Random unitaries give quantum expanders

M. B. Hastings
Center for Nonlinear Studies and Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA
(Received 6 June 2007; published 13 September 2007)

We show that randomly choosing the matrices in a completely positive map from the unitary group gives a quantum expander. We consider Hermitian and non-Hermitian cases, and we provide asymptotically tight bounds in the Hermitian case on the typical value of the second largest eigenvalue. The key idea is the use of Schwinger-Dyson equations from lattice gauge theory to efficiently compute averages over the unitary group.

DOI: 10.1103/PhysRevA.76.032315 PACS number(s): 03.67.−a

Recently, two papers [1,2] introduced the idea of *expander maps*: quantum analogs of expander graphs. An expander graph [4] may be defined in several ways. One is the property of having a large number of vertices, a small coordination number for each vertex, and also having a gap in the spectrum of the diffusion equation on the graph, so that a particle classically diffusing on an expander graph rapidly loses memory of where it started. In the quantum case, we replace the random process of diffusion by a completely positive, trace preserving map $\mathcal{E}(M)$. We define a quantum expander to be such a map from the space of N-by-N matrices M to the same space with the following properties. First, N is large, in analogy to the large number of vertices. Second, the map has eigenvalues $\lambda_1,\lambda_2,\lambda_3,\ldots,\lambda_{N^2}$, with $\lambda_1=1$ and $|\lambda_a|\leq 1-\delta$ for all $a>1$ so that the eigenvalue spectrum has a gap. Finally, the map can be written as

$$\mathcal{E}(M)=\sum_{s=1}^{D} A^{\dagger}(s)MA(s) \tag{1}$$

for some relatively small value of D, with $\Sigma_{s=1}^{D}A(s)A^{\dagger}(s)=\mathbb{1}$ so that the map is trace preserving, and with $\Sigma_{s=1}^{D}A^{\dagger}(s)A(s)=\mathbb{1}$, so that the eigenvector corresponding to eigenvalue λ_1 is $(1/\sqrt{N})\mathbb{1}$. Here, $\mathbb{1}$ is the N-by-N unit matrix. This requirement of small D is in analogy to the low coordination number.

These maps were applied in [1] to construct many-body states in one dimension with the property of having a short correlation length (this corresponds to the gap δ in the spectrum of eigenvalues of $\mathcal{E}$), small Hilbert space dimension on each site (this corresponds to the small D), and yet large entanglement entropy (this corresponds to the large entropy of the eigenvector of $\mathcal{E}$ with unit eigenvalue). Since expander graphs have a large number of applications in problems dealing with classical statistics, such as in error-correcting codes [5], derandomization, and the PCP theorem [6], to name a few, it seems worth further exploring the quantum case.

A number of possible forms of an expander map are possible: in [2] an expander was defined as having

$$A(s)=\frac{1}{\sqrt{D}}U(s), \tag{2}$$

for some unitary matrices $U(s)$, and in fact this is the form of $A(s)$ considered in this paper. However, the more general definition with arbitrary $A(s)$ constrained by $\Sigma_{s=1}^{D}A^{\dagger}(s)A(s)=\mathbb{1}$, $\Sigma_{s=1}^{D}A(s)A^{\dagger}(s)=\mathbb{1}$ seems useful also; in fact, although we do not consider it in this paper, it may be useful to weaken this constraint further, and explore the properties of completely positive, trace preserving maps, with no other constraint on the A, requiring only that the entropy of the density matrix, which is the eigenvector with unit eigenvalue, is large [3].

Our goal is to try to find families of maps with arbitrarily large N, such that the gap δ is bounded below by some N-independent constant and such that D does not grow too rapidly with N. The first paper [1] provided an explicit construction of such a family of maps with D of order $\ln(N)$ and provided numerical evidence for an alternate construction with D *independent* of N. The second paper [2] gave yet a different construction with D of order $\ln(N)$ but also provided a construction that had D independent of N and succeeded in proving an N-independent lower bound on the gap δ in this case.

Experience with expander graphs suggests that, while finding deterministic constructions of them is difficult [7], with high probability a *random* graph of fixed coordination number greater than 2 is an expander [8]. Thus, the natural question is to investigate whether Eqs. (1) and (2) will give an expander map if the matrices U are chosen randomly from the unitary group $U(N)$ using the Haar measure. We consider two cases. In the first case, the map $\mathcal{E}$ is non-Hermitian and the D matrices are chosen independently at random. In the second case the matrices $U(s)$ are chosen independently at random for $s=1,\ldots,D/2$ and we pick $U(s+D/2)=U(s)^{\dagger}$. In this case, D is even, and the map $\mathcal{E}$ is Hermitian and has real eigenvalues. In this paper we begin in generality with the non-Hermitian case, but then restrict to the Hermitian case for simplicity of notation.

In the Hermitian case, we consider $D\geq 4$, while in the non-Hermitian case we consider $D\geq 2$, as otherwise we would clearly not have an expander. Let λ_2 be the eigenvalue with the second largest absolute value of all eigenvalues other than λ_1. Let

$$\lambda_H=\frac{2\sqrt{D-1}}{D}. \tag{3}$$

The main result of this paper is that, in the Hermitian case, for any $\epsilon>0$ the probability that $|\lambda_2|$ is within ϵ of λ_H approaches unity as $N\to\infty$. Interestingly, this is the same as the recently proven tight bound [9] in the classical case, but the proof in the quantum case is much simpler.

 ©2007 The American Physical Society

The proof is based on a version of the trace method. We begin by introducing the trace method and describing its application to the Hermitian and non-Hermitian cases. We then give lower bounds on $|\lambda_2|$ based on the return probability of a random walk on a Cayley tree and discuss some numerical results. We next introduce a set of Schwinger-Dyson equations, analogous to those used in lattice gauge theory [10]. This is the key step, which enables us to take averages over the unitary group efficiently. We will use these equations to develop a *convergent* perturbation theory in $1/N$ for various traces of unitary matrices, and bound the correction terms in this perturbation theory. We start with a loose bound, giving a loose bound on $|\lambda_2|$, and then tighten to get the tight bound above. Finally, in the Appendix we discuss a related problem of "quantum edge expanders," which gives an analog in the quantum case of the combinatorial definition of a classical expander graph.

The space of N-by-N complex matrices M has a natural inner product: $(M,N)=\mathrm{tr}(M^\dagger N)$. With respect to this inner product, an orthonormal basis of matrices consists of the matrices $M(i,j)$, defined to have a 1 in the ith row and jth column, and zeroes everywhere else. Given this inner product, we can consider the space of N-by-N matrices as an N^2-dimensional vector space, with $\mathcal{E}$ acting as a linear operator on this space. Then, in the Hermitian case, it is possible to find a linear operator V, which is unitary with respect to this inner product, such that $\mathcal{E}=V^\dagger \Lambda V$, with Λ being a diagonal matrix with entries λ_a. Note that here $\mathcal{E}$, V, and Λ are all N^2-by-N^2 dimensional matrices. In the non-Hermitian case, we can write $\mathcal{E}=V^\dagger T V$, with T an upper triangular matrix whose diagonal entries are the eigenvalues λ_a. Thus,

$$\sum_{i,j}(\mathcal{E}^m(M(i,j)),\mathcal{E}^m(M(i,j)))=\sum_{i,j}(T^m(M(i,j)),T^m(M(i,j)))\geq\sum_{a=1}^{N^2}|\lambda_a|^{2m},\tag{4}$$

where $\mathcal{E}^m(M)$ denotes acting with the map $\mathcal{E}$ successively m times on M, and similarly for $T^m(M)$. In the case where $\mathcal{E}$ is Hermitian, Eq. (4) is an equality.

To simplify notation, we now restrict to the Hermitian case. In this case, Eq. (4) can be replaced by

$$\sum_{i,j}(M(i,j),\mathcal{E}^m(M(i,j)))=\sum_{a=1}^{N^2}|\lambda_a|^m\geq 1+|\lambda_2|^m,\tag{5}$$

where we pick m to be an even integer. Then,

$$\begin{aligned}\sum_{a=1}^{N^2}|\lambda_a|^m&=\sum_{i,j}(M(i,j),\mathcal{E}^m(M(i,j)))\\&=\left(\frac{1}{D}\right)^m\sum_{s_1=1}^{D}\sum_{s_2=1}^{D}\cdots\sum_{s_m=1}^{D}\mathrm{tr}[U(s_m+D/2)\cdots U(s_2+D/2)\\&\quad\times U(s_1+D/2)]\mathrm{tr}[U(s_1)U(s_2)\cdots U(s_m)].\end{aligned}\tag{6}$$

For notational convenience, we identify s_i+D with s_i throughout this paper, so that s_i is a periodic variable with period D.

Averaging $U(1),\ldots,U(D)$ over the unitary group we find that $E[\sum_{i,j}(M(i,j),\mathcal{E}^m(M(i,j)))]=E[\sum_{a=1}^{N^2}|\lambda_a|^m]$, where $E[\cdots]$ denotes the given average. Averaging Eq. (6) we find

$$E_1\equiv\left(\frac{1}{D}\right)^m\sum_{s_1=1}^{D}\sum_{s_2=1}^{D}\cdots\sum_{s_m=1}^{D}E_0(s_1,\ldots,s_m)=E\left[\sum_{a=1}^{N^2}|\lambda_a|^m\right],\tag{7}$$

where

$$\begin{aligned}&E_0(s_1,\ldots,s_m)\\&\quad\equiv E[\mathrm{tr}[U^\dagger(s_m)\cdots U^\dagger(s_2)U^\dagger(s_1)]\mathrm{tr}[U(s_1)U(s_2)\cdots U(s_m)]]\\&\quad=E[\mathrm{tr}[U(s_m+D/2)\cdots U(S_2+D/2)U(S_1+D/2)]\\&\qquad\times\mathrm{tr}[U(S_1)U(S_2)\cdots U(s_m)]].\end{aligned}\tag{8}$$

I. LOWER BOUNDS AND NUMERICAL RESULTS

In this section we present lower bounds on $|\lambda_2|$ based on random walks on a Cayley tree and then provide some numerical results. In the Hermitian case, it is possible, for certain choices of $s_1,\ldots,s_m$ in either Eq. (7) or Eq. (6), that the trace $\mathrm{tr}[U(s_1)U(s_2)\cdots U(s_m)]$ can be reduced to a trivial trace of the identity matrix by canceling successive appearances of $U(s)U(s+D/2)$ and replacing them with $\mathbb{1}$. The contribution of such choices to E_1 is proportional to a return probability of a random walk on a Cayley tree as will be seen.

We begin with an upper bound on the number of such choices: consider the unitaries $U(s_1)U(s_2)\cdots U(s_k)$ for some k, $0\leq k\leq m$. After making all possible cancellations of successive terms, $U(s)U(s+D/2)$, this sequence of unitaries may be reduced to another sequence of unitaries $U(s_1'(k))U(s_2'(k))\cdots U(s_{l(k)}'(k))$, for some $l(k)\leq k$. Then, consider the sequences of unitaries $U(s_1)U(s_2)\cdots U(s_{k+1})$. After making the same cancellations, and then possibly canceling $U(s_{k+1})$ against $U(s_{l(k)}'(k))$, we find a new sequence of unitaries, $U(s_1'(k+1))U(s_2'(k+1))\cdots U(s_{l(k+1)}'(k+1))$ with $l(k+1)=l(k)\pm 1$ and $s_j'(k+1)=s_j'(k)$ for $j<l(k)$. When $l(k+1)=l(k)-1$, then s_{k+1} is determined by s_k. When $l(k+1)=l(k)+1$, then if $l(k)>0$ there are $D-1$ possible values of s_{k+1}, while if $l(k)=0$ there are D possible values. Note that $l(k)\geq 0$ for all k. We define $N(l(m),m)$ to be the number of choices of $s_1,s_2,\ldots,s_m$, which give rise to the given $l(m)$. This is precisely the number of random walks of length m, on a tree with D daughters at the root and $D-1$ daughters for every other node, that end at a distance $l(m)$ from the root. Note that $N(0,m)$ is equal to D^m times the return probability of a random walk of length m on the Cayley tree.

An upper bound on $N(0,m)$ is given by

$$N(0,m)\leq(D-1)^{m/2}\frac{m!}{(m/2)!(m/2)!}\leq(D-1)^{m/2}2^m.\tag{9}$$

To show Eq. (9), we consider a related problem: consider sequences of $l(k)$ in which $l(k)$ may become *negative*, while the number of choices of s_m is considered to be $D-1$ whenever $l(k+1)=l(k)+1$, and the number of choices is consid-

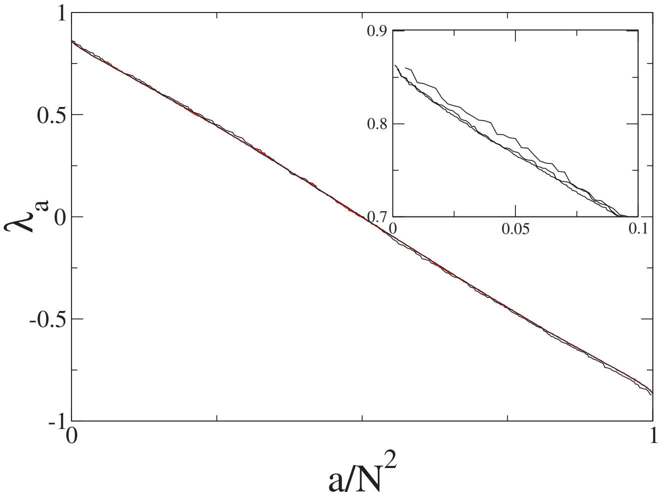

FIG. 1. (Color online) Eigenvalues from numerical diagonalization of a completely positive map based on the construction in [1] using expander graphs, for $N=20,30,50$. The eigenvalue with eigenvalue unity is not shown. The second largest eigenvalue is at roughly $\sqrt{3}/2$. Only a single realization is shown for each N. The inset shows a detail of the behavior at small a. Curves in the inset are $N=20,30,50$ from top to bottom; the curves in the main figure are not distinguishable.

ered to be 1 whenever $l(k)=l(k)-1$. This gives an overestimate of the number of sequences, and gives the value in Eq. (9).

On the other hand, a lower bound on $N(0,m)$ is given by assuming that if $l(k+1)=l(k)+1$ there are only $D-1$ possible choices of s_{k+1}, regardless of $l(k)$, in which case we find that

$$N(0,m) \geq c(2\sqrt{D-1})^m/(m+1)^{3/2}, \tag{10}$$

for some constant c of order unity. These bounds, Eqs. (9) and (10), are completely standard bounds [4], and we only repeat their derivation for completeness.

We now use Eq. (10) to get a lower bound on $|\lambda_2|$ for any completely positive map where the matrices $A(s)$ are given by Eq. (2). We emphasize that, while the upper bounds elsewhere in this paper are upper bounds on the typical behavior of $|\lambda_2|$, the present result is valid for any such map given by Eqs. (1) and (2), and is a quantum analog of the Alon-Boppana bound [11]. First, $\sum_{a=1}^{N^2}|\lambda_a|^m = 1+\sum_{a=2}^{N^2}|\lambda_a|^m \leq 1+(N^2-1)|\lambda_2|^m < 1+N^2|\lambda_2|^m$. Note that the product of traces in Eq. (6) is equal to $|\mathrm{tr}[U(s_1)U(s_2)\cdots U(s_m)]|^2$ and so is positive for all choices of $s_1,\ldots,s_m$. If $l(m)=0$, then $|\mathrm{tr}[U(s_1)U(s_2)\cdots U(s_m)]|^2=N^2$, and so the contribution of terms with $l(m)=0$ to the sum in Eq. (6) is equal to $N^2N(0,m)/D^m$. Therefore,

$$1+N^2|\lambda_2|^m \geq \sum_{a=1}^{N^2}|\lambda_a|^m \geq N^2\frac{N(0,m)}{D^m} \geq cN^2\lambda_H^m/m^{3/2}. \tag{11}$$

Thus, $|\lambda_2| \geq \lambda_H(c/m^{3/2})^{1/m}[1-m^{3/2}/(c\lambda_H^m N^2)]^{1/m}$. Picking $m=\{\ln(cN^2/2)-(3/2)\ln[\ln(cN^2/2)]\}/\ln(1/\lambda_H)$, we find

$$|\lambda_2| \geq \lambda_H[1-O(\ln(\ln(N))/\ln(N))]. \tag{12}$$

A very interesting question is to see whether a bound such as Eq. (12) still holds for arbitrary trace preserving, completely positive Hermitian maps $\mathcal{E}(M)$. As a partial step toward this more general bound, note that the bound of Eq. (12) can be readily generalized to the following case: let $A(s)=\sqrt{P(s)}U(s)$, with the numbers $P(s)$ obeying $\sum_{s=1}^{D}P(s)=1$, and with $U(s)=U(s+D/2)^\dagger$ and $P(s)=P(s+D/2)$. Equation (2) is a special case of this with $P(s)=1/D$.

As stated before, the main result of this paper is that, for any ϵ, when the unitary matrices are chosen randomly, the probability that $|\lambda_2|$ is within ϵ of λ_H approaches unity when N becomes large. Interestingly, this seems to be true in more generality than just for unitary matrices chosen with the Haar measure. Using the construction in [1], in which we pick a random graph with constant coordination number and derive unitary matrices from that graph and from certain random phases, numerical studies also show that $|\lambda_2|$ is close to λ_H. We show in Fig. 1 the results of numerical diagonalization of systems of size $N=20,30,50$, so that there are 400, 900, and 2500 eigenvalues, respectively. The second largest eigenvalue is indeed very close to $\sqrt{3}/2$. After sorting the eigenvalues by λ_a, from most positive to most negative, we plotted the eigenvalues as a function of a/N^2: the scaling collapse of the curves is extremely good.

II. BOUNDS ON EIGENVALUES

A. Schwinger-Dyson equations

We will develop a perturbation theory in $1/N$ to estimate the average (7), which is a product of two traces. To do this, we will develop general machinery for computing the average over the unitary group of products of an arbitrary number of traces. Consider such a product of the form

$$E[L_1L_2\cdots L_k], \tag{13}$$

where

$$L_1=\mathrm{tr}[U(s_{1,1})U(s_{1,2})\cdots U(s_{1,m_1})],$$

$$L_2=\mathrm{tr}[U(s_{2,1})U(s_{2,2})\cdots U(s_{2,m_2})],\ \ldots\ . \tag{14}$$

Here we have an average of k traces, $L_1,\ldots L_k$, each of which is a product of m_k unitary matrices. We now present the Schwinger-Dyson equations.

Let T^a, for $a=1,\ldots,N^2$, be Hermitian matrices such that

$$\sum_{a=1}^{N^2} T^a_{\mu\nu}T^a_{\rho\sigma}=\delta_{\mu\sigma}\delta_{\nu\rho}. \tag{15}$$

Then

$$\sum_{a=1}^{N^2}(T^aT^a)_{\mu\nu}=N\delta_{\mu\nu}. \tag{16}$$

To compute the average in Eq. (13), we begin with

$$E[\mathrm{tr}[T^a U(s_{1,1})U(s_{1,2})\cdots U(s_{1,m_1})]L_2\cdots L_k]. \qquad (17)$$

We then use the invariance of the average over the unitary group under an infinitesimal change in variables as follows:

$$U(s_1)\rightarrow(1+i\epsilon T^a)U(s_1),$$

$$U(s_1+D/2)\rightarrow U(s_1+D/2)(1-i\epsilon T^a), \qquad (18)$$

where we recall that $U(s+D/2)=U(s)^\dagger$. Applying the change in variables given in Eq. (18) to Eq. (17), and summing over a and dividing by N, we find that

$$\begin{aligned}
&E[\mathrm{tr}[U(s_{1,1})U(s_{1,2})\cdots U(s_{1,m_1})]L_2\cdots L_k]\\
&\quad=-\frac{1}{N}\sum_{j=2}^{m_1}\delta_{s_{1,1},s_{1,j}}E[\mathrm{tr}[U(s_{1,1})\cdots U(s_{1,j-1})]\mathrm{tr}[U(s_{1,j})\cdots U(s_{1,m_1})]L_2\cdots L_k]\\
&\qquad+\frac{1}{N}\sum_{j=2}^{m_1}\delta_{s_{1,1},s_{1,j+D/2}}E[\mathrm{tr}[U(s_{1,1})\cdots U(s_{1,j})]\mathrm{tr}[U(s_{j+1,1})\cdots U(s_{1,m_1})]L_2\cdots L_k]\\
&\qquad-\frac{1}{N}\sum_{l=2}^{k}\sum_{j=1}^{m_l}\delta_{s_{1,1},s_{l,j}}E[\mathrm{tr}[U(s_{1,1})\cdots U(s_{1,m_1})U(s_{l,j})U(s_{l,j+1})\cdots U(s_{l,j-1})]L_2\cdots L_{l-1}L_{l+1}\cdots L_k]\\
&\qquad+\frac{1}{N}\sum_{l=2}^{k}\sum_{j=1}^{m_l}\delta_{s_{1,1},s_{l,j}+D/2}E[\mathrm{tr}[U(s_{1,1})\cdots U(s_{1,m_1})U(s_{l,j+1})U(s_{l,j+2})\cdots U(s_{l,j-1})U(s_{l,j})]L_2\cdots L_{l-1}L_{l+1}\cdots L_k].
\end{aligned} \qquad (19)$$

We simplify the second and fourth lines after the equality sign of the above equation using $U(s)U(s+D/2)=\mathbb{1}$ to get

$$\begin{aligned}
&E[\mathrm{tr}[U(s_{1,1})U(s_{1,2})\cdots U(s_{1,m_1})]L_2\cdots L_k]\\
&\quad=-\frac{1}{N}\sum_{j=2}^{m_1}\delta_{s_{1,1},s_{1,j}}E[\mathrm{tr}[U(s_{1,1})\cdots U(s_{1,j-1})]\mathrm{tr}[U(s_{1,j})\cdots U(s_{1,m_1})]L_2\cdots L_k]\\
&\qquad+\frac{1}{N}\sum_{j=2}^{m_1}\delta_{s_{1,1},s_{1,j+D/2}}E[\mathrm{tr}[U(s_{1,2})\cdots U(s_{1,j-1})]\mathrm{tr}[U(s_{j+1,1})\cdots U(s_{1,m_1})]L_2\cdots L_k]\\
&\qquad-\frac{1}{N}\sum_{l=2}^{k}\sum_{j=1}^{m_l}\delta_{s_{1,1},s_{l,j}}E[\mathrm{tr}[U(s_{1,1})\cdots U(s_{1,m_1})U(s_{l,j})U(s_{l,j+1})\cdots U(s_{l,j-1})]L_2\cdots L_{l-1}L_{l+1}\cdots L_k]\\
&\qquad+\frac{1}{N}\sum_{l=2}^{k}\sum_{j=1}^{m_l}\delta_{s_{1,1},s_{l,j}+D/2}E[\mathrm{tr}[U(s_{1,2})\cdots U(s_{1,m_1})U(s_{l,j+1})U(s_{l,j+2})\cdots U(s_{l,j-1})]L_2\cdots L_{l-1}L_{l+1}\cdots L_k].
\end{aligned} \qquad (20)$$

These Schwinger-Dyson equations are quite long when written out, but in reality are quite simple. Let us apply them to compute the average of $\mathrm{tr}(U)\mathrm{tr}(U^\dagger)$ over unitary matrices U. We find after one iteration of Eq. (20) that this is equal to $(1/N)\mathrm{tr}(\mathbb{1})=1$. Now consider a more complicated example, to compute the average of $\mathrm{tr}(UU)\mathrm{tr}(U^\dagger U^\dagger)$ over unitary matrices U. Then, the Schwinger-Dyson equations give after the first iteration, $(1/N)(-E[\mathrm{tr}(U)\mathrm{tr}(U)\mathrm{tr}(U^\dagger U^\dagger)]+2E[\mathrm{tr}(UU^\dagger)]=-(1/N)E[\mathrm{tr}(U)\mathrm{tr}(U)\mathrm{tr}(U^\dagger U^\dagger)]+2$. We then apply the equations again to the average $E[\mathrm{tr}(U)\mathrm{tr}(U)\mathrm{tr}(U^\dagger U^\dagger)]$, giving $E[\mathrm{tr}(U)\mathrm{tr}(U)\mathrm{tr}(U^\dagger U^\dagger)]=(1/N)(-E[\mathrm{tr}(UU)\mathrm{tr}(U^\dagger U^\dagger)]+2E[\mathrm{tr}(U)\mathrm{tr}(U^\dagger)])$. Since we have already worked out $E[\mathrm{tr}(U)\mathrm{tr}(U^\dagger)]=1$, we have $E[\mathrm{tr}(U)\mathrm{tr}(U)\mathrm{tr}(U^\dagger U^\dagger)]=-(1/N)E[\mathrm{tr}(UU)\mathrm{tr}(U^\dagger U^\dagger)]+2$. Thus, putting it all together, we find that

$$E[\mathrm{tr}(UU)\mathrm{tr}(U^\dagger U^\dagger)]=2+(1/N^2)E[\mathrm{tr}(UU)\mathrm{tr}(U^\dagger U^\dagger)]-(2/N^2), \qquad (21)$$

and hence $E[\mathrm{tr}(UU)\mathrm{tr}(U^\dagger U^\dagger)]=2$.

We now describe the general algorithm for reducing traces of the form (13). We initially cancel all pairs of matrices $U(s)U(s+D/2)$ appearing successively in the same trace, replacing them with $\mathbb{1}$. We then apply Eq. (20). Then, we cancel all pairs of matrices $U(s)U(s+D/2)$ appearing successively in the resulting traces, replace the trace $\mathrm{tr}(\mathbb{1})$ by N, and repeat this procedure for each term. After the first application of Eq. (20), the number of terms on the right-hand side will be at most $m_{\mathrm{total}}-1$, where $m_{\mathrm{total}}\equiv m_1+m_2+\cdots+m_k$. Applying the equations repeatedly will generate more and more terms at each application. We regard this as a branching process: each term on the right-hand side can be then fed back into the left-hand side of the equation to generate new terms on the right-hand side. Note that at every stage, each term will produce at most $m_{\text{total}}-1$ terms on the right-hand side since the total number of unitary matrices, which appear in the traces $m_1+m_2+\cdots+m_k$, will always be at most m_{total}. If the number of unitary matrices becomes equal to zero in some term after n iterations of Eq. (20), then we are left with only trivial traces and we say that this term "terminates" at level n.

This algorithm generates an infinite series, where the nth term in the series is equal to the sum of all terms terminating at the nth level. We claim (and will show later when we discuss the convergence of the series) that, if $m_{\text{total}} \leq N$, then this series is absolutely convergent and also the average of the original trace is equal to the sum over all levels $n \geq 1$ of the terms which terminate at each level, so that the series converges to the desired answer.

This series is in fact an infinite series for many simple examples. In fact, for $E[\text{tr}(UU)\text{tr}(U^\dagger U^\dagger)]$ we find that after two repetitions of the process, the same average $E[\text{tr}(UU)\text{tr}(U^\dagger U^\dagger)]$ has reappeared, as can be seen on the right-hand side of Eq. (21), and thus the algorithm above does not ever finish because there are always some terms with nontrivial traces. In this particular case, however, although the algorithm does not ever finish, the sum of the terms terminating at any level $n>1$ is equal to zero; in other cases [12] this is not true and the given series has an infinite number of nonvanishing coefficients. We will later see how this infinite series is related to an infinite series in $1/N$ for the given trace.

We will apply this procedure to the trace $E_0=E[\text{tr}[U(s_m+D/2)\cdots U(s_2+D/2)U(s_1+D/2)]\text{tr}[U(s_1)U(s_2)\cdots U(s_m)]]$. Thus, $L_1=\text{tr}[U(s_m+D/2)\cdots U(s_2+D/2)U(s_1+D/2)]$, and $L_2=\text{tr}[U(s_1)U(s_2)\cdots U(s_m)]$. Begin by reducing all successive pairs of a unitary matrix followed by its Hermitian conjugate. What is left is two traces L_1, L_2 such that $m_1=m_2$ and $s_{1,i}=s_{2,m_2+1-i}+D/2$. We will proceed by estimating the probability of different values of m_1 given a random choice of $s_1,\ldots,s_m$, and then estimating the behavior of E_0 for the given $m_1=m_2$.

B. Length of the reduced trace

In this subsection, we will estimate the number of choices of $s_1,\ldots,s_m$ such that the reduced traces L_1, L_2 have a given $m_1=m_2$.

We start with the case $m_1=m_2=0$ in which case $E_0=N^2$. The number of different choices of $s_1,\ldots,s_m$ with $m_1=m_2=0$ is given by Eq. (9) so the contribution of all such choices to E_1 is bounded by

$$N^2 D^{-m}(D-1)^{m/2}2^m = N^2\lambda_H^m. \tag{22}$$

We can also bound the number of choices of $s_1,\ldots,s_m$, which give a given $m_1>0$. In this case, $l(m)=m_1$ and $s'_j(m)=s_{1,j}$. Using the same argument as given for Eq. (9), the number of such choices is bounded by

$$(D-1)^{m_1/2}(D-1)^{m/2}2^m. \tag{23}$$

This number is independent of the particular values of $s_{1,1},\ldots,s_{1,m_1}$. There are $[D/(D-1)](D-1)^{m_1}$ different possible values of $s_{1,1},\ldots,s_{1,m_1}$ and therefore the total number of choices of $s_1,\ldots,s_m$, which give rise to a given choice of $s_{1,1},\ldots,s_{1,m_1}$, is bounded by

$$\frac{D-1}{D}\left(\frac{1}{\sqrt{D-1}}\right)^{m_1}(D-1)^{m/2}2^m. \tag{24}$$

C. Nontrivial words

We now consider the case $m_1>0$. After the first application of Eq. (20), the term on the fourth line with $l=2$ and $j=m$ reduces the trace to $(1/N)E[\text{tr}[U(s_{1,2})\cdots U(s_{1,m_1})U(s_{2,1})\cdots U(s_{2,m_1-1})]]=(1/N)E[\mathbb{1}]=1$, so that the series terminates at level $n=1$. There may also be other terms, which reduce the trace to a trivial one after a single application of Eq. (20) if the sequence of values $s_{1,1}s_{1,2}\cdots s_{1,m_1}$ has a symmetry under a shift: $s_{j,1}=s_{j+m_1/o,1}$ for some $o>1$, which we refer to as the period of the shift. Here, we treat the index j as periodic with period m_1. For example, the problem studied in Eq. (21) has such a symmetry under a shift with $o=2$. In the event that there is such a shift symmetry, then the sum of terms terminating at level $n=1$ is equal to o. For a given m_1, the number of choices of $s_{1,1},\ldots,s_{1,m_1}$, which have a shift symmetry with period o, is bounded by $[D/(D-1)](D-1)^{m_1/o}$. Thus, from Eq. (24), the total number of choices of $s_1,\ldots,s_m$, which give rise to a given m_1, o, is bounded by $(D-1)^{m_1/o}(\frac{1}{\sqrt{D-1}})^{m_1}(D-1)^{m/2}2^m$. Thus, the contribution to E_1 of terms terminating at level $n=1$ is bounded by

$$1+\sum_{m_1\leq m}\sum_{o=2}^{m_1} o(D-1)^{m_1/o}\left(\frac{1}{\sqrt{D-1}}\right)^{m_1}\lambda_H^m. \tag{25}$$

The term in Eq. (25) with $o=2$ is is bounded by $2m\lambda_H^m$, while the sum of terms in Eq. (25) with $o>2$ is bounded by $(D-1)^{-1/6}/[1-(D-1)^{-1/6}]^2\lambda_H^m$, which for $D\geq 4$ is bounded by $30\lambda_H^m$ so that Eq. (25) is bounded by

$$1+2m\lambda_H^m+30\lambda_H^m. \tag{26}$$

We will now bound the sum of all terms terminating at level $n>1$. Assuming that the sequence $s_{1,2}\cdots s_{1,m_1}$ lacks the shift symmetry discussed above, this is the only term that terminates at level 1, and the other terms that appear after the first iteration do not terminate and continue to branch, but some of their descendents will terminate at lower levels.

We can estimate the value of a term that terminates at a given level $n>1$ as follows. First, there is a sign equal to plus or minus 1. Next, there is a factor of $(1/N)^n$. Finally, there is a factor of N for each trace of the form $\text{tr}(\mathbb{1})$ that appeared in this process. Suppose there are p such traces, giving a factor of N^p. How big can p be? Initially we have $k=2$ different traces. The given term at level n arose from a specific choice of terms on the right-hand side of Eq. (20) on the first iteration. This specific choice has k_1 different traces in it, with k_1 equal to either 1 or 3. After the second iteration there are k_2 traces, then k_3, and so on. The number of traces $k_2, k_3, \ldots$ can be determined as follows: an application of Eq. (20) may increase the number of traces by one if the term arises from the first or second line on the right-hand side, or may decrease the number of traces by one if the term arises from the third or fourth line on the right-hand side of Eq. (20). Next, some of the traces may be trivial. In the event that the term arose from the first, second, or third line of Eq. (20) it is not possible for any of the traces to be trivial, under the assumption that any repetitions of the form $U(s)U(s+D/2)$ have been replaced by $\mathbb{1}$ in the trace. However, in the event that the term arose from the fourth line, then it is possible for one of the traces to be trivial, increasing p by one. Thus, for each $b \leq n$, $k_b - k_{b-1}$ is equal to either $+1$, -1, or -2. Let q be equal to the number of times the first or second line was used from Eq. (20) and $n-q$ equal the number of times the third or fourth line was used. Then, in order for all traces to be trivial in this particular term resulting from n iterations of Eq. (20),

$$2 + q - (n - q) - p = 0. \tag{27}$$

Also, since p can only increase when a term from the fourth line is used,

$$p \leq n - q. \tag{28}$$

Thus,

$$p \leq \lfloor (2 + n)/3 \rfloor. \tag{29}$$

Therefore, the value of a term terminating at the nth level, $n > 0$, is bounded in absolute value by

$$N^{\lfloor (2+n)/3 \rfloor - n}. \tag{30}$$

The number of terms terminating at the nth level is bounded by

$$(2m - 1)^n. \tag{31}$$

Note also that there are no terms terminating at level $n=2$: if the term does not terminate at level 1, then there are either 1 or 3 traces after the first iteration of Eq. (20), and then there is no way to have the term terminate at level 2. Thus, the sum of terms terminating at level $n > 1$ is bounded in absolute value by

$$\begin{aligned}(2m)^3 N^{-2} &+ (2m)^4 N^{-2} + (2m)^5 N^{-3} \\ &+ (2m)^6 N^{-4} + (2m)^7 N^{-5} + \cdots \\ &\leq 8m^3 N^{-2} + 16 m^4 N^{-2}(1 + 2mN^{-1} + 4m^2 N^{-2}) \frac{1}{1 - 8m^3 N^{-2}}.\end{aligned} \tag{32}$$

D. Convergence of series

We now show the claim that, for $m_{\text{total}} \equiv m_1 + m_2 + \cdots + m_k \leq N$, the average $E[L_1 L_2 \cdots L_k]$ is indeed equal to the sum over all levels $n \geq 1$ of the number of terms terminating at each level and that the series is absolutely convergent. After n iterations of Eq. (20) some of the terms have terminated. There are at most $(m_{\text{total}} - 1)^n$ terms, which have not terminated, since there are at most $(m_{\text{total}} - 1)^n$ terms. Each of these terms is equal to plus or minus one times N^{-n} times N^{p_n}, where p_n is the number of times a trivial trace appeared in the process, times the average of a product of traces. There are at most $m_{\text{total}} - p_n$ different traces in the product since there were originally at most m_{total} unitary matrices. Thus, since each trace is bounded in absolute magnitude by N, the sum of all terms which have not terminated after n applications of Eq. (20), is bounded in absolute value by $(m_{\text{total}} - 1)^n N^{-n} N^{m_{\text{total}}}$, which converges to zero as $n \to \infty$ for $m_{\text{total}} \leq N$. Thus, the difference between the sum of the terms terminating at the first n levels and the actual value of the average $E[L_1 L_2 \cdots L_k]$ converges to zero as $n \to \infty$. The sum of all terms terminating at a given level is bounded in absolute value by the number of such terms, times $N^{-n} N^{p_n}$, and so is bounded by $(m_{\text{total}} - 1)^n N^{-n} N^{m_{\text{total}}}$ and so the series is absolutely convergent for $m_{\text{total}} \leq N$. This shows the desired claim.

E. Loose bound

Adding the results in Eqs. (22), (26), and (32), we find that for $2m < N$,

$$\begin{aligned}E_1 &\leq 1 + [N^2 + 2m + 30]\lambda_H^m + 8m^3 N^{-2} \\ &\quad + 16 m^4 N^{-2}(1 + 2mN^{-1} + 4m^2 N^{-2}) \frac{1}{1 - 8m^3 N^{-2}}.\end{aligned} \tag{33}$$

We now pick $m = \ln(N^4)/\ln(1/\lambda_H)$, so

$$E_1 \leq 1 + 16[1 + o(1)][\log(N^4)/\log(1/\lambda_H)]^4 N^{-2}, \tag{34}$$

where $o(1)$ denotes terms asymptotically tending to zero as $N \to \infty$. Thus, the average of $|\lambda_2|$ over the unitary group is bounded by

$$\begin{aligned}&\{16[1 + o(1)][\ln(N^4)/\ln(1/\lambda_H)]^4 N^{-2}\}^{\ln(1/\lambda_H)/\ln(N^4)} \\ &\quad = [1 + o(1)]\lambda_{\text{loose}}(D),\end{aligned} \tag{35}$$

where

$$\lambda_{\text{loose}}(D) \equiv \sqrt{\lambda_H} = \sqrt{\frac{2\sqrt{D-1}}{D}}. \tag{36}$$

Further, using Markov's inequality, the probability that $|\lambda_2|$ is greater than $c\lambda_{\text{loose}}(D)$, for any $c \geq 1$, is bounded by $[1 + o(1)]c^{-\ln(N^4)/\ln(1/\lambda_H)}$, so that for large N it is very rare for $|\lambda_2|$ to be significantly above the loose bound $\lambda_{\text{loose}}(D)$.

F. Tight bound

We now tighten the bound. On a given iteration of the Schwinger-Dyson equations, we go from a product of k traces to a product of $k+1$, $k-1$, or $k-2$ traces. We will keep track of how the matrices move under this iteration process using a function $f_n((l,i))$ from pairs of integers to pairs of integers. We say that the matrix $U(s_{l,i})$ in the given product of traces, $L_1 L_2 \cdots L_k$, is in position (l,i). Let us consider the case of a term on the first line, where m increases by one. Then, for any given j in the sum on the first line, we say that the matrix in position $(i,1)$, for $i<j$, on the $n+1$st iteration corresponds to the matrix in position $(1,i)$ on the nth iteration, and so $f_n((1,i))=(1,i)$, while the matrix in position $(2,i)$ on the $n+1$st iteration corresponds to the matrix in position $(1,i+j-1)$ on the nth iteration, so $f_n((1,i+j-1))=(2,i)$. The matrix in position (l,i), for $2<l\le k+1$, on the $n+1$st iteration corresponds to the matrix $(l-1,i)$ on the nth iteration, so $f_n(l-1,i)=(l,i)$. We follow a similar procedure for the other lines of Eq. (20) and if there are cancellations, we keep track of how the matrix moves under the cancellations.

We then keep track of which matrix after n iterations corresponds to a given matrix before any iterations, by defining $F_n((l,i))=f_n(f_{n-1}(\cdots f_1((l,i))))$ for $l=1,2$. Let us say that the matrix at position (l,i) for $l=1,2$ is "trivially moved" under the nth iteration of the Schwinger-Dyson equations if we are considering a term in the equations, which did not arise from $T^a U((s_{l,i}))$; that is, the matrix is trivially moved if it is not in position $(1,j)$ using a term on the first or second line, or in position (l,j) using a position from the third or fourth line, or in position (1,1). Let us define a "rung cancellation of matrix i" to be the case in which, for some n, after the nth iteration of the Schwinger-Dyson equation we perform a series of cancellations such that the following hold [13]. First, a matrix in position (l,j) is canceled against a matrix in position (l',j') such that $(l,j)=F_{n-1}((1,i))$ and $(l',j')=F_{n-1}((2,m_1+1-i))$. Second, at all previous iterations up to the $n-1$th iteration, the given matrix was trivially moved. If there is a rung cancellation of matrix 1 on the first iteration, then all matrices cancel and the trace is equal to unity; this is precisely the case with $l=2$, $j=m$ discussed at the start of the section "Nontrivial words." Note that the matrix in position $(l',j')=F_{n-1}((2,m_1+1-i))$ is equal to $U(s_{2,m_1+1-i})=U(s_i+D/2)$, which is why the matrix in position $(l,j)=F_{n-1}((1,i))$ can be canceled against this matrix.

We now make a stronger claim: for any given i, the sum of all terms with a rung cancellation of matrix i is equal to unity. To show this, consider the trace $\mathrm{tr}[U(s_m+D/2)\cdots U(s_{i+1}+D/2)X^\dagger U(s_{i-1}+D/2)\cdots U(s_1+D/2)]\mathrm{tr}[U(s_1)\cdots U(s_{i-i})XU(s_{i+1})\cdots U(s_m)]$, where X is some arbitrary unitary matrix. Averaging this trace over all unitary matrices $U(s)$ and over all unitary matrices X with the Haar measure, we find that the trace is equal to unity. However, applying the Schwinger-Dyson equations to this trace generates precisely the sum of terms mentioned above, those in which there is a rung cancellation of matrix i. Thus, this sum equals unity. We further claim that for any given $i_1,i_2,\ldots,i_d$, the sum of all terms with rung cancellations of matrices $i_1,i_2,\ldots,i_d$ is equal to unity, as may be shown by considering a trace in which matrices $U(s_{i_1}),U(s_{i_2}),\ldots$ are replaced by $X_1,X_2,\ldots$, and the trace is averaged over the different $X_1,X_2,\ldots$.

On the other hand, if a term terminates at level n and matrix i does not have a rung cancellation, then at some previous iteration n either the matrix was not trivially moved or was canceled against a matrix in position (l,j) such that $(l,j)=F_{n-1}(l',j')$ with $(l',j')\neq(2,m_1+1-i)$. In the latter case, for $l'=2$ we know that $s_{m+1-j'}+D/2=s_i$, while for $l'=1$ we know that $s_{j'}=s_i$, thus in both cases identifying some $k\neq i$ such as either $s_{1,i}=s_{1,k}$ or $s_{1,i}=s_{1,k}+D/2$. If the matrix was not trivially moved, we can also identify some $k\neq i$ with the same properties. Let us write $k=\tau(i)$ in both cases, for some function $\tau(i)$.

Now consider the sum of terms in which there is no rung cancellation for i of matrix i. By the inclusion-exclusion principle in combinatorics, this is equal to the sum of all terms, minus the sum over i of the sum of terms in which there is a rung cancellation of matrix i, plus one half the sum over $i_1\neq i_2$ of the sum of terms in which there are rung cancellations of matrices i_1,i_2, and so on. This is equal to the sum of all terms minus the sum

$$m_1 - m_1(m_1-1)/2 + m_1(m_1-1)(m_1-2)/6 - \cdots = 1. \tag{37}$$

Thus, the sum of all terms is equal to one plus the sum of terms in which for no i is there a rung cancellation of matrix i. So, we now focus on the sum of terms with no rung cancellation, which we define to be $E_0'(s_1,\ldots,s_m)$. If $s_1,\ldots,s_m$ has a shift symmetry as above, then there may be terms in this sum terminating at the first level; the sum of these terms is $o-1$.

Each E_0 we are averaging over the unitary group results from a particular set of choices of $s_1,\ldots,s_m$ in the sum in Eq. (7). There are D^m different terms in this sum in Eq. (7). We begin by bounding, for any given level n, the number of choices of $s_1,\ldots,s_m$, which give rise to an E_0, which produces a term in the Schwinger-Dyson equations, which terminates at level n with no rung cancellations. Suppose for a given choice of $s_1,\ldots,s_m$ there is such a term, which terminates at level n with no rung cancellations. There were two traces of m unitaries in the definition of E_0; then, after canceling successive pairs $U(s)U(s+D/2)$, we have $m_1\le m$ unitaries for some m_1. The number of different initial choices of $s_1,\ldots,s_D$, which produce a given m_1 after these cancellations, is given in Eq. (24).

Then we iterate the Schwinger-Dyson equations with a particular choice of l,j at each level, where $1\le l\le 2m_1$ and $1\le j\le 2m_1$ as given in Eq. (20); if we pick a term from the first or second line of the Schwinger-Dyson equations, we set $l=1$ at that level. At each such iteration of the Schwinger-Dyson equations, there may be cancellations in two different traces if the term came from the second line of Eq. (20), with at most m_1 cancellations in each trace, or cancellations in two different places of a single trace, if the term came from the fourth line of Eq. (20), with at most m_1 cancellations in each place. Let us call the number of cancellations c_1,c_2 with $0\le c_1\le m_1$ and $0\le c_2\le m_1$. Then, by specifying l, j, c_1, c_2 for each iteration, we succeed in fully specifying how the matrices move under the n iterations of the Schwinger-Dyson equation; this requires specifying $2n$ numbers ranging from $1\cdots 2m_1$, and $2n$ numbers ranging from $0\cdots m_1$. In particular, since there are no rung cancellations, we succeed in specifying for each i, $1\le i\le m_1$, some $j\neq i$ such that either

$s_{1,i}=s_{1,j}$ or $s_{1,i}=s_{1,j}+D/2$, giving the function $\tau(i)$. Having specified this function, there are now only at most $[D/(D-1)](D-1)^{m_1/2}$ possible values of $s_{1,1},\ldots,s_{1,m_1}$. To show this, we start by specifying the value of $s_{1,1}$, which can assume any of D different values. By specifying $s_{1,1}$ we have fixed the value of $s_{1,\tau(1)}$, as well as the value of any j such that $\tau(j)=1$, so that there are now at most m_1-2 different values of $s_{1,i}$, which remain undetermined. We then find the smallest j_1 such that s_{1,j_1} is undetermined and specify its value. Note that there are only $D-1$ possible values of this s_{1,j_1} since, by assumption, $s_{1,j_1}\neq s_{1,j_1-1}+D/2$. Having specified this s_{1,j_1}, we have fixed the value of $s_{1,\tau(j_1)}$ as well as the value of any j such that $\tau(j)=j_1$. We then find the smallest j_2 such that s_{1,j_2} is undetermined and specify that value. Proceeding in this way, we succeed in specifying $s_{1,1},\ldots,s_{1,m_1}$ by specifying one of at most $[D/(D-1)](D-1)^{m_1/2}$ different choices. Thus, there are at most

$$[D/(D-1)](D-1)^{m_1/2}(2m_1)^{2n}(m_1+1)^{2n} \leq [D/(D-1)](D-1)^{m_1/2}(2m_1)^{4n} \tag{38}$$

such choices of $s_1,\ldots,s_{m_1}$ which can produce a term that terminates at level n. Using Eq. (24), the number of choices of $s_1,\ldots,s_m$, which can produce a term that terminates at level n, is at most

$$\sum_{m_1=0}^{m}(D-1)^{m/2}2^m(2m_1)^{4n} \leq (D-1)^{m/2}2^m\frac{(2m+1)^{4n+1}}{4n+1}. \tag{39}$$

For any $s_1,\ldots,s_m$, we define $n_{\min}(s_1,\ldots,s_m)$ to be the smallest level at which a term terminates with no rung cancellations. We rewrite the sum in Eq. (7) as

$$E_1 = 1+\left(\frac{1}{D}\right)^m\sum_{n=0}^{\infty}\sum_{s_1=1}^{D}\sum_{s_2=1}^{D}\cdots\sum_{s_m=1}^{D}\delta_{n_{\min}(s_1,\ldots,s_m),n} \times E_0'(s_1,\ldots,s_m), \tag{40}$$

so that the second sum is over the set of all values of $s_1,\ldots,s_m$ with the given $n_{\min}=n$. We note that the bound of Eq. (30) continues to apply to the terms terminating with no rung cancellations, and the bound of Eq. (31) continues to bound the number of such terms terminating with no rung cancellations. From Eq. (30), a bound on the value of the term terminating at the nth level, for any $n\geq 0$, is

$$N^2N^{-(2/3)n}. \tag{41}$$

Therefore, for any $s_1,\ldots,s_m$,

$$E_0'(s_1,\ldots,s_m) \leq N^2\sum_{n\geq n_{\min}(s_1,\ldots,s_m)}N^{-(2/3)n}(2m-1)^n = N^2\frac{[N^{-2/3}(2m-1)]^{n_{\min}}}{1-N^{-2/3}(2m-1)}. \tag{42}$$

From Eqs. (39), (40), and (42),

$$E_1 \leq 1+N^2\lambda_H^m\sum_{n=0}^{\infty}\frac{(2m+1)^{4n+1}}{4n+1}\frac{[N^{-2/3}(2m-1)]^n}{1-N^{-2/3}(2m-1)} \leq 1+N^2\lambda_H^m\sum_{n=0}^{\infty}\frac{2m+1}{(4n+1)[1-N^{-2/3}(2m-1)]} \times[N^{-2/3}(2m+1)^5]^n. \tag{43}$$

We then pick $m=(1/4)N^{2/15}$, so that $N^{-2/3}(2m+1)^5\leq 1/2$ and

$$|\lambda_2| \leq (E_1-1)^{1/m} \leq N^{2/m}\lambda_H(1+O(1))^{1/m} = \lambda_H[1+O(\log(N))N^{-2/15}]. \tag{44}$$

As before, using Markov's inequality, the probability that $|\lambda_2|$ is greater than $c\lambda_H(D)$, for any $c\geq 1$, is bounded by $c^{-(1/4)N^{2/15}}[1+O(\log(N))N^{-2/15}]$.

This shows that for any ϵ, the probability that $\lambda_2\leq\lambda_H+\epsilon$ approaches unity as $N\to\infty$. Combined with the previous lower bound (12), this proves the main result.

III. DISCUSSION

We consider some analogies between these results and lattice gauge theory, some applications of these results, and some extensions. We begin with analogies between the random construction of quantum expanders and lattice gauge theory and the Eguchi-Kawai construction [10].

A. Gauge theory analogies

Consider a lattice gauge $U(N)$ theory in $D/2$ dimensions on a hypercubic lattice, with unitary matrices $U_{\hat{d}}(x)$ defined for each link of the lattice. Here, x represents a point on the lattice, and $\hat{d}$ represents the direction of the link: if $d\leq D/2$, it points in the direction of increasing the dth coordinate by unity, while if $D/2<d\leq D$, then it points in the direction of decreasing $d-D/2$th coordinate by unity. Then, for a given choice of $s_1,\ldots,s_m$ we can define a path, starting at the origin, and then moving in direction $\hat{s}_1,\hat{s}_2,\ldots$ until m steps have been taken. We can define a product of traces associated with this path: $\mathrm{tr}[U_{s_1}(0)U_{s_2}(0+\hat{s}_1)\ldots]\mathrm{tr}[\ldots U_{s_2}(0+\hat{s}_1)^\dagger U_{s_1}(0)^\dagger]$. For certain choices of the $s_1,\ldots$ this path returns to the origin after m steps, in which case the product of traces is a product of two Wilson loops. If, however, the path does not return to the origin, the product of traces is not invariant under non-Abelian gauge transformations, and hence the average of the product of traces is equal to unity.

At infinite coupling, all of the unitary matrices are independent, except for the constraint $U_{\hat{d}}(x)=U_{d+\hat{D}/2}(x)$, and even if the path does return to the origin, the average of this product of traces is equal to unity, unless, by chance, the path of length m exactly retraces itself. The probability of this retracing, for a random path, is precisely the Cayley tree return probability discussed previously. Thus, this lattice gauge theory at infinite coupling has $\mathrm{tr}[U_{s_1}(0)U_{s_2}(0+\hat{s}_1)\ldots]\mathrm{tr}[\ldots U_{s_2}(0+\hat{s}_1)^\dagger U_{s_1}(0)^\dagger]=1+N^2D^{-m}N(0,m)$. The Eguchi-Kawai construction is an approximation to large N gauge theory, which replaces the infinite lattice by a single site: this turns $\mathrm{tr}[U_{s_1}(0)U_{s_2}(0+\hat{s}_1)\ldots]\mathrm{tr}[\ldots U_{s_2}(0$

$+\hat{s}_1)^\dagger U_{s_1}(0)^\dagger]$ into $E_0(s_1,s_2,\ldots)$, the quantity we considered before. Thus, this paper can be seen as an estimation of corrections to the Eguchi-Kawai construction in the infinite coupling limit. There are a number of interesting terms in these corrections: for example, the average $\mathrm{tr}[U(1)U(1)]\mathrm{tr}[U^\dagger(1)U^\dagger(1)]$ is equal to 2 as calculated before, but the corresponding average in the lattice gauge theory is equal to 1.

B. Applications

The general properties of ground states of local Hamiltonians with an excitation gap have become of great interest recently. A basic result [14,15] is that correlations decay exponentially in such systems. One application of quantum expanders is to finding matrix product states of one-dimensional quantum systems with the following properties: the correlation length is of order unity, the Hilbert space dimension on a single site is small, also of order unity, and yet the entanglement entropy across any cut is large. As an example, consider a matrix product state of the form

$$\Psi(s_1,s_2,\ldots,s_N) = \sum_{\alpha,\beta,\ldots} A_{\alpha,\beta}(s_1)A_{\beta,\gamma}(s_2)A_{\gamma,\delta}(s_3)\ldots, \tag{45}$$

where $s_1,s_2,\ldots,s_N$ are spin variables in a one-dimensional quantum system of N sites. Associated with the matrix product state is a completely positive map as in Eq. (1). If this map has a gap in its eigenspectrum to the second largest eigenvalue, then the state Ψ has exponentially decaying correlations [16], so that if operator A has support on sites $1,\ldots,j$ and operator B has support on sites $j+l,\ldots,N$, then $\langle\Psi,AB\Psi\rangle-\langle\Psi,A\Psi\rangle\langle\Psi,B\Psi\rangle$ is exponentially small in l, as required for the ground state of a gapped, local quantum system. However, as discussed in [1], this means that the existence of quantum expanders implies that the mere fact of exponentially decaying correlations does not suffice to prove bounds on entanglement entropy. Instead, bounds on the entanglement entropy [17] proceed through a different route and currently give weak bounds.

However, in [17], a conjecture was developed regarding properties of quantum expanders that may help in proving tighter bounds on entanglement entropy. Consider the following different correlation function. Let A be an operator with support on the sites $1,2,\ldots,j-l$ and $j+l,j+l+1,\ldots,N$. Let $\Psi=\Sigma_{\alpha=1}A(\alpha)\Psi_L(\alpha)\otimes\Psi_R(\alpha)$, where $\Psi_L(\alpha)$ are orthonormal states on sites $1,\ldots,j$ and $\Psi_R(\alpha)$ are orthonormal states on $j+1,\ldots,N$. Let $B_L=\Sigma_{\alpha=1}O(\alpha)\langle\Psi_{L,0}(\alpha)\rangle\times\langle\Psi_{L,0}\otimes\mathbb{1}_R\rangle$, where $\mathbb{1}_R$ is the unit operator on $X_{j+1,N}$, for some function $O(\alpha)$. Then, it was shown that for a gapped local Hamiltonian,

$$\langle\Psi_0,AB_L\Psi_0\rangle-\langle\Psi_0,A\Psi_0\rangle\langle\Psi_0,B_L\Psi_0\rangle \leq \|A\|\,\|B\|O(\exp[-l/l_0]), \tag{46}$$

for some l_0.

It was conjectured in [17] that there is a function $f(D_{\mathrm{eff}})$ such that if Eq. (46) holds for a state Ψ for some l_0, then the entanglement entropy of Ψ is bounded by $f(D^{l_0})$. Interestingly, it seems that an expander where the $A(s)$ are random unitaries is unlikely to satisfy Eq. (46). If this could be shown to be a general property of expanders, showing the conjecture, this would provide another way of studying area laws in quantum systems.

C. Extensions

The method of Schwinger-Dyson equations used here is fairly general and could be applied to other groups, such as O(N) or Sp($2N$). We have not done the calculation, but it seems that random choices from these groups will also give quantum expanders. Always, the unit matrix is an eigenvector of the map $\mathcal{E}(M)$ with eigenvalue unity. Any matrix in the center of the group is also an eigenvector of $\mathcal{E}(M)$ with eigenvalue unity, but for these cases, all elements of the center are proportional to the identity matrix, and thus do not give rise to additional eigenvectors with unit eigenvalue.

The method can be directly extended to the non-Hermitian case. Some of the combinatorics become slightly easier here. From Eq. (4), the average of $\Sigma_{a=1}^{N^2}|\lambda_a|^{2m}$ over the unitary group is bounded by the average of the trace as follows:

$$\left(\frac{1}{D}\right)^{2m}\sum_{s_1=1}^{D}\cdots\sum_{s_m=1}^{D}\sum_{\bar{s}_1=1}^{D}\cdots\sum_{\bar{s}_m=1}^{D}E[\mathrm{tr}[U(\bar{s}_1)\times U(\bar{s}_2)\cdots U(\bar{s}_m)U^\dagger(s_m)\cdots U^\dagger(s_2)U^\dagger(s_1)]\times\mathrm{tr}[U(s_1)U(s_2)\cdots U(s_m)U^\dagger(\bar{s}_m)\cdots U^\dagger(\bar{s}_2)U^\dagger(\bar{s}_1)]]. \tag{47}$$

The probability of having $U(s_1)U(s_2)\cdots U(s_m)\times U^\dagger(\bar{s}_m)\cdots U^\dagger(\bar{s}_2)U^\dagger(\bar{s}_1)$ cancel to the identity matrix is equal to $1/D^m$. Let

$$\lambda_{nH}=\frac{1}{\sqrt{D}}. \tag{48}$$

Carrying through the calculation one finds that, for any $\epsilon>0$, the probability that $|\lambda_2|\leq\lambda_{nH}+\epsilon$ approaches unity as $N\to\infty$. Note that $\lambda_{nH}<\lambda_H$ and also note that in the non-Hermitian case even a tight estimate on the average of the trace only provides an upper bound on the eigenvalue, due to the inequality in Eq. (4). However, numerical work suggests that the eigenvalue is asymptotically equal to λ_{nH} with high probability in this case.

We can also provide a lower bound on the trace in the non-Hermitian case. For any choice of unitaries $U(s)$, the sum of terms in Eq. (47) with $s_i=\bar{s}_i$ is bounded below by $1/D^m$, while the other terms are all positive, so that the given average (47) is bounded below by N^2D^{-m}. This result extends readily to an arbitrary choice of $A(s)$, constrained only by the trace-preserving condition $\Sigma_{s=1}^{D}A(s)A^\dagger(s)=\mathbb{1}$.

ACKNOWLEDGMENTS

I thank T. J. Osborne for many useful discussions on this topic. This work was inspired by discussions with A. Harrow, T. J. Osborne, and F. Verstraete at the workshop on "Lieb-

Robinson Bounds and Applications" at the Erwin Schrödinger Institute. This work was supported by U.S. DOE Contract No. DE-AC52-06NA25396.

APPENDIX: QUANTUM EDGE EXPANDERS

In this Appendix, we discuss the relationship between quantum expanders and another concept, a quantum version of an edge expander. We define a map to be a "quantum edge expander" if the following condition holds: for any N-by-N Hermitian matrix P such that $P^2=P$ and such that P has l nonzero eigenvalues, $l \le N/2$,

$$\mathrm{tr}[P\mathcal{E}(P)] \le \lambda_e \,\mathrm{tr}(P), \tag{A1}$$

for some λ_e less than one. We then prove a relation between λ_e and $|\lambda_2|$, showing that a quantum edge expander is a quantum expander. This is a quantum analog of a theorem of Tanner [18] and Alon and Milman [19], which shows that an edge expander has a spectral gap. We consider only the Hermitian case in this Appendix, leaving the behavior in the non-Hermitian case open. We also assume that the second largest eigenvalue is positive; the case where it is negative can be considered by looking at the square of the map $\mathcal{E}(M)$.

Let X be the eigenvector of $\mathcal{E}$ with eigenvalue λ_2. We work in a basis in which X is diagonal,

$$X = \begin{pmatrix} e_1 & & \\ & e_2 & \\ & & \cdots \end{pmatrix}, \tag{A2}$$

and such that $e_1 \ge e_2 \ge \cdots \ge e_N$. Since X is orthogonal to the unit matrix, using the inner product $(X,N)=\mathrm{tr}(XN)$, we have $\mathrm{tr}(X)=0$. Define m such that $e_i>0$ for $i \le m$ and $e_i \le 0$ for $i > m$. Without loss of generality we may suppose that $m \le N/2$, as otherwise we could have considered the matrix $-X$, which has the same eigenvalue. Define $M(i,j)$ to be the matrix with a unit entry in the ith row and jth column and zero everywhere else. Define

$$P_{ij} = \mathrm{tr}[M(i,i)\mathcal{E}(M(j,j))]. \tag{A3}$$

Then, since the map $\mathcal{E}$ is trace preserving, we have

$$\sum_i P_{ij} = 1 \tag{A4}$$

for all j. Also, we have $P_{ij}=P_{ji}$. Finally, since $\mathcal{E}$ is completely positive, we have $P_{ij} \ge 0$ for all i,j. Then,

$$\lambda_2 = \mathrm{tr}[X\mathcal{E}(X)] = \sum_{i=1}^{N}\sum_{j=1}^{N} e_i e_j P_{ij}. \tag{A5}$$

Define f_i by

$$f_i = e_i \Big/ \sqrt{\sum_{i=1}^{m} e_i^2}, \quad i \le m,$$

$$f_i = 0, \quad i > m. \tag{A6}$$

Then,

$$\lambda_2 \le \sum_{i=1}^{m}\sum_{j=1}^{m} f_i f_j P_{ij}. \tag{A7}$$

Then,

$$\begin{aligned}\sum_{i=1}^{m}\sum_{j>i}^{m} (f_i^2 - f_j^2) P_{ij} &= \sum_{i=1}^{m}\sum_{j>i}^{m} [\sqrt{P_{ij}}(f_i - f_j)][\sqrt{P_{ij}}(f_i + f_j)] \\ &\le \sqrt{\sum_{i=1}^{m}\sum_{j>i}^{m} P_{ij}(f_i - f_j)^2}\sqrt{\sum_{i=1}^{m}\sum_{j>i}^{m} P_{ij}(f_i + f_j)^2} \\ &\le \sqrt{\sum_{i=1}^{m}\sum_{j>i}^{m} P_{ij}(f_i - f_j)^2}\sqrt{(1/2)\sum_{i=1}^{m}\sum_{j=1}^{m} P_{ij}2(f_i^2 + f_j^2)} \\ &= \sqrt{2}\sqrt{\sum_{i=1}^{m}\sum_{j>i}^{m} P_{ij}(f_i - f_j)^2} \le \sqrt{2}\sqrt{(1-\lambda_2)},\end{aligned} \tag{A8}$$

where the first inequality uses Cauchy-Schwarz, the last equality uses Eq. (A4), and the last inequality uses Eq. (A7).

Let P_i be the projector onto the vector space spanned by the first i eigenvectors of M. Then,

$$\sum_{i=1}^{m}\sum_{j>i}^{m} (f_i^2 - f_j^2) P_{ij} = \sum_{i=2}^{m} (f_i^2 - f_{i-1}^2)\mathrm{tr}[(\mathbb{1} - P_i)\mathcal{E}(P_i)]. \tag{A9}$$

Using the property of a quantum edge expander (A1), we have

$$\begin{aligned}\sum_{i=1}^{m} (f_i^2 - f_{i+1}^2)\mathrm{tr}[(\mathbb{1} - P_i)\mathcal{E}(P_i)] &\ge \sum_{i=1}^{m} (f_i^2 - f_{i+1}^2)\mathrm{tr}(P_i)(1-\lambda_e) \\ &= \sum_{i=1}^{m} (f_i^2 - f_{i+1}^2) i (1-\lambda_e) \\ &= \sum_{i=1}^{m} f_i^2 (1-\lambda_e) = (1-\lambda_e).\end{aligned} \tag{A10}$$

Combining Eqs. (A8)–(A10), we find that

$$1 - \lambda_e \le \sqrt{2(1-\lambda_2)}. \tag{A11}$$

We finally show the converse result, that a quantum expander is a quantum edge expander. The normalized eigenvector with unit eigenvalue is $v_1 \equiv (1/\sqrt{N})\mathbb{1}$. We have $(v_1,P)=\mathrm{tr}(P)/\sqrt{N}$. So, $P=\mathrm{tr}(P)\mathbb{1}/N+P'$, where $P'=P-\mathrm{tr}(P)\mathbb{1}/N$. Then $\mathrm{tr}[P\mathcal{E}(P)] \le |\lambda_2|\mathrm{tr}(P'^2)+(v_1,P)^2 = |\lambda_2|[\mathrm{tr}(P)-\mathrm{tr}(P)^2/N]+\mathrm{tr}(P)^2/N$. If $\mathrm{tr}(P)=l \le N/2$, then $\mathrm{tr}(P)^2 \le (N/2)\mathrm{tr}(P)$ and $\mathrm{tr}[P\mathcal{E}(P)] \le |\lambda_2|\mathrm{tr}(P)+(1-|\lambda_2|)\mathrm{tr}(P)/2$ so

$$\lambda_e \le |\lambda_2|/2 + 1/2. \tag{A12}$$

[1] M. B. Hastings, Phys. Rev. B **76**, 035114 (2007).
[2] A. Ben-Aroya and A. Ta-Shma, arXiv:quant-ph/0702129.
[3] An appropriate formal definition may be to define a $(D,|\lambda_2|,N,S)$ expander, where S is the entropy of the eigenvector with eigenvalue unity.
[4] S. Hoory, N. Linial, and A. Wigderson, Bull., New Ser., Am. Math. Soc. **43**, 439 (2006).
[5] M. Sipser and D. Spielman, IEEE Trans. Inf. Theory **42**, 1710 (1996).
[6] I. Dinur, *Proceedings of the Annual ACM Symposium on the Theory of Computing*, p. 241 (2006).
[7] G. A. Margulis, Problemy Peredaci Informacii 9, 71 (1973); A. Lubotzky, R. Phillips, and P. Sarnak, Combinatorica **9**, 261 (1988).
[8] M. Pinsker, *7th International Teletraffic Conference*, p. 318/1 (1973); A. Broder and E. Shamir, *28th Annual Symposium on Foundations of Computer Science*, p. 286 (1987).
[9] J. Friedman, *Conference Proceedings of the Annual ACM Symposium on Theory of Computing*, p. 720 (2003); J. Friedman, e-print arXiv:cs/0405020.
[10] T. Eguchi and H. Kawai, Phys. Rev. Lett. **48**, 1063 (1982).
[11] N. Alon, Combinatorica **6**, 83 (1986); A. Nilli, Discrete Math. **91**, 207 (1991).
[12] An example where the series has an infinite number of nonvanishing coefficients is the average of

$$\mathrm{tr}(UUVV)\mathrm{tr}(V^\dagger V^\dagger U^\dagger U^\dagger)$$

over all unitary matrices U, V. A single application of Eq. (20) gives

$$1+(1/N)E[\mathrm{tr}(UVVU^\dagger V^\dagger V^\dagger)]$$
$$-(1/N)E[\mathrm{tr}(U)\mathrm{tr}(UVV)\mathrm{tr}(V^\dagger V^\dagger U^\dagger U^\dagger)].$$

Applying Eq. (20) again gives

$$1+(1/N^2)E[\mathrm{tr}(VV)\mathrm{tr}(V^\dagger V^\dagger)]$$
$$+(1/N^2)E[\mathrm{tr}(UUVV)\mathrm{tr}(V^\dagger V^\dagger U^\dagger U^\dagger)]$$
$$-(2/N^2)E[\mathrm{tr}(UVV)\mathrm{tr}(V^\dagger V^\dagger U^\dagger)].$$

This is equal to

$$1+(2/N^2)+(1/N^2)E[\mathrm{tr}(UUVV)\mathrm{tr}(V^\dagger V^\dagger U^\dagger U^\dagger)]-(2/N^2).$$

Thus,

$$E[\mathrm{tr}(UUVV)\mathrm{tr}(V^\dagger V^\dagger U^\dagger U^\dagger)]$$
$$=1+(1/N^2)E[\mathrm{tr}(UUVV)\mathrm{tr}(V^\dagger V^\dagger U^\dagger U^\dagger)]$$
$$=1+(1/N^2)+(1/N^4)+\cdots.$$

[13] We refer to these as "rung cancellations" because in certain diagrammatic expansions they correspond to a single rung of a ladder diagram from field theory.
[14] M. B. Hastings, Phys. Rev. Lett. **93**, 140402 (2004).
[15] M. B. Hastings, Phys. Rev. B **69**, 104431 (2004).
[16] M. Fannes, B. Nachtergaele, and R. F. Werner, Commun. Math. Phys. **144**, 443 (1992); D. Pérez-García, F. Verstraete, M. M. Wolf, and J. I. Cirac, QIC **7**, 401 (2007).
[17] M. B. Hastings, e-print arXiv:07052024.
[18] R. M. Tanner, SIAM J. Algebraic Discrete Methods **5**, 287 (1984).
[19] N. Alon and V. D. Milman, J. Comb. Theory, Ser. B **38**, 73 (1985).