I. INTRODUCTION

The principle of equal a priori probabilities, which is at the foundation of statistical mechanics, postulates subjective ignorance at the core of our understanding of macroscopic systems. It assumes that an isolated system is described by the microcanonical ensemble: an equal mixture over all possible microscopic configurations. On the other hand, quantum mechanics provides us with a powerful theory for understanding microscopic systems. It states that an isolated system is described by a pure state, which leaves no room for subjective ignorance. In the attempt to provide a consistent foundation of quantum statistical mechanics a purely quantum mechanical explanation of the effectiveness of statistical ensembles has been debated for quite a long time.

Recently an attempt to provide an alternative foundation to statistical mechanics was suggested, in which entanglement is viewed at the origin of generalized canonical ensembles. In this approach, there is no need to assume the principle of equal a priori probability, and the isolated system can be in a pure state, consistent with quantum mechanics [1, 2]. The mathematical justification for the effectiveness of this approach is based on the concentration-of-measure phenomenon, which has also appeared in the literature under the name of typicality [3]. Simplifying the results of [1, 2], one can prove that the vast majority of pure states picked uniformly at random in a sufficiently large Hilbert space will be almost indistinguishable at the level of sufficiently small subsystems, stemming from the appropriate statistical mechanical ensemble. Therefore it is not necessary to assume subjective ignorance over the state of the system, describing it with a mixed state, since any pure state will typically give very similar results.

Despite the elegance and simplicity of this approach there are aspects which need to be addressed in order to understand its concrete effectiveness. For example, it has already being pointed out in other contexts [4, 5] that it is extremely inefficient to reach or create a random pure state in a large system. This is due to the fact that the uniformly distributed random unitary generating the state requires an exponential number of parameters to be specified, and any dynamical process corresponding to it will take an exponentially long time. This is in contrast with the relatively fast process of equilibration, which naturally leads the subsystem to its generalized canonical ensemble. Therefore, if typicality has to provide an explanation for the statistical mechanical ensembles corresponding to equilibrium states, then it should also be able to account for fast equilibration processes. One way of solving this problem is to restrict the class of allowed states in the isolated system. This kind of approach has also been considered in Ref. [6].

In previous work, we reported first results on the occurrence of typicality in random matrix product states [7]. Constraining the states of the system to be MPS-like is one way to avoid the efficiency problem, mentioned above, in the realization of the concentration-of-measure phenomenon. MPSs are an example of physically meaningful states. They can be ground states of gapped Hamiltonians [8, 9], and they are at the core of some computational techniques, based on quantum information theory, which have been recently introduced [10]. The principal reason for the effectiveness of MPS is that the ground states of many relevant systems can be well approximated by them (see [12] for a recent review and original references). Here we present more details about typicality in MPSs giving further evidence of its relevance.

In Sec. II we provide some background on the literature concerning typicality in random quantum states. In Sec. III we define the particular set of random MPS we are interested in and review previous results on the concentration-of-measure phenomenon in this context [8]. In Sec. IV we provide a detailed numerical analysis of the initial results discussed in Ref. [8] and present an analytical derivation of the average non-homogeneous RMPS.
We also study specific statistical properties of RMPS and show how they can be used to efficiently approximate some characteristics of general random pure states. Sec. V is devoted to conclusions.

II. TYPICALITY BACKGROUND

Arguments related to typicality have been presented throughout the quantum and statistical mechanics literature since the very beginning of this field (e.g. [13] and chapter 6 in [14]). More recently, Lebowitz has emphasized the importance of typicality in classical statistical mechanics in the context of the second law of thermodynamics [12]. In Refs. [1] [2] the importance of typicality was recognized also in the context of quantum statistical mechanics. A different and more heuristic approach was used in [19]. A nice review with original contributions can be found in [14]. In the following, we provide a brief overview of some of these works.

The work in [2] makes direct use of the concentration-of-measure phenomenon and focuses on the properties of the subsystem’s states. The concentration-of-measure phenomenon is a well known topic in the mathematical literature and appears in many different contexts [17–19]. Roughly speaking, suppose we have a function $f: V^n \to \mathbb{C}$, whose domain is an n-dimensional vector space equipped with a probability measure. If $f$ does not oscillate too much then, for $n$ sufficiently large, $f$ is almost constant, with very high probability. More formally:

$$\text{Prob}\{ |f - \bar{f}| \geq \epsilon \} \leq k_1 \exp(-k_2 \epsilon^2 n/\eta^2),$$

where $\bar{f}$ is the average value of $f$, $k_{1,2}$ are some universal positive constants, and $\eta$ is the Lipschitz constant of $f$. In this sense $f$ is concentrated around its average value. The Lipschitz constant can be defined as the supremum of the gradient. Within this formalism one can prove, for example, that the area of a hypersphere concentrates around the equator [17]. When the domain of $f$ is a hypersphere the concentration-of-measure phenomenon is also referred to as Levi’s lemma [17].

A. Local typicality

Suppose we have a system $R$, whose Hilbert space $\mathcal{H}_R$ belongs to $\mathcal{H}_S \otimes \mathcal{H}_B$, the tensor product of a subsystem $S$ and a bath $B$. We denote the dimensions of the corresponding Hilbert spaces with $d_R$, $d_S$ and $d_B$ respectively. The states in the isolated system $R$ may satisfy some constraints or restrictions, for example belonging to a small energy-shell with respect to a Hamiltonian. We now consider a uniformly random state $|\psi\rangle \in \mathcal{H}_R$. This can be achieved by generating $|\psi\rangle$ with a random unitary matrix, distributed according to the Haar measure, acting on some reference state in $\mathcal{H}_R$. Consider now the state of the subsystem $\rho_S = \text{Tr}_B |\psi\rangle \langle \psi|$. The results in [2] show that, for almost all $|\psi\rangle$, $\rho_S$ will be very close in trace norm to the generalized canonical distribution. This is proved by defining the function $D_1 \equiv \|\rho_S - \bar{\rho}_S\|_1$, for which the Lipschitz constant is 2, and using Levi’s lemma to obtain a concentration-of-measure result on the function $D_1$. Then one has to show that also the average value $\bar{D}_1$ is small: most states are almost at the same distance from the average state, and this distance is small. This last result comes from the following inequalities [2]

$$\bar{D}_1 \leq \sqrt{d_S \text{Tr}[\rho_B^2]} \leq \sqrt{d_S^2/d_R}. \quad (2)$$

From this, using Chebyshev’s inequality and the fact that $D_1 \geq 0$, one obtains a first concentration result, i.e. a bound on the probability of fluctuations. However, Levi’s lemma gives a much stronger bound [2]

$$\text{Prob}\left[ D_1 \geq \epsilon + \sqrt{d_S \text{Tr}[\rho_B^2]} \right] \leq 2 e^{-Cdn^2}, \quad (3)$$

with $(C = 18\pi^3)^{-1}$. Note that the probability is exponentially suppressed in the dimension of the Hilbert space and not just in the size of $R$, which would be the result obtained using Chebyshev’s inequality.

From the bound on the state distance one can easily derive a similar result for any observable, in this way establishing typicality for the subsystem. It is worth emphasizing that this derivation of a generalized canonical ensemble is fully consistent with the quantum formalism and makes no use of the principle of equal a priori probability. On the other hand, the entanglement between the subsystem and the bath is seen as the main mechanism through which incomplete knowledge over the state of the subsystem emerges.

B. Global typicality

An alternative approach to typicality [12] aims to establish it at the level of the isolated global system, and to generalize it further to probability distributions other than uniform. Let us consider a basis decomposition of a random state

$$|\psi\rangle = \sum_n c_n |n\rangle, \quad (4)$$

where the coefficients $c_n$ are random. We make the following two assumptions: $c_n$ are independent (but do not need to be identically distributed) random variables, and they are phase independent, i.e.

$$\mathcal{P}(\psi) \to \mathcal{P}(c_1, c_2, ...) = \prod_n \mathcal{P}(|c_n|). \quad (5)$$

This kind of probability distribution gives an average state which is diagonal and coincides with the dephased
Defining $\rho \equiv |\psi\rangle \langle \psi|$, and assuming that the purity of the average state is low ($\text{Tr} [\rho^2] \ll 1$), one can show that the variance in the expectation value of $A = \langle \psi | A | \psi \rangle$ is small

$$\sigma_A^2 \equiv \langle (A - \bar{A})^2 \rangle \leq ||A||^2_2 (\max_n q_n) \text{Tr}[\rho^2]$$

with $q_n$ the normalized variance of the random variables $|c_n|^2$. Note that the operator norm $||A||^2_2$ quantifies the range of possible values for the observable $A$. From the above inequality it follows that a random pure state is likely to yield expectation values for $A$ which are very close to the ensemble average $\text{Tr}[\rho A]$. This behavior occurs for probability distributions such that $\text{Tr} [\rho^2] \ll 1$, and for observables, where $||A||^2_2$ does not increase linearly with the dimension of the Hilbert space (which can be argued to be the physically most interesting, see also the comments in [24]). A trivial way to guarantee that $||A||^2_2$ does not increase is to restrict to local observables, supported on finitely many subsystems. From the bound on the variance one can use Chebyshev’s inequality to bound the probability that $A$ will be far away from its average value $\bar{A}$.

C. Remarks

The use of non-uniform probability distributions for the random state of the isolated system could be considered unphysical, since they are not invariant under a unitary transformation and consequently they depend on the chosen basis. But, in Ref. [10] it was argued that also assuming that all eigenstates are populated with the same probability is rather unphysical. In this sense, these results are interesting since they show that typicality does not depend on the details of the probability distribution. The canonical ensemble, for example, can be recovered under much more general and realistic conditions than the ones usually considered, such as the principle of equal a priori distribution. Another advantage of this approach is that it establishes typicality at a global level. From this it follows that typicality is a concept independent of entanglement. Typicality at the subsystem level follows naturally from typicality at the global level. However, since there are some observables for which typicality does not hold at the global level, it would be interesting to understand their physical relevance.

It is important to underline that the above discussion does not show thermalization emerging from a dynamical process. Nevertheless, they show that most of the states of the Hilbert space are thermalized. Therefore one expects that most of the dynamics will give thermalization, the exception being some particular dynamics that keep the states within an exponentially small atypical subspace. In fact, in Refs. [21,22] it has been shown in great generality that most of the time the state is close to some fixed state, the time averaged one, and in this sense there is equilibration of the system. An interesting area of research in this respect is the unitary dynamics of isolated systems after a sudden quench [23, 24].

The typicality arguments reviewed above provide a consistent foundation for quantum statistical mechanics, although some questions can still be raised. In particular, one can argue about the exponential amount of resources necessary to generate uniformly distributed random states and the fact that apparently no dynamical process can generate such a state efficiently. In this sense, one might also wonder whether most of the states in the Hilbert space are physically realizable. These arguments are also relevant in condensed matter, where one is not able to use all possible states because of the exponential number of parameters that would be needed. Due to the locality of the interactions only a subset of states will be effectively useful in the description of the system. The problem then is to find an efficient representation for such physical states.

Therefore it is important to study if typicality is still valid for a smaller set of more relevant states. Hence, we decide to focus on MPSs, relying on their importance for Hamiltonians with local interactions.

III. RANDOM MATRIX PRODUCT STATES

A matrix product state is a pure quantum state whose coefficients are specified by a product of matrices. For the case of periodic boundary conditions (PBC) a MPS can be written as

$$\sum_{i_1,\ldots,i_N} \text{Tr} (A^{i_1}[1] \cdots A^{i_N}[N]) |i_1 \cdots i_N\rangle,$$

whereas with open boundary conditions (OBC) one has

$$\sum_{i_1,\ldots,i_N} \langle \phi_F | A^{i_1}[1] \cdots A^{i_N}[N] | \phi_F \rangle |i_1 \cdots i_N\rangle,$$

with $|\phi_{F}\rangle$ specifying the states at the boundaries, and $|i_k\rangle$ is a local basis at site $k$. The matrices $\{A^1[k], A^2[k], \ldots, A^D[k]\}$, with $k \in \{1, \ldots, N\}$, are $\chi$-dimensional complex matrices, where $D$ is the local Hilbert space dimension. We denote as homogeneous MPSs the states for which the set $\{A^1[k], A^2[k], \ldots, A^D[k]\}$ is the same for all sites $k$. By definition, a MPS is specified by the set $\{A^1[k], A^2[k], \ldots, A^D[k]\}$. However there may be different sets of matrices that originate the same MPS. In [25] it was shown that this gauge degree of freedom can be fixed using a canonical form. The fundamental parameter characterizing the properties of MPS states is $\chi$, the size of the $A$-matrices. Note that any non-homogeneous MPS is parametrized by $N D \chi^2$ numbers, which can be much less than the $D^N$ values needed for a general state. Furthermore one can prove that the maximum entanglement a subsystem can have with its environment depends on
\[ \chi \]. It can also be shown that any state can be described as a MPS for sufficiently large \( \chi \propto D^N \), but there is no advantage in such a representation. In [10] it was proved that the ground state of any one-dimensional local Hamiltonian is well approximated by a MPS. At the computational level, MPSs are also very useful in algorithms based on the density matrix renormalization group [12].

For all these reasons MPSs can be considered to be good representations of physical states of one-dimensional systems, and it is also possible to generalize this formalism to higher-dimensional systems [12].

In the present work, using the sequential generation of MPSs introduced in [23, 24], we consider random MPSs with a clear operational definition. Consider a quantum spin chain initially in a product state \( |0\rangle^\otimes N \in \mathcal{H}_B^\otimes N \) (with \( \mathcal{H}_B \simeq \mathbb{C}^D \)) and an ancillary system in the state \( |\phi_I\rangle \in \mathcal{H}_A \simeq \mathbb{C}^\chi \). Let \( U[k] \) be a unitary operation on \( \mathcal{H}_A \otimes \mathcal{H}_B \), acting on the ancillary system and the \( k \)th site of the chain (see Fig. 1). The \( A[k] \) matrices are then defined by

\[ A_{\alpha,\beta}[k] \equiv \langle i, \alpha|U[k]|0, \beta \rangle, \tag{9} \]

where the Greek indices refer to the ancilla space and the Latin indices to the physical space (see [23, 24] for related works on truncated random unitaries). For homogeneous MPSs the index \( k \) is removed, implying that the unitary interaction is the same for all sites in the spin chain. Due to unitarity we have \( \sum_i A_i^\dagger[k] A_i[k] = I_{\chi} \) for all \( k \) in the bulk (see [3] for more details). This property, together with a proper normalization of the boundaries, corresponds to a MPS of unit-norm. Letting the ancilla interact sequentially with the \( N \) sites of the chain and assuming that the ancilla decouples in the last step (this can be done without loss of generality, as shown in [24]), the state on \( \mathcal{H}_B^\otimes N \) is described by

\[ |\psi\rangle = \sum_{i_1, \ldots, i_N} \langle \phi_F|A_{i_N}^\dagger \cdots A_{i_1}^\dagger|\phi_I\rangle|i_N \cdots i_1\rangle, \tag{10} \]

which is a homogeneous MPS with open boundary conditions. It can be proved [23, 24] that the set of states generated in this way is equal to the set of OBC-MPSs. We choose the interaction characterizing the RMPS ensemble to be represented by random unitary matrices \( U[k] \) distributed according to the Haar measure.

Since any state can be described by a MPS when \( \chi \propto D^N \), typicality naturally holds true for MPSs with this scaling of \( \chi \). What is not obvious is whether one can have typicality in the set of RMPSs where \( \chi \) increases at most polynomially with the number of particles \( \chi \propto N^a \), with some constant \( a > 0 \). In [8], using a concentration-of-measure result for the group of unitary matrices [13], we were able to prove typicality for the expectation values of local observables over \( L \) sites

\[ O \equiv \left( \bigotimes_{k=1}^{L} O[k] \right) \left( \bigotimes_{k=L+1}^{N} I[k] \right), \tag{11} \]

with respect to normalized RMPS \( |\psi\rangle \) generated according to the sequential construction. Let us define

\[ f \equiv \langle \psi|O|\psi\rangle = Tr \left( \prod_{k=1}^{L} E_O[k] \prod_{k=L+1}^{N} E_I[k] \right), \tag{12} \]

where the transfer operator \( E \) is defined as

\[ E_O[k] \equiv \sum_{i_k, j_k=1}^{D} \langle i_k|O[j_k]\rangle A_{i_k}[k] \otimes A_{j_k}^*[k], \tag{13} \]

\[ E_I[k] \equiv \sum_{i_k=1}^{D} A_{i_k}[k] \otimes A_{i_k}^*[k]. \tag{14} \]

Since the \( A \)-matrices are sub-blocks or random unitaries \( U \), they define a random variable \( f : U(\chi D) \rightarrow \mathbb{R} \) satisfying [8] (\( c_{1,2} \) are positive constants)

\[ \Pr \left[ |f - \overline{f}| \geq \epsilon \right] \leq c_1 \exp \left( -c_2 \epsilon^2 D \frac{\chi}{N^2} \right). \tag{15} \]

This means that if \( \chi \propto N^a \), with \( a > 2 \), increasing the size of the system renders the expectation values more concentrated around their averages, i.e. statistical fluctuations will be suppressed [37]. From this result one can also derive a weaker concentration result for the probability of fluctuation of the trace distance at the sub-system level. Defining \( \rho_s \equiv Tr_{\text{Env}}|\psi\rangle \langle \psi| \), and analogously for the average state, we have (see Appendix C in [8] for a derivation)

\[ \Pr \left[ \|\rho_s - \overline{\rho_s}\|_1 \geq 4\epsilon \right] \leq 4^{L} c_1 \exp(-c_2 \epsilon^2 D \frac{\chi}{N^2}). \tag{16} \]

This shows that although the number of MPS is much smaller than the total number of states, picking random matrix product states \( |\psi\rangle \) according to the above construction still provides a concentration-of-measure result for suitable random variables defined through \( |\psi\rangle \). This fact could be exploited numerically for the efficient simulation of quantum systems. Ref. [23] provides a possible indication of the role that the concentration of measure may play in numerical simulations. In this work a differ-

---

Figure 1: Sequential generation of a MPS \(|\psi\rangle\).
ent class of random MPS is constructed and used for the efficient sampling of local observables. We suspect that the efficiency of the sampling is due to a concentration of measure phenomenon.

We point out that the above result concerning the subsystem state is weaker than the one obtained in [2], where the probability of fluctuations is exponentially suppressed in the dimension of the total Hilbert space, i.e. doubly-exponentially in \( N \). Similar weaker bounds have been obtained in the context of k-designs [2].

IV. STATISTICAL PROPERTIES

A. Haar distributed random states and RMPS

In this section, we discuss numerical signatures of typicality for the case of random pure states distributed according to the Haar measure. These simulations are useful in illustrating typicality and the tightness of the theoretical bounds, and they also serve as a reference point for understanding the properties of RMPS.

We first want to understand the effect of the number \( r \) of sampled states used in the averages. In Fig. 2 we evaluate the trace distance between the empirical average random state and its exact value for the circular unitary ensemble (CUE) which is the ensemble of unitary matrices corresponding to the Haar measure, as a function of \( r \). \( D_1 \equiv \|\bar{\rho} - \mathbb{1}/DN\|_1 \). The figure shows the results for different system sizes: 3, 6 and 8 qubits. As expected, it is observed that more states are needed as the dimension of the Hilbert space is increased. At the level of the subsystem, we also check how statistical fluctuations are suppressed when increasing the number of sampled states. For this purpose we consider the average Hilbert-Schmidt distance for a sub-system of one qubit, \( D_2 \equiv \|\bar{\rho}_s - \overline{\rho}_s\|_2 \), which is easier to evaluate numerically than the trace distance. From Fig. 3 we see that 500 states seem already sufficient to suppress the statistical fluctuations in the subsystem.

Now, using 500 random pure states, we compare the bound given in [2] for \( D_1 \) and its empirical value obtained from the simulations. Fig. 4 shows how the analytical value and the numerical results are pretty close and scale in the same way with the bath size \( n_B = N - L \). In the figure we also plot the actual value of the average distance using the trace norm and Hilbert-Schmidt norm.

We now consider RMPS originated from random unitaries appearing in the sequential generation scheme. The method used to obtain random unitaries distributed according to the Haar measure is documented in the literature [30]. It makes use of the orthonormalization of a random matrix where all the elements are i.i.d. Gaussian random numbers of zero average and unit variance.

Again we need to estimate the size \( r \) of the sample, in order to avoid statistical fluctuations that would be too strong. We look at how the average trace distance from the average RMPS for the subsystem,
Figure 4: (Color online) Illustration of typicality for general states. The figure shows the analytical value of the bound, $\sqrt{d_S/d_B}$ (with $d_S$ and $d_B$ the dimension of the system and the bath), and the numerical average value of $D^1_r$ and $D^2_r$ for general random states distributed with the Haar measure.

Figure 5: (Color online) Average trace distance from the average state for a subsystem of one qubit, $D^1_r = |\|\rho_s - \rho_r^s\|_1|$, for homogeneous MPSs and OBC as the number of states used in the average increases for $\chi = N - L = 50, 60, 70, 80, 90, 100, 110, 150$, and $200$ (from top to bottom).

Figure 6: (Color online) Average trace distance from the average state for a subsystem of one qubit, $D^1_r = |\|\rho_s - \rho_r^s\|_1|$, for homogeneous MPSs and OBC as the number of states used in the average increases for $\chi = N - L = 50, 60, 70, 80, 90, 100, 110, 150$, and $200$ (from top to bottom).

Figure 7: (Color online) $D^1_r$ as a function of the difference between the size of the system and the subsystem, for fixed but different values of $\chi$ and for $L = 1$. The symbols denote homogeneous MPS, whereas the symbols joined by lines denotes non-homogeneous MPS with the same value of $\chi$. Periodic boundary conditions are imposed.

In Fig. 6 we investigate the behavior of $D^1_r$ for non-homogeneous MPSs and homogeneous MPSs with OBC. When we fix the value of $\chi$ and increase the number of qubits in the bath, $D^1_r$ starts to decrease, but soon reaches a constant value. This value depend on $\chi$ and decreases as we increase $\chi$. Such behavior is consistent with the previously mentioned analytic result [8]. The same behavior is observed for larger subsystems (up to 5 qubits), but with higher saturation values. For PBC the simulations are slower, since we have a scaling with $\chi^5$, but we checked that until values around $\chi = 20$ the behavior is similar. This is shown in Fig. 8.

In Fig. 7 we consider the dependence of $D^1_r$ in $\chi$, for fixed system and bath sizes. In Fig. 9 it can be seen that $D^1_r$ also decreases with $\chi$, until a value that depends
on the bath size. For small system (up to 8 qubits) we checked that this limiting value is the same as the one for the CUE. Since χ can be viewed as a sort of correlation length, this behavior is easily explained considering that when the correlation length is of the order of the size of chain, $D_1$ cannot decrease anymore.

We now allow χ to scale linearly with the size of the bath: $\chi = N - L$, see Fig. 10. In this case we observe that until $\chi = 200$ the variance is decreasing monotonically, which indicates that typicality can emerge already for a linear scaling of χ with the number of particles.

B. Analytical Results for the average MPS

An important aspect in the study of the concentration-of-measure for RMPS is the characterization of the average state. Below we provide an analytical expression for the averaged MPS state and numerical verifications of the results obtained.

We use bold faced letters to denote vectors: $\mathbf{i} \equiv (i_1, \ldots, i_N)$. Let

$$|\psi\rangle = \sum_{\{i\}} \text{Tr}(A_{i_1}[1]A_{i_2}[2] \cdots A_{i_N}[N]) |i_1i_2\cdots i_N\rangle \quad (17)$$

be an instance of RMPS, and the $A$-matrices are defined as $A^{ij}[k] \equiv \text{Tr}_{\mathcal{F}}[(|i_k\rangle\langle 0| \otimes \mathcal{I}_A) \cdot U[k]]$, where A is the ancillary Hilbert space and $\mathcal{F}$ is the physical Hilbert space. For brevity of notation we will not always specify the site index of the $A$-matrices and $U$-matrices. The density matrix $\psi \equiv |\psi\rangle\langle\psi|$ is given by

$$\psi(U) = \sum_{\{i,j\}} \text{Tr}(A^{i_1} \cdots A^{i_n}) \text{Tr}(A^{j_1} \cdots A^{j_N})^*$$

$$\times |i_1 \cdots i_n\rangle \langle j_1 \cdots j_n|, \quad (18)$$

where $^*$ denotes complex conjugation. The average density matrix is given by

$$\overline{\psi} = \int_{\text{Haar}} \psi(U) \, dU. \quad (19)$$

The factor in the coefficient of the density matrix can also be written as

$$\text{Tr}_{\mathcal{A}} (A^{i_1} \cdots A^{i_N}) =$$

$$\text{Tr}_{\mathcal{A} \otimes N} [S_{\mathcal{A} \otimes N} \cdot (A^{i_1} \otimes \cdots \otimes A^{i_N})] =$$
\[ \text{Tr}_{\mathcal{A}^\otimes N \otimes F^\otimes N} \left[ (i) (0)_{\mathcal{F}^\otimes N} \otimes S_{\mathcal{A}^\otimes N} \right] \left( \bigotimes_{k=1}^{N} U_k \right) \]  

(20)

where \( S \) is the operator which cyclically permutes the states in \( \mathcal{A}^\otimes N \)
\[ S |\alpha\rangle_1 \cdots |\alpha\rangle_N = |\alpha\rangle_N |\alpha\rangle_1 \cdots |\alpha\rangle_{N-1} \]  

(21)

The coefficient of the density matrix (18) can then be rewritten as
\[ \text{Tr}_{\mathcal{A}^\otimes 2N \otimes F^\otimes 2N} \left[ (i) (0) \otimes S \otimes |j\rangle (0) \otimes S \right] \times \left( \bigotimes_{k=1}^{N} U[k] \right) \otimes \left( \bigotimes_{k=1}^{N} U[k]^* \right) \]

(22)

In a more compact form we can define a new averaged density matrix in \( \mathcal{F}^\otimes 2N \)
\[ \Psi = \text{Tr}_{\mathcal{A}^\otimes 2N} \left[ (\mathbb{I} \otimes S \otimes \mathbb{I} \otimes S) \cdot \left( \bigotimes_{k=1}^{N} U[k] \right) \otimes \left( \bigotimes_{k=1}^{N} U[k]^* \right) \right] \]

(23)

In the previous expression the quantity
\[ \left( \bigotimes_{k=1}^{N} U[k] \right) \otimes \left( \bigotimes_{k=1}^{N} U[k]^* \right) \]

(24)

carries information about the average over the Haar measure. Since, for non-homogeneous RMPS, permutations of the factors inside the integral are allowed, one sees that
\[ \left( \bigotimes_{k=1}^{N} U[k] \right) \otimes \left( \bigotimes_{k=1}^{N} U[k]^* \right) = \otimes_{k=1}^{N} U[k] \otimes U[k]^* = \Pi_\chi^\otimes N \]

(25)

where \( \Pi_\chi^\otimes \) is the density matrix of the maximally entangled state, (see the appendix in [31])
\[ \Pi_\chi^\otimes \equiv \frac{1}{\chi^\otimes} \sum_{l=1}^{\chi^\otimes} \sum_{l'=1}^{\chi^\otimes} |l, l\rangle \langle l', l'| \]

(26)

Now let us go back to the expression for \( \Psi \). Since the operator \( S \) that cyclically permutes the ancilla spaces acts on \( N \)-tensor copies of the same state, it does not change the state
\[ \text{Tr}_{\mathcal{A}^\otimes 2N} \left[ (\mathbb{I} \otimes S \otimes \mathbb{I} \otimes S) \cdot \left( \bigotimes_{k=1}^{N} U[k] \right) \otimes \left( \bigotimes_{k=1}^{N} U[k]^* \right) \right] = \text{Tr}_{\mathcal{A}^\otimes 2N} \left[ \Pi_\chi^\otimes \right] \]

(27)

and the trace will just restrict the projector to the maximally entangled state over the physical subspace
\[ \text{Tr}_{\mathcal{A}^\otimes 2N} \left[ \Pi_\chi^\otimes \right] = \Pi_\chi^\otimes \]

(28)

Now it is easy to see that
\[ \langle i,j | \Pi_\chi^\otimes \rangle = \frac{1}{D^N} \sum_{l=1}^{D^N} \sum_{l'=1}^{D^N} (i,j) |l,l\rangle \langle l', l'| (0,0) = \frac{1}{D^N} \delta_{i,j} \]

(29)

This proves that the average non-homogeneous RMPS is the completely mixed state for any value of \( \chi \)
\[ \Psi^\text{RMPS} = \Psi^\text{CUE} \]

(30)

Since the average state is the same for the two ensembles, all the functional depending only on the first moment of the distribution will be identical in the non-homogeneous RMPS ensemble and CUE ensemble. This is the case for the expectation value of observables
\[ \text{Tr}(O \rho)^\text{RMPS} = \text{Tr}(O \rho)^\text{CUE} \]

(31)

The case of homogeneous RMPS is more complicated, although a close expression can be obtained starting from Eq. (22). The problem comes from the fact that now it is not anymore possible to factorize the average in the tensor product of unitaries, but nevertheless the integral can be expressed as [31]
\[ \int_{\text{Haar}} (U \otimes U^*)^\otimes N dU = \sum_{\sigma} |\hat{P}_{\sigma}\rangle |\hat{P}_{\sigma}\rangle \]

(32)

where the vector \( |\hat{P}_{\sigma}\rangle \) is obtained superimposing the column of the matrix \( P_{\sigma} \), or an orthonormal representation of the permutation \( \sigma \) acting on \( N \) elements. In general it is hard to find such a representation for large \( N \) and \( D \).

C. Numerical Results for the average MPS

In this section we numerically check the previous results. In Fig. [11] we examine the distance between the numerically averaged MPS state and its analytical value \( \Pi/D^N \), for systems of two and eight qubits, as the number \( r \) of sampled states increases. It can be seen that the average non-homogeneous MPS converges to the mixed state as well as the average general state (Fig. [2]). The simulation shows that for these small systems the average MPS state does not depend on \( \chi \), as expected from Eq. (20) (see also Fig. [12]). Note that this \( \chi \)-independence of the average MPS is a-priori not at all obvious.

Now we want to obtain more insight into the average MPS, studying its behavior for larger system. We use
As the number of states used in the average is increased, for general states (CUE), however the different plots can barely be distinguished in the plot scale.

The purity of the average MPS is given by

\[ \Tr[(\bar{\rho})^2] = \frac{1}{r^2} \sum_{i,j=1}^{r} \frac{|\langle \psi_i | \psi_j \rangle|^2}{|\langle \psi_i | \psi_j \rangle|^2}. \]

and the overlap between different MPSs can be efficiently evaluated. Before showing the results, let us rewrite the previous formula, remembering that we want the purity to attain its minimum value \(1/D^N\)

\[ \Tr[(\bar{\rho})^2] = \frac{1}{r^2} + \frac{1}{r^2} \sum_{i \neq j} \frac{|\langle \psi_i | \psi_j \rangle|^2}{|\langle \psi_i | \psi_j \rangle|^2}. \]

In the limit of a large sample, the first term vanishes, and the second term should converge to \(1/D^N\). It is clear that for large systems, \(D^N \gg r\), the principal limitation comes from the finite size of the sample. Note that these expressions are also valid for general states. Let us study how much the second term on the left hand side of the previous equation differs from \(1/D^N\). This is shown in Fig. 13 as a function of the system size, and for different values of \(\chi\). An exponential decrease of the purity is observed for systems of up to 50 qubits. In the plot we considered values of \(\chi = 2, 4, 20\) and 50. However, they all collapse to the same point. In fact, Fig. 13 shows that there is no relevant dependence on \(\chi\) with a very high accuracy. We also study the relative error of the purity in relation to the analytical value. This is an upper bound for the trace distance (again the term \(1/r\) is neglected):

\[ \frac{\Tr[(\bar{\rho})^2] - \frac{1}{D^N}}{D^N} = \frac{1}{D^N} - \frac{1}{D^N} |\bar{\rho} - \frac{1}{D^N}||^2 - \frac{D^N}{r}. \]

Fig. 13 shows this expression as a function of the number of qubits and for different values of \(\chi = 2, 4, 8, 20\) and 50 (from top to bottom). Note that already for \(\chi = 2\) the error is around 10% and does not increase with the system size. Some fluctuations are observed, and we
D. Random MPS and uniformly distributed states

The scheme for the sequential generation of RMPS can be viewed as an algorithm for the efficient generation of random states. The computational efficiency is given by the number of random unitary matrices needed, which equals the size of the system, and by the size of these matrices, which depends on \( \chi \). In this section we study how well the ensemble of non-homogeneous RMPS mimics a subset of statistical features associated with the Haar measure. In particular we consider the average bipartite entanglement, the minimum eigenvalue distribution and higher moments of the reduced density matrix.

The distribution of entanglement produced by the present scheme for the generation of random states can be a relevant quantity for quantum information tasks. The average bipartite entanglement (ABE) or global entanglement \[32, 33\], here denoted with \( Q \), is a measure of multipartite entanglement. It is defined as

\[
Q \equiv 2 - \frac{2}{N} \sum_{i=1}^{N} Tr[\rho_i^2],
\]

where \( \rho_i \) is the reduced density matrix of the i-th qubit. \( Q \) can have values between 0 and 1, ranging from a product state to a maximally entangled state. In Fig. 16 we show histograms of \( Q \) obtained over \( 10^6 \) realizations of RMPS for different system sizes. The distribution shows the same features as the ones obtained with other random circuit schemes \[4\]. Increasing the dimension of the Hilbert space the values of \( Q \) concentrate around the average. This is a manifestation of the concentration of measure phenomenon.

In Fig. 17(a) we plot the difference between the average value of \( Q \) for RMPS and the exact value known for the CUE ensemble: \( |\bar{Q}_{RMPS} - \bar{Q}^{CUE}| \). The simulation
shows that the average $\overline{Q}_{\text{RMP S}}$, for $\chi$ sufficiently large, depends only on the size of the system and the larger the system the closer the average value is to the CUE exact result: $\frac{d}{d + 1}$, with $n$ the size of the system. Although from the numerical simulations the difference between the two is always finite. This behavior is similar to other random circuit constructions \[3], but in our case the size of the system plays the role of the depth of the random circuit and, for a fixed $\chi$, determines the computational cost of the simulation. This is also an indication of the fact that the second moment of the RMPS $\overline{(\rho \otimes \rho)_{\text{RMP S}}}^2$ is different from $\overline{(\rho \otimes \rho)_{\text{CUE}}}^2$. This can be seen from the fact that the average purity is a functional of two copies of the random state: $\overline{Tr(\rho^2)} = Tr(S \cdot \rho \otimes \rho)$, where $S$ is the swap operator. The discrepancy between the RMPS and CUE ensembles can then be detected from the second moment of the distributions.

In Fig. 17(b) we plot the decrease of the standard deviation of $Q$ as a function of the system size $n$, with $\chi = 64$. The exponential decrease is again a signature of the concentration-of-measure phenomenon for RMPS: for large system size almost all RMPS have $Q$ exponentially close to the average value $\overline{Q}_{\text{RMP S}}$.

The other quantities we want to compare with the ones obtained from the CUE are the higher moments of the reduced density matrix evaluated for a partition of the system in two parts, $A$ and $B$, of dimension $d_A$ and $d_B$ respectively. Calculations in the CUE provide the following exact results \[3, 33\]

$$\overline{Tr(\rho^2)}_{\text{CUE}} = \frac{d_A + d_B}{d_A d_B + 1}.$$  \hspace{1cm} (39)

$$\overline{Tr(\rho^2)}_{\text{RMP S}} = \frac{d_A^2 + 3d_A d_B + d_B^2 + 1}{(d_A d_B + 1)(d_A d_B + 2)}.$$  \hspace{1cm} (40)

$$\overline{Tr(\rho^2)}_{\text{CUE}} = \frac{d_A^3 + 6d_A^2 d_B + 6d_A d_B^2 + d_B^3 + 5d_A + 5d_B}{(d_A d_B + 1)(d_A d_B + 2)(d_A d_B + 3)}.$$  \hspace{1cm} (41)

Fig. 17(c) shows how the distance of the averaged RMPS value from the CUE results depends on $\chi$. As can be seen a finite and relatively small value of $\chi$ is sufficient to guarantee a very good approximation of the CUE value. Again, after some point there is no improvement in increasing the value of $\chi$, as already observed for the average bipartite entanglement in Fig. 17(a).

We also consider the statistical properties of the minimum eigenvalue of the reduced density matrix of a subsystem $A$ of dimension $d_A$. This is a quantity related to the entanglement of the subsystem, and its values can range from 0, for product states, to $1/d_A$, for maximally entangled states \[39\]. Fig. 17(d) shows the dependence on $\chi$ of $|\lambda_{\text{min}}\text{RMP S} - \lambda_{\text{min}}\text{CUE}|$, where $\lambda_{\text{min}}\text{CUE} = 1/d_A$ and $d_A = 4$ (in a system of 6 qubits). The figure indicates an exponential convergence to the CUE value. The fluctuations seen in the plot are due to the finite size of the sampling set and to the small value of the quantity that we want to estimate (of the order of $10^{-5}$). This result is again an indication of the good accuracy that can be obtained in approximating some properties of the CUE with the RMPS states.
V. CONCLUSIONS

In conclusion, in this work we have studied in detail a set of random matrix product states (RMPS) introduced in Ref. [8]. As already pointed out in that reference, RMPSs can be a useful tool to address foundational problems of quantum statistical mechanics. In particular here we have proved that the set of non-homogeneous RMPS and the set of uniformly distributed general random states have the same average state. This property, together with the validity of the concentration-of-measure phenomenon, implies that any generalized canonical state can be approximated by the reduced density matrix of a random matrix product state, as long as the average random MPS coincide with the associated averaged microcanonical ensemble. Let us call \( \Omega = \frac{1}{d_R} \) the totally mixed state of a global Hilbert space \( \mathcal{H}_R \) satisfying some set of restrictions denoted with \( R \) (e.g. having a fixed energy), and whose MPS states are denoted with \( \psi_R \equiv \langle \psi_R | \psi_R \rangle_{\text{MPS}} \). Assuming that the average random MPS satisfies \( \bar{\psi}_R = \Omega \) and substituting this in Eq. (16) it follows

\[
\Pr \left[ \| \rho_s - Tr_{Env}[\psi] \|_1 \geq 4^{3L/2} \right] \leq 4^L C_L \exp \left( -c_2 \epsilon^2 D \frac{X}{N^2} \right),
\]

(42)

with \( \rho_s \equiv Tr_{Env}\psi_R \). Since \( Tr_{Env}\Omega \) is nothing but the generalized canonical state this prove our statement. Notice that this is true only assuming measure concentration in the restricted space and the equality \( \psi_{\text{MPS}} = \Omega \). The present work focused on the case when the Hilbert space has no restrictions. An interesting future direction of research would be to check the identity between \( \psi_R \) and \( \Omega \) when some kind of constraints are imposed.

Another interesting application of RMPSs is in the field of pseudo-random quantum circuits. We show that statistical properties of general quantum random states, which are computational resources for some quantum information tasks [4], are very well approximated by RMPSs. Since these states can be generated efficiently, as long as \( \chi \) scales polynomially in the size of the system, they constitute an efficient tool for the approximate simulation of random quantum states.

Acknowledgments

This work has been supported by NSF grants: PHY-083304, DMR-0804914.
[35] I. Nechita, Annales Henri Poincare 8, 1521 (2007).
[36] S. N. Majumbar, O. Bohigas, and A. Lakshminarayan, Journal of Statistical Physics 131, 33 (2008).
[37] Note that in the case of non-homogeneous RMPS the domain of the function $f$ is $N\chi D$–dimensional, implying a concentration result for $\chi \propto N^a$ and $a > 1$. 