PROCEEDINGS OF THE ROYAL SOCIETY A

Proc. R. Soc. A (2009) **465**, 3289–3308
doi:10.1098/rspa.2009.0232
Published online 5 August 2009

Large deviation bounds for *k*-designs

By Richard A. Low*

Department of Computer Science, University of Bristol, Bristol BS8 1UB, UK

We present a technique for de-randomizing large deviation bounds of functions on the unitary group. We replace the Haar measure with a pseudo-random distribution, a k-design. k-Designs have the first k moments equal to those of the Haar measure. The advantage of this is that (approximate) k-designs can be implemented efficiently, whereas Haar random unitaries cannot. We find large deviation bounds for unitaries chosen from a k-design and then illustrate this general technique with three applications. We first show that the von Neumann entropy of a pseudo-random state is almost maximal. Then we show that, if the dynamics of the universe produces a k-design, then suitably sized subsystems will be in the canonical state, as predicted by statistical mechanics. Finally we show that pseudo-random states are useless for measurement-based quantum computation.

Keywords: large deviation bounds; measure concentration; designs; de-randomization

1. Introduction

There are many results in quantum information theory that show generic properties of states or unitaries (e.g. Hayden *et al.* 2004, 2006). Often, these results say that, with high probability, a random state or unitary has some property, for example high entropy. However, simple parameter counting shows that random unitaries cannot be obtained efficiently. This limits the usefulness of such results since no physical systems will behave truly randomly. To make such results more physically relevant, it would be desirable to show that these properties are generic properties of unitaries from some natural distribution that can be implemented efficiently. Only then could we conclude that we would expect to see such properties in natural systems.

In many cases, the generic properties of unitaries are desirable but randomized constructions given by the large deviation bounds are inefficient. We would like to come up with distributions which can be implemented efficiently that have similar generic properties. One example where the best known construction is an inefficient randomized one is the ∞-norm randomizing map (Hayden *et al.* 2006; Aubrun 2008), which is a quantum map that takes any state to a state close to the identity, as measured by the ∞-norm. Another example is locking of classical correlations (DiVincenzo *et al.* 2004; Hayden *et al.* 2006), which is a quantum phenomenon whereby a small amount of communication can greatly enhance the classical correlation between two parties. To prove the randomized constructions,

*low@cs.bris.ac.uk

Received 29 April 2009
Accepted 8 July 2009

the authors show that, with some non-zero probability, random unitaries have the required property. However, there are no known efficient constructions of unitaries with these properties. If, on the other hand, we could show that unitaries drawn randomly from a set that can be implemented efficiently have the property with non-zero probability, we could move an important step closer to finding efficient constructions. (It would not actually provide an efficient construction unless we could find an efficient sampling method.) In fact, for the case of ∞-norm randomization, this was done by Aubrun (2008).

By random unitaries, we mean unitary matrices distributed according to the unitarily invariant Haar measure, which is the unique measure on the unitary group with the property of unitary invariance. In this paper, we will consider replacing the Haar measure with a k-design. A unitary k-design is an ensemble of unitaries such that the kth moments are the same as for the Haar measure (Dankert *et al.* 2006) (k-designs are formally defined in §2). In particular, this means that the expectation of a polynomial in the elements of the unitary matrices of degree at most k is the same whether the distribution is the Haar measure or a k-design. We will also consider replacing Haar random states with a state k-design, which is an ensemble of states such that the kth moments are the same as for the Haar measure (Ambainis & Emerson 2007).

The reason for using k-designs is twofold. Firstly, because the first k moments are the same we would expect similar (although weaker) measure concentration results. Secondly, for $k = \mathrm{poly}(n)$ (when the design is on n qubits), we might expect to be able to implement the k-design efficiently (i.e. in $\mathrm{poly}(n)$ time). Indeed, for $k = O(n/\log n)$, Harrow & Low (2008*a*) provide an efficient k-design construction, provided we allow for approximate designs. However, in the applications we consider here we can always make the approximation good enough to make the error negligible.

Not only can k-designs be constructed efficiently, they may even be the product of generic dynamics. In Harrow & Low (2008*b*), it is shown that random quantum circuits quickly converge to a 2-design for a quite general model of such circuits. It is also conjectured in Harrow & Low (2008*b*) that random circuits give k-designs for $k > 2$ and $k = \mathrm{poly}(n)$ in polynomial time. If a physical system can be accurately modelled by a random circuit then, assuming this conjecture, the naturally occurring states will be k-designs rather than fully random states.

We now summarize some related results in this area. Smith & Leung (2006) and Dahlsten & Plenio (2006) found large deviation bounds for stabilizer states. They showed that, in certain regimes, stabilizer states are very likely to have large entanglement. Stabilizer states are state 2-designs, so our results can be seen as a generalization of this to k-designs for $k > 2$ and to other problems. There are also some recent classical results related to the present work. Alon & Nussboim (2008) consider replacing full randomness with k-wise independence, a classical analogue of k-designs, in random graph theory. They show that k-wise independent random graphs with $k = \mathrm{poly}(\log N)$ (N is the number of vertices) have similar generic properties to fully random graphs.

(*a*) *Introductory problem: entanglement of a 2-design*

We now illustrate our main idea by showing a large deviation bound for the entanglement of a 2-design, but in a different way to Smith & Leung (2006) and Dahlsten & Plenio (2006).

It has been known for a long time that random states are highly entangled across any bipartition (Page 1993; Foong & Kanno 1994; Sanchez-Ruiz 1995). Further, in Hayden *et al.* (2006), it is shown that random unitaries generate almost maximally entangled states with high probability. However, generating random states is inefficient, so it is an interesting question to ask if random efficiently obtainable states are highly entangled.

Let the system be $\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_E$, where we label the two systems S and E. Let the dimensions be d_S and d_E and $d = d_S d_E$. Let the overall initial state be any fixed state ρ_0. Then consider applying a random unitary U to SE to get the state $\psi = U\rho_0 U^\dagger$. Then the von Neumann entropy $S(\psi_S)$ of the reduced state $\psi_S = \mathrm{tr}_E \psi$ is close to $\log_2 d_S$ (the maximal) with high probability:

Theorem 1.1 (Hayden *et al.* (2006), theorem 3.3). *Let $d_E \geq d_S \geq 3$. Then for unitaries chosen from the Haar measure*

$$\mathbb{P}(S(\psi_S) \leq \log_2 d_S - \alpha - \beta) \leq \exp\left(-\frac{(d-1)C\alpha^2}{(\log_2 d_S)^2}\right), \tag{1.1}$$

where $C = 1/8\pi^2$ and $\beta = (1/\ln 2)(d_S/d_E)$.

Now, consider choosing the unitary from a 2-design instead. Later on (lemma 4.1), we show that $\mathbb{E}\,\mathrm{tr}\,\psi_S^2 = (d_S + d_E)/(d+1) =: \mu$. Since purity is a polynomial of degree 2, it does not matter if we take the expectation over the Haar measure or the 2-design. We now apply Markov's inequality:

$$\mathbb{P}(\mathrm{tr}\,\psi_S^2 \geq \mu\gamma) \leq \frac{\mathbb{E}\,\mathrm{tr}\,\psi_S^2}{\mu\gamma} = \frac{1}{\gamma}.$$

Using the bound $S(\psi_S) \geq -\log_2 \mathrm{tr}\,\psi_S^2$ and some manipulations (the details are in §4), this can be written as

$$\mathbb{P}(S(\psi_S) \leq \log_2 d_S - \alpha - \beta) \leq 2^{-\alpha}, \tag{1.2}$$

where β is as in theorem 1.1. This bound is much weaker than the bound in theorem 1.1 and, in particular, does not show stronger concentration as d increases. Later in the paper, we will show that choosing unitaries from a k-design with larger k will give a much stronger bound that does give sharp concentration results for large d.

(*b*) *Main results*

We will now state our main results. In the remainder of the paper we will use the following notation to identify the measure we are using. When the unitaries are chosen from a measure ν, we will write $\mathbb{P}_\nu$ to mean $\mathbb{P}_{U\sim\nu}$, the probability when U is chosen from ν. Similarly for $\mathbb{E}_\nu$, the expectation. Usually ν will be a k-design. When the measure is the Haar measure, we will write a subscript H. So for the Haar average we write $\mathbb{E}_H$ for $\mathbb{E}_{U\sim\mathcal{U}(d)}$.

Our most general result is:

Theorem 1.2. *Let f be a polynomial of degree K. Let $f(U) = \sum_i \alpha_i M_i(U)$ where $M_i(U)$ are monomials and let $\alpha(f) = \sum_i |\alpha_i|$. Suppose that f has probability*

concentration

$$\mathbb{P}_H(|f-\mu|\geq\delta)\leq C\mathrm{e}^{-a\delta^2}, \tag{1.3}$$

and let ν *be an* ϵ*-approximate unitary* k*-design. Then*

$$\mathbb{P}_\nu(|f-\mu|\geq\delta)\leq\frac{1}{\delta^{2m}}\left(C\left(\frac{m}{a}\right)^m+\frac{\epsilon}{d^k}(\alpha+|\mu|)^{2m}\right), \tag{1.4}$$

for integer m *with* $2mK\leq k$.

We therefore take a bound for Haar random unitaries of the form equation (1.3) and turn it into a bound for k-designs. For our definition of ϵ-approximate designs, see §2. Often, we will use Levy's Lemma (lemma 3.2) to give the initial concentration bound in equation (1.3). In this case, $a=\Theta(d)$ (provided the Lipschitz constant (see later) is constant).

We then apply this to entropy, as a generalization of §1*a*. We go via the 2-norm since the entropy function is not a polynomial. We find

Theorem 1.3. *Let* ν *be a* 4^{-n^2}*-approximate unitary* $n/10\log_2 n$*-design on dimension* 2^n *with* $n\geq 19$. *Let* $d_S d_E=2^n$ *and* $2\leq d_S\leq 2^{n/10}$ *and* $\alpha\geq 2$. *Then,*

$$\mathbb{P}_\nu(S(\psi_S)\leq\log_2 d_S-\alpha-\beta)\leq 8\exp_2\left(-\frac{n}{80\log_2 n}\left(\frac{n}{5}+\alpha\right)\right), \tag{1.5}$$

where $\beta=(1/\ln 2)(d_S/d_E)$ *and* $\exp_2$ *is the exponential function base 2.*

We choose a k-design for $k=n/10\log_2 n$ since this is (up to constants) the largest k for which we have an efficient unitary k-design construction (see §2*a*).

We then move on to apply our results to ideas in statistical mechanics from Popescu *et al.* (2006). In this paper, the authors show that, for almost all pure states of the universe, any subsystem is very close to the canonical state, which is the state obtained by assuming a uniform distribution over all allowed states of the universe (defined in equation (5.2)). This could be achieved if the dynamics of the universe produced a random unitary, but this would take exponential time in the size of the universe. We show that the random unitary can be replaced by a k-design, showing that the canonical state can be reached in polynomial time:

Theorem 1.4. *Let* Ω_S *be the canonical state of the system (defined in equation* (5.2)*) and* ρ_S *be the state after choosing a unitary from an* ϵ*-approximate* k*-design. Let* d_R *be the dimension of the universe's Hilbert space subject to the arbitrary constraint* R *(normally this will be a total energy constraint). Then for* $\epsilon\leq(3/2)(4d_S^3/d_R)^{k/8}$, $k\leq 4d_S^2/(9\pi^3)$

$$\mathbb{P}_\nu(||\rho_S-\Omega_S||_1\geq\delta)\leq 6\left(\frac{4d_S^3}{d_R\delta^2}\right)^{k/8}. \tag{1.6}$$

Finally, we use results from Gross *et al.* (2008) to show that most states in an $O(1)$-approximate state n^2-design on n qubits are useless for measurement-based quantum computing (MBQC), in the sense that any computation using such states could be simulated efficiently on a classical computer. We do this, following Gross *et al.* (2008), by showing that the states are so entangled that the measurement outcomes are essentially random.

(c) *Optimality of results*

An important question is how close our results are to optimal, in terms of their scaling with dimension d. In theorem 1.2, we will normally have $a = \Theta(d)$ so for m constant, we obtain polynomial bounds, rather than the exponential bounds for full randomness. This is to be expected:

Theorem 1.5. *Let ν be an ϵ-approximate unitary k-design. Suppose also that it is discrete, i.e. contains a finite number S of unitaries. Let $f(U)$ be any function on matrix elements of U and μ be any constant. Then either $f(U) = \mu$ for all U in ν or for some $\delta > 0$,*

$$\mathbb{P}_\nu(|f - \mu| \geq \delta) \geq p_{\min}, \tag{1.7}$$

where $p_{\min}$ is the probability of choosing the least probable unitary from ν. If the probability is uniform, $p_{\min} = 1/S$.

Proof. There exists at least one U such that $|f(U) - \mu| \geq \delta$ for some $\delta > 0$; the probability of selecting one such U is at least $p_{\min}$. ■

Corollary 1.6. *Our results are polynomially related to the optimal (i.e. the optimal bounds can be obtained by raising ours to a constant power).*

Proof. Our results apply for any design, so must obey the bound in theorem 1.5 for all designs. The unitary design construction we use (lemma 2.7) has $p_{\min} = d^{-O(k)}$, hence the bounds cannot scale better than this. ■

We can also almost recover the tail bound for full randomness in theorem 1.2. Suppose for simplicity that we have an exact design (i.e. $\epsilon = 0$), so that

$$\mathbb{P}_\nu(|f - \mu| \geq \delta) \leq C\left(\frac{m}{a\delta^2}\right)^m.$$

The optimal m is $a\delta^2/\mathrm{e}$, which gives

$$\mathbb{P}_\nu(|f - \mu| \geq \delta) \leq C\mathrm{e}^{-a\delta^2/\mathrm{e}}.$$

So our result allows us to interpolate from Markov's inequality, which gives weak bounds, all the way to full Haar randomness and is within a polynomial correction of optimal for the full range.

The remainder of the paper is organized as follows. In §2 we formally define k-designs and what we mean by approximate designs. Then in §3 we present our main technique for finding large deviation bounds for k-designs. We then apply this to entropy in §4, to ideas in statistical mechanics in §5 and to using k-designs for MBQC in §6. We then conclude in §7.

2. *k*-Designs

Here we formally define k-designs.

Definition 2.1. Let ν be a distribution on the unitary group. ν is a unitary k-design if

$$\mathbb{E}_\nu\left[U^{\otimes k}\rho(U^\dagger)^{\otimes k}\right] = \mathbb{E}_H\left[U^{\otimes k}\rho(U^\dagger)^{\otimes k}\right], \tag{2.1}$$

for all $d^k \times d^k$ complex matrices ρ (not necessarily valid states).

We can write this as an equivalent, and for our purposes more useful, definition in terms of monomials of the elements of the matrices. We will first define what we mean by degree of a monomial (or polynomial):

Definition 2.2. A monomial in elements of a matrix U is of degree (k_1, k_2) if it contains k_1 conjugated elements and k_2 unconjugated elements. We call it balanced if $k_1 = k_2$ and will simply say a balanced monomial has degree k if it is degree (k, k). A balanced polynomial is of degree k if it is a sum of balanced monomials of degree at most k.

So that, in this definition, $U_{ij}U^*_{pq}$ is a balanced monomial of degree $(1, 1)$ and $U_{ij}U_{kl}$ is a monomial of degree $(2, 0)$ and is unbalanced. We now state an equivalent definition of unitary k-designs in terms of monomials:

Definition 2.3. ν is a unitary k-design if, for all balanced monomials M of degree k,

$$\mathbb{E}_\nu M(U) = \mathbb{E}_H M(U). \tag{2.2}$$

That definitions 2.1 and 2.3 are equivalent can be seen by considering matrices ρ of the form $|i_1, i_2, \ldots, i_k\rangle\langle j_1, j_2, \ldots, j_k|$ in definition 2.1. Then each element of $U^{\otimes k}\rho(U^\dagger)^{\otimes k}$ is a balanced monomial of degree k. Further, each balanced monomial appears for some choice of ρ.

We will use state k-designs, which are related to unitary k-designs although less general:

Definition 2.4. Let ν be a distribution on states and let ν_H be the uniform distribution on states, which can be thought of as a random unitary being applied to any fixed state. Then ν is a state k-design if

$$\mathbb{E}_{|\psi\rangle\sim\nu}\left[(|\psi\rangle\langle\psi|)^{\otimes k}\right] = \mathbb{E}_{|\psi\rangle\sim\nu_H}\left[(|\psi\rangle\langle\psi|)^{\otimes k}\right]. \tag{2.3}$$

By considering unitaries acting on a fixed state, it can be seen that a unitary k-design can provide a state k-design, although the reverse is not necessarily true.

(*a*) *Approximate k-designs*

There are no known efficient constructions of exact unitary k-designs. However, for our purposes, only approximate designs are required. In Ambainis & Emerson (2007), the authors define an ϵ-approximate state k-design using the ∞-norm:

Definition 2.5 (Ambainis & Emerson 2007). ν is an ϵ-approximate state k-design if

$$\left\|\mathbb{E}_{|\psi\rangle\sim\nu}\left[(|\psi\rangle\langle\psi|)^{\otimes k}\right] - \mathbb{E}_{|\psi\rangle\sim\nu_H}\left[(|\psi\rangle\langle\psi|)^{\otimes k}\right]\right\|_\infty \leq \frac{\epsilon}{\binom{k+d-1}{d-1}}. \tag{2.4}$$

$\binom{k+d-1}{d-1}$ appears because it is the dimension of the symmetric subspace.

We will need a definition of an approximate unitary design and will use a slightly different form to the approximate state design definition above that is simpler for our purposes:

Definition 2.6. ν is an ϵ-approximate unitary k-design if, for all balanced monomials M of degree $\leq k$,

$$|\mathbb{E}_\nu M(U) - \mathbb{E}_H M(U)| \leq \frac{\epsilon}{d^k}. \tag{2.5}$$

While this definition is different to previous approximate k-design definitions (e.g. Dankert *et al.* 2006; Harrow & Low 2008*b*), it is equivalent up to multiplying ϵ by a polynomial (or inverse polynomial) in dimension. Since any reasonable construction of such a design will use resources that scale with $\log 1/\epsilon$, this leads to only $\log d$ differences in resource requirements between the definitions.

Finally, we will show how to construct an approximate unitary design according to definition 2.6. We would like to be able to have an ϵ-approximate k-design from which we can sample and implement the unitaries using $\mathrm{poly}(\log d, k, \log 1/\epsilon)$ resources. Firstly, Ambainis & Emerson (2007) provide an efficient construction of an ϵ-approximate state k-design for all $k \leq d/2$. For unitary designs, we can use the efficient tensor product expander construction by Harrow & Low (2008*a*). A (d, D, λ, k) tensor product expander (TPE) is an ensemble of D unitaries ν in dimension d with, for all ρ,

$$\left\| \mathbb{E}_\nu \left[U^{\otimes k} \rho (U^\dagger)^{\otimes k} \right] - \mathbb{E}_H \left[U^{\otimes k} \rho (U^\dagger)^{\otimes k} \right] \right\|_2 \leq \lambda \|\rho\|_2, \tag{2.6}$$

where $\lambda < 1$. In Harrow & Low (2008*a*), an efficient construction is presented with D and λ constant for $k = O(\log d / \log\log d)$. In particular, we can obtain an efficient construction for $k = \log_2 d/(10 \log_2 \log_2 d)$. To obtain a design according to definition 2.6, we can iterate the expander:

Lemma 2.7. *Iterating a* (d, D, λ, k)*-TPE* $O(k \log d + \log 1/\epsilon)$ *times gives an* ϵ*-approximate unitary* k*-design.*

The slightly technical proof is presented in appendix A*a*. Using the efficient TPE construction from Harrow & Low (2008*a*), we have an efficient construction of an ϵ-approximate k-design for $k = O(\log d / \log\log d)$.

3. Main technique

The main idea in this paper can be summarized in three steps. Let $f : \mathcal{U}(d) \to \mathbb{C}$ be a balanced polynomial of degree K in the matrix elements of a unitary U. Then to obtain a concentration bound on f when U is chosen from a k-design:

(1) Find some measure concentration result for $|f(U) - \mu|$ when the unitaries are chosen uniformly at random from the Haar measure. Normally, μ will be the expectation of f.
(2) Use this to bound the moments $\mathbb{E}|f(U) - \mu|^{2m}$ for some integer $m \leq k/2K$.
(3) Then use Markov's inequality and the fact that for a (approximate) k-design the moments are (almost) the same as for uniform randomness. We then optimize the bound for m, which will often involve setting m close to the maximum, $\lfloor k/2K \rfloor$.

We will now work through each of these steps and finish with a proof of theorem (1.2).

(*a*) *Step 1: concentration for uniform randomness*

For the first step, we will often start with Levy's Lemma. This states, roughly speaking, that slowly varying functions in high dimensions are approximately constant. We quantify 'slowly varying' by the Lipschitz constant:

Definition 3.1. The Lipschitz constant η (with respect to the Euclidean norm) for a function f is

$$\eta = \sup_{U_1, U_2} \frac{|f(U_1) - f(U_2)|}{||U_1 - U_2||_2}. \tag{3.1}$$

Then we have Levy's Lemma:

Lemma 3.2 (Levy, see Ledoux 2001). *Let f be an η-Lipschitz function on $U(d)$ with mean $\mathbb{E}f$. Then*

$$\mathbb{P}(|f - \mathbb{E}f| \geq \delta) \leq 4 \exp\left(-\frac{C_1 d \delta^2}{\eta^2}\right), \tag{3.2}$$

where C_1 can be taken to be $2/(9\pi^3)$.

(*b*) *Step 2: a bound on the moments*

Levy's Lemma says that f is close to its mean. This means that $\mathbb{E}|f - \mathbb{E}f|^m$ should be small. We will bound the moments for slightly more general concentration results:

Lemma 3.3. *Let X be any random variable with probability concentration*

$$\mathbb{P}(|X - \mu| \geq \delta) \leq C \mathrm{e}^{-a\delta^2}. \tag{3.3}$$

(*Normally μ will be the expectation of X, although the bound does not assume this.*) *Then*

$$\mathbb{E}|X - \mu|^m \leq C \Gamma(m/2 + 1) a^{-m/2} \leq C \left(\frac{m}{2a}\right)^{m/2}, \tag{3.4}$$

for any $m > 0$.

Proof. This proof is based on the proof of an analogous result by Bellare & Rompel (1994), lemma A 1.

Note that, for any random variable $Y \geq 0$,

$$\mathbb{E}\, Y = \int_0^\infty \mathbb{P}(Y \geq y)\, \mathrm{d}y. \tag{3.5}$$

Therefore

$$\begin{aligned} \mathbb{E}|X - \mu|^m &= \int_0^\infty \mathbb{P}(|X - \mu|^m \geq x)\, \mathrm{d}x \\ &= \int_0^\infty \mathbb{P}(|X - \mu| \geq x^{1/m})\, \mathrm{d}x \\ &\leq C \int_0^\infty \exp(-a x^{2/m})\, \mathrm{d}x, \end{aligned}$$

where in the last line we used the assumed large deviation bound equation (3.3). To evaluate this integral, use the change of variables $y = ax^{2/m}$ to get

$$\begin{aligned}\mathbb{E}|X-\mu|^m &\le \frac{Cm}{2}a^{-m/2}\int_0^\infty \mathrm{e}^{-y}y^{m/2-1}\,\mathrm{d}y\\ &= Ca^{-m/2}\Gamma(m/2+1)\\ &\le C\left(\frac{m}{2a}\right)^{m/2}.\end{aligned}$$

■

(*c*) *Step 3: a concentration bound for a k-design*

Now we show how to obtain a measure concentration result for polynomials when the unitaries are selected from an approximate k-design. We first show that the moments of $|f-\mu|$ for f, a polynomial, are close to the Haar measure moments:

Lemma 3.4. *Let f be a balanced polynomial of degree K and μ be any constant. Let $f=\sum_{i=1}^t \alpha_i M_i$ where each M_i is a monomial. Let $\alpha(f)=\sum_i|\alpha_i|$. Then for m an integer with $2mK\le k$ and ν an ϵ-approximate k-design,*

$$\mathbb{E}_\nu|f-\mu|^{2m} \le \mathbb{E}_H|f-\mu|^{2m} + \frac{\epsilon}{d^k}(\alpha+|\mu|)^{2m}. \tag{3.6}$$

Proof. For simplicity, we assume that f and μ are real. Our proof easily generalizes to the complex case.

Firstly we calculate $|\mathbb{E}_\nu f^i - \mathbb{E}_H f^i|$ using the multinomial theorem:

$$\begin{aligned}&|\mathbb{E}_\nu f^i - \mathbb{E}_H f^i|\\ &= \left|\sum_{k_1+\ldots+k_t=i}\binom{i}{k_1,\ldots,k_t}\alpha_1^{k_1}\ldots\alpha_t^{k_t}\left(\mathbb{E}_\nu M_1^{k_1}\ldots M_t^{k_t}-\mathbb{E}_H M_1^{k_1}\ldots M_t^{k_t}\right)\right|\\ &\le \sum_{k_1+\ldots+k_t=i}\binom{i}{k_1,\ldots,k_t}|\alpha_1|^{k_1}\ldots|\alpha_t|^{k_t}\left|\mathbb{E}_\nu M_1^{k_1}\ldots M_t^{k_t}-\mathbb{E}_H M_1^{k_1}\ldots M_t^{k_t}\right|\\ &\le \frac{\epsilon}{d^k}\sum_{k_1+\ldots+k_t=i}\binom{i}{k_1,\ldots,k_t}|\alpha_1|^{k_1}\ldots|\alpha_t|^{k_t}\\ &= \frac{\epsilon}{d^k}\alpha^i.\end{aligned}$$

We now calculate $\mathbb{E}_\nu|f-\mu|^{2m}$:

$$\begin{aligned}|\mathbb{E}_\nu|f-\mu|^{2m} - \mathbb{E}_H|f-\mu|^{2m}| &= |\mathbb{E}_\nu(f-\mu)^{2m} - \mathbb{E}_H(f-\mu)^{2m}|\\ &= \left|\sum_{i=0}^{2m}\binom{2m}{i}(\mathbb{E}_\nu f^i - \mathbb{E}_H f^i)(-\mu)^{2m-i}\right|\\ &\le \sum_{i=0}^{2m}\binom{2m}{i}|\mathbb{E}_\nu f^i - \mathbb{E}_H f^i||\mu|^{2m-i}\end{aligned}$$

$$\leq \frac{\epsilon}{d^k}\sum_{i=0}^{2m}\binom{2m}{i}\alpha^i|\mu|^{2m-i}$$
$$= \frac{\epsilon}{d^k}(\alpha+|\mu|)^{2m}. \qquad \blacksquare$$

Now we can simply apply Markov's inequality to prove theorem 1.2.

Proof of theorem 1.2. Apply Markov's inequality and lemmas 3.3 and 3.4:

$$\mathbb{P}_\nu(|f-\mu|\geq\delta) = \mathbb{P}_\nu(|f-\mu|^{2m}\geq\delta^{2m})$$
$$\leq \frac{\mathbb{E}_\nu|f-\mu|^{2m}}{\delta^{2m}}$$
$$\leq \frac{1}{\delta^{2m}}\left(C\left(\frac{m}{a}\right)^m+\frac{\epsilon}{d^k}(\alpha+|\mu|)^{2m}\right). \qquad \blacksquare$$

We finish this section with two remarks. Firstly, provided $\alpha(f)$ (the sum of the absolute value of all the coefficients) is at most polynomially large in d, we can choose ϵ to be polynomially small to cancel this at no change to the asymptotic efficiency. Secondly, when applying the theorem we will optimize the choice of m (and normally choose $k=2mK$). Often $a=\Theta(d)$ and the optimal choice of m is often $\Theta(d)$ as well. However, we will not take m so large because we can only implement an efficient k-design for $k=O(\log d/\log\log d)$.

4. Application 1: entropy of a k-design

We now apply the above to show that most unitaries in a k-design generate large amounts of entropy across any bipartition, provided the dimensions are sufficiently far apart. This means that, for any initial state, for most choices of a unitary from a k-design applied to the state, the resulting state will be highly entangled. We go via the purity of the reduced density matrix, since the entropy function is not a polynomial.

We will call the two systems S (the 'system') and E (the 'environment') and calculate the purity of the reduced state. That the purity, $\mathrm{tr}[(\mathrm{tr}_E\, U\rho U^\dagger)^2]$, is a balanced polynomial of degree 2 is easily seen by noting that the trace is linear and the reduced state is squared. However, we should check that there are not too many terms or terms with large coefficients. To do this, we should calculate α to apply theorem 1.2.

There is a general method for calculating $\alpha(f)$ which we will use. Write $f(U)=\sum_i \alpha_i M_i(U)$ for monomials M_i. To evaluate $\alpha(f)=\sum_i|\alpha_i|$, calculate $f(A)$ where A is the matrix with all entries equal to 1 (so that $M_i(A)=1$) and replace α_i with $|\alpha_i|$. Using this here we find

$$\alpha = d^2\left(\sum_{ij}|\rho_{ij}|\right)^2$$
$$\leq d^4\sum_{ij}|\rho_{ij}|^2$$

$$= d^4 ||\rho||_2^2$$
$$\leq d^4.$$

We now calculate the expected purity:

Lemma 4.1. *The expected purity of the reduced state is* $(d_S + d_E)/(d+1)$*, where* d_S *is the dimension of subsystem* S *and* $d_E = d/d_S$ *is the dimension of subsystem* E*.*

Proof. We have

$$\mathbb{E}_H ||\psi_S||_2^2 = \mathbb{E}_H [\mathrm{tr}\, \mathcal{F}_{S_1 S_2} (\mathrm{tr}_E\, U\rho U^\dagger \otimes \mathrm{tr}_E\, U\rho U^\dagger)], \tag{4.1}$$

where $\mathcal{F}_{S_1 S_2}$ is the swap acting between systems S_1 and S_2. By linearity of the trace, we can commute the $\mathbb{E}_H$ through and use $\mathbb{E}_H[U\rho U^\dagger \otimes U\rho U^\dagger] = (I_{12} + \mathcal{F}_{12})/d(d+1)$ to find

$$\mathbb{E}_H ||\psi_S||_2^2 = \mathrm{tr}\left[\frac{\mathcal{F}_{S_1 S_2}}{d(d+1)} (d_E^2 I_{S_1 S_2} + d_E \mathcal{F}_{S_1 S_2})\right]$$
$$= \frac{d_S + d_E}{d+1}.$$

■

Working out the higher moments in this way is difficult (although has been done in Giraud 2007), so we use Levy's Lemma and lemma 3.3. To use Levy's Lemma, all we have to do is find the Lipschitz constant for the purity:

Lemma 4.2. *The Lipschitz constant for purity is* ≤ 2*.*

Proof.

$$\eta = \sup_{\psi,\phi} \frac{\left| ||\psi_S||_2^2 - ||\phi_S||_2^2 \right|}{||\psi - \phi||_2}$$
$$= \sup_{\psi,\phi} \frac{|\,||\psi_S||_2 - ||\phi_S||_2|\,(||\psi_S||_2 + ||\phi_S||_2)}{||\psi - \phi||_2}.$$

Now we use $|\,||S||_2 - ||T||_2| \leq ||S - T||_2$ to find

$$\eta \leq \sup_{\psi,\phi} (||\psi_S||_2 + ||\phi_S||_2) \leq 2,$$

using the fact that the purity is upper bounded by 1. ■

Lemma 4.3. *For* $\mu = (d_S + d_E)/(d+1)$ *and* m *an integer with* $m \leq k/4$ *and* ν *an* ϵ*-approximate* k*-design,*

$$\mathbb{P}_\nu (S(\psi_S) \leq -\log_2 \mu - \alpha) \leq \frac{1}{(\mu(2^\alpha - 1))^{2m}} \left(4\left(\frac{4m}{C_1 d}\right)^m + \frac{\epsilon}{d^k}(d^4 + \mu)^{2m}\right). \tag{4.2}$$

Proof. We use the fact that von Neumann entropy is lower bounded by the Renyi 2-entropy, i.e. $-\log_2 ||\psi_S||_2^2$

$$S(\psi_S) \geq S_2(\psi_S) = -\log_2 ||\psi_S||_2^2. \tag{4.3}$$

Then

$$\begin{aligned}\mathbb{P}_\nu(S(\psi_S) \le -\log_2(1+\delta)\mu) &\le \mathbb{P}_\nu(S_2(\psi_S) \le -\log_2(1+\delta)\mu)\\ &= \mathbb{P}_\nu(||\psi_S||_2^2 \ge (1+\delta)\mu)\\ &\le \mathbb{P}_\nu(|||\psi_S||_2^2 - \mu| \ge \delta\mu)\\ &\le \frac{1}{(\mu\delta)^{2m}}\left(4\left(\frac{4m}{C_1 d}\right)^m + \frac{\epsilon}{d^k}(d^4+\mu)^{2m}\right),\end{aligned}$$

using theorem 1.2 in the last line. ■

We have written this in a more convenient form in theorem 1.3, which is proved in the appendix A*b*. This is to be compared with the Haar random version theorem 1.1. As expected, we have $n = \log_2 d$ appearing in the exponent rather than d. Note also that our bound does not work well for $d_S \approx d_E$. In fact, in this case, we do not get a bound that improves with dimension. To get this, d_S must be polynomially smaller than d_E.

5. Application 2: *k*-designs and statistical mechanics

We can also apply these ideas to partially de-randomize some of the arguments on the foundations of statistical mechanics in Popescu *et al.* (2006). In this paper, the authors develop the idea that the uncertainty in statistical mechanics comes from entanglement rather than the traditional assumption of the principle of equal *a priori* probabilities. They consider the universe being in a pure quantum state and that the uncertainty in the state of a subsystem comes from the entanglement between this system and the rest of the universe.

The setting is that there is an arbitrary global linear constraint R. Often this will be a total energy constraint although this is not assumed. Let the Hilbert space of states satisfying R be $\mathcal{H}_R$. Then let the system and environment Hilbert spaces be $\mathcal{H}_S$ and $\mathcal{H}_E$, respectively. Then

$$\mathcal{H}_R \subseteq \mathcal{H}_S \otimes \mathcal{H}_E. \tag{5.1}$$

Let the dimensions be d_R, d_S and d_E and let $\mathcal{E}_R = I_R/d_R$. Note that $d_R \le d_S d_E$, unlike in the above where we took $d = d_S d_E$. Normally, we will have $d_S \ll d_R$. The principle of equal *a priori* probabilities says that the state of the universe is $\mathcal{E}_R$, which implies that the subsystem state is the canonical state, given by

$$\Omega_S = \mathrm{tr}_E(\mathcal{E}_R). \tag{5.2}$$

The main result of Popescu *et al.* (2006) (the 'principle of *apparently* equal *a priori* probabilities') is that, for almost all pure states of the universe, the subsystem state is almost exactly the canonical state.

Theorem 5.1 (Popescu *et al.* (2006), theorem 1). *For a randomly chosen state* $|\phi\rangle \in \mathcal{H}_R \subseteq \mathcal{H}_S \otimes \mathcal{H}_E$ *and arbitrary* $\epsilon > 0$, *the distance between the reduced density matrix of the system* $\rho_S = \mathrm{tr}_E(|\phi\rangle\langle\phi|)$ *and the canonical state* Ω_S *(equation* (5.2)*)*

is given probabilistically by

$$\mathbb{P}_H\left(||\rho_S - \Omega_S||_1 \geq \epsilon + \sqrt{\frac{d_S}{d_E^{\text{eff}}}}\right) \leq 2\exp(-C_2 d_R \epsilon^2), \tag{5.3}$$

where $C_2 = 1/(18\pi^3)$ *and* $d_E^{\text{eff}} = 1/\operatorname{tr}\Omega_E^2 \geq d_R/d_S$.

This result gives compelling evidence to replace the principle of equal *a priori* probabilities with the principle of apparently equal *a priori* probabilities, but it does not address the problem of how the system reaches this state. It will take an extremely (exponentially) long time for the universe to reach a randomly pure state, in contrast to the observed fact that thermalization occurs quickly. Here, we show that for almost all unitaries in a k-design applied to the universe, the subsystem state is close to the canonical state. Since these unitaries can be implemented and sampled efficiently, this means that equilibrium could be reached quickly to match observations.

We are now ready to show that a k-design gives a small $||\rho_S - \Omega_S||_1$. Firstly, we have to modify lemma 3.3 slightly.

Lemma 5.2. *Let X be any non-negative random variable with probability concentration*

$$\mathbb{P}(X \geq \delta + \eta) \leq C\mathrm{e}^{-a\delta^2}, \tag{5.4}$$

where $\eta \geq 0$. Then

$$\mathbb{E}X^m \leq C\left(\frac{2m}{a}\right)^{m/2} + (2\eta)^m, \tag{5.5}$$

for any $m > 0$.

The proof is very similar to the proof of lemma 3.3.

Now we state and prove the main result in this section:

Theorem 5.3. *Let ν be an ϵ-approximate unitary k-design. Then*

$$\mathbb{P}_\nu(||\rho_S - \Omega_S||_1 \geq \delta) \leq \left(\frac{d_S}{\delta^2}\right)^{k/8}\left(2\left(\frac{k}{2C_2 d_R}\right)^{k/8} + \left(\frac{4d_S^2}{d_R}\right)^{k/8} + \frac{\epsilon}{d_R^k}(d_R^2+1)^{k/2}\right). \tag{5.6}$$

In particular, with $\epsilon = (3/2)(4d_S^3/d_R)^{k/8}$, $\quad k \leq 8C_2 d_S^2$,

$$\mathbb{P}_\nu(||\rho_S - \Omega_S||_1 \geq \delta) \leq 6\left(\frac{4d_S^3}{d_R\delta^2}\right)^{k/8}. \tag{5.7}$$

Again, we need d_S to be polynomially smaller than d_R to obtain non-trivial bounds.

Proof. We go via the 2-norm and use lemmas 5.2 and 3.4.

We have from theorem 5.1 that

$$\mathbb{P}_H(||\rho_S - \Omega_S||_1 \geq \delta + \eta) \leq 2\mathrm{e}^{-C_2 d_R \delta^2}, \tag{5.8}$$

where $\eta = \sqrt{d_S/d_E^{\mathrm{eff}}} \leq d_S/\sqrt{d_R}$. Since $||\rho_S - \Omega_S||_2 \leq ||\rho_S - \Omega_S||_1$,

$$\mathbb{P}_H(||\rho_S - \Omega_S||_2 \geq \delta + \eta) \leq 2\mathrm{e}^{-C_2 d_R \delta^2}. \tag{5.9}$$

We now apply lemma 5.2 to get

$$\mathbb{E}_H||\rho_S - \Omega_S||_2^{2m} \leq 2\left(\frac{4m}{C_2 d_R}\right)^m + (2\eta)^{2m}. \tag{5.10}$$

So for $m \leq k/4$, using Markov's inequality and lemma 3.4 (with $\mu = 0$) on the polynomial $||\rho_S - \Omega_S||_2^2$

$$\mathbb{P}_v(||\rho_S - \Omega_S||_2 \geq \delta) \leq \frac{1}{\delta^{2m}}\left(2\left(\frac{4m}{C_2 d_R}\right)^m + (2\eta)^{2m} + \frac{\epsilon}{d_R^k}(d_R^2+1)^{4m}\right). \tag{5.11}$$

Here, we used an estimate of α, the sum of the moduli of the coefficients

$$\alpha \leq (d_R^2+1)^2, \tag{5.12}$$

which we obtain via a calculation similar to that in §4.

Now we go back to the 1-norm, using $||\rho_S - \Omega_S||_1 \leq \sqrt{d_S}||\rho_S - \Omega_S||_2$ to get

$$\begin{aligned}\mathbb{P}_v(||\rho_S - \Omega_S||_1 \geq \delta) &\leq \mathbb{P}_v(||\rho_S - \Omega_S||_2 \geq \delta/\sqrt{d_S}) \\ &\leq \left(\frac{d_S}{\delta^2}\right)^m \left(2\left(\frac{4m}{C_2 d_R}\right)^m + (2\eta)^{2m} + \frac{\epsilon}{d_R^k}(d_R^2+1)^{4m}\right).\end{aligned} \tag{5.13}$$

To obtain the result in equation (5.6), we just use $\eta \leq d_S/\sqrt{d_R}$ and set $m = k/8$.

To prove the simplified version, first use, as in §4, that $(d_R^2+1)^{4m} \leq 2d_R^{8m}$ for $m \leq d_R^2/8$. This is implied by $k \leq 8C_2 d_S^2$. We then set $m = k/8$ to find

$$\mathbb{P}_v(||\rho_S - \Omega_S||_1 \geq \delta) \leq 2\left(\frac{kd_S}{2C_2 d_R \delta^2}\right)^{k/8} + \left(\frac{4d_S^3}{d_R\delta^2}\right)^{k/8} + 2\frac{\epsilon}{\delta^{k/4}}. \tag{5.14}$$

Then, using $k \leq 8C_2 d_S^2$, with $\epsilon \leq (3/2)(4d_S^3/d_R)^{k/8}$, we obtain the simplified result equation (5.7). ■

6. Application 3: using *k*-designs for measurement-based quantum computing

Here we apply our ideas to partially de-randomize some results of Gross *et al.* (2008) and Bremner *et al.* (2008). The main result in these two papers is that most states do not offer any advantage over classical computation when used in the MBQC model. In MBQC, a classical computer is given access to a large quantum state on which it can do single qubit measurements. Some states allow for universal quantum computation, whereas others do not add any extra power to the classical computer. These results are concerned with the question of

characterizing which states do and do not work. Showing that random states do not give any speed up shows that useful states for MBQC are not generic and so must be carefully constructed.

While the results in these two papers are similar, we will concentrate on the methods from Gross *et al.* (2008) since their methods are simpler to apply here. They prove their result by showing that most states are very entangled in the geometric measure (see definition 6.1). They then use this to show that the measurement outcomes of even the best possible measurement scheme are almost completely random. In fact, the state could be thrown away and the measurement outcomes replaced with random numbers to solve the computational problem just as efficiently. This shows that you can classically simulate any quantum computation that uses these highly entangled states. The measure of entanglement they use is the geometric measure:

Definition 6.1. The geometric measure of entanglement of a state $|\Psi\rangle$ is (Shimony 1995; Barnum & Linden 2001)

$$E_g(|\Psi\rangle) = -\log_2 \sup_{\alpha\in\mathcal{P}} |\langle\alpha|\Psi\rangle|^2, \tag{6.1}$$

where $\mathcal{P}$ is the set of all product states.

They show that any MBQC using a state $|\Psi\rangle$ with $E_g(|\Psi\rangle) = n - O(\log_2 n)$ can be efficiently simulated classically. They then show that (we abuse notation slightly by writing $\mathbb{P}_H$ for $\mathbb{P}_{|\Psi\rangle\sim\nu_H}$, etc.)

Theorem 6.2 (Gross *et al.* (2008), theorem 2). *For $n \geq 11$,*

$$\mathbb{P}_H(E_g(|\Psi\rangle) \leq n - 2\log_2 n - 3) \leq \mathrm{e}^{-n^2}. \tag{6.2}$$

This shows that most states are useless. We partially de-randomize this result to show that most states in an ϵ-approximate (ϵ can be taken as a constant) state n^2-design have high geometric measure of entanglement and thus are useless in the same way.

We could apply our technique and use theorem 1.2, but in this case it is simpler to directly bound the probability using Markov's inequality.

Lemma 6.3.

$$\mathbb{P}_\nu(|\langle\Phi|\Psi\rangle|^2 \geq \delta) \leq (1+\epsilon)\frac{m!}{(d\delta)^m} \leq (1+\epsilon)\left(\frac{m}{d\delta}\right)^m, \tag{6.3}$$

where $|\Psi\rangle$ is chosen from an ϵ-approximate state k-design ν, $m \leq k$ and a positive integer and $|\Phi\rangle$ is any fixed state.

Proof. We prove this bound directly using Markov's inequality

$$\begin{aligned}\mathbb{P}_\nu(|\langle\Phi|\Psi\rangle|^2 \geq \delta) &= \mathbb{P}_\nu(|\langle\Phi|\Psi\rangle|^{2m} \geq \delta^m)\\ &\leq \frac{\mathbb{E}_\nu|\langle\Phi|\Psi\rangle|^{2m}}{\delta^m}\end{aligned}$$

$$= \frac{\langle \Phi|^{\otimes m} \mathbb{E}_\nu \left[|\Psi\rangle^{\otimes m} \langle \Psi|^{\otimes m} \right] |\Phi\rangle^{\otimes m}}{\delta^m}$$
$$\leq \frac{\langle \Phi|^{\otimes m} (1+\epsilon) \Pi_m^{\mathrm{sym}} / \binom{m+d-1}{d-1} |\Phi\rangle^{\otimes m}}{\delta^m}$$
$$= \frac{1+\epsilon}{\binom{m+d-1}{d-1}\delta^m} \leq (1+\epsilon) \left(\frac{m}{d\delta}\right)^m. \qquad \blacksquare$$

We now prove the main result in this section:

Theorem 6.4. *For $|\Psi\rangle$ randomly drawn from an ϵ-approximate state k-design with $d = 2^n$*

$$\mathbb{P}_\nu(E_g(|\Psi\rangle) \leq n - \delta) \leq (1+\epsilon) \exp_2(k \log_2 2k + 4n \log_2 10n - k\delta + 4n(n-\delta)). \tag{6.4}$$

In particular, for $k = n^2$, $\delta = 3\log_2 n + 5$ and $\epsilon = 1$,

$$\mathbb{P}_\nu(E_g(|\Psi\rangle) \leq n - 3\log_2 n - 5) \leq 2 \cdot n^{-n^2}. \tag{6.5}$$

We note that this bound is almost the same as in theorem 6.2. It only works for slightly larger deviations from n, which is why we obtain a slightly better probability bound. Note also that we can obtain an exponential bound in n (not $d = 2^n$) because the design is exponentially large in n.

Proof. This proof closely mirrors the proof of theorem 2 in Gross *et al.* (2008). We use the idea of a γ-net. $\mathcal{N}_{\gamma,n}$ is a γ-net on product states if

$$\sup_{|\alpha\rangle \in \mathcal{P}} \inf_{|\tilde{\alpha}\rangle \in \mathcal{N}_{\delta,n}} \left|\left| |\alpha\rangle - |\tilde{\alpha}\rangle \right|\right|_2 \leq \frac{\gamma}{2}. \tag{6.6}$$

In Gross *et al.* (2008), it is shown that such a net exists with $|\mathcal{N}_{\gamma,n}| \leq (5n/\gamma)^{4n}$. We then proceed by showing that most states in the state design have small overlap with every state in the net using the union bound and lemma 6.3. Finally, since every state is close to one in the net, we can show that most states in the design have small overlap with every product state.

We now formalize the above. Using lemma 6.3 and the union bound,

$$\mathbb{P}_\nu \left(\sup_{|\tilde{\alpha}\rangle \in \mathcal{N}_{\gamma,n}} |\langle \tilde{\alpha}|\Psi\rangle|^2 \geq \delta'/2 \right) \leq |\mathcal{N}_{\gamma,n}| (1+\epsilon) \left(\frac{2k}{d\delta'}\right)^k \leq \left(\frac{5n}{\gamma}\right)^{4n} (1+\epsilon) \left(\frac{2k}{2^n\delta'}\right)^k. \tag{6.7}$$

Now, we need to bound

$$\mathbb{P}_\nu(E_g(|\Psi\rangle) \leq n - \delta) = \mathbb{P}_\nu \left(-\log_2 \sup_{|\alpha\rangle \in \mathcal{P}} |\langle \alpha|\Psi\rangle|^2 \leq n - \delta \right)$$
$$= \mathbb{P}_\nu \left(\sup_{|\alpha\rangle \in \mathcal{P}} |\langle \alpha|\Psi\rangle|^2 \geq 2^{-(n-\delta)} \right).$$

We now claim that

$$\sup_{|\alpha\rangle\in\mathcal{P}} |\langle\alpha|\Psi\rangle|^2 \geq \delta' \Rightarrow \sup_{|\tilde{\alpha}\rangle\in\mathcal{N}_{\delta'/2,n}} |\langle\tilde{\alpha}|\Psi\rangle|^2 \geq \frac{\delta'}{2}. \tag{6.8}$$

To prove this claim, let $|\alpha\rangle$ be the state that achieves the supremum on the left hand side, and let $|\tilde{\alpha}\rangle$ be the state closest to it in the $\delta'/2$-net. It is shown in Gross *et al.* (2008) that this implies for any $|\Psi\rangle$

$$\left||\langle\alpha|\Psi\rangle|^2 - |\langle\tilde{\alpha}|\Psi\rangle|^2\right| \leq \frac{\delta'}{2}. \tag{6.9}$$

Therefore

$$|\langle\tilde{\alpha}|\Psi\rangle|^2 \geq |\langle\alpha|\Psi\rangle|^2 - \frac{\delta'}{2} \geq \frac{\delta'}{2}. \tag{6.10}$$

This implies that the supremum over all states in the net must be at least $\delta'/2$ to prove the claim.

We can now finish the proof. Set $\delta' = 2^{-(n-\delta)}$ in equation (6.8) and use equation (6.7) with $\gamma = \delta'/2$ to find

$$\begin{aligned}
&\mathbb{P}_\nu\left(\sup_{|\alpha\rangle\in\mathcal{P}} |\langle\alpha|\Psi\rangle|^2 \geq 2^{-(n-\delta)}\right) \\
&\quad\leq \mathbb{P}_\nu\left(\sup_{|\tilde{\alpha}\rangle\in\mathcal{N}_{2^{-(n-\delta)-1},n}} |\langle\tilde{\alpha}|\Psi\rangle|^2 \geq 2^{-(n-\delta)-1}\right) \\
&\quad\leq (1+\epsilon)\exp_2(k\log_2 2k + 4n\log_2 10n - k\delta + 4n(n-\delta)).
\end{aligned}$$

■

Combining this with the arguments of Gross *et al.* (2008) shows that most states in a state n^2-design on n qubits are useless for MBQC. This shows that even many efficiently preparable states are useless.

7. Conclusions

We have seen how to turn large deviation bounds for Haar-random unitaries into bounds for k-designs. The main technique was applied to show that unitaries from k-designs generate large amounts of entanglement. Then we showed that, if the dynamics of the universe produced a k-design, the entanglement generated would be sufficient to reproduce the principle of equal *a priori* probabilities. Finally, we showed that most states in sufficiently large state designs are useless for MBQC, in the sense that computation using them can be efficiently simulated classically.

However, there are other bounds for which our technique does not work. Since we cannot obtain exponential bounds for polynomially sized designs, our technique cannot directly de-randomize some bounds. Some results, for example showing that the ∞-norm of the reduced state of a random pure state is close to $1/d_S$ (Harrow *et al.* 2004), are proven by using an ϵ-net of states and the union bound. Since the ϵ-net is exponentially large, exponentially small bounds

are required. We do not know how to apply our idea to results of this kind and still have $k = \mathrm{poly}(\log d)$. (Note that we could cope with the ϵ-net in §6 since it was just a net on product states which is considerably smaller.)

It is also possible that our ideas could be used to completely de-randomize some constructions, for example locking (DiVincenzo *et al.* 2004; Hayden *et al.* 2004). If we could show that unitaries drawn from a k-design work with non-zero probability, and come up with an efficient sampling method, then we could obtain efficient randomized constructions.

I am grateful for funding from the UK Engineering and Physical Science Research Council through 'QIP IRC.' I thank Aram Harrow for many useful discussions on this topic, comments on earlier drafts of this manuscript and for suggesting the use of existing large deviation bounds to bound the high moments. I also thank Toby Cubitt for suggesting applying this method to the results of Popescu *et al.* (2006), Ashley Montanaro for useful discussions and comments on drafts of this manuscript as well as Andreas Winter and Michael Bremner for useful discussions and comments.

Appendix A

Here we present some miscellaneous proofs.

(a) Proof of lemma 2.7

Proof of lemma 2.7. We claim that, if for all $d^k \times d^k$ matrices ρ,

$$\|\mathbb{E}_\sigma[U^{\otimes k}\rho(U^\dagger)^{\otimes k}] - \mathbb{E}_H[U^{\otimes k}\rho(U^\dagger)^{\otimes k}]\|_2 \leq \frac{\epsilon}{d^{3k/2}}\|\rho\|_2, \tag{A 1}$$

then σ is an ϵ-approximate k-design. To prove this claim, let $m \leq k$ and take M to be any balanced monomial of degree m. Write $M = U_{p_1 q_1} \dots U_{p_m q_m} U^*_{r_1 s_1} \dots U^*_{r_m s_m}$. Then let $\rho_m = |q_1, \dots, q_m\rangle\langle s_1, \dots, s_m|$. Let $\mathcal{E}_{\sigma,k}(\rho) = \mathbb{E}_\sigma[U^{\otimes k}\rho(U^\dagger)^{\otimes k}]$, $\mathcal{E}_{H,k}(\rho) = \mathbb{E}_H[U^{\otimes k}\rho(U^\dagger)^{\otimes k}]$ and $\rho_k = \rho_m \otimes I^{\otimes k-m}/d^{k-m}$. Then

$$\begin{aligned}\frac{\epsilon}{d^{3k/2}} &\geq ||\mathcal{E}_{\sigma,k}(\rho_k) - \mathcal{E}_{H,k}(\rho_k)||_2 \\ &= \left|\left|\left(\mathcal{E}_{\sigma,m}(\rho_m) - \mathcal{E}_{H,m}(\rho_m)\right) \otimes \frac{I^{\otimes k-m}}{d^{k-m}}\right|\right|_2 \\ &= \frac{1}{\sqrt{d^{k-m}}}||\mathcal{E}_{\sigma,m}(\rho_m) - \mathcal{E}_{H,m}(\rho_m)||_2.\end{aligned}$$

We then use the fact that the largest matrix element is upper bounded by the 2-norm. For any matrix A,

$$|A_{ij}| \leq \sqrt{\sum_{i'j'} |A_{i'j'}|^2} = \sqrt{\mathrm{tr}\, A^\dagger A} = ||A||_2.$$

For us, this implies

$$|(\mathcal{E}_{\sigma,m}(\rho_m) - \mathcal{E}_{H,m}(\rho_m))_{p_1 \dots p_m, r_1 \dots r_m}| \leq ||\mathcal{E}_{\sigma,m}(\rho_m) - \mathcal{E}_{H,m}(\rho_m)||_2, \tag{A 2}$$

which gives

$$|\mathbb{E}_\nu M - \mathbb{E}_H M| \leq \frac{\epsilon}{d^k}, \tag{A 3}$$

to prove the claim.

Then we just have to show how to obtain equation (A 1) from equation (2.6). Iterating the TPE t times gives

$$\|\mathbb{E}_{\nu^t}[U^{\otimes k}\rho(U^\dagger)^{\otimes k}] - \mathbb{E}_H[U^{\otimes k}\rho(U^\dagger)^{\otimes k}]\|_2 \leq \lambda^t, \tag{A 4}$$

where ν^t is the ensemble obtained by applying t unitaries from ν. Now choose t such that $\lambda^t \leq \epsilon/d^{3k/2}$. ■

(*b*) *Proof of theorem 1.3*

Here we prove the more convenient form of lemma 4.3 stated as theorem 1.3.

Proof of theorem 1.3. Firstly, we will write the left hand side of equation (4.2) in a more useful way. Using $\ln(1+x) \leq x$, we find

$$-\log_2 \mu \geq \log_2 d_S - \beta,$$

where $\beta = (1/\ln 2)(d_S/d_E)$, following the notation in Hayden *et al.* (2006). This means

$$\begin{aligned}\mathbb{P}_\nu(S(\psi_S) \leq \log_2 d_S - \alpha - \beta) &\leq \mathbb{P}_\nu(S(\psi_S) \leq -\log_2 \mu - \alpha) \\ &\leq \frac{1}{(\mu(2^\alpha - 1))^{2m}}\left(4\left(\frac{4m}{C_1 d}\right)^m + \frac{\epsilon}{d^k}(d^4+\mu)^{2m}\right).\end{aligned}$$

We now simplify the right hand side. Let $\delta = 2^\alpha - 1$. For $d_S \geq 2$, we have $\mu \geq 1/d_S$. We shall also assume that $m = k/8$. This gives us (using $\mu \leq 1$)

$$\mathbb{P}_\nu(S(\psi_S) \leq \log_2 d_S - \alpha - \beta) \leq \left(\frac{d_S}{\delta}\right)^{k/4}\left(4\left(\frac{k}{2C_1 d}\right)^{k/8} + \epsilon\left(1+\frac{1}{d^4}\right)^{k/4}\right). \tag{A 5}$$

Now, one can easily show (e.g. by induction on n) that

$$(1+\delta)^n \leq 2, \tag{A 6}$$

for $2n\delta \leq 1$. We use this for $n = k/4$ and $\delta = 1/d^4$. The condition is then $k \leq 2d^4$, which we shall assume (we will set $k = \log d/\log\log d$ later). We now obtain

$$\mathbb{P}_\nu(S(\psi_S) \leq \log_2 d_S - \alpha - \beta) \leq \left(\frac{d_S}{\delta}\right)^{k/4}\left(4\left(\frac{k}{2C_1 d}\right)^{k/8} + 2\epsilon\right). \tag{A 7}$$

We will now take $\epsilon = 2(k/(2C_1 d))^{k/8}$, so that the two terms are the same. $\log 1/\epsilon$ is poly $\log d$, so this remains efficient. Now

$$\mathbb{P}_\nu(S(\psi_S) \leq \log_2 d_S - \alpha - \beta) \leq 8\left(\frac{d_S^2 k}{2C_1 d\delta^2}\right)^{k/8}. \tag{A 8}$$

Assuming that $\delta^2 > kd_S^2/(2C_1 d)$, we should take k as large as possible up to $2C_1\delta^2 d/(\mathrm{e}d_S^2)$, when the right hand side is maximized. We then find the result after further simplification. ■

References

Alon, N. & Nussboim, A. 2008 k-wise independent random graphs. *49th Annual IEEE Symp. Foundations of Computer Science, Philadelphia, PA, 25–28 October 2008*, pp. 813–822. (doi:10.1109/FOCS.2008.61)

Ambainis, A. & Emerson, E. 2007 Quantum t-designs: t-wise independence in the quantum world. *IEEE Conf. Comput. Complex.* **2007**, 129–140. (doi:10.1109/CCC.2007.26)

Aubrun, G. 2008 On almost randomizing channels with a short Kraus decomposition. (http://arxiv.org/abs/0805.2900).

Barnum, H. & Linden, N. 2001 Monotones and invariants for multi-particle quantum states. *J. Phys. A* **34**, 6787–6805. (doi:10.1088/0305-4470/34/35/305)

Bellare, M. & Rompel, J. 1994 Randomness-efficient oblivious sampling. *35th Annual IEEE Symp. Foundations of Computer Science, Santa Fe, NM, 20–22 November 1994*, pp. 276–287. (doi:10.1109/SFCS.1994.365687)

Bremner, M. J., Mora, C. & Winter, A. 2008 Are random pure states useful for quantum computation? *Phys. Rev. Lett.* **102**, 190 502. (doi:10.1103/PhysRevLett.102.190502)

Dahlsten, O. & Plenio, M. 2006 Entanglement probability distribution of bipartite randomised stabilizer states. *Quant. Inf. Comp.* **6**, 527–538.

Dankert, C., Cleve, R., Emerson, J. & Livine, E. 2006 Exact and approximate unitary 2-designs: constructions and applications. (http://arxiv.org/abs/quant-ph/0606161).

DiVincenzo, D. P., Horodecki, M., Leung, D. W., Smolin, J. A. & Terhal, B. M. 2004 Locking classical correlations in quantum states. *Phys. Rev. Lett.* **92**, 067 902. (doi:10.1103/PhysRevLett.92.067902)

Foong, S. K. & Kanno, S. 1994 Proof of Page's conjecture on the average entropy of a subsystem. *Phys. Rev. Lett.* **42** 1148–1151. (doi:10.1103/PhysRevLett.72.1148)

Giraud, O. 2007 Distribution of bipartite entanglement for random pure states. *J. Phys. A.* **40**, 2793–2801. (doi:10.1088/1751-8113/40/11/014)

Gross, D., Flammia, S. & Eisert, J. 2008 Most quantum states are too entangled to be useful as computational resources. *Phys. Rev. Lett.* **102**, 190 501. (doi:10.1103/PhysRevLett.102.190501)

Harrow, A. W. & Low, R. A. 2008*a* Efficient quantum tensor product expanders and k-designs. (http://arxiv.org/abs/0811.2597).

Harrow, A. W. & Low, R. A. 2008*b* Random quantum circuits are approximate 2-designs. (http://arxiv.org/abs/0802.1919).

Harrow, A., Hayden, P. & Leung, D. 2004 Superdense coding of quantum states. *Phys. Rev. Lett.* **92**, 187 901. (doi:10.1103/PhysRevLett.92.187901)

Hayden, P., Leung, D., Shor, P. W. & Winter, A. 2004 Randomizing quantum states: constructions and applications. *Commun. Math. Phys.* **250**, 371–391. (doi:10.1007/s00220-004-1087-6)

Hayden, P., Leung, D. W., & Winter, A. 2006 Aspects of generic entanglement. *Commun. Math. Phys.* **265**, 95–117. (doi:10.1007/s00220-006-1535-6)

Ledoux, M. 2001 *The concentration of measure phenomenon.* RI, USA: American Mathematical Society.

Page, D. N. 1993 Average entropy of a subsystem. *Phys. Rev. Lett.* **71**, 1291. (doi:10.1103/PhysRevLett.71.1291)

Popescu, S., Short, A. J. & Winter, A. 2006 Entanglement and the foundations of statistical mechanics. *Nat. Phys.* **2**, 754–758. (doi:10.1038/nphys444)

Sanchez-Ruiz, J. 1995 Simple proof of Page's conjecture on the average entropy of a subsystem. *Phys. Rev. E* **52**, 5653–5655 (doi:10.1103/PhysRevE.52.5653)

Shimony, A. 1995 Degree of entanglement. *Ann. NY Acad. Sci.* **755**, 675. (doi:10.1111/j.1749-6632.1995.tb39008.x)

Smith, G. & Leung, D. 2006 Typical entanglement of stabilizer states. *Phys. Rev. A* **74**, 062 314. (doi:10.1103/PhysRevA.74.062314)