Random quantum expanders

Ethan Lake

These notes are essentially a rambling solution to a problem assigned by John McGreevy to his QI class (I found the problem statement online). We consider a family of quantum channels $\mathcal{E}_N$ which act as

$$\mathcal{E}_N \in \mathrm{End}^2(\mathbb{C}^d), \qquad \mathcal{E}_N : \rho \mapsto \sum_{a=1}^{N} \frac{1}{N} U_a \rho U_a^\dagger, \tag{1}$$

where the U_a are drawn from a particular sampling of the Haar measure on $U(d)$. We will be looking at what happens under repeated application of this channel, by computing the vN entropy

$$S(N,n) = -\langle \mathrm{Tr}[\mathcal{E}_N^n(\rho_0) \ln \mathcal{E}_N^n(\rho_0)] \rangle, \tag{2}$$

where the brackets denote the Haar average, and where ρ_0 is some initial pure state (we will take $\rho = |0\rangle\langle 0|$).

* * *

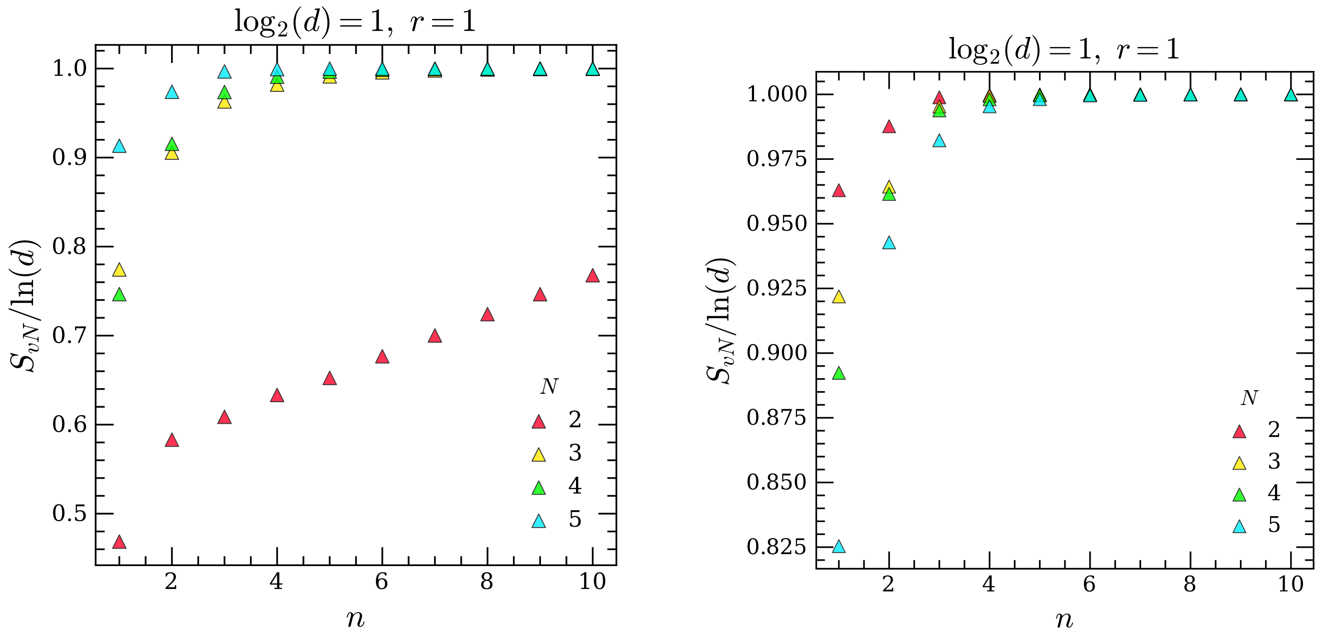
First, note that as $n \to \infty$, we expect that regardless of N and r, we will have (this intuitively obvious fact will be proved below)

$$\mathcal{E}_N^n(\rho) \xrightarrow{n\to\infty} \mathbf{1}. \tag{3}$$

Therefore in all of the plots which follow, we will compute the difference between the vN entropy of $\mathcal{E}_N^n(\rho)$ and $\ln d$, with d the Hilbert space dimension.

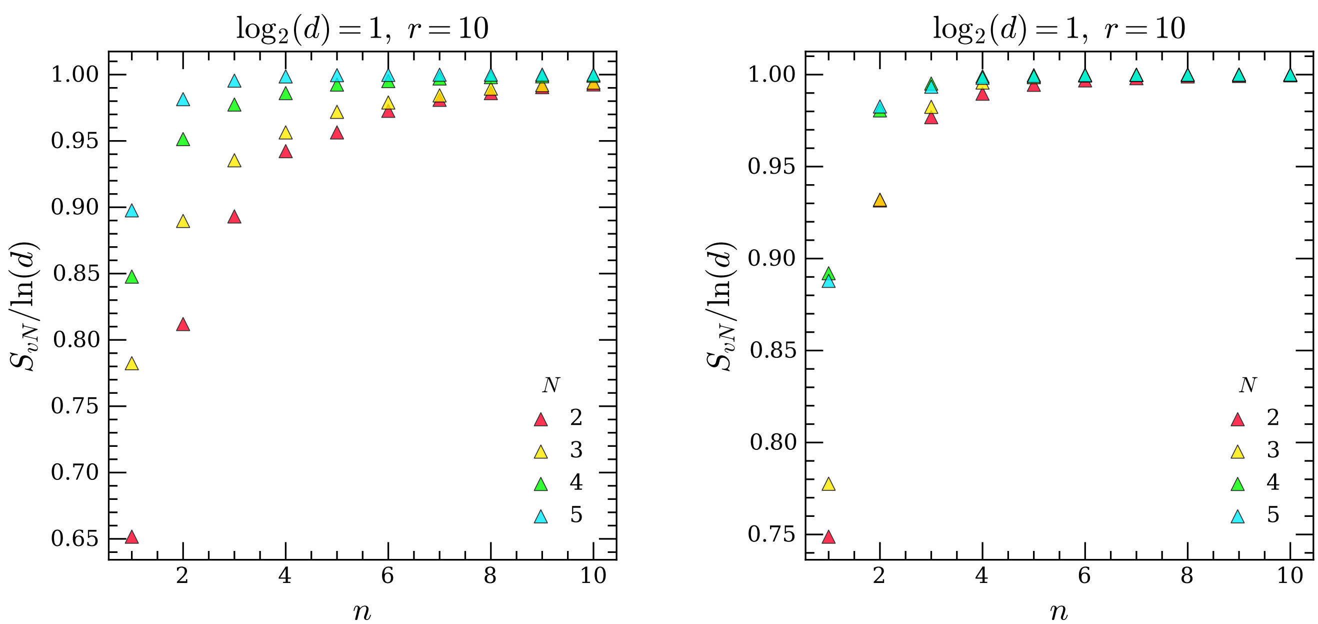
To compute the vN entropy of $\mathcal{E}_N^n(\rho)$ numerically, we approximate the Haar average by averaging over $r \in \mathbb{N}$ instances of the vN entropies computed for fixed realizations of the U_as, with $r \to \infty$ being the Haar average.

Let us first look at a single qubit. The variations in each realization of the sampling from $U(2)$ are huge. For example, two such one-shot realizations give (here the matrices used in each sampling for $N = k$ are a subset of those for $N > k$)



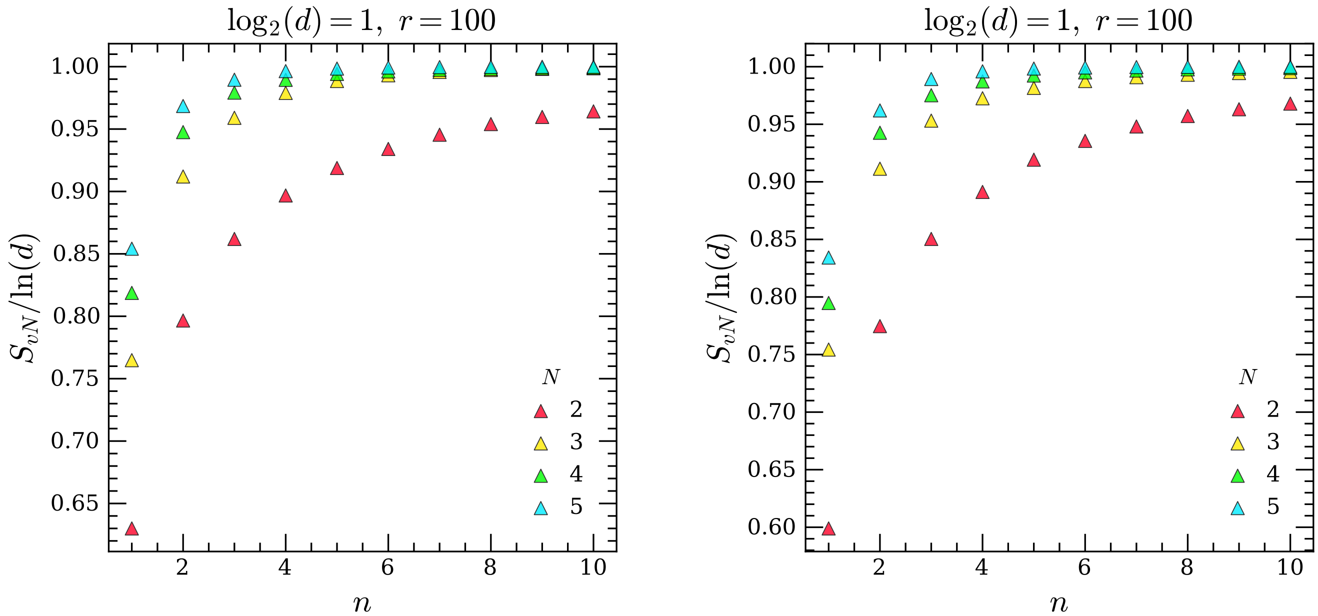
(4)

$r = 10$ is not enough large enough to effectively implement the average:



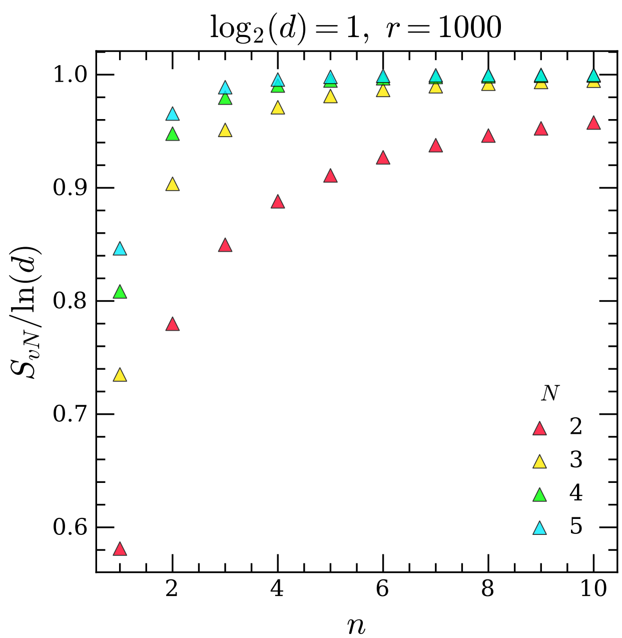
(5)

For $r = 100$, we have



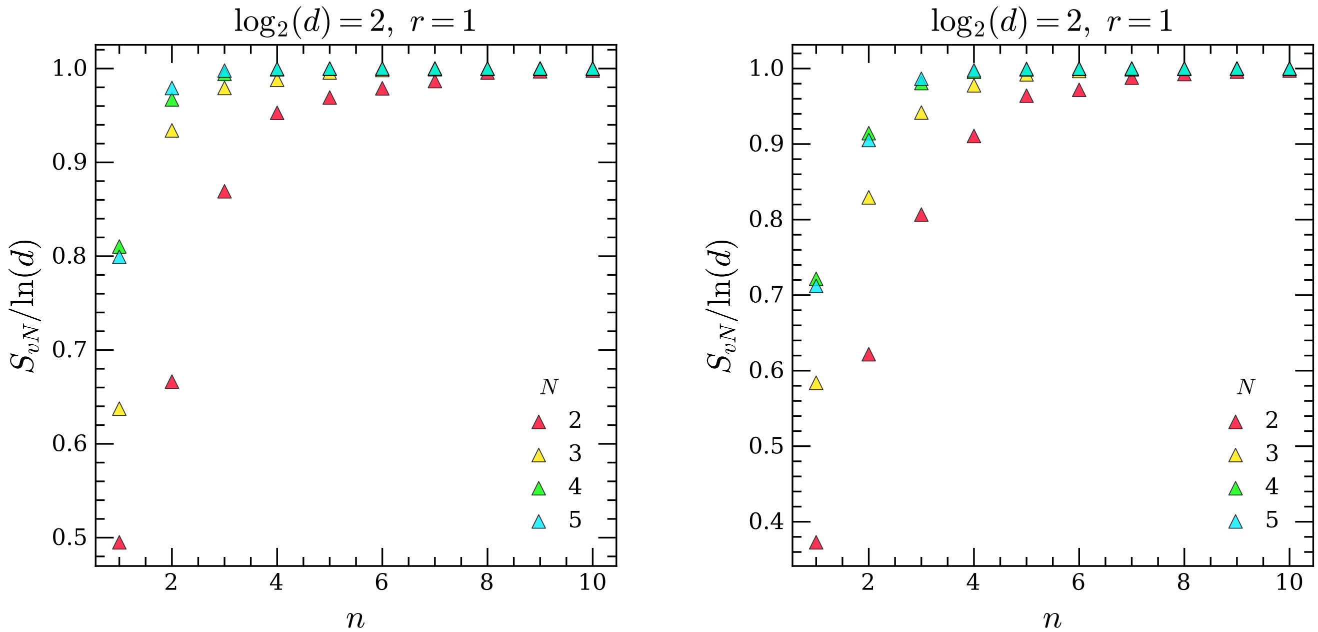
(6)

Therefore for 1 qubit, $r = 100$ comes pretty close to implementing the full Haar average. Indeed, increasing by another factor of 10 doesn't change the plots appreciably:



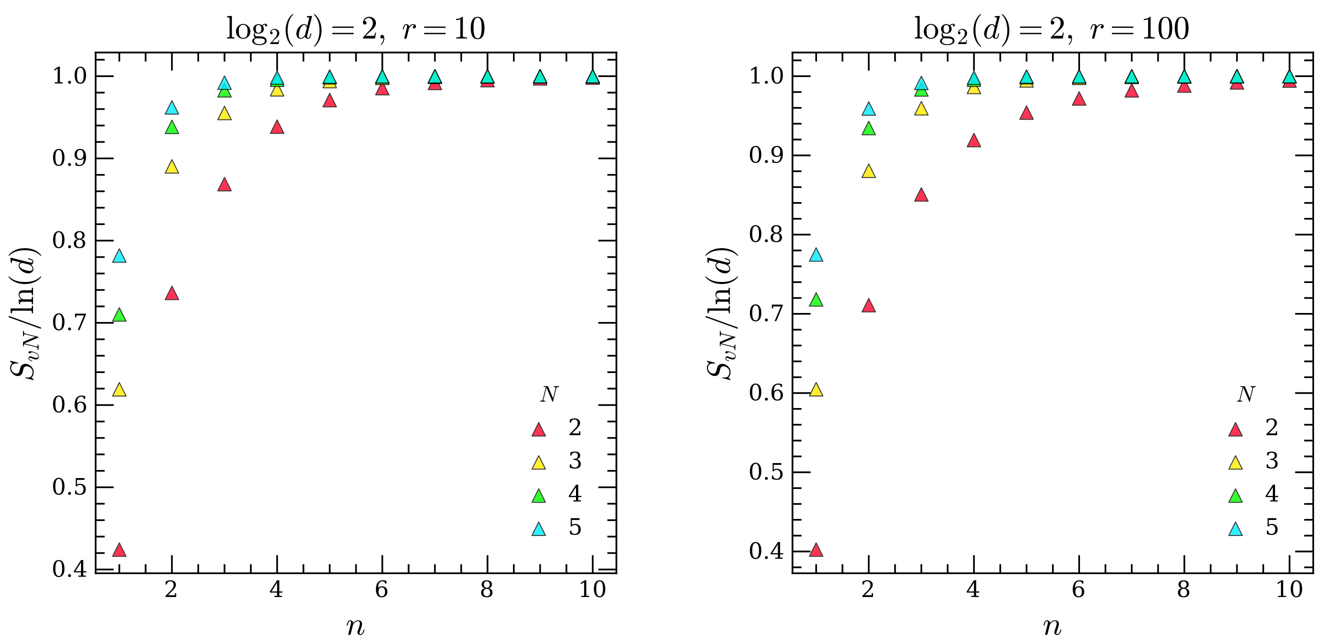
(7)

Now let us look at what happens when we increase the number of qubits. For two qubits, we still have appreciable variations between single-shot realizations of the unitaries:



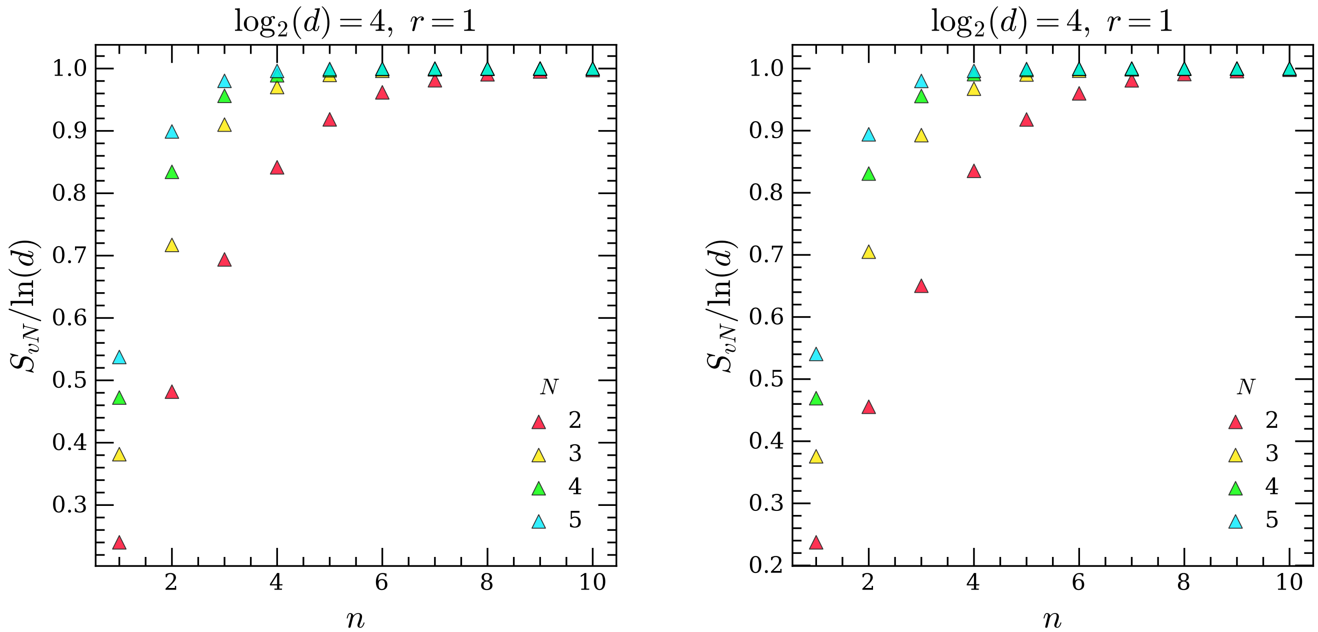
(8)

However, the convergence to the full Haar average is rather quick, with the $r = 10, 100$ plots looking almost identical:



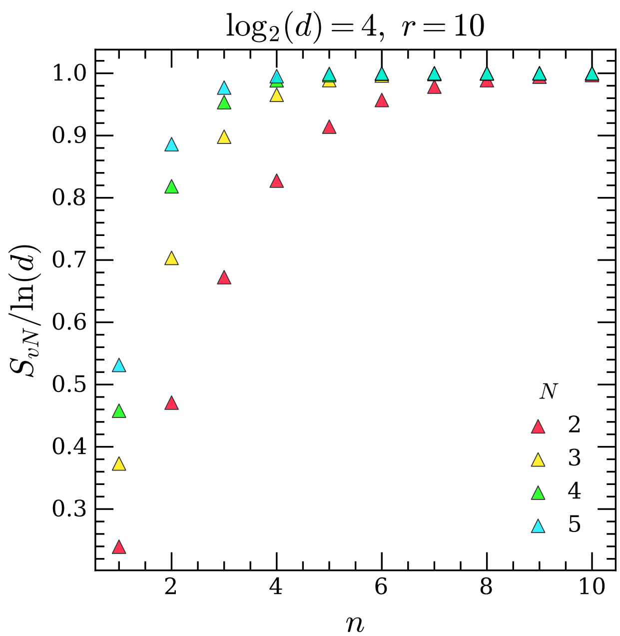
(9)

For four qubits, even two different $r = 1$ realizations look very similar:



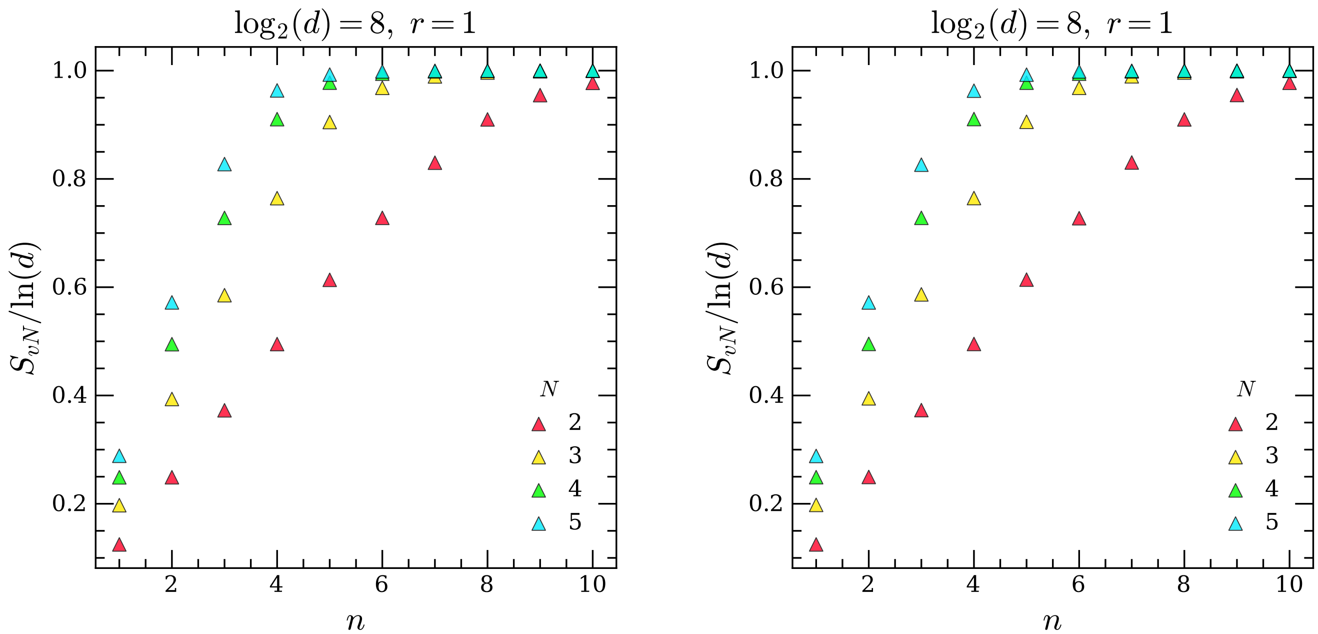
(10)

For $r = 10$,



(11)

where there is no real reason to go beyond $r = 10$ due to the fast convergence. Two different one-shot realizations for eight qubits gives



(12)

with the two plots being essentially indistinguishable.

Let's now try to get an analytic understanding of these results. The basic fact that $S(n \to \infty, N)$ monotonically increases with n and asymptotically approaches the maximal value of $\ln d$ is a consequence of the result that the output of a unital channel is majorized by its input, i.e. that unital channels increase the mixed-ness of states on which they act. This statement is proved in a subsection below. Since $\mathcal{E}_N$ is obviously unital, the limit as $n \to \infty$ of $\mathcal{E}_N^n$ will be the most mixed possible state, viz. $\mathbf{1}/d$.

Unfortunately since S_{vN} has a log, it is hard to make any more precise analytic statements about it directly. We can however make statements about e.g. things involving powers of ρ, like the Renyi entropies. For simplicity, we will look at the purity[1] $\mathcal{P} = \mathrm{Tr}[\rho^2]$. Again for simplicity, we will just look at its average under a single application of the channel (we know that it must be monotonically decreasing as a function of n since $\rho \prec \sigma$ implies that $\mathcal{P}_\rho \leq \mathcal{P}_\sigma$). Therefore we would like to compute

$$\langle \mathcal{P} \rangle = \langle \mathrm{Tr}[\mathcal{E}_N(\rho)^2] \rangle. \tag{13}$$

Now each U_a in the quantum channel is averaged over as an independent Haar unitary. Therefore

$$\langle \mathcal{E}_N(\rho)^2 \rangle = \frac{1}{N^2} \left(\sum_{a \neq b} \langle U_a \rho U_a^\dagger \rangle \langle U_b \rho U_b^\dagger \rangle + \sum_a \langle U_a \rho^2 U_a^\dagger \rangle \right). \tag{14}$$

The individual matrix elements of the square of the quantum channel are then

$$\langle [\mathcal{E}_N(\rho)^2]_{ad} \rangle = \frac{N-1}{N} \langle U_{aa'} U^*_{cc'} \rangle \langle U_{bb'} U^*_{dd'} \rangle \rho_{a'c'} \rho_{b'd'} \delta_{cb} + \frac{1}{N} \langle U_{aa'} U^*_{bb'} \rangle \rho_{a'e} \rho_{eb'} \tag{15}$$

[1] This is a simple measure of purity because it is 1 for pure states and minimal on maximally-mixed ones.

We know this will be proportional to the identity by Schur's lemma and shifting the Haar integral, but in order to get the coefficient we need to recall how to average over Haar unitaries. This is explained in gory detail in a separate diary entry; here we just need that

$$\int dU\, U_{ab}U^*_{cd} = \frac{1}{d}\delta_{ac}\delta_{bd}. \tag{16}$$

The above will work for computing the purity under one application of the channel; for larger values of n one also needs expressions like

$$\int dU\, [U_{a'a}U_{b'b}U^*_{c'c}U^*_{d'd}] = \frac{1}{d^2-1}\left(\delta_{a'c'}\delta_{b'd'}\delta_{ac}\delta_{bd} + \delta_{a'b'}\delta_{c'd'}\delta_{ad}\delta_{bc} - \frac{1}{d}\left(\delta_{a'c'}\delta_{b'd'}\delta_{ad}\delta_{bc} + \delta_{a'b'}\delta_{c'd'}\delta_{ac}\delta_{bd}\right)\right) \tag{17}$$

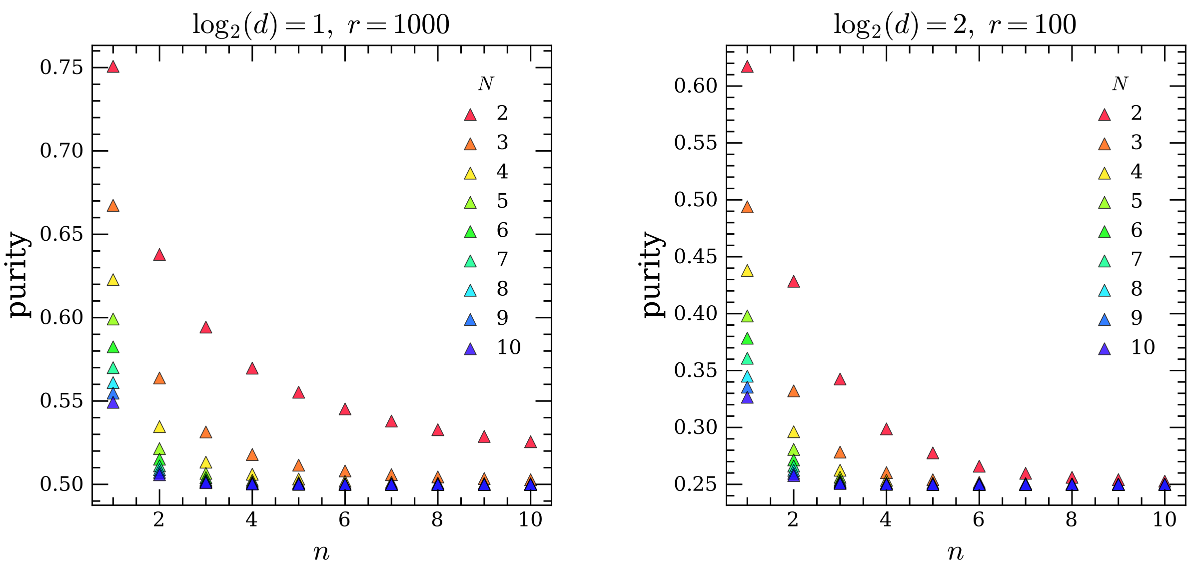
Using the above formula for the second moment, we find

$$\langle \mathcal{E}_N(\rho)^2\rangle = \left(\frac{N-1}{d^2N}\mathrm{Tr}[\rho]^2 + \frac{1}{dN}\mathrm{Tr}[\rho^2]\right)\mathbf{1}. \tag{18}$$

Since we are choosing an initial ρ which is pure, we obtain

$$\langle \mathcal{P}\rangle = \frac{N-1}{d^2N} + \frac{1}{Nd}. \tag{19}$$

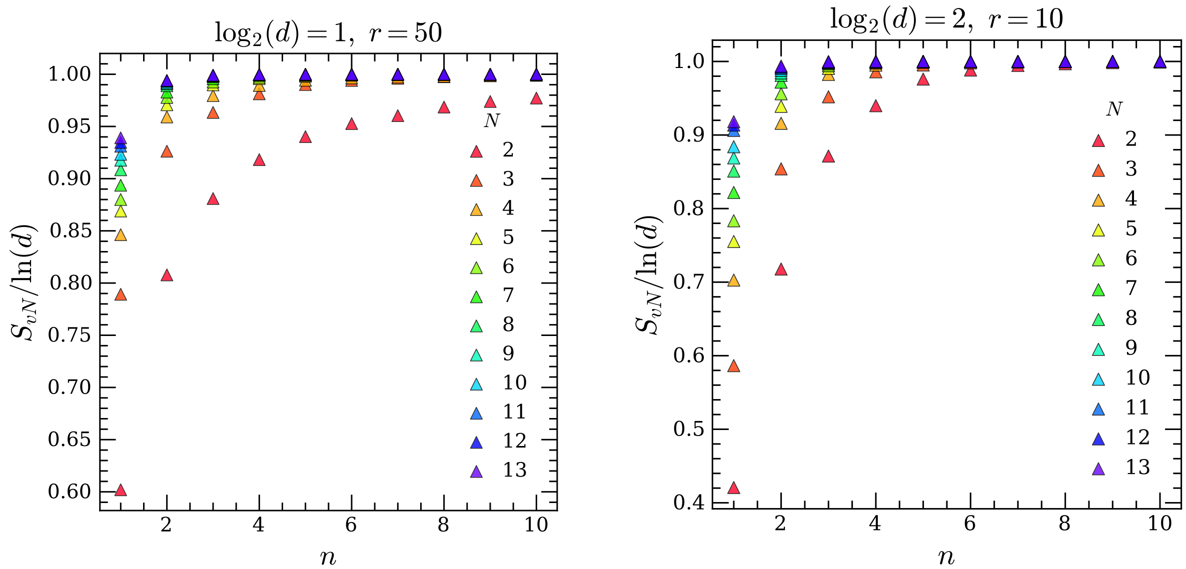
We can check this by computing the purities numerically: for one- and two-qubit systems, we find



(20)

Plugging in numbers to (19) verifies that this agrees exactly with the values obtained from the plots shown above.

Note that the purity (19) asymptotes to a fixed nonzero value of $1/d^2$ as $N \to \infty$. We therefore expect that the vN entropy after one application of the channel should also asymtote as $N \to \infty$ to some fixed value less than 1. This can be confirmed by increasing the range of N: for one- and two-qubit systems, we have



(21)

Diversion on majorization and doubly stochastic maps

We will now do a problem in Preskill's QI notes on unital channels and majorization (a channel $\mathcal{C}$ is unital if $\mathcal{C}(\mathbf{1}) = \mathbf{1}$; clearly the $\mathcal{E}_N$ considered in today's diary entry are unital), which will give the result mentioned above on the increasing mixedness of ρ under application of $\mathcal{E}_N$.

Theorem 1. *If $\mathcal{C}$ is a unital quantum channel, then for any density matrix ρ, $\mathcal{C}(\rho)$ is majorized by ρ:*

$$\mathcal{C}(\rho) \prec \rho. \tag{22}$$

Recall that this means that if $\Delta_{\mathcal{C}(\rho)}, \Delta_\rho$ are the eigenvalues of $\mathcal{C}(\rho), \rho$ respectively, then for all $1 \leq n \leq \dim \Delta_\rho$, we have

$$\mathcal{C}(\rho) \prec \rho \quad \Longrightarrow \quad \sum_i^n [\Delta^\downarrow_{\mathcal{C}(\rho)}]_i \ \leq \ \sum_i^n [\Delta^\downarrow_\rho]_i, \tag{23}$$

where the $\downarrow$ on each set of eigenvalues indicates that they are ordered from greatest to least. $\mathcal{C}(\rho) \prec \rho$ is pronounced as "$\mathcal{C}(\rho)$ is at least as mixed as ρ". Hence the claim is that if $\mathcal{C}$ is unital, the purity of ρ decreases monotonically under repeated application of the channel.

`<proof>`

The proof is pretty simple. Since $\mathcal{C}(\rho)$ and ρ are both Hermitian, they are diagonalized by unitaries U and V, respectively. If $\mathcal{C}$ has a representation in terms of Kraus operators $\mathcal{K}_a$, then

$$\Delta_{\mathcal{C}(\rho)} = \sum_a U^\dagger \mathcal{K}_a V \Delta_\rho V^\dagger \mathcal{K}_a^\dagger U, \tag{24}$$

where both Δ matrices are diagonal. We can then write the RHS as

$$\sum_a U^\dagger \mathcal{K}_a V \Delta_\rho V^\dagger \mathcal{K}_a^\dagger U = \mathcal{M}\Delta_\rho \tag{25}$$

with the matrix $\mathcal{M}$ given by

$$\mathcal{M}_{ij} = \sum_a [U^\dagger \mathcal{K}_a V]_{ij} [V^\dagger \mathcal{K}_a^\dagger U]_{ji} = \sum_a [U^\dagger \mathcal{K}_a V]_{ij} [U^T \mathcal{K}_a^* V^*]_{ij} = \sum_a |[V^\dagger \mathcal{K}_a^\dagger U]_{ji}|^2. \tag{26}$$

All the entries of $\mathcal{M}$ are obviously positive. Furthermore all the rows of $\mathcal{M}$ sum to 1:[2]

$$\sum_i \mathcal{M}_{ij} = \sum_a [V^\dagger \mathcal{K}_a^\dagger \mathcal{K}_a V]_{jj} = 1. \tag{27}$$

Since $\mathcal{C}$ is unital, we must have $\mathcal{C}(\mathbf{1}) = \sum_a \mathcal{K}_a \mathcal{K}_a^\dagger = \mathbf{1}$. This implies that all the columns of $\mathcal{M}$ sum to 1 as well:

$$\sum_j \mathcal{M}_{ij} = \sum_a [U^\dagger \mathcal{K}_a \mathcal{K}_a^\dagger U]_{ii} = 1. \tag{28}$$

Therefore we have $\Delta_{\mathcal{C}(\rho)} = \mathcal{M}\Delta_\rho$, with $\mathcal{M}$ a doubly stochastic matrix.

The fact that the two vectors of eigenvalues are related by a doubly stochastic matrix means that the eigenvalues of $\mathcal{C}(\rho)$ are convex combinations of the eigenvalues of ρ. This follows from the Brikhoff-von Neumann theorem, viz. that the set of doubly stochastic matrices is the convex hull[3] of the set of permutation matrices. Intuitively, this means that the eigenvalues of $\mathcal{C}(\rho)$ are more "smoothed out" and "uniform" than those of ρ. The precise statement is of course that $\mathcal{C}(\rho) \prec \rho$, although we won't go into the details of the proof here. We won't show here that the double stochasticity of $\mathcal{M}$ implies the majorization; this can be looked up in chapter 12 of N&C.

`</proof>`

[2]If we have $p = Dq$ for two probability distributions p, q (as we do here), then the rows of D generically sum to one — this is required if D is to preserve the normalization $\sum_i q_i = 1$, and hence is why proving it requires using $\sum_a \mathcal{K}_a^\dagger \mathcal{K}_a = \mathbf{1}$. The property of the columns of D summing to one is something extra, and is what determines the double stochasticity of D.

[3]The fact that convex combinations of doubly stochastic matrices are doubly stochastic is easy to see. The fact that the corners of this convex hull are given by the permutation matrices is also intuitively reasonable — taking convex combinations "smoothens out" the eigenvalues while permutation matrices leave the eigenvalues unchanged, hence the permutation matrices cannot be formed from convex combinations of matrices which all "smooth out" the eigenvalues.