, \varphi, \chi, \psi\} \ni p < \beta/9$ [63] depends on $i_j$ and $o_j$.

Step 2: We now argue that the spectral radius of the matrix

$$\left(I^{\otimes(j-1)} \otimes \langle o_j|\right) \left(M_{j-1} \cdots M_1\right) \left(|+\rangle \otimes I^{\otimes(j-1)}\right) \tag{N13}$$

is strictly bounded by 1 for $2 \leq j \leq h$. It holds that the spectral radius of $M$ is bounded by 1. While the first column of $|M|$ sums to 1, the sums of the other columns are strictly bounded by 1 [63]. As a result, the spectral radius of $M_{j-1} \cdots M_1$ is bounded by 1. The boundary condition $|o\rangle$ does not affect this because it specifies a subset of columns of the matrix $M_{j-1} \cdots M_1$. In fact, it imposes a strict bound because this subset does not include the only column summing to 1. Also the boundary condition $\langle +|$ does not affect the bound on the spectral radius. The boundary condition means that Eq. (N13) is a sum of matrices. Each of these matrices comprises a disjoint subset of rows of the matrix $M_{j-1} \cdots M_1$. Because the spectral radius of $M_{j-1} \cdots M_1$, the matrix comprising the whole set of rows, is bounded by 1, so is the sum of the matrices comprising the disjoint subsets.

Step 3: 1 and $\beta$ are the two leading eigenvalues of the first diagonal block $N$. They are nondegenerate. Because the spectral radius of the matrix in Eq. (N13) is strictly bounded by 1, the eigenvalues of any other diagonal block are strictly smaller than $\beta$.

The statement follows. ∎

## APPENDIX O: PROOF OF RESULT 4

*Result 4.*—The average of $I_2(A : B)$ with respect to the random isoTNS ensemble and subsystems $A$ and $B$ as sketched in Fig. 4(a) decays exponentially as specified in Definition 1 with the average correlation length $\xi_{2D}$ defined in Eq. (35b).

*Proof.*—We split the proof into four steps, following the structure of the proof of Result 1. The steps are overall very similar to those of that proof.

Step 1: We rewrite $\mathbb{E}I_2(A : B)$ in terms of expressions of the form of Eq. (99). As in one dimension, we make the assumption that $\mathbb{E}\log(X) = \log(\mathbb{E}X)$. Then,

$$\mathbb{E}I_2(A : B) = \log\left[\mathbb{E}\operatorname{tr}\left(\varrho_{AB}^2\right)\right] - \log\left[\mathbb{E}\operatorname{tr}\left(\varrho_A^2\right)\right] - \log\left[\mathbb{E}tr\left(\varrho_B^2\right)\right]. \tag{O1}$$

$\mathbb{E}\operatorname{tr}\left(\varrho_A^2\right)$, $\mathbb{E}\operatorname{tr}\left(\varrho_B^2\right)$, and $\mathbb{E}\operatorname{tr}\left(\varrho_{AB}^2\right)$ can be written in the desired form [see Eqs. (66) and (67)].

Step 2: We express $\mathbb{E}I_2(A : B)$ in terms of the transfer tensors defined in Sec. V A and use Eq. (106) to map contractions of two-dimensional tensor networks to multiplications of matrices. The latter is enabled by our definition of subsystems $A$ and $B$ [see Fig. 4(a)]. We have done this for $\mathbb{E}\operatorname{tr}\left(\varrho_A^2\right)$ in graphical notation in Sec. V C [see Eqs. (110) and (111)]. It is easy to confirm that

$$\mathbb{E}I_2(A : B) = \log\left(\langle \mathcal{I}_2| \mathcal{T}_e^c \mathcal{T}_{(12)}^a \mathcal{T}_e^r \mathcal{T}_{(12)}^b |\mathcal{F}_2\rangle\right) - \log\left(\langle \mathcal{I}_2| \mathcal{T}_e^c \mathcal{T}_{(12)}^a |\mathcal{F}_2\rangle\right) - \log\left(\langle \mathcal{I}_2| \mathcal{T}_e^{c+a+r} \mathcal{T}_{(12)}^b |\mathcal{F}_2\rangle\right). \tag{O2}$$

Step 3:    We expand $\mathbb{E}I_2(A : B)$ in terms of the spectrum of $\mathcal{T}_e$ with $e \in S_2$, which we consider in Appendix M. Because we know $\lambda_1$ and $\lambda_2$ as well as their algebraic and geometric multiplicities, we do not need $\mathcal{T}_e$ to be diagonalizable. Expanding $\mathcal{T}_e^c$ and taking the limit of $c \to \infty$ yields

$$\mathbb{E}I_2(A : B) = \log\left(\langle L_1|\mathcal{T}_{(12)}^a\mathcal{T}_e^r\mathcal{T}_{(12)}^b|\mathcal{F}_2\rangle\right) - \log\left(\langle L_1|\mathcal{T}_{(12)}^a|\mathcal{F}_2\rangle\right) - \log\left(\langle L_1|\mathcal{T}_{(12)}^b|\mathcal{F}_2\rangle\right), \tag{O3}$$

where we have used that $\langle \mathcal{I}_2|R_1\rangle = 1$. After expanding $\mathcal{T}_e^r$ and using that $|\mathcal{F}_2\rangle = |R_1\rangle$, we have

$$I_2(A : B) = \log\left[\langle L_1|\mathcal{T}_{(12)}^a|R_1\rangle\langle L_1|\mathcal{T}_{(12)}^b|R_1\rangle + \lambda_2^r\langle L_1|\mathcal{T}_{(12)}^a|R_2\rangle\langle L_2|\mathcal{T}_{(12)}^b|R_1\rangle + O\left(r^{v-1}\lambda_3^r\right)\right]$$
$$- \log\left(\langle L_1|\mathcal{T}_{(12)}^a|R_1\rangle\right) - \log\left(\langle L_1|\mathcal{T}_{(12)}^b|R_1\rangle\right), \tag{O4}$$

where $v$ denotes the size of the largest Jordan block with respect to $\lambda_3$.

Step 4:    Finally, we can write $\mathbb{E}I_2(A : B)$ in the form of Definition 1. With $\Lambda = \max\left(\{\lambda_2^{2r}, r^{v-1}\lambda_3^r\}\right)$,

$$\mathbb{E}I_2(A : B) = \log\left[1 + \lambda_2^r\frac{\langle L_1|\mathcal{T}_{(12)}^a|R_2\rangle\langle L_2|\mathcal{T}_{(12)}^b|R_1\rangle}{\langle L_1|\mathcal{T}_{(12)}^a|R_1\rangle\langle L_1|\mathcal{T}_{(12)}^b|R_1\rangle} + O\left(r^{v-1}\lambda_3^r\right)\right] \tag{O5}$$

$$= \lambda_2^r\frac{\langle L_1|\mathcal{T}_{(12)}^a|R_2\rangle\langle L_2|\mathcal{T}_{(12)}^b|R_1\rangle}{\langle L_1|\mathcal{T}_{(12)}^a|R_1\rangle\langle L_1|\mathcal{T}_{(12)}^b|R_1\rangle} + O(\Lambda) \tag{O6}$$

$$\equiv K\exp\left(-\frac{r}{\xi}\right) + O\left[\exp\left(-\frac{r}{\chi}\right)\right], \tag{O7}$$

where

$$K = \frac{\langle L_1|\mathcal{T}_{(12)}^a|R_2\rangle\langle L_2|\mathcal{T}_{(12)}^b|R_1\rangle}{\langle L_1|\mathcal{T}_{(12)}^a|R_1\rangle\langle L_1|\mathcal{T}_{(12)}^b|R_1\rangle} \tag{O8}$$

and

$$\xi = -\frac{1}{\log(\lambda_2)} = -\left[\log\left(\frac{dD^3 - dD}{d^2D^4 - 1}\right)\right]^{-1} = \xi_{2D} > \chi. \tag{O9}$$

This concludes the proof.                                                                                                ∎

## APPENDIX P: PROOF OF RESULT 5

*Result 5.*—The average of $N(A : B)$ with respect to the random isoTNS ensemble and subsystems $A$ and $B$ as sketched in Fig. 4(a) decays exponentially as specified in Definition 1 with the average correlation length $\xi_{2D}$ defined in Eq. (35b).

*Proof.*—We split the proof into four steps, following the structure of the proof of Result 1. The steps are overall very similar to those of the proof of Result 2 (see Appendix H).

Step 1:    We rewrite $\mathbb{E}N(A : B)$ in terms of expressions of the form of Eq. (99). As in one dimension, with the Hilbert-Schmidt inner product,

$$\mathbb{E}N(A : B) = \mathbb{E}\,\mathrm{tr}\left(\varrho_{AB}^2\right) + \mathbb{E}\,\mathrm{tr}\left(\varrho_A^2\right)\mathrm{tr}\left(\varrho_B^2\right) - 2\mathbb{E}\,\mathrm{tr}\left[\varrho_{AB}\left(\varrho_A \otimes \varrho_B\right)\right]. \tag{P1}$$

The right-hand side can be written in the desired form [see Eq. (H2)].

Step 2:    We express $\mathbb{E}N(A : B)$ in terms of the transfer tensors defined in Sec. V A and use Eq. (106) to map contractions of two-dimensional tensor networks to multiplications of matrices. The latter is enabled by our definition of subsystems

$A$ and $B$ [see Fig. 4(a)]. It is easy to confirm that

$$\mathbb{E}N(A:B) = \langle \mathcal{I}_4 | \mathcal{T}_e^c \mathcal{T}_{(12)}^a \mathcal{T}_e^r \mathcal{T}_{(12)}^b | \mathcal{F}_4 \rangle + \langle \mathcal{I}_4 | \mathcal{T}_e^c \mathcal{T}_{(34)}^a \mathcal{T}_e^r \mathcal{T}_{(12)}^b | \mathcal{F}_4 \rangle - 2\langle \mathcal{I}_4 | \mathcal{T}_e^c \mathcal{T}_{(12)}^a \mathcal{T}_e^r \mathcal{T}_{(13)}^b | \mathcal{F}_4 \rangle \tag{P2}$$

$$= \langle \mathcal{I}_4 | \mathcal{T}_e^c \mathcal{T}_{(12)}^a \mathcal{T}_e^r \left( \mathcal{T}_{(12)}^b + \mathcal{T}_{(34)}^b - 2\mathcal{T}_{(13)}^b \right) | \mathcal{F}_4 \rangle \tag{P3}$$

$$\equiv \langle \mathcal{I}_4 | \mathcal{T}_e^c \mathcal{A} \mathcal{T}_e^r \mathcal{B} | \mathcal{F}_4 \rangle, \tag{P4}$$

where we have defined

$$\mathcal{A} = \mathcal{T}_{(12)}^a \qquad \text{and} \qquad \mathcal{B} = \mathcal{T}_{(12)}^b + \mathcal{T}_{(34)}^b - 2\mathcal{T}_{(13)}^b. \tag{P5}$$

**Step 3:** We expand $\mathbb{E}N(A:B)$ in terms of the spectrum of $\mathcal{T}_e$ with $e \in S_4$, which we consider in Appendix N. Because we know $\lambda_1$ and $\lambda_2$ as well as their algebraic and geometric multiplicities, we do not need $\mathcal{T}_e$ to be diagonalizable. Expanding $\mathcal{T}_e^c$ and taking the limit of $c \to \infty$ yields

$$\mathbb{E}N(A:B) = \langle L_1 | \mathcal{A} \mathcal{T}_e^r \mathcal{B} | \mathcal{F}_4 \rangle, \tag{P6}$$

where we have used that $\langle \mathcal{I}_4 | R_1 \rangle = 1$. After expanding $\mathcal{T}_e^r$ and using that $|\mathcal{F}_4\rangle = |R_1\rangle$, we have

$$\mathbb{E}N(A:B) = \langle L_1 | \mathcal{A} | R_1 \rangle \langle L_1 | \mathcal{B} | R_1 \rangle + \lambda_2^r \sum_{\mu=1}^{w_2} \langle L_1 | \mathcal{A} | R_2^{(\mu)} \rangle \langle L_2^{(\mu)} | \mathcal{B} | R_1 \rangle + O\left(\lambda_3^r\right) \tag{P7}$$

$$= \lambda_2^r \sum_{\mu=1}^{w_2} \langle L_1 | \mathcal{A} | R_2^{(\mu)} \rangle \langle L_2^{(\mu)} | \mathcal{B} | R_1 \rangle + O\left(\lambda_3^r\right), \tag{P8}$$

where, in the second line, we have used that

$$\langle L_1 | \mathcal{B} | R_1 \rangle = 0, \tag{P9}$$

which we prove in the following.

As in the proof of Result 2 (see Appendix H), we prove that $\langle L_1 | \mathcal{T}_t^b | R_1 \rangle$ does not depend on the two elements the transposition $t \in S_4$ acts upon. In fact, the proof follows from the proof of Eq. (H12) of that Appendix. We just need two additional considerations. First, the proof of Lemma 8 is not specific to the trivial permutation $e \in S_4$. In particular, $\mathcal{T}_\rho$ exhibits an upper block triangular structure for any $\rho \in S_4$. The first diagonal block is given by $T_\rho$ with $\rho \in S_4$,

$$\left( \langle s_i | \mathcal{T}_\rho | s_j \rangle \right)_{\substack{1 \leq i \leq 24 \\ 1 \leq j \leq 24}} = T_\rho, \tag{P10}$$

which implies that

$$\left( \langle s_i | \mathcal{T}_\rho^b | s_j \rangle \right)_{\substack{1 \leq i \leq 24 \\ 1 \leq j \leq 24}} = T_\rho^b. \tag{P11}$$

Second, the eigenvectors $\langle L_1 |$ and $| R_1 \rangle$ of $\mathcal{T}_e$ with $e \in S_4$ arise from those of $T_e$ with $e \in S_4$. That is,

$$|R\rangle_1 = s_1 \qquad \text{and} \qquad \left( \langle L_1 | s_i \rangle \right)_{1 \leq i \leq 7} = \left( 1 \quad \frac{\alpha}{\beta - 1} \quad \frac{\alpha}{\beta - 1} \quad \frac{\alpha}{\beta - 1} \quad \frac{\alpha}{\beta - 1} \quad \frac{\alpha}{\beta - 1} \quad \frac{\alpha}{\beta - 1} \right). \tag{P12}$$

$\langle L_1 | \mathcal{T}_t^b | R_1 \rangle$ thus does not depend on the two elements the transposition $t \in S_4$ acts upon because $\langle L_1 | T_t^b | R_1 \rangle$ does not. This implies that

$$\langle L_1 | \mathcal{B} | R_1 \rangle = \langle L_1 | \left( \mathcal{T}_{(12)}^b + \mathcal{T}_{(34)}^b - 2\mathcal{T}_{(13)}^b \right) | R_1 \rangle = 0. \tag{P13}$$

Step 4:    Finally, we can write $\mathbb{E}N(A:B)$ in the form of Definition 1. That is,

$$\mathbb{E}N(A:B) \equiv K \exp\left(-\frac{r}{\xi}\right) + O\left[\exp\left(-\frac{r}{\chi}\right)\right], \tag{P14}$$

where

$$K = \sum_{\mu=1}^{w_2} \langle L_1|\mathcal{A}|R_2^{(\mu)}\rangle\langle L_2^{(\mu)}|\mathcal{B}|R_1\rangle \tag{P15}$$

and

$$\xi = -\frac{1}{\log(\lambda_2)} = -\left[\log\left(\frac{dD^3 - dD}{d^2D^4 - 1}\right)\right]^{-1} = \xi_{2D} > \chi. \tag{P16}$$

This concludes the proof.                                                              ∎

[1] D. N. Page, Average Entropy of a Subsystem, Phys. Rev. Lett. **71**, 1291 (1993).

[2] S. K. Foong and S. Kanno, Proof of Page's Conjecture on the Average Entropy of a Subsystem, Phys. Rev. Lett. **72**, 1148 (1994).

[3] J. Sánchez-Ruiz, Simple proof of Page's conjecture on the average entropy of a subsystem, Phys. Rev. E **52**, 5653 (1995).

[4] P. Hayden, D. W. Leung, and A. Winter, Aspects of generic entanglement, Comm. Math. Phys. **265**, 95 (2006).

[5] I. Bengtsson and K. Zyczkowski, *Geometry of Quantum States* (Cambridge University Press, Cambridge, 2006).

[6] D. Poulin, A. Qarry, R. Somma, and F. Verstraete, Quantum Simulation of Time-Dependent Hamiltonians and the Convenient Illusion of Hilbert Space, Phys. Rev. Lett. **106**, 170501 (2011).

[7] J. I. Cirac, D. Pérez-García, N. Schuch, and F. Verstraete, Matrix product states and projected entangled pair states: Concepts, symmetries, theorems, Rev. Mod. Phys. **93**, 045003 (2021).

[8] F. Verstraete and J. I. Cirac, Matrix product states represent ground states faithfully, Phys. Rev. B **73**, 094423 (2006).

[9] M. B. Hastings, An area law for one-dimensional quantum systems, J. Stat. Mech. **2007**, P08024 (2007).

[10] I. Arad, A. Kitaev, Z. Landau, and U. Vazirani, An area law and sub-exponential algorithm for 1D systems, ArXiv:1301.1162.

[11] X. Chen, Z.-C. Gu, and X.-G. Wen, Classification of gapped symmetric phases in one-dimensional spin systems, Phys. Rev. B **83**, 035107 (2011).

[12] N. Schuch, D. Pérez-García, and I. Cirac, Classifying quantum phases using matrix product states and projected entangled pair states, Phys. Rev. B **84**, 165139 (2011).

[13] D. Pérez-García, F. Verstraete, J. I. Cirac, and M. M. Wolf, PEPS as unique ground states of local Hamiltonians, Quant. Inf. Comp. **8**, 650 (2008).

[14] F. Verstraete, M. M. Wolf, D. Pérez-García, and J. I. Cirac, Criticality, the Area Law, and the Computational Power of Projected Entangled Pair States, Phys. Rev. Lett. **96**, 220601 (2006).

[15] M. A. Levin and X.-G. Wen, String-net condensation: A physical mechanism for topological phases, Phys. Rev. B **71**, 045110 (2005).

[16] Z.-C. Gu, M. Levin, B. Swingle, and X.-G. Wen, Tensor-product representations for string-net condensed states, Phys. Rev. B **79**, 085118 (2009).

[17] O. Buerschaper, M. Aguado, and G. Vidal, Explicit tensor network representation for the ground states of string-net models, Phys. Rev. B **79**, 085119 (2009).

[18] S. Garnerone, T. R. de Oliveira, and P. Zanardi, Typicality in random matrix product states, Phys. Rev. A **81**, 032336 (2010).

[19] S. Garnerone, T. R. de Oliveira, S. Haas, and P. Zanardi, Statistical properties of random matrix product states, Phys. Rev. A **82**, 052312 (2010).

[20] J. Haferkamp, C. Bertoni, I. Roth, and J. Eisert, Emergent Statistical Mechanics from Properties of Disordered Random Matrix Product States, PRX Quantum **2**, 040308 (2021).

[21] B. Collins, C. E. González-Guillén, and D. Pérez-García, Matrix product states, random matrix theory and the principle of maximum entropy, Comm. Math. Phys. **320**, 663 (2013).

[22] L. Chen, R. J. Garcia, K. Bu, and A. Jaffe, Magic of random matrix product states, ArXiv:2211.10350.

[23] R. Movassagh, Generic Local Hamiltonians are Gapless, Phys. Rev. Lett. **119**, 220504 (2017).

[24] R. Movassagh and J. Schenker, An ergodic theorem for quantum processes with applications to matrix product states, Comm. Math. Phys. **395**, 1175 (2022).

[25] C. Lancien and D. Pérez-García, Correlation length in random MPS and PEPS, Ann. Henri Poincaré **23**, 141 (2021).

[26] J. Bensa and M. Žnidarič, Purity decay rate in random circuits with different configurations of gates, Phys. Rev. A **107**, 022604 (2023).

[27] P. Svetlichnyy, S. Mittal, and T. A. B. Kennedy, Matrix product states and the decay of quantum conditional mutual information, ArXiv:2211.06794.

[28] R. Movassagh and J. Schenker, Theory of Ergodic Quantum Processes, Phys. Rev. X **11**, 041001 (2021).

[29] M. Lemm, Gaplessness is not generic for translation-invariant spin chains, Phys. Rev. B **100**, 035113 (2019).

[30] I. Jauslin and M. Lemm, Random translation-invariant Hamiltonians and their spectral gaps, Quantum **6**, 790 (2022).

[31] C. Schön, E. Solano, F. Verstraete, J. I. Cirac, and M. M. Wolf, Sequential Generation of Entangled Multiqubit States, Phys. Rev. Lett. **95**, 110503 (2005).

[32] M. P. Zaletel and F. Pollmann, Isometric Tensor Network States in Two Dimensions, Phys. Rev. Lett. **124**, 037201 (2020).

[33] Z.-Y. Wei, D. Malz, and J. I. Cirac, Sequential Generation of Projected Entangled-Pair States, Phys. Rev. Lett. **128**, 010607 (2022).

[34] T. Soejima, K. Siva, N. Bultinck, S. Chatterjee, F. Pollmann, and M. P. Zaletel, Isometric tensor network representation of string-net liquids, Phys. Rev. B **101**, 085117 (2020).

[35] M. M. Wilde, *Quantum Information Theory* (Cambridge University Press, Cambridge, 2013).

[36] U. Schollwöck, The density-matrix renormalization group in the age of matrix product states, Ann. Phys. **326**, 96 (2011).

[37] R. Haghshenas, M. J. O'Rourke, and G. K.-L. Chan, Conversion of projected entangled pair states into a canonical form, Phys. Rev. B **100**, 054404 (2019).

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

[39] C. T. Asplund and A. Bernamonti, Mutual information after a local quench in conformal field theory, Phys. Rev. D **89**, 066015 (2014).

[40] C. A. Agón and T. Faulkner, Quantum corrections to holographic mutual information, J. High Energy Phys. **2016**, 118 (2016).

[41] B. Chen, P.-X. Hao, and W. Song, Rényi mutual information in holographic warped CFTs, J. High Energy Phys. **2019**, 37 (2019).

[42] H. Bernigau, M. J. Kastoryano, and J. Eisert, Mutual information area laws for thermal free fermions, J. Stat. Mech. **2015**, P02008 (2015).

[43] S. Murciano, V. Alba, and P. Calabrese, in *Entanglement in Spin Chains* (Springer International Publishing, 2022), p. 397.

[44] M. McGinley, S. Leontica, S. J. Garratt, J. Jovanovic, and S. H. Simon, Quantifying information scrambling via classical shadow tomography on programmable quantum simulators, Phys. Rev. A **106**, 012441 (2022).

[45] S. Khatri and M. M. Wilde, Principles of Quantum Communication Theory: A Modern Approach, ArXiv:2011.04672.

[46] M. Kormos and Z. Zimborás, Temperature driven quenches in the Ising model: appearance of negative Rényi mutual information, J. Phys. A: Math. Theor. **50**, 264005 (2017).

[47] S. O. Scalet, Á. M. Alhambra, G. Styliaris, and J. I. Cirac, Computable Rényi mutual information: Area laws and correlations, Quantum **5**, 541 (2021).

[48] J. Watrous, *The Theory of Quantum Information* (Cambridge University Press, Cambridge, 2018).

[49] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2012).

[50] D. DiVincenzo, D. Leung, and B. Terhal, Quantum data hiding, IEEE Trans. Inf. Theory **48**, 580 (2002).

[51] B. Collins and P. Śniady, Integration with respect to the Haar measure on unitary, orthogonal and symplectic group, Comm. Math. Phys. **264**, 773 (2006).

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

[53] F. G. S. L. Brandão, W. Chemissany, N. Hunter-Jones, R. Kueng, and J. Preskill, Models of Quantum Complexity Growth, PRX Quantum **2**, 030316 (2021).

[54] B. Collins, Moments and cumulants of polynomial random variables on unitary groups, the Itzykson-Zuber integral, and free probability, Int. Math. Res. Not. **2003**, 953 (2003).

[55] Although the Weingarten matrix $W = G^{-1}$ exits only if $k \leq q$ [79], the Weingarten function can easily be extended to $k > q$ [51].

[56] D. Pérez-García, F. Verstraete, M. M. Wolf, and J. I. Cirac, Matrix product state representations, Quant. Inf. Comp. **7**, 401 (2007).

[57] A. Y. Kitaev, Fault-tolerant quantum computation by anyons, Ann. Phys. **303**, 2 (2003).

[58] I. MacCormack, A. Galda, and A. L. Lyon, Simulating large PEPS tensor networks on small quantum devices, ArXiv:2110.00507.

[59] L. Slattery and B. K. Clark, Quantum circuits for two-dimensional isometric tensor networks, ArXiv:2108.02792.

[60] M. Hauru, M. Van Damme, and J. Haegeman, Riemannian optimization of isometric tensor networks, SciPost Phys. **10**, 040 (2021).

[61] M. S. J. Tepaske and D. J. Luitz, Three-dimensional isometric tensor networks, Phys. Rev. Res. **3**, 023236 (2021).

[62] R. Haghshenas, J. Gray, A. C. Potter, and G. K.-L. Chan, Variational Power of Quantum Circuit Tensor Networks, Phys. Rev. X **12**, 011047 (2022).

[63] D. Haag, F. Baccari, and G. Styliaris, sequential GitHub repository.

[64] M. Fukuda, R. König, and I. Nechita, RTNI—A symbolic integrator for Haar-random tensor networks, J. Phys. A: Math. Theor. **52**, 425303 (2019).

[65] M. Fannes, B. Nachtergaele, and R. F. Werner, Finitely correlated states on quantum spin chains, Comm. Math. Phys. **144**, 443 (1992).

[66] C. W. von Keyserlingk, T. Rakovszky, F. Pollmann, and S. L. Sondhi, Operator Hydrodynamics, OTOCs, and Entanglement Growth in Systems without Conservation Laws, Phys. Rev. X **8**, 021013 (2018).

[67] T. Zhou and A. Nahum, Emergent statistical mechanics of entanglement in random unitary circuits, Phys. Rev. B **99**, 174205 (2019).

[68] B. Bertini and L. Piroli, Scrambling in random unitary circuits: Exact results, Phys. Rev. B **102**, 064305 (2020).

[69] Note that the entries of $T_e$ with $e \in S_2$ have previously appeared in Ref. [20].

[70] A. E. Rastegin, Relations for certain symmetric norms and anti-norms before and after partial trace, J. Stat. Phys. **148**, 1040 (2012).

[71] The Gram matrix $G$ is positive definite only if $k \leq dD$ because the set $\{P_\sigma^{(dD)}\}$ is linearly independent only if $k \leq dD$ [79]. As noted in Footnote [55], the Weingarten matrix $W = G^{-1}$ thus exists only if $k \leq dD$.

[72] R. A. Horn and C. R. Johnson, *Matrix Analysis* (Cambridge University Press, Cambridge, 2012).

[73] M. M. Wolf, Quantum Channels & Operations, Lecture notes (2012).

[74] M. Sanz, D. Perez-Garcia, M. M. Wolf, and J. I. Cirac, A quantum version of Wielandt's inequality, IEEE Trans. Inf. Theory **56**, 4668 (2010).

[75] B. Collins and S. Matsumoto, Weingarten calculus via orthogonality relations: new applications, Lat. Am. J. Probab. Math. Stat. **14**, 631 (2017).

[76] C. Akers, T. Faulkner, S. Lin, and P. Rath, Reflected entropy in random tensor networks, J. High Energy Phys. **2022**, 162 (2022).

[77] P. Biane, Some properties of crossings and partitions, Discrete Math. **175**, 41 (1997).

[78] T. Castellani and A. Cavagna, Spin-glass theory for pedestrians, J. Stat. Mech. **2005**, P05012 (2005).

[79] B. Collins, S. Matsumoto, and J. Novak, The Weingarten calculus, ArXiv:2109.14890.